Université Polytechnique HAUTS-DE-FRANCE

CC BY NC SA

Thèse de doctorat

Pour obtenir le grade de Docteur de

L'UNIVERSITÉ POLYTECHNIQUE HAUTS-DE-FRANCE

Discipline : Electronique

Présentée et soutenue par : Tarek Saad Omar ESEHOLI

Le 17/12/2018, à Valenciennes

Ecole doctorale : Sciences Pour l'Ingénieur (ED SPI 072)
Equipe de recherche, Laboratoire : Institut d'Electronique de Microélectronique et de Nanotechnologie - Département Opto-Acousto-Electronique (IEMN DOAE – UMR 8520)

Optimisation des techniques de compression d'images fixes et de vidéo en vue de la caractérisation des matériaux : Applications à la mécanique

---

JURY

Président de jury ;

M. Yannis POUSSET, Professeur des Universités, Laboratoire XLIM-SIC UMR 7252, Poitiers.

Rapporteurs :

M. Pierre Emmanuel MAZERAN, Maître de Conférences HDR, UTC, Laboratoire ROBERVAL UMR 7337, Compiègne.

Examinateurs :

Mme Anne-Sophie DESCAMPS, Maître de Conférences, Laboratoire IETR UMR 6164, Nantes.

M. Maxence BIGERELLE, Professeur des Universités, UPHF, Laboratoire LAMIH UMR 8201, Valenciennes.

Directeur de Thèse :

M. Patrick CORLAY, Professeur des Universités, UPHF, Laboratoire IEMN DOAE UMR 8520, Valenciennes.

Co-Directeur de Thèse :

M. François-Xavier COUDOUX, Professeur des Universités, UPHF, Laboratoire IEMN DOAE UMR 8520, Valenciennes.

Membres Invités :

Mme Delphine NOTTA, Maître de Conférences, UPHF, Laboratoire LAMIH UMR 8201, Valenciennes.

Acknowledgements

"Say: Are those equal, those who know and those who do not know?"

The Noble Quran [39 :9].

I would like to dedicate this work to spirit of my pure mother.

It is a pleasure to thank my supervisors Prof. Patrick CORLAY and Prof. François-Xavier COUDOUX for their guidance and encouragement.

I would like to thank my PhD. thesis jury members: my reviewers Prof. Yannis POUSSET, Prof. Pierre Emmanuel MAZERAN, as well as my examinators Mme Anne-Sophie DESCAMPS, Mme Delphine NOTTA, and Prof. Maxence BIGERELLE, for their time to review this manuscript and for their valuable remarks and comments which rich my work.

My father, thanks for your support and unconditional love. You can take all the credit for much of what I have achieved and what I will achieve in the future.

I would also like to thank Libyan Research, Science and Technology Authority for their financial support.

I would like to thank my colleagues and brothers in IEMN and LAMIH laboratories for their help and love.

I would like to thank Prof. Maxence BIGERELLE and Mme Delphine NOTTA for their time and patience to explain me the mechanical engineering concepts and techniques used in this research.

This thesis would have never been possible without my family Reima, Bailasan and Mohammed as well as my brothers, Sokina, Hesham, Salah and Bothina and their loving kids.

To all of you, I shall be forever indebted

Tarek Saad Omar Eseholi

January 2019

Résumé

Cette thèse porte sur l'optimisation des techniques de compression d'images fixes et de vidéos en vue de la caractérisation des matériaux pour des applications dans le domaine de la mécanique, et s'inscrit dans le cadre du projet de recherche MEgABIt (MEchAnic Big Images Technology) soutenu par l'Université Polytechnique Hauts-de-France. L'objectif scientifique du projet MEgABIt est d'investiguer dans l'aptitude à compresser de gros volumes de flux de données issues d'instrumentation mécanique de déformations à grands volumes tant spatiaux que fréquentiels. Nous proposons de concevoir des algorithmes originaux de traitement dans l'espace compressé afin de rendre possible au niveau calculatoire l'évaluation des paramètres mécaniques, tout en préservant le maximum d'informations fournis par les systèmes d'acquisitions (imagerie à grande vitesse, tomographie 3D). La compression pertinente de la mesure de déformation des matériaux en haute définition et en grande dynamique doit permettre le calcul optimal de paramètres morpho-mécaniques sans entraîner la perte des caractéristiques essentielles du contenu des images de surface mécaniques, ce qui pourrait conduire à une analyse ou une classification erronée. Dans cette thèse, nous utilisons le standard HEVC (High Efficiency Video Coding) à la pointe des technologies de compression actuelles avant l'analyse, la classification ou le traitement permettant l'évaluation des paramètres mécaniques. Nous avons tout d'abord quantifié l'impact de la compression des séquences vidéos issues d'une caméra ultra-rapide. Les résultats expérimentaux obtenus ont montré que des taux de compression allant jusque 100 :1 pouvaient être appliqués sans dégradation significative de la réponse mécanique de surface du matériau mesurée par l'outil d'analyse VIC-2D. Finalement, nous avons développé une méthode de classification originale dans le domaine compressé d'une base d'images de topographie de surface. Le descripteur d'image topographique est obtenu à partir des modes de prédiction calculés par la prédiction intra-image appliquée lors de la compression sans pertes HEVC des images. La machine à vecteurs de support (SVM) a également été introduite pour renforcer les performances du système proposé. Les

résultats expérimentaux montrent que le classificateur dans le domaine compressé est robuste pour la classification de nos six catégories de topographies mécaniques différentes basées sur des méthodologies d'analyse simples ou multi-échelles, pour des taux de compression sans perte obtenus allant jusque 6: 1 en fonction de la complexité de l'image. Nous avons également évalué les effets des types de filtrage de surface (filtres passe-haut, passe-bas et passe-bande) et de l'échelle d'analyse sur l'efficacité du classifieur proposé. La grande échelle des composantes haute fréquence du profil de surface est la mieux appropriée pour classer notre base d'images topographiques avec une précision atteignant 96%.

Mots-clés : Big Data – mécanique - science des matériaux - compression et analyse des données - traitement de l'information - codage vidéo à haute efficacité (HEVC) - machine à vecteurs de support (SVM).

Abstract

This PhD. thesis focuses on the optimization of fixed image and video compression techniques for the characterization of materials in mechanical science applications, and it constitutes a part of MEgABIt (MEchAnic Big Images Technology) research project supported by the Polytechnic University Hauts-de-France (UPHF). The scientific objective of the MEgABIt project is to investigate the ability to compress large volumes of data flows from mechanical instrumentation of deformations with large volumes both in the spatial and frequency domain. We propose to design original processing algorithms for data processing in the compressed domain in order to make possible at the computational level the evaluation of the mechanical parameters, while preserving the maximum of information provided by the acquisitions systems (high-speed imaging, tomography 3D). In order to be relevant image compression should allow the optimal computation of morpho-mechanical parameters without causing the loss of the essential characteristics of the contents of the mechanical surface images, which could lead to wrong analysis or classification. In this thesis, we use the state-of-the-art HEVC standard prior to image analysis, classification or storage processing in order to make the evaluation of the mechanical parameters possible at the computational level. We first quantify the impact of compression of video sequences from a high-speed camera. The experimental results obtained show that compression ratios up to 100: 1 could be applied without significant degradation of the mechanical surface response of the material measured by the VIC-2D analysis tool. Then, we develop an original classification method in the compressed domain of a surface topography database. The topographical image descriptor is obtained from the prediction modes calculated by intra-image prediction applied during the lossless HEVC compression of the images. The Support vector machine (SVM) is also introduced for strengthening the performance of the proposed system. Experimental results show that the compressed-domain topographies classifier is robust for classifying the six different mechanical topographies either based

on single or multi-scale analyzing methodologies. The achieved lossless compression ratios up to 6:1 depend on image complexity. We evaluate the effects of surface filtering types (high-pass, low-pass, and band-pass filter) and the scale of analysis on the efficiency of the proposed compressed-domain classifier. We verify that the high analysis scale of high-frequency components of the surface profile is more appropriate for classifying our surface topographies with accuracy of 96 %.

Keywords: Big Data - Mechanics, Materials Sience - Data Compression and analysis - High Efficiency Video Coding (HEVC) - Support Vector Machine (SVM).

TABLE OF CONTENTS

LIST OF FIGURES ……..XII

LIST OF TABLES……XVI

CHAPTER 1 INTRODUCTION 1

1.1 Context and Motivation 1

1.2 Challenges 3

1.3 Contributions 4

1.4 Structure of the Manuscript 5

CHAPTER 2 STATE-OF-THE-ART 7

2.1 Digital Images and Video Compression 7

2.1.1 Basics of Image Compression Techniques 8

2.1.2 Illustrative Example of the JPEG Still Image Compression Standard 14

2.1.3 Motion Compensation and Video Compression 17

2.2 Material Surface Engineering 22

2.2.1 Surface Topography 22

2.2.2 Surface Topography Measurement 25

2.2.3 Mechanical Image Deformation Analysis 26

2.2.4 Surface Topographical Images Classification 30

2.3 Support Vector Machine (SVM) 33

2.3.1 Mathematical Linear SVM 34

2.3.2 Nonlinear SVM 37

2.3.3 K-Fold Cross-Validation 40

2.3.4 Multiclass SVM 41

2.3.4.1 One-Against-All (OAA) 41

2.3.4.2 One-Against-One (OAO) 41

2.4 Conclusion 42

CHAPTER 3 HIGH EFFICIENCY VIDEO CODING (HEVC) 43

3.1 Improvements in HEVC Coding Stages 43

3.2 HEVC Intra Prediction Coding 47

3.3 Lossless Coding 52

3.4 High Bit Depth Still Picture Coding 53

3.5 Conclusion 53

CHAPTER 4 PERFORMANCE EVALUATION OF STRAIN FIELD MEASUREMENT BY DIGITAL IMAGE CORRELATION USING HEVC COMPRESSED ULTRA-HIGH-SPEED VIDEO SEQUENCES. 55

4.1 Context of the study 55

4.2 Methodology 56

4.3 Methods and Materials 56

4.3.1 High-speed test device 56

4.3.2 HEVC Lossy and Lossless Compression 58

4.4 Results 59

4.4.1 Tensile Test of Polypropylene (PP) Specimen 59

4.4.2 Sikapower Arcan test 60

4.4.3 Discussion 61

4.5 Conclusion 73

CHAPTER 5 SVM CLASSIFICATION OF MULTI-SCALE TOPOGRAPHICAL MATERIALS IMAGES IN THE HEVC-COMPRESSED DOMAIN 75

5.1 Context of the study 75

5.2 Methodology 76

5.2.1 Methods and Materials 77

5.2.2 Surface Processing 77

5.2.3 Topographical Materials Texture Image Dataset 78

5.2.4 IPHM-Based Classification 79

5.2.5 HEVC Lossless 4x4 PU Compression 83

5.2.6 SVM Classification 83

5.3 Results 84

5.3.1 The Impact of Surface Topography Filtering Types on Achieved Compression Ratios 85

5.3.2 Evaluating IPMH As Texture Feature Descriptor 87

5.3.3 The Impact of Surface Topography Filtering Types on Topographical Images Classification Accuracy 89

5.3.4 The Impact of Scale of Analysis on Topographical Images Classification Accuracy……95

5.4 Conclusion 99

CHAPTER 6 CONCLUSION AND PERSPECTIVES 100

REFERENCES 103

LIST OF FIGURES

Figure 2-1 Mobile video will represent 78% of the world's mobile data traffic by 2021, according to Cisco [18]……7
Figure 2-2 Representation of a digital video signal……9
Figure 2-3 The general image compression framework [23]……11
Figure 2-4 Simplified block diagram of the JPEG DCT-based encoder [38]……15
Figure 2-5 Classic motion-compensated codec scheme……18
Figure 2-6 Block matching algorithm [16]……18
Figure 2-7 MPEG GOP example……19
Figure 2-8 Pictorial display of surface texture [51] ……24
Figure 2-9 Schematic Diagram of Stylus Instrument [52]……25
Figure 2-10 Sample of LAMIH topographical image databases with size of [1024x1024 16-bit depth]……26
Figure 2-11 A schematic of the DIC system……28
Figure 2-12 Corresponding relation of deformed and undeformed sub-image……29
Figure 2-13 The optimal separation hyperplane (OSH)……35
Figure 2-14 Transformation of the data set by Φ [104]……38
Figure 2-15 5-Fold Cross-Validation [106]……40
Figure 3-1 Structure of HEVC encoder and decoder (with elements shaded in light gray) [110].43
Figure 3 2 HEVC Intra/Inter partitioning modes of a CU to PUs [49], [112]……45
Figure 3-3 Example for the partitioning of a 64x64 coding tree unit (CTU) into coding units (CUs) with different coding depths……46
Figure 3-4 (a) HEVC intra-prediction modes (b) Prediction principle for 4x4 PU [119]……48
Figure 3-5 The reference sample locations relative to the current sample for Horizontal and Vertical angular intra prediction (with positive and negative prediction angles) respectively (the idea is [121])……49
Figure 3-6 Representation of the Planar prediction……51
Figure 4-1 Reference image for 2D-DIC specimen measurements with Subset size of 18×18 pixels……57

Figure 4-2 First image at undeformed stage of dynamic PP tensile specimen (global size of 512x472 pixels vs. useful part of 128x384 pixels)……60
Figure 4-3 First image at undeformed stage of Arcan test at 45° of Sikapower glue joint……...61
Figure 4-4 R-D curves for the two video sequences……..63
Figure 4-5 Illustration of HEVC high quality performances for compressed sequence1 (QP=25, PSNR =44.4dB and SSIM=0.99) compared with the original sequence1……..63
Figure 4-6 Illustration of HEVC high quality performances for compressed sequence2 (QP=25, PSNR =39.2dB and SSIM=0.98) compared with the original sequence2……..64
Figure 4-7 Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP0 (All data)……...66
Figure 4-8 Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP0 (Focus on relative gaps between -10% and 10%.....66
Figure 4-9 Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP5 (All data) ……....67
Figure 4-10Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP5 (Focus on relative gaps between -10% and 10) ……...….67
Figure 4-11 PP - Case Lossy - QP20 (All data) ……68
Figure 4-12 Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP20 (Focus on relative gaps between -10% and 10%) ……....……68
Figure 4-13 Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP25 (All data) ……....69
Figure 4-14 Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP25 (Focus on relative gaps between -10% and 10%)……...……69
Figure 4-15 Evolution of relative gaps between strains computed from Lossy images and Lossless images of Sequence 2 (Arcan shear test of glue joint), in the ZOI of maximal shear strain (Axial strain)……...……...70
Figure 4-16 Evolution of relative gaps between strains computed from Lossy images and Lossless images of Sequence 2 (Arcan shear test of glue joint), in the ZOI of maximal shear strain (Shear strain)……..71
Figure 4-17 Evolution upon loading of strains computed from Lossy images of Sequence 2 (Arcan shear test of glue joint), in the ZOI of maximal shear strain (Axial strain) ...……72

Figure 4-18 Evolution upon loading of strains computed from Lossy images of Sequence 2 (Arcan shear test of glue joint), in the ZOI of maximal shear strain (Shear strain) ...……72
Figure 5-1 Nomenclature used to represent the collected mechanical topographic images…….…..78
Figure 5-2 Represents one image (Resolution of 1024x1024 pixels) from six mechanical material categories, with two different zooming and three filtered images……...79
Figure 5-3 Original Image 512x512 (A), selected modes to predict the original image presented with 35 colors (B) Intra Predicted Image (C) and The Residual image (D)……...80
Figure 5-4 First five retrieved images for six images tests (categories 1 to 6) using IPMH, which indicate classification accuracy of 30 %.....82
Figure 5-5 Block diagram depicting the procedure for learning and testing the SVM model. ……84
Figure 5-6 Relationship between the scale of analysis and the six surface categories compression performance by using the multi-scale LP-datasets…….85
Figure 5-7 Relationship between the scale of analysis and the six surface categories compression performance by using the multi-scale BP-datasets…….86
Figure 5-8 Relationship between the scale of analysis and the six surface categories compression performance by using the multi-scale HP-datasets…….87
Figure 5-9 Original Image 1024x1024 (A), selected modes to predict the original image presented with 35 colors (B) Intra Predicted Image (C) and The Residual image (D)……...87
Figure 5-10 Comparison between the IPMHs averages for three different filtered image data sets; LP, BP, and HP data set……..88
Figure 5-11 The effect of increasing the training set size on the classification accuracy while using mixed multi-scale HP, LP, and BP datasets……..89
Figure 5-12 The Confusion matrix for classifying the six surface categories (Mixed)……..90
Figure 5-13 The effect of increasing the training set size on the classification accuracy while using HP- datasets……...90
Figure 5-14 Confusion matrix for six surface categories classification by using 60 % of multi-scale HP data set for training……..91
Figure 5-15 The relation between the size of the training data set and the six surface categories classification performance by using the mixed multi-scale LP- datasets……...92
Figure 5-16 Confusion matrix for six surface categories classification by using 60 % of multi-scale LP data set for training……...92

Figure 5-17 The relation between the size of the training data set and the six surface categories classification performance by using the mixed multi-scale BP-datasets……93
Figure 5-18 Confusion matrix for six surface categories classification by using 60 % of multi-scale BP data set for training……...94
Figure 5-19 The relation between the scale of analysis and the six surface categories compression and classification performance by using the multi-scale LP-datasets…….….95
Figure 5-20 The relation between the scale of analysis and the six surface categories compression and classification performance by using the multi-scale BP-datasets……..…96
Figure 5-21 The relation between the scale of analysis and the six surface categories compression and classification performance by using the multi-scale HP-datasets……..…97
Figure 5-22 Confusion matrix for six surface categories classification by using 60 % of highest-scale HP data set for training……..…...97

LIST OF TABLES

Table 3-1 Displacement Angle corresponding to Angular prediction Mode [41]. ..... .48
Table 4-1 HM 10.1 Encoder Parameters..... 58
Table 4-2 HEVC Compression Performances 62

LIST OF ACRONYMS

AVC	Advanced Video Coding
ANNs	Artificial Neural Networks
bpp	bits per pixel
CABAC	Context Adaptive Binary Arithmetic Coding
CBR	Constant Bit Rate
CT	Computed Tomography
CTU	Coding Tree Unit
CU	Coding Unit
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DIP	Digital Image Processing
DPCM	Differential Pulse Code Modulation
DST	Discrete Sine Transform
GOP	Group of Pictures
HEVC	High-Efficiency Video Coding
HM	HEVC test Model
HVS	Human Visual System
IDCT	Inverse Discrete Cosine Transform

IDR	Instantaneous Decoding Refresh
IEC	International Electrotechnical Commission
ISO	International Standard Organization
ITU-T	International Telecommunication Union
i2i	integer-to-integer
JPEG	Joint Photographic Experts Group
JVT	Joint Video Team
KLT	Karhunen-Loeve Transform
LCU	Large Coding Unit
MAD	Mean Absolute Difference
MSE	Mean Square Error
PPS	Picture Parameter Set
PSNR	Peak Signal to Noise Ratio
PCM	Pulse Code Modulation
PU	Prediction Unit
QP	Quantization Parameter
RDO	Rate-Distortion Optimization
RExt	Range Extension
RLE	Run-Length Encoding
ROI	Region of Interest
SAP	Sample Adaptive intra-Prediction
SAP-G	Sample-based Angular intra-Prediction with Gradient-based
SAP-ME	Sample-based Angular intra-Prediction with Median and Edge
SWP	Sample-based Weighted Prediction
SAO	Sample Adaptive Offset
SSIM	Structural Similarity Index
SVM	Support Vector Machine
TU	Transform Unit
UHD	Ultra High Definition

VCEG	Video Coding Expert Group
VQA	Video Quality Assessment
WPP	Wavefront Parallel Processing

CHAPTER 1

INTRODUCTION

Context and Motivation

During the last decade, several advanced technological techniques emerged in the fields of material science engineering that allow establishing the links between the structure, dynamics and functioning of materials [1], [2]. Materials science engineering studies the characteristics of materials: biological, chemical, physical, optical, and mechanical properties. The mechanical properties are defined from the surface topography when exposed to different types of loads and stresses such as tension, compression, bending, torsion and drawing at micro- or nanometer scale dimensions [3], [4]. This analysis scale is recommended for: aquistition and analysis of material surface for improving the understanding of material surface functionality [5], [6]. Surface topography or surface texture is one of the most relevant characteristics of any material surface that has widely exploited in many mechanical machining processes such as: grinding, shaping and milling [7]. Indeed, several imaging methods for characterizing the physical-mechanical properties exist including advanced optical and scanning electron microscopy, X-ray imaging, spectroscopy, high-speed imaging, electron or micro-topography. Topography measurement systems allow us to obtain specific images for the surface particles represented in three dimensions: height, width, and depth (surface profile) [8]. Obviously, many of these material imaging techniques generate big image databases with high spatial and temporal resolution, large size and pixel depth, which represent a significant amount of data to be stored or transmitted [4], [9]. It would be of great interest to apply lossy or lossless compression prior to classify or store images, where the decompressed image is used as an input for any image processing material algorithm [9]–[11]. Unfortunately, compression artifacts introduced by these algorithms not only affect the visual quality of an image but can also distort the features that one computes for subsequent tasks related to image analysis or pattern recognition in material science

engineering. Since imaging technologies are widely used to analyze the mechanical properties of materials, considering the image quality is essential. The reason behind is that the materials properties is largely dependent on microstructures, which could be affected if the decompressed image is different from the original one. Materials science engineering is a scientific field that requires the acquisition, processing, and analysis of a tremendous amount of image and video information data [12]. Today, the two main topics in material surface imagery fields are: (1) Measuring the similarity or matching between two surface topographies images for comparing the characteristics of two different engineering surfaces. Surface similarity measurement is a promising operation for solving many problems in different material engineering fields such as: industrial surface inspection, defect detection, remote sensing, material classification and biomedical image analysis [13], [14]. (2) Determining which surface filtering range and analyse scale should be used for multi-scale surface topography analyzing and classification. Multi-scale surface filtering decomposition techniques have proven their efficiency in roughness functional analysis [15]. It decomposes the surface topography profile into three different filtered images: high-pass (HP), low-pass (LP) and band-pass (BP) filtered images, which represent the surface roughness, the primary form, and the waviness, respectively. Most of previously surface engineering studies were based on roughness which is represented by the high frequency (HP) component of surface profile.

This thesis addresses the challenges of MEgABIt (MEchAnic Big Images Technology) project. MEgABIt is a project that aims to measure the surface topography and study the surface functionality in the compressed-domain in order to reduce the computation cost. The MEgABIt database is composed of high-resolution 8-bpp and16-bpp images of size 1024x1024 pixels. The similarity between image pairs in the database is too high to the extent that makes the classification problem very challenging.

- The objectives of this thesis are:

1. Study the impact of HEVC (lossless and lossy) compression on the characterization of material mechanical response image sequences for crash and impact loading processes captured by an ultra-high-speed camera.

2. Evaluate the impact of the HEVC lossy and lossless compression algorithms by: (1) Analyzing the mechanical loading response by Vic-2D Software after applying HEVC compression techniques. (2) Considering the compression ratios as well as the quality of the reconstructed video.
3. Implementing HEVC lossless compression algorithm to evaluate the impact of surface filtering types and the scale of analysis on the compression and classification efficiency by considering the compression ratios as well as the classification accuracy for different study conditions.

Challenges

1. The processing task is very challenging when considering mechanical image databases of high-bit resolution, high frame rate, large size and pixel depth, which represent a big amount of data in terms of storage, analysis, classification or transmission over networks.
2. Using lossless compression techniques compression ratio results is limited (or reduced) compared with lossy compression techniques. However, lossy compression techniques give high compression ratios that could cause the loss of the essential characteristics in the mechanical surface image, which leads to wrong analysis or classification. Therefore, we should smartly select the CR for the lossy techniques in order to preserve the original mechanical contents needed for mechanical analysis.
3. Multi-scale surface topographies classification is difficult because the topographical image feature descriptors must be invariant to transformation of surface images like filtering range of surface profile and scale of analysis.
4. Images are stored and transmitted in compressed form. So, they have to be reconstructed prior to be analyzed or classified. This process might be time consuming for retrieval and classification applications.

Contributions

In this work, we implemented HEVC compression standard to reduce the computation complexity of analysis and classification of mechanical surface topographies. The internal bit depth is extended to 16-bpp and the prediction unit is fixed to the size of to 4x4 samples block to enhance the quality of the extracted features for accurate mechanical response image analysis and classification. Our contributions can be summarized in the following points:

1. We evaluate the effect of HEVC lossy and lossless compression on characterizing material surface mechanical response when subject to severe loading conditions over a wide range of strain rates. The impact of HEVC compression was evaluated by analyzing the compressed mechanical loading response sequences by Vic-2D Software. In addition, we evaluated the efficiency of the proposed compression techniques by considering the compression ratios as well as the quality of the reconstructed video. After using two different image sequences, the results demonstrated that HEVC provided very high coding efficiency as well as high visual quality. In addition to that, we succeeded to retrieve the original mechanical data from the HEVC compressed sequences at Quantization Parameter ranging from 0 to 20 as indicated after analyzis of mechanical response using the Digital Image Correlation (DIC) software.
2. We evaluated the effects of surface filtering types and the scale of analysis on the efficiency of the proposed lossless compressed-domain topographical images classifier by considering the compression ratios as well as the classification accuracy for different study conditions. Each surface profile was decomposed into three multi-scale filtered image types: high-pass HP, low-pass LP, and band-pass BP filtered image datasets. Furthermore, the collected database was lossless compressed using HEVC, then the compressed-domain Intra Prediction Mode Histogram (IPMH) feature descriptor was extracted from each predicted image. Simultaneously, we need to keep the visual quality good for visual analysis of mechanical image by experts.
3. We used the support vector machine (SVM) algorithm to classify the high similarity image pairs of the collected topographical image databases (LAMIH databases), i.e. decide if they were taken either from the same category or from different categories.

The model has evaluated 13608 multi-scale topographical images by considering the compression ratios as well as the classification accuracy for each study condition. The experimental results showed that robust compressed-domain topographies classifier was either based on single or multi-scale analyzing methodologies. The high-frequency components (HP-dataset) of the surface profile were the most appropriate for characterizing our surface topographies with achieved accuracy of 96 %.

Structure of the Manuscript

This thesis is organized as following: **Chapter 2** reviews the proposed approaches in the literature for: Firstly, the basics of image and video compression including a brief explanation about digital image and video compression including JPEG, motion compensation and video compression. Secondly, we briefly explain the fundamentals of mechanical surface measuring, analyzing and classification. Thirdly, we give the SVM basics: linear, non-linear and the extension of the binary SVMs to the multiclass case.

Chapter 3 presents the fundamentals of the state-of-the-art HEVC digital video coding standard. We mention the main enhancements introduced to HEVC standard compared to previous coding solutions like Intra Prediction coding technique, or high-bit depth still-image compression. We detail our contribution to improve HEVC lossy and lossless techniques by fixing the prediction unit to 4x4 samples and increasing the internal bit depth to 16 bits.

Chapter 4 describes the implementation of our proposed methods for testing the influence of image compression on mechanical response analysis as well as presenting the obtained results.

Chapter 5 presents an original method based on the SVM algorithm for multi-scale surface classification in the compressed domain. In the experimental results we discuss the effect of surface filter type and scale of analysis on the compression and the classification accuracy.

Finally, we give conclusions and perspectives of our research work.

CHAPTER 2

STATE-OF-THE-ART

Digital Images and Video Compression

Digital images and videos constitute the first pillar of multimedia technologies like broadcast TV, online gaming, mobile communications or multimedia streaming [16]. Every minute, a huge amount of images and videos is created in a great number of domains such as medical imaging, entertainment, earth monitoring or industrial applications [9]. Indeed, telecom operators like Cisco or Nokia predict that video traffic will represent about 80% of all consumer Internet traffic in the coming years [17].

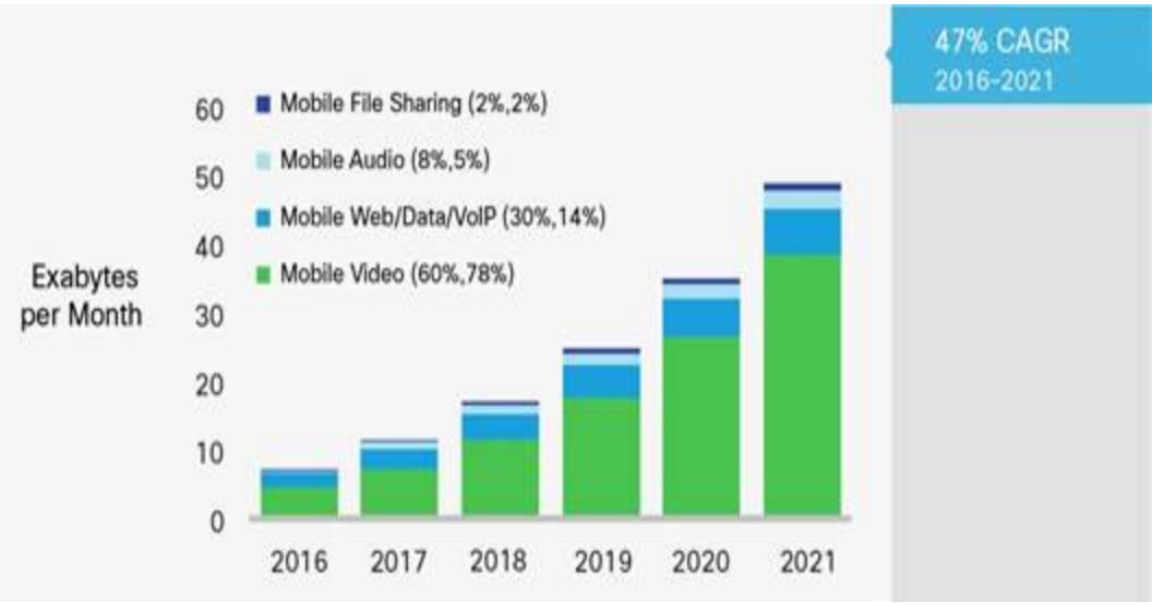


Figure 2-1 Mobile video will represent 78% of the world's mobile data traffic by 2021, according to Cisco [18].

Consequently, the processing, storage and transmission of images and videos over networks constitute a very challenging task. In order to overcome such problem, several digital image and video compression techniques have been developed during the last twenty years in order to reduce the image and video data size while keeping good video quality. In what follows, we first introduce the basics of image and video compression. Then we describe the JPEG still image compression standard, as an illustrative example. Finally, we describe motion estimation/compensation technology and give an overview of first-generation video compression standards.

2.1.1 Basics of Image Compression Techniques

Digital image and video compression is the science of coding the image content to reduce the number of bits required in representing it, aiming facilitate the storage or transmission of images with a level of quality required for given application (digital cinema, mobile video streaming) [19]. Typically, a digital image signal contains visual information in a two-dimensional matrix of size equal to N rows by M columns. Each spatial sample also known as a *pixel* is represented digitally with a finite number of bits called *bit-depth* [20].

For example, each pixel in a grayscale image is typically represented by a byte word, i.e. 8 bit-depth. A standard RGB color image is represented by three byte words, i.e. 24 bit-depth corresponding to 8 bits for the red component, 8 bits for the green one, and 8 bits for the blue one [21].

Digital video signals are represented as a collection of successive still images separated by a fixed interval time, which determines the so-called *frame rate* [16].

The compression process can be performed by exploited many duplicated information in the digital image or video signals. For example, it is possible to exploit the fact that the human eye is more sensitive to brightness than color for reducing the size of an image. To do that, the RGB components of the color image are first converted into the three YUV color components, where Y corresponds to the luminance (brightness) and U and V are the chrominance components, respectively. Then, the chrominance components are usually reduced by a factor of 1,5 or 2 by appropriate spatial down sampling [22].

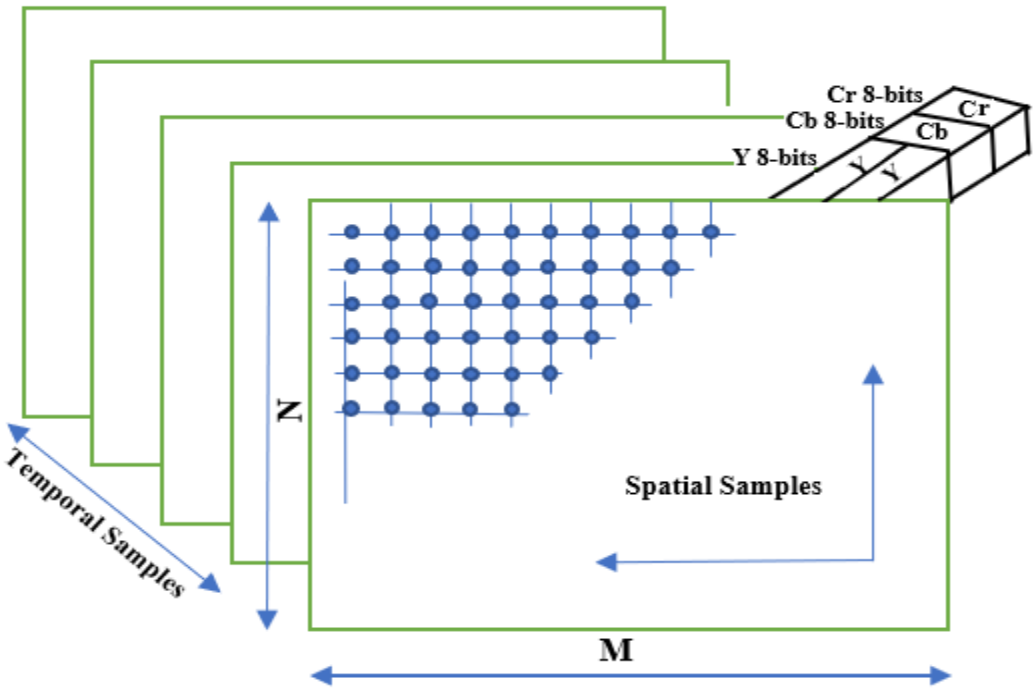


Figure 2-2 Representation of a digital video signal.

For instance, a high-definition broadcast video signal typically consists in a sequence of successive frames at a frame rate of 25 fps. Each frame corresponds to a matrix of

1920x1080 pixels with 4:2:0 sample format (i.e. 1920x1080 luminance samples and 960x540 samples for each chrominance component), and 8 bit-depth precision. Hence, the corresponding bit rate around 620 Mbit /s.

Indeed, *chroma subsampling* does not reduce the image data size in a sufficient way and other compression operations are needed to fulfill bandwidth constraints. Efficient image and video compression can be achieved by eliminating two main types of redundancies known as *statistical redundancy* and *psychovisual redundancy* [16], [19], [20], [23]:

- **Statistical Redundancy** can be divided into two categories:
  - first, the pixel-to-pixel redundancy traduces the correlation which exists between pixels both in the spatial and temporal domains:
- **Spatial redundancy** is related to statistical correlation between the intensity values of neighbor pixels very closed to each other. Spatial redundancy can be eliminated using the differential coding or intra-predictive coding.
- **Temporal redundancy** is related to the statistical correlation between pixels belonging to two successive video frames, which is as high as the time interval between two consecutive video frames is short. Often, this kind of redundancy can be eliminated using inter-predictive coding between consecutive frames.
  - **The Coding Redundancy** is coming from the information redundancy between coded symbols, it can eliminate by using so-called binary *entropy coding* techniques.
- **Psychovisual Redundancy** is based on the characteristics of the Human Visual System (HVS). Indeed, some visual informations are less relevant than others in a frame content due to so-called *masking* phenomenon which may occur in luminance, contrast, texture and frequency domain. Consequently, these irrelevant visual data can be suppressed without degrading visual quality. However, in this case, it should be noted that the reconstructed signal is mathematically different from the original one.

In fact, digital image compression techniques are broadly classified into two categories; Lossless (reversible) and Lossy (or irreversible) compression techniques [20], [21], as generalized in Figure 2-3.

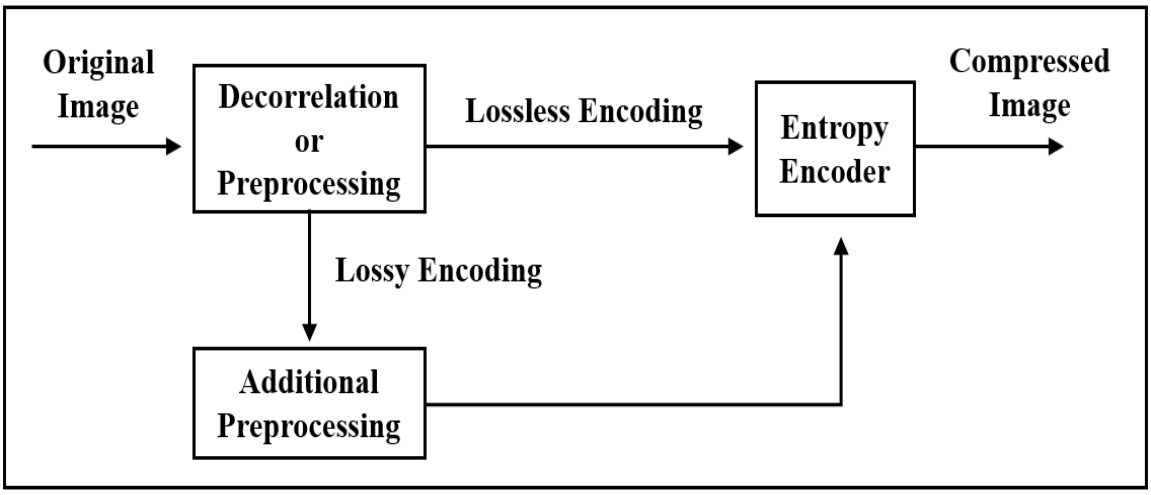


Figure 2-3 The general image compression framework [23]**.**

The Lossless image compression category is widely used in specific-imagery applications requiring the reconstructed image to be exact compared to the original one, such as medical imaging, document archiving and scientific imagery. Here, the decoder is perfect inverse of the encoding process and the original image can be fully retrieved from the compressed file. However, the so-called *compression ratio* remains moderate ranging from 2:1 to 10:1 on average, based on image complexity [24]. Here, the compression ratio (CR) is defined as the ratio of the size of the original image in bits, to the size of the compressed stream expressed also in bits:

$$\text{Compression Ratio (CR)} = \frac{\text{Total number of bits in the original image}}{\text{Total number of bits in the compressed image}} \qquad (2.1)$$

For example, CR=2:1 means that the compressed file is twice as small as its original version. The spatial redundancy is reduced in the decorrelation stage by using the prediction-based methods [21], [25], or transform-based methods such as discrete cosine transform DCT [26] or the reversible wavelet transformation (DWT) [27]. Followed by Entropy coding for

reducing the coding redundancy. The number of bits required to represent a sequence of symbols is reduced to a minimum length as in binary arithmetic coding, Huffman Coding and Variable-length Coding [28].

The Lossy image compression category allows achieving higher compression ratios (up to more than 100:1) while losing part of image information. Further, the reconstructed image will not be mathematically identical to the original image, so it is necessary to determine the minimum data required for retrieving all necessary information [24]. The spatial redundancy is reduced by using one of the existing pixel decorrelation techniques as predictive coding, transform coding, sub-band coding [29]. The residual of transform data is computed and subject to an additional non-reversible process known as *quantization* to increase significantly coding efficiency. Finally, the quantized coefficients are lossless compressed by entropy coding [19]–[21], [25], [30], [31].

The optimal compression scheme will be able to obtain the highest compression ratio as well as best image quality with least computation complexity.

In the case of lossless compression, there is no distortion, so the reconstructed image is mathematically and by consequence visually identical to the original one. In the case of lossy compression, however, there is a need to additionally measure decoded image quality as well as the achieved CR [21].

Digital image quality can be evaluated both subjectively and objectively [29], [32].

- **The subjective quality measurement** is based on observations performed by human viewers in a controlled test environment. Human viewers are asked to give a score to the processed images according to different quality or degradation scales [33]. Subjective tests represent the ground truth, but they are often time consuming and expensive. Hence, objective image and video quality metrics are frequently used because of their ease of implementation.
- **The objective quality measurement** is generally based on the image statistical properties and permits to evaluate the rate-distortion (RD) performances of digital image and video compression algorithms. Different quality metrics exist depending on the

knowledge or not of the original visual signal [34]. Among these, Full Reference (FR) quality metrics are calculated from the original image and its compressed version.

The best-known FR objective video metric is undoubtedly the Peak Signal-to-Noise Ratio (PSNR). The PSNR metric, expressed in decibels, is defined as:

$$PSNR_{dB} = 10\, log_{10} \frac{(Maximum\ Pixel\ Intensity)^2}{MSE} \tag{2.2}$$

Where the Mean Squared Error (MSE) is defined as:

$$MSE = \frac{1}{\mathrm{M\,N}} \sum_{i=1}^{M} \sum_{j=1}^{N} \lfloor I(i,j) - R(i,j) \rfloor^{2} \tag{2.3}$$

and

- I (i, j) represents the pixel value at position (i,j) in the original image of size MxN pixels.
- R (i, j) represents the pixel value at position (i,j) in the reconstructed image of same size.
- Maximum Pixel Intensity is equal to 255 for 8-bit resolution.

The higher PSNR value is, the better visual quality is. In the case of lossless compression, the PSNR value is equal to infinity. However, even if the PSNR metric is easy to use, it is well known that PSNR (or in an equivalent way, the MSE) is poorly correlated with the human visual judgment [35]. To overcome this problem, many other quality metrics derived from the PSNR have been proposed in the literature that try to mimic the Human Visual System (HVS) [29]. We can cite for example the Weighted Signal to Noise Ratio (WSNR), Noise Quality Measure (NQM) and Visual Signal to Noise Ratio (VSNR). The Structural Similarity Index, known as SSIM, and its variants constitute another alternative to the PSNR metric. SSIM is a full reference quality metric. It measures the visibility of any error in the structural information of the image and incorporates HVS properties like luminance and contrast masking. SSIM varies between 0 (poor quality) and 1 (perfect). It is commonly

accepted that SSIM clearly outperforms PSNR or MSE and SSIM is nowadays widely used by the image processing community.

To conclude about image quality evaluation, the term «image compression» should be understood in a mechanical application-specific sense in this thesis. This makes a great difference with “classical” image compression, where the decompressed images are intended to be viewed by a human observer. In this case, the quality of decompressed image is evaluated based on perceptual considerations obtained according to HVS properties.

In our work, the decompressed image is used as an input of an image processing algorithm. The decompressed image is not viewed by a human viewer, which makes unusable applying of considerations obtained according to HVS properties. In the present case, the error introduced by lossy compression should not affect the accuracy of information data extracted from the decompressed image that is needed for further material analysis or classification, while keeping high compression ratio. It is not strictly necessary that the decompressed image looks visually close to the original one. Rather, the decompressed image should contain as minimum information data as needed to guarantee material imaging processes of high quality.

2.1.2 Illustrative Example of the JPEG Still Image Compression Standard

The JPEG compression standard is one of the most well-known image compression standards. It takes its name from the working group called the Joint Photographic Expert Group that developed it in the early 1990s. Today, the JPEG standard is still widely used [26] in a broad range of digital imaging applications like digital photography, medical imaging, or video recording (using Motion JPEG) [26]. Moreover, it provides the basis for future standards including JPEG2000, and High Efficiency Video Coding (HEVC)-Intra. JPEG is designed to handle color and grayscale image compression with an achieved compression ratio of up to 1: 100 [36] . It is based on the Discrete Cosine Transform (DCT) which analyses the image as the human eye does. The human eye does not see all the colored details present in the image, consequently the fine details corresponding to high spatial frequencies can be removed with no effect for the human viewer [37]. The encoding process is started by

dividing the original image into squared blocks of 8x8 samples. Each block is transformed by Forward DCT or DCT from the pixel domain to the frequency domain in order to reduce the spatial redundancy. After DCT, the block energy is generally concentrated in few low frequency transform coefficients. Then, the 64 DCT coefficients are quantized hence reducing the number of non-null values. Finally, the quantized coefficients are sent to the entropy coder that delivers the output stream of compressed image data as illustrated in Figure 2-4.

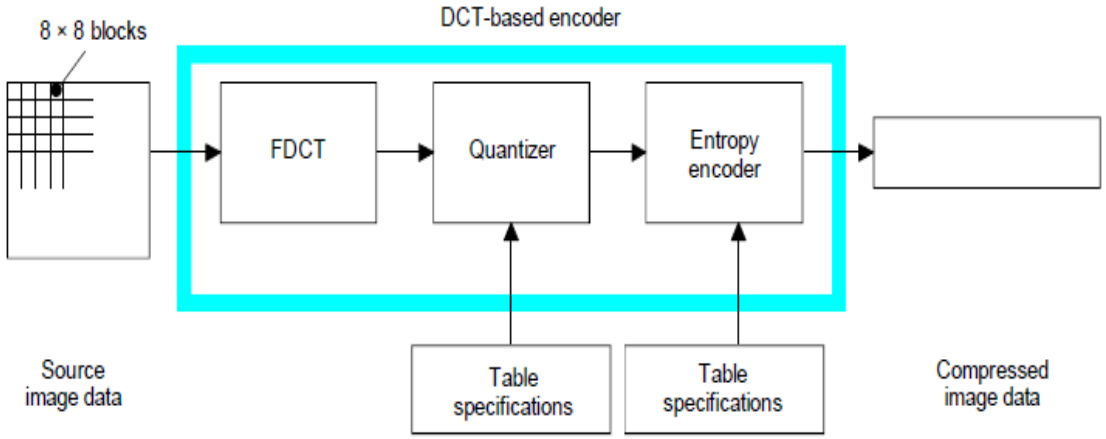


Figure 2-4 Simplified block diagram of the JPEG DCT-based encoder [38].

The main JPEG encoding process steps are briefly illustrated below:

1. **Block segmentation:** the full compression of the image makes non-ideal compression results. For this reason, JPEG standard suggests that image is dividing into 8x8 blocks and starting from this stage each of these 64-pixel blocks is processed separately at all codec stages.
2. **Discrete Cosine Transform (DCT):** transforming each block of 8x8 pixels into the spatial frequency domain allows the algorithm to ignore less critical pixels of the original block by removing the inter-pixel redundancy inside the original image. This process makes the quantization process easier to know which parts of the block are less important. Typically, the highest AC coefficients are deleted during the quantization process. JPEG calculates the Forward Discrete Cosine Transform (FDCT) and Inverse Discrete Cosine Transform (IDCT) by two following equations [38]:

FDCT: $$S_{vu} = \frac{1}{4} C_u C_v \sum_{x=0}^{7}\sum_{y=0}^{7} S_{yx} \cos\frac{(2x+1)u\pi}{16} \cos\frac{(2y+1)v\pi}{16} \quad (2.4)$$

IDCT: $$S_{yx} = \frac{1}{4} \sum_{u=0}^{7}\sum_{v=0}^{7} C_u C_v S_{vu} cos\frac{(2x+1)u\pi}{16} cos\frac{(2y+1)v\pi}{16} \quad (2.5)$$

where: $C_u, C_v = {}^{1}/_{\sqrt{2}}$ For $u, v = 0$

$C_u, C_v = 1$ Otherwise

3. **Quantization:** the quantization stage is centered at the core of any lossy encoding algorithm for reducing the psychovisual redundancy. It is a non-reversible operation, and it must be bypassed in lossless compression mode [19], [31], [37], [39], [40]. In order to remove the less significant DCT coefficients in the transformed block, every element in the 8x8 FDCT matrix S_{VU} is divided by a corresponding step size Q_{VU} from a previously calculated 8x8 quantization table (Q-table) [23] and rounded to the nearest integer as shown in Eq. (2.6):

$$Sq_{vu} = round\left(\frac{S_{vu}}{Q_{vu}}\right) \quad (2.6)$$

Moreover, the magnitude of the non-zero coefficient values is limited after division and rounding to smaller values close to zero.

4. **Entropy Coding:** the quantization operation produces a block consisting of 64 values, most of which are zeros. Normally, the best way to compress this type of data is to combine zeros with each other. That is what JPEG does for the 63 quantized AC coefficients using Run-length Encoding (RLE). The DC component S_{00} corresponding to the null frequency range is encoded separately.

- **Huffman Encoding of DC Coefficients:** the difference DIFF between the quantized DC coefficient values of two adjacent blocks is encoded independently using DPCM using the following Eq. (2.7):

$$\text{DIFF} = \text{DC}_\text{i} - \text{DC}_{\text{i}-1} \quad (2.7)$$

- **Zig-zag:** After computing DPCM for the DC coefficient, the AC coefficients are converted to vector thanks to zig-zag scanning.
- **Run-Length Encoding for AC coefficients:** as we have indicated above, the quantization operation used to reduce the high-frequency components to be more likely to be zeros. RLE represents these 63 coefficients with sequence (**Run, Length**) pairs where **Run** indicates the number of zeros preceding a non-zero coefficient and **Length** indicates the magnitude (indeed, the so-called *category*) of the non-zero coefficient. Finally, we calculate the total number of bits that represent each pair of (**Run**, **Length**) using Huffman tables for the AC coefficients [24].

Usually, JPEG Lossless compression is a two-step algorithm as illustrated in [24]. The first step consists in exploit the inter-pixel redundancy present in the original image. JPEG Lossless considers the well-known DPCM (differential pulse coded modulation) coding technique to predict each pixel from its neighbors and then compute the residual error. The second step uses a Huffman encoder to remove the coding redundancy.

2.1.3 Motion Compensation and Video Compression

Digital video compression is the process that aims to reduce the spatiotemporal redundancy contained in successive video frames to achieve a given bit rate [41]. The primary constraints concern the quality of the decoded video must satisfy specific requirements and the computational complexity involved in the operation. In order to exploit temporal redundancy, a video coder incorporates an additional Motion Estimation (ME)/motion compensation (MC) process. ME aims at estimating the displacement parameters of moving objects between two consecutive frames, while MC exploits these parameters to match the objects along the temporal axis. ME/MC has proven its efficiency in digital video processing and has become the core component of digital video compression technologies such as MPEG, H.264/AVC, and HEVC for removing the temporal redundancy. The concept of motion-compensated codec presents in following classic codec scheme in Figure 2-5.

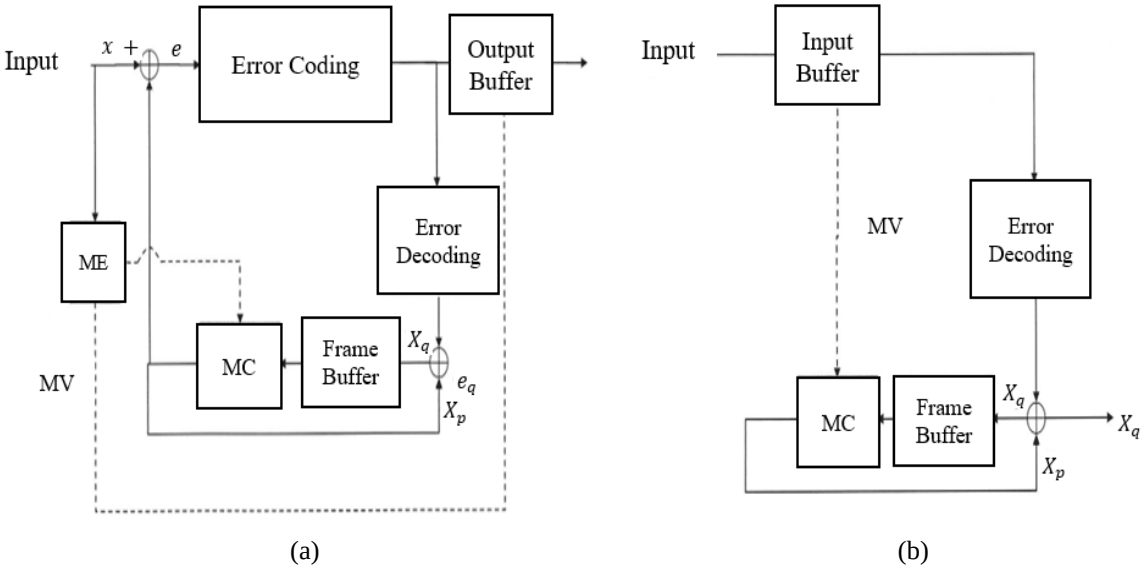


Figure 2-5 Classic motion-compensated codec scheme.

In practice, ME is performed using the well-known block matching algorithm.

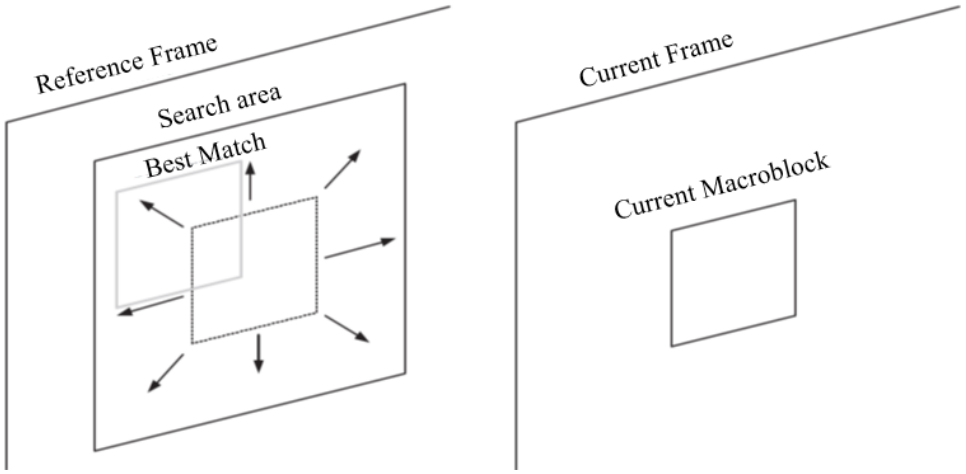


Figure 2-6 Block matching algorithm [16].

First the video frame is partitioned into fixed M x N rectangular sections known as *macroblocks*. Then the Motion Estimation (ME) stage search for each macroblock in the current frame to be encoded the best correspondence with a macroblock in the previously encoded frame which serves as a reference frame. The best candidate is the one which

minimizes the so-called displaced frame difference or DFD [16]. The displacement coordinates between the macroblock to be encoded and the reference macroblock is represented by a Motion Vector (MV). This MV is transmitted to the decoder in the output compressed bitstream as illustrated in Figure 2-5 (a). Then the MC stage computes the residual between the current macroblock and the estimated one in the reference image. Finally, the residual signal is quantized and entropy coded prior to be sent to the decoder; it is also used to reconstruct the decoded macroblock necessary for the next encoding step at the encoder side [42]. The decoder uses the received motion vector MV as well as the decoded residual macroblock to recreate the decoded macroblock.

In order to organize the video stream, the video sequence is divided into Groups of Pictures noted as (GOP) and each GOP includes a given number N of coded frames. Three different coding frames can be included within a GOP [16], [43]:

- The **I-frame** uses an Intra-prediction allows initiating the compression process as it is independent of other encoded images. I frames are used as references for inter prediction.
- The **P-frame** uses the Inter prediction with a unique previous reference I or P frame.
- The **B-frame** uses the Inter-prediction with two reference images that can be previous or next frames.

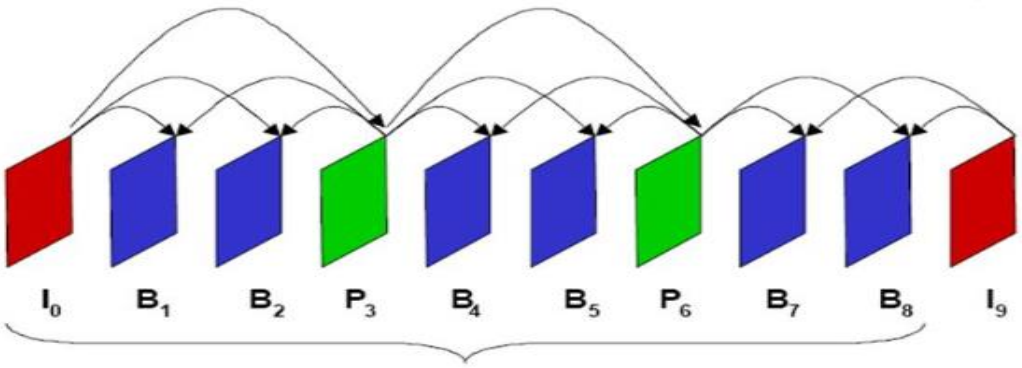


Figure 2-7 MPEG GOP example.

The hybrid motion-compensated DCT-based video compression scheme described above constitutes the basis of all existing digital video coding standards. From one standard

to its successor, however, the performances of each processing step are improved, and additional tools are introduced to further increase coding efficiency. Thus, it is common to say that a new video coding standard outperforms its predecessor by doubling the coding efficiency for the same video quality. Several international organizations are involved in the standardization of digital video coding schemes, among which[44]:

- International Telecommunication Union - Telecommunication Video Coding Experts Group (**ITU-T VCEG**): the organization that has developed a series of compression standards for videotelephony such as H261, H263.

- International Organization for Standardization / International Electrotechnical Commission (**ISO / IEC**): the international body whose best-known group is the Moving Experts Group (MPEG). Founded in 1988 to develop video compression standards, this group has developed the MPEG1, MPEG2, and MPEG4 standards.

- Joint Video Team (**JVT**) which results from the association of the first two groups. JVT created the famous H264 / AVC, which always known as MPEG-4 AVC or MPEG-4 Part 10.

To conclude, we give a brief overview of the most popular video coding standards as [45];

-MPEG-1 Standard [46], [47]: is the first standard developed by MPEG group to compress a digital video. MPEG-1 considers a frame resolution of 352 pixels by 240 pixels with video compression ratios over 100:1. MPEG 1 has been finalized in 1993. The first three parts of the standard were accepted by ISO and deal with video coding (Part1), audio coding (Part2) and system including multiplexing and packetisation (Part3). Part 4 (1995) describes a testing platform for verifying compatibility on all media, and Part 5 (1998) is a reference implementation of algorithms.

-**MPEG-2 Standard:** in order to overvome the limitations of MPEG1 standard in the face of the rapid evolution of computer and digital resources, the MPEG2 appeared and finalized in 1994 introducing a wide range of choices regarding resolution and bit rate control. MPEG-2 allows the compression of progressive or interlaced video at rates ranging from 1.5Mb / s to

30Mb / s [48]. The MPEG-2 standard really exploded with the wide deployment of digital terrestrial television.

-MPEG-4 Standard: in 1995, the MPEG4 standard began to emerge from the theoretical point of view with the aim of producing a bit rate below 64 kbit / s. Further, it has finished in early 1998 with a new dimension allowing a much more flexible and much more efficient standard. The standard allows the encoding of a wide variety of video formats (size, resolution, frame rate) but also the coding of arbitrarily shaped video objects, still images as well as 3D synthetic objects [48]. As a result, this standard addresses a wide range of audiovisual applications ranging from video conferencing to audiovisual production via internet streaming.

-H.264 / AVC Advanced Video Coding is a widespread standard. The JVT group developed a high-performance video coding standard for both low and high bitrate applications in collaboration with ITU-T. H.264/AVC recommendations have been finalized in 1999, and H.264/AVC rests the most powerful coding standard until the end of the year 2012 [43]. H.264/AVC introduces many new efficient coding tools including:

- Intra-prediction coding with 11 modes that can be implemented with flexible block sizes (16 × 16, 8 × 8 and 4 × 4 pixels).
- 2D discrete cosine transform (DCT) of different sizes (4 × 4 and 8 × 8 pixels), and integer transform.
- For reducing the temporal redundancy, inter-prediction is applied on macro-blocks with variable size partition of 16 × 16, 16 × 8, 8 × 16 and 8 × 8 samples that can themselves be partitioned into 8 × 4, 4 × 8 and 4 × 4 pixels. Also, it uses a sub-pixel representation of the movement that can be realized until even quarter 1⁄4 or half 1⁄2 pixel samples motion compensation [16].

Thanks to these innovations, H.264/AVC succeeded in improving the coding efficiency by a factor of two compared to the MPEG-2 standard, for the same video quality.

In January 2013, a draft of the successor named High-Efficiency Video Coding (HEVC) standard was announced. It can greatly improve the decoded video quality compared

to H.264/ AVC for the same video bit rate, but with a significant increase of the encoding time complexity [49].

The current state-of-the-art High Efficiency Video Coding (HEVC) standard will be discussed in details in Chapter 3.

Material Surface Engineering

When considering surface analysis, materials science engineering studies the characteristics of materials: biological, chemical, physical, optical, and mechanical properties extracted from surface topography, e.g. Mechanical material engineering focuses on the study of evolution of material properties when subjected to different types of loads and stresses such as tension, compression, bending, torsion and drawing from macro to micro- or nanometer scale dimensions[3], [4]. Micro or nano- scale analysis improves the understanding of material surface functionality. These improvements are generalized for manufacturing many different analysis and acquisition systems at various physical scales [6]. Micro and nano- scale analysis is widely used in advanced science sectors including: environmental changes, renewable energies, metallurgy, materials science, biology, healthcare and biotechnology [5]. Materials science engineering is a wide area of research and we will not cover all its scientific aspects in the present work. So, in the following section, we will focus on studing specific material imaging techniques that will be at the heart of our research, namely deformation analysis and topographies classification.

2.1.4 Surface Topography

The topographical measurement system allows us to obtain specific images for the surface structures represented in three dimensions: height, width, and depth which is known as surface profile [8]. Surface topography or surface texture is one of the most relevant characteristics of any material surface that has been widely exploited in many mechanical machining processes such as: grinding, shaping and milling [7]. Surface topography is defined as the random repetitive forms of the nominal surface to represent: roughness, waviness, lay, and flaws in 3D topography as illustrated in Figure 2-8 [50], [51]. The

roughness (nano- and micro-roughness) is defined as the vertical and horizontal deviations and irregular depth that are incorporated into the general surface curves. It is characterized by the local maxima (asperities, hills *or* peaks) and local minima (valleys) with varying amplitude and spacing [52]. It has been measured for decades via 2D cross section and recently via 3D cross section [15].

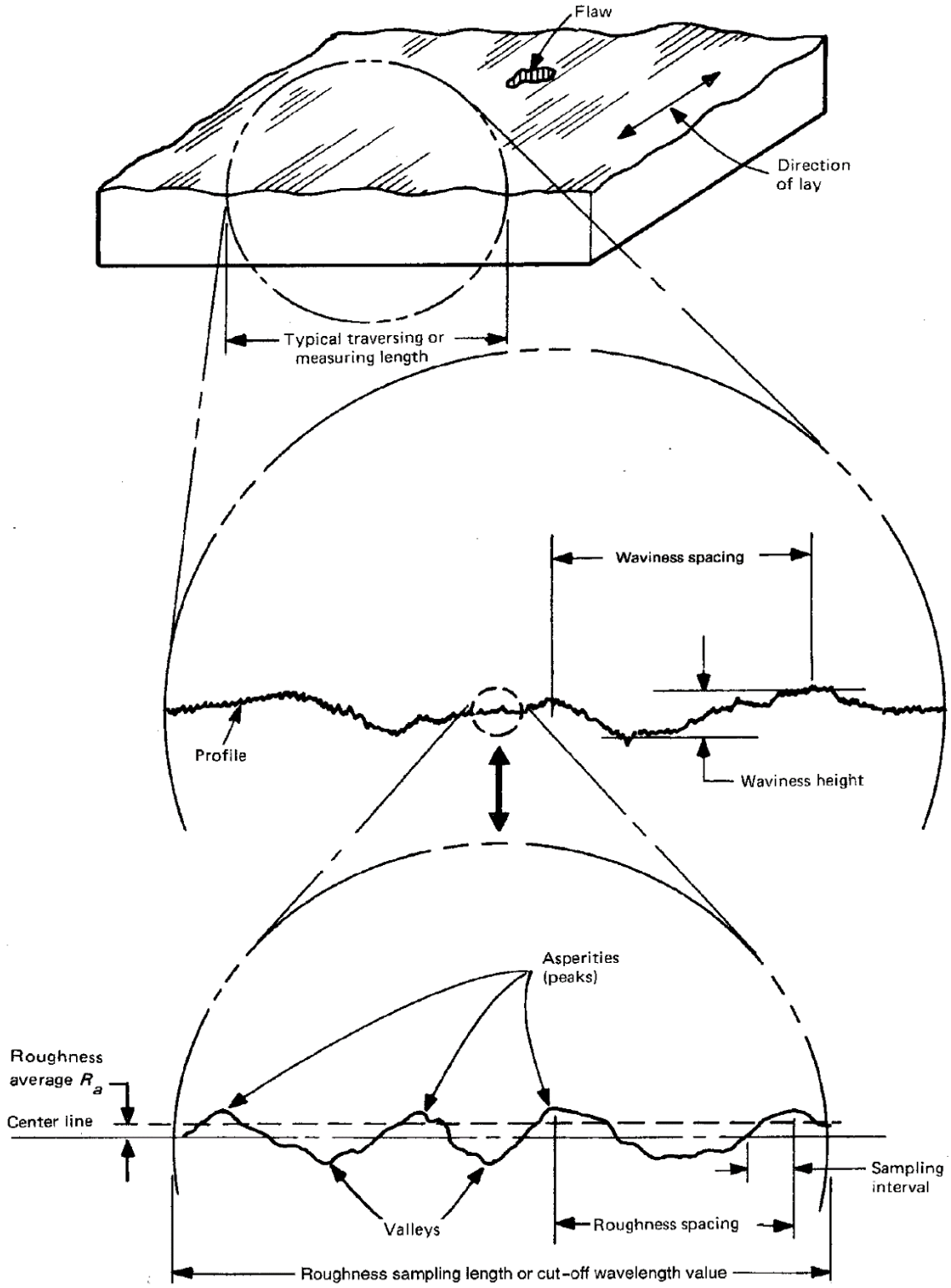


Figure 2-8 Pictorial display of surface texture [51].

Flaws are benefitless and unwanted interruptions in the surface profile analysis. Waviness (macro-roughness) is the surface irregularity with longer wavelength which is greater than roughness wavelength. While the primary form (lay) results from removing the short wavelets except the shortest wavelength components which represent the roughness (nano- and-micro roughness) [51], [53].

2.1.5 Surface Topography Measurement

In the past, surface topography measurement techniques were based on microscale analysis where the measurement principle relied on contact and near contact technology techniques like using capacitance, electrical, hydraulic and pneumatic instruments [54]. Stylus instrument is the most common contact metrology technique used for measuring the surface roughness as depicted in Figure 2-9. It is a diamond pointed end probe which scans accurately in straight lines the surface heights from one point to another at a constant speed to show the surface height variation [8], [55]. Normally, the transducer will convert the measured movement into electrical signal to generate 2D profile. This technique is difficult to calibrate and could cause damages to the tested surface [56].

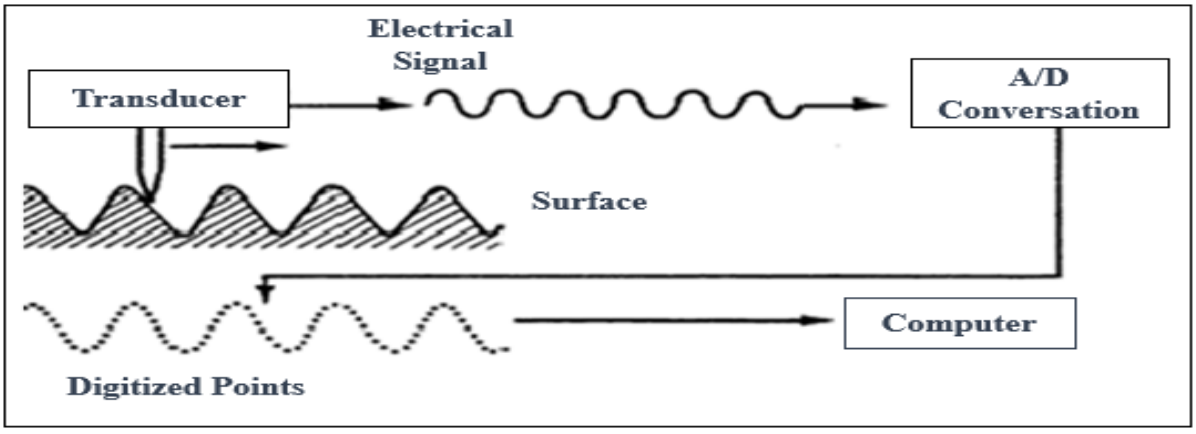


Figure 2-9 Schematic Diagram of Stylus Instrument [52].

In the recent century, advanced computer technologies and new optical acquisition devices allow further development of non-contact optical imaging systems with high-quality topographical image reconstruction. Thanks to these innovative solutions, the fundamental

intrinsic properties of materials are studied with the aim to establish links between the structure, dynamics and materials functioning [3], [51], [57]. Topography imaging analysis is a challenging task because of the significant changing of the material surface texture due to both analysis scales and local physical properties [58], [59]. Six methods exist that differ on the underlying physical principle used: mechanical stylus, optical, scanning probe microscopy (SPM), fluid, electrical, and electron microscopy [51]. Obviously, all these material imaging techniques produce very big image databases. Usually the obtained images have high spatial resolution with large number of pixels and high pixel depth precision. For example, the LAMIH image database used in our research work has been generated using an optical topography imaging system [51]. This system will be presented in detail in Chapter 5. The LAMIH database consists in more than 53000 images of size 1024x1024 pixels available with two different bit-depths: 8 and 16 bits per pixel, respectively. As an illustration, Figure 2-10 shows one of the images (1024x1024 pixels, 16 bits/pixel) from the LAMIH database.

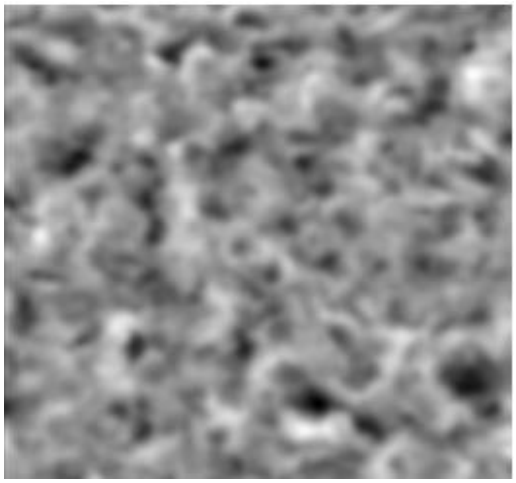


Figure 2-10 Sample of LAMIH topographical image databases with size of [1024x1024 16-bit depth].

Hence applying lossy or lossless compression appears as a great solution to store or transmit the images in an efficient way [10], [11].

2.1.6 Mechanical Image Deformation Analysis

Digital Image Processing has several aims in material science engineering as: Image acquisition, enhancement, filtering, segmentation and analysis [12]. For instance, similarity

or matching measurements in surface images are very useful for comparing the characteristics of two different engineering surfaces. This comparision is used to control different machining processes such as tensile, compression etc. [58]. The surface similarity is mostly used in industrial surface inspection, remote sensing, material classification and biomedical image analysis [13]. The images matching measurement is a promising operation for solving many problems in different material engineering fields [14]. In mechanical engineering, imaging may be a powerful tool for measurement of displacement and strain fields on a specimen during testing. If images can be recorded at high enough frame rate, they can then be used to characterized material behaviour (i.e. allow the identification of stress vs true strain curve) over a wide range of strain rate (possibly up to that encountered during crash or impact, e.g.). Digital Image Correlation (DIC) is frequently used in material mechanical tests for computation of in-plane displacement and strain fields [60]. Digital Image Correlation (DIC) is a non-contact optical full-field measurement technique developed in the 1980s [61]. It requires computer software and a camera with suitable frame rate to record the material surface under loading as presented in Figure 2-11. The recorded images of deformed surface upon loading are compared with the initial image of undeformed surface to calculate in plane displacement and strain fields. More precisely, a random pattern (spray of black paint on surface painted in white, e.g.) is created on specimen surface or Region of Interest (ROI) that allows its division into sub-surfaces, called facets, characterized by a unique signature in terms of grey level. The DIC software then tracks the displacement of each facets between a deformed image and the underformed one, thanks to its unique signature. Displacement and strain components are therefore obtained locally on the specimen, i.e. in each facets of the ROI, thus allowing the extraction of enriched data compared to “simple” ROI’s elongation measurement for instance.



Figure 2-11 A schematic of the DIC system.

DIC was also used for visualizing the deformations in beams in order to understand the hypotheses of the beam theory and to better define their area of validity [62]. We present hereafter a DIC illustrative example.

Let us assume, we need to use DIC to measure the correlation between the initial surface image I (x, y) and the deformed surface image D (x, y). Both images are in grayscale with dimensions $M \times$ N. Each image is subdivided into several subset regions (facets). Each point in the subset region is projected to a certain pixel in the surface image plane. Then, the strain field displacement for the deformed image is compared with the original surface by subtracting the end position to obtain the displacement in the x [u (x, y)] and y [v (x, y)] directions as illustrated in Figure 2-12.

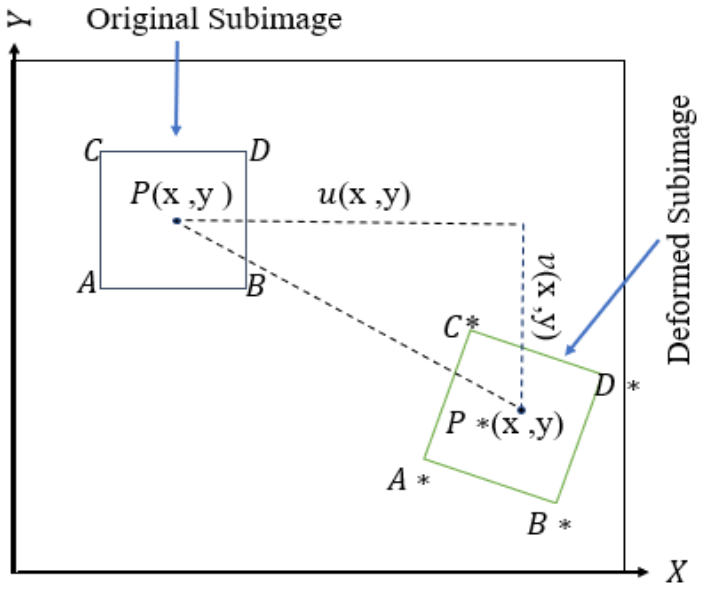


Figure 2-12 Corresponding relation of deformed and undeformed sub-image.

The in-plane displacements of all facet centers are determined with respect to their position in the reference image. For example, the center point P is the target point to track. After deformation P is shifted to point P∗. The mathematical definition for the displacement field measurement in the spatial domain can be computed by using the following the two following equations [63]:

$$xP* = xP + u(x,y) \tag{2.8}$$

$$yP* = yP + v(x,y) \tag{2.9}$$

According to the theory, strains can be simply estimated by computing the partial derivatives of the displacement's fields. Consequently, the strain in the x direction and y direction as well as the shear strain of ABCD are respectively expressed as follows:

$$\varepsilon_{xx} = \frac{du_x}{dx} \tag{2.10}$$

$$\varepsilon_{yy} = \frac{du_y}{dy} \tag{2.11}$$

$$\varepsilon_{xy} = \frac{1}{2}\left(\frac{du_x}{dy} + \frac{du_y}{dx}\right) \tag{2.12}$$

Typically, the displacement field measurement can be evaluated in the spectral domain mainly based on 2D-Fourier Transform, phase shift and the convolution theorem [14] [65] which is mathematically defined as:

$$F`(f_{x,}f_{y,}) = F(f_{x,}f_{y,}) \exp[i\Delta\varphi(f_x) + i\Delta\varphi(f_y)] \tag{2.13}$$

$$\left.\begin{matrix}\Delta\varphi(f_x) = 2\pi\Delta x f_x \\ \Delta\varphi(f_y) = 2\pi\Delta y f_y\end{matrix}\right\} \tag{2.14}$$

Where $F\ (f_{x,}f_{y,})and\ F`(f_{x,}f_{y,})$ are the Fourier transforms of the original image I (x, y) and deformed image D (x, y) respectively. f_x and f_y are the spatial frequency in the x and y directions, respectively.

To conclude, we can note from the above discussion that the DIC process is also used during the motion estimation process in inter-frame video coding. Indeed, the block matching algorithm used in the actual video coding standards like H.264/AVC or HEVC consists in finding in a past or next frame the macroblock that is the best correlated to the current macroblock to be encoded. *This suggests the possibility of establishing a bridge between image compression and image analysis since both use the same image processing tools.* Moreover, the quality of the results obtained by image correlation should depend directly on the intrinsic quality of the images used: contrast, edge sharpness, etc. So, it would be interesting to study the performances of the image correlation if the images are compressed now. This will be the topic of Chapter **4**.

2.1.7 Surface Topographical Images Classification

The image classification is the process to distinguish between different images by inferring a link between the characteristic content features of each image and their category. It aims to assign an unknown image to one set between the different categories. Classification

was defined in [65] as “the attribution of the correct class label to a set of measurements”. Image classification techniques generally exploit a rich set of various image features which are able to characterize various aspects of image visual information [66]. Indeed, the success of a classification problem is strongly related to the quality of data and the variables that characterize them [67]. Generally, image classification can be performed in both pixel and frequency domain. In the literature, most researches were interested in using Content-Based Image Retrieval (CBIR) methods. It is the most known method which has been utilized since the 1980s for searching and retrieving images from big database based on similarity measurement. The image feature descriptors are extracted either in the pixel or the frequency domain [68], [69].

2.2.4.1 What is an image descriptor?

An image descriptor describes the visual content of an image from a set of significant attributes or features of this image [70]. The feature extraction process uses an algorithm to extract these significant attributes from the image to be further manipulated and used in different imagery applications. This processing has a great impact on the quality of the classification result [68],[34].

Generally, image feature descriptors can be extracted either directly from the visual information in the pixel domain: color, texture, and shape, or in the frequency domain after image spectral transformation by using FFT, DCT, Gabor or wavelets [71]. There are many different image feature descriptors which are determined by the type of the used data and by the addressed application. For example:

1. Scene classification: the descriptor represents the structural characteristic of image such as flat surfaces, tiles, rectangular shapes [72].
2. Object recognition and tracking: the descriptor is related local geometric properties [73].
3. Texture image classification: the descriptor represents the surface microstructure properties (physical variation) in some image areas [68].

Texture image classification is mostly used in industrial surface inspection, remote sensing, material classification and biomedical image analysis [74], [75]. In the last decade, the combination between feature extraction and machine learning improves the classification performance.

2.2.4.2 Texture Feature

The texture feature is one of the most important visual descriptors in the field of image classification [76], pattern recognition and computer vision [77], [78]. To the best of our knowledge, in literature there is no unique definition or mathematical model to synthesize texture. The more general definition of texture in many languages is that texture refers to surface variations. The more clear definition that describes texture come from Hawkins [46]: "The notion of texture appears to depend upon three ingredients: (i) some local 'order' is repeated over a region which is large in comparison to the order's size, (ii) the order consists in the nonrandom arrangement of elementary parts and (iii) the parts are roughly uniform entities having approximately the same dimensions everywhere within the textured region". Typically, the surface texture is characterized by local pixel variations repeated in regular or random spatial distribution of intensity 'patterns' [65]. It provides useful information about spatial distribution and can be used for numerous digital imaging processes including image analysis, segmentation and classification [79], [65], [70].

In fact, the extracted texture feature must be invariant to image position, scale and rotation [66]. Traditionally, several methods have been studied in the literature to extract and characterize the texture feature descriptors [80], [81]. These methods can be categorized into four fundamental extraction methods [82]:

1. **Statistical Methods:** it derives a set of statistical features represents the related gray value distribution for each image pixel based on first order, second order or higher order statistics. Most widely used approach based on n co-occurrence matrices of gray level statistics.
2. **Geometrical methods:** it characterizes the geometric properties of textures and find

the rules that govern their spatial organization.

3. **Model-based methods:** estimates a parametric model based on the intensity distribution of the calculated texture descriptors.
4. **Signal processing methods:** it is exploiting the filtered image in frequency-domain bases on multi-channel Gabor filtering, Fourier transforms and wavelets including different scales and orientations.

For instance, statistical methods are based on the computation of mathematical indicators like eigenvectors (using Principal Component Analysis) or Fourier power spectra. Transform-based methods apply multi-resolution filtering, based on Gabor filters, Wavelet transform, or fractals [83], [84]. Several studies have been already proposed in the literature for image classification based on the combination of machine learning tools with the texture descriptors such as locally binary pattern (LBP) features [85], filter bank features [86], or co-occurrence matrix-based features [87]. Indeed, the success of a pattern recognition and classification problem is strongly related to the quality of the data and the features that characterize them. A good feature/image descriptor should reflect as much as possible the local content of the image. In a certain sense, the same constraint exists when performing predictive image coding: the so-called best candidate after intra or inter image prediction must best reflect the local content of the image to be coded in order to minimize spatial redundancy. In Chapter 5, we propose to apply classification in the compressed domain by considering intra prediction modes as texture image feature descriptors.

Support Vector Machine (SVM)

During starting this century, Machine Learning (ML) becomes one of the essential fields in computer science. It shapes our future by doing enormous and complex tasks over human capabilities [88]. Machine Learning (ML) trains computer to learn from the training data and to make predictions on other data by using expensive intelligence computation algorithms. The more clear definition for Machine Learning comes by Tom Mitchell “A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by improves with

experience E" [89]. Machine Learning categorizes into three categories: supervised learning (the label is known), unsupervised learning where the label is unknown and reinforcement learning where the system doesn't has any information about the datasets and that has to perform some specific processing to solve the classification problem [90]. Most of neural networks use at least two layers of neurons that they can approximate any function and perform automatic feature extraction without human intervention [91]. Well-known machine learning approaches are: (1) Artificial Neural Networks (ANN) which comprises of unites named (neuron). The neuron is the basic computational unit which is center of computation to produce the output signal using an activation function which controls input mapping to neurons output [135], [136]. The ANN basic architecture consists of three different layers: input, output and hidden layer [18]. ANN can handle very high-dimensional datasets [92]. (2) Convolution Neural Network (CNN) is a multi-layer perception, which includes three layers: the input layer, the output layer, and many functional hidden layers [93]. The main feature of using CNN is its ability to develop an internal representation of a two-dimensional image.

Support Vector Machine (SVM) is one of important supervised machine learning tools initiated by V. Vapnik in 1990. It has been proposed in many scientific classification fields, such as bioinformatics [94], medical diagnosis [95], environment monitoring [96], and material scientific classification and other many efficient technologies [97]. SVM uses training data (Features) to give the computers acknowledge without previous programming based on recent advances in statistical learning theory. SVM aims to find the optimal separation (Hyperplane) of a labeled dataset [88]. This separator can be a linear or nonlinear for any dimension data sets: it could be a line (2D), plane (3D), or hyperplane (4D+) [98].

2.1.8 Mathematical Linear SVM

The classification model assigns an input x pattern (set of features) to its correct label or category y based on N training dataset examples $(x_{i,}y_{i,})$ for i = 1 … to … N, with $y_i \in \{1,$ n} is the label of the sample $x_{i,} \in R^P$ [98]. SVM based on binary classification between two labels y ∈ {1 or -1}. It aims to find the separation between these two assigned classes. Figure

2-13 illustrates the best hyperplane (w^T+ b = 0) that separates the red data points for the category (y_i = -1), from the blue data points for the category (y_i = +1). The region which lies between the two hyperplanes w^T x + b = -1 and w^T x + b = +1 is called the Margin.

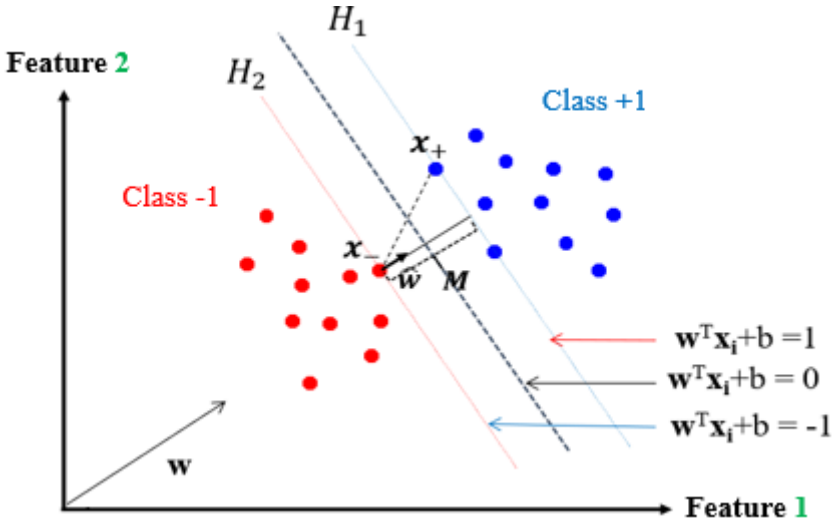


Figure 2-13 The optimal separation hyperplane (OSH).

The classifier can mathematically express as a separating hyperplane 'H' is parameterizing by a weight vector (w) and bias (b) and represented as:

Positive class: $w^T . X_+ + b \geq 0$ for all i with y_i = +1 (2.15)

Negative class: $w^T . X_- + b \leq 0$ for all i with y_i = -1 (2.16)

We can rewrite the above two equations to be equivalent to:

$$y_i(w^T x_i + b) \geq 1 \quad (2.17)$$

In order to calculate the margin, we subtract the two equations to get the distance between the two nearest class points (Support Vectors).

$$M = \frac{2}{\|\boldsymbol{w}\|} \tag{2.18}$$

So, we need to minimize norm ||w|| for maximizing the margin region.

The classical methods for quadratic programming such as quasi-Newton methods [99] or interior point methods [100] used to identify which training points x_i are the support vectors with non-zero Lagrangian multipliers (α_i). The non-negative value α_i ($\alpha_i \geq 0$) will correspond to the support vectors. For solving the optimization problem, we insert Lagrangian Multiplier (α_i) to determine the optimal αi to solve the optimization problem to find the optimal w and b.

$$L = f(x) - \alpha \,.\, g(x) \tag{2.19}$$

Where

$$f(x) = \frac{1}{2}\|\boldsymbol{w}\|^2 \tag{2.20}$$

$$g(x) = y_i(\boldsymbol{w}^T \boldsymbol{x}_i + b) - 1 \tag{2.21}$$

Now, we can formulate a quadratic programming (QP) optimization problem as:

$$\mathcal{L}_p(\boldsymbol{w}, \boldsymbol{b}, \propto) = \frac{1}{2}\|\boldsymbol{w}\|^2 - \sum_{i=1}^{n} \propto_i [y_i(\boldsymbol{w}^T \boldsymbol{x}_i + b) - 1] \tag{2.21}$$

After substituting the two equations 2.22 and 2.23.

$$w = \sum_{i=1}^{n} \propto_i y_i \boldsymbol{x}_i \tag{2.22}$$

$$\sum_{i=1}^{n} \propto_i y_i = 0 \tag{2.23}$$

We get the primal dual form:

$$\mathcal{L}_{pd}(\propto) = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \propto_i \propto_j y_i y_j \boldsymbol{x}_i \boldsymbol{x}_j + \sum_{i=1}^{n} \propto_i \tag{2.24}$$

And the results are $\propto = \propto_{\boldsymbol{1}}, \ldots \propto_{\boldsymbol{n}}$ and vector $\mathbf{w}$ can be calculated by eq. 2.22. Finally, based on Karush Kuhn Tucker (KKT) condition:

$$\propto_{\boldsymbol{i}} \left(y_i(\boldsymbol{w}^T \boldsymbol{x}_i + b) - 1 \right) = 0 \quad (2.25)$$

we can solve for the constant b.

$$b = y_i - \boldsymbol{w}^T \boldsymbol{x}_i \quad (2.26)$$

The classifier $\boldsymbol{f}$ **(x)** will return 1 (for +1 class) or -1 (for -1 class) for a given an unknown point u expressed as:

$$f(x) = sign\ (\boldsymbol{w}^T\ u + b) \quad (2.27)$$

Or $$f(x) = sign\ (\sum_{i=1}^{n} \propto_i\ y_i x_i u + b) \quad (2.28)$$

2.1.9 Nonlinear SVM

Most of the pattern classification problems are non-linear pattern distribution. SVM solves the non-linear classification problem, by increasing the dimensionality to find the optimal hyperplane in kernel space F [26] that made it a robust and powerful tool to solve many classification applications [101]. To do that, non-linear mapping is used to map the input data to high-dimensional space as:

$$R^P \quad \ni \quad x \quad \rightarrow \quad \Phi(x) \quad \in \quad F \quad (2.29)$$

Indeed, its complexity depends on the number of training samples, which increases the computation of the quadratic programming problem. While it does not depend on the kernel space dimensionality [102], [103]. The kernel (a specific feature transform) can be built based on the algebraic properties as (sum or multiplication). The kernel function computes the inner-product between two projected vectors to make new data form for each input sample pair x (i, j) as:

$$k \quad (x^i, x^j) \quad = \quad \Phi(x^i)^T . \Phi \quad (x^j) \quad (2.30)$$

For example, if we consider the input vector x = (x_1, x_2) in R^2 as presented in Figure 2-14 which can't linearly separate [103]. By using the following explicit Φ (x) = $(x_1^2, \sqrt{2x}_{1,}x_2, x_2^2)$ for each feature sample pair x (i, j) searching for separation in R^3 and the input vector x will be transformed in three dimensions as:

$$\Phi([x_1\ x_2]^{\mathrm{T}}) = \left[x_1^2\ \sqrt{2x}_1 x_2\ \ x_2^2\right]^T \quad (2.31)$$

$$\Phi\left(\left[x_1^{\mathrm{i}}\ x_2^{\mathrm{i}}\right]^{\mathrm{T}}\right)^T \quad \Phi\left(\left[x_1^{\mathrm{j}}\ x_2^{\mathrm{j}}\right]^{\mathrm{T}}\right)$$

$$= (x_1^{(\mathrm{i})^2} x_1^{(\mathrm{j})^2} + 2x_1^{(\mathrm{i})} x_2^{(\mathrm{i})} x_1^{(\mathrm{j})} x_2^{(\mathrm{j})} + x_2^{(\mathrm{i})^2} x_2^{(\mathrm{j})^2})$$

$$= (x_1^{(\mathrm{i})} x_1^{(\mathrm{j})} + x_2^{(\mathrm{i})} x_2^{(\mathrm{i})})^2$$

$$= (x^{(i)^T} x^j)^2$$

$$= \mathrm{k}\,(x^{(i)}, x^{(j)}).$$

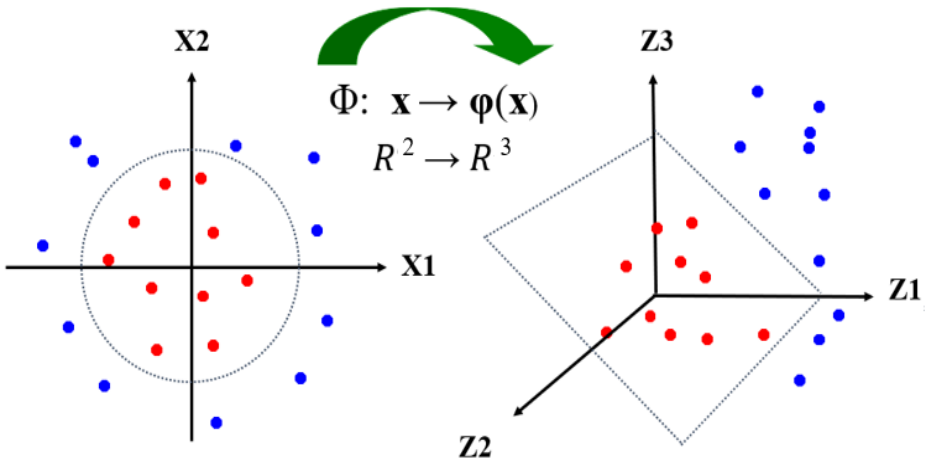


Figure 2-14 Transformation of the data set by Φ [104].

Now the non-linear input data becomes linearly separable by using non-linear mapping. The implementation of SVM for an unknown given point consists in solving the dual quadratic programming equation (2.32) to determine the maximum margin hyperplane.

$$\sum_{i=1}^{n} \propto_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \propto_i \propto_j \; y_i y_j \boldsymbol{x}_i \boldsymbol{x}_j \; \mathrm{K}\,(x_{i,} x_j)$$

(2.32)

Where the x_{i} are the training examples, n their number and $y_{\mathrm{i}} = \pm 1$ their respective classes, the α_{i} is the Lagrange multipliers to be determined, and K is the used kernel.

Moreover, many existing kernels explained in [12] including Fourier, splines, B-splines, additive kernels and tensor products.

In practice, when the number of the input vector is high, there are two problems arise: (1) the size of the kernel matrix becomes unbearable by the main memory, (2) the needed time for searching for optimal α_{i} becomes exhaustive.

Some SVM popular Kernels are:

1. Linear*:* If the data is linearly separable, we do not need to change space, and the scalar product is enough to define the decision function.

$$\mathrm{K}\,(x_{i,} x_j) = x_{\mathrm{i}}^{\mathrm{T}} x_j. \tag{2.33}$$

2. *Polynomial**:*** is a popular method for non-linear mapping. The polynomial kernel elevates the dot product to natural power to has more hyperparameters than the Radial Basis Function RBF kernel:

$$(x^{(i)}, x^{(j)}) = \; (\gamma + c x^{i^T} x^{(j)})^d \; . \tag{2.34}$$

Where: *d* is the polynomial degree, and (γ, c) are kernal tuning parameters.

3. ***Radial Basis Function (RBF):*** is finding the features separation into an infinite dimension as:

$$(x^{(i)}, x^{(j)}) = \exp(-\gamma \left\| x^{(i)}, x^{(j)} \right\|_2^2 . \tag{2.35}$$

Where γ is a parameter that sets "spread "of the kernel.

SVM kernel has two critical parameters (*C* and γ). Both parameters are used to control the over-fit weights and biases. Thus, before performing training or testing, it is essential to have the best values for *C* and γ. The tuning of the hyperparameter *C* is a delicate task. A

larger *C* implies a smaller training error, but also a smaller margin and possibly a lower generalization performance [105]. While, increasing the γ is increase the SVM Complexity [103].

The goal of any supervised machine learning tasks is to use the training set and validation set to minimize error measure evaluated on the test set which must be unknown for the model [12].

2.1.10 K-Fold Cross-Validation

Before training the system, we do a statistical method of evaluating and comparing learning algorithms (Cross-Validation). Cross-validation process is a common strategy to estimate the performance of different algorithm functions as (***Polynomial, RBF, or Linear SVM***). It finds out the best algorithm for the available data [106]. The training dataset has to be randomized to avoid overfitting. After that, the randomized training dataset is separated into *k* equal size bins. *k*-1 bins are used as the training data, while one bin is used as a validation dataset for testing the model. This process is repeated k times (the folds), where in each round one of the k bins is used as the validation dataset. This technique is called k-fold cross-validation. The k results from the folds can then be averaged to produce a single estimation as illustrated in Figure 2-15.

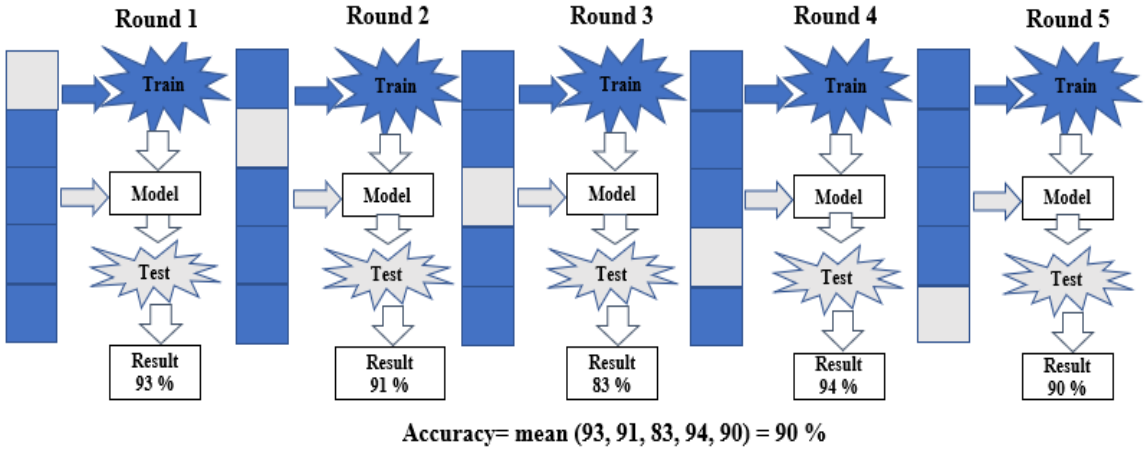


Figure 2-15 5-Fold Cross-Validation [106].

The advantage of this method is that all the observations are used for both training and validation, where each observation is used for validation exactly once. In this work, we will use Cross-Validation to obtain a better Kernel that matches to our topographical image datasets.

2.1.11 Multiclass SVM

The problems of the real world are in most cases multiclass. In such cases, a single hyperplane is not enough anymore. Support vector machine are designed initially to solve two-class binary classification. However, numerous improvements have been developed to make it possible to work in multi-class classification with two different approaches: One-Against-All (OAA), One-Against-One (OAO)[88], [95], [101]. Both techniques searching for solving non-linear classification problem of multi-class dataset.

2.3.4.1 One-Against-All (OAA)

This is the simplest and the oldest method according to Vapnik's formulation [107], for each class k it consists in determining a hyperplane $H_k(w_k, b_k)$ separating it from other classes. This class k is considered to be a positive class (+1) and the other classes as the negative class (-1) [108]. Where, we do it K times and all the decision functions are combined for computing the final decision function.

2.3.4.2 One-Against-One (OAO)

This method invented by **Kner** in [137] was proposed for neural networks. It is also known as "pairwise," based on using one classifier for each pair of classes. Instead of modeling K decision functions, K (K - 1)/2 decision functions (Hyperplanes) are needed to discriminate between these k classes. Then, the voting is used to determine which decision function to be selected. Several libraries for implementing an SVM are freely available on the internet, the most used one are SVMlight [17] and an open-source library for large-scale known as LIBSVM packages [109]. In our work, we will use One vs. One for solving our classification problem.

Conclusion

In this first chapter, we have presented the basic theoretical notions that will be used in the context of our interdisciplinary research. We first recalled the basics of digital image and video compression, and then briefly described the two main problems of materials engineering that we will address in the rest of our work. Finally, we introduced the SVM method since it will be further at the heart of one of our contributions. In the next chapter, we propose to focus on the state of the art HEVC (High Efficiency Video Coding) video compression standard that we will use to move from the pixel domain to the compressed domain.

CHAPTER 3

HIGH EFFICIENCY VIDEO CODING (HEVC)

High-Efficiency Video Coding (HEVC) is the current state-of-the-art digital video coding standard. It is based on a hybrid motion-compensated block-based transform video coding scheme. It includes spatial and temporal prediction, transform and error prediction quantization, as well as entropy coding [43]. Several improvements have been introduced to the main structure of HEVC compared to its predecessor H.264/AVC. HEVC allows roughly to double the compression ratio at the same perceptual quality but at the expense of increasing the computation complexity [43] [49].

Improvements in HEVC Coding Stages

HEVC standard works in a closed loop which integrates the encoder and the decoder. As illustrated in Figure 3-1, the already coded Prediction Unit (PU) should be available at the decoder to encode the current frame.

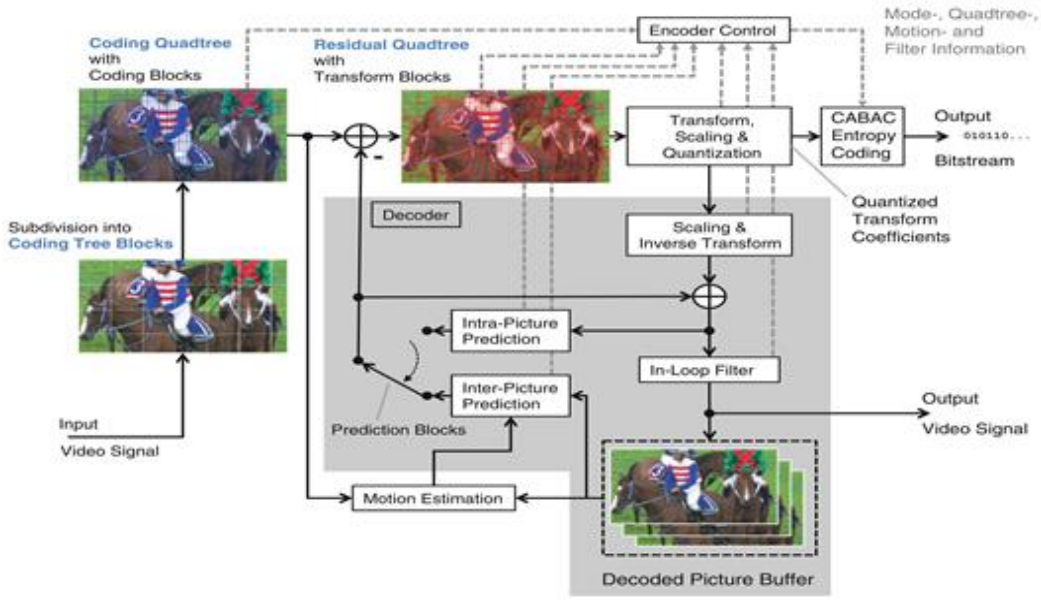


Figure 3-2 Structure of HEVC encoder and decoder (with elements shaded in light gray) [110]**.**

The main innovative features of HEVC are:

- **HEVC allows two data partition structure forms: slice and tiles** to facilitate the resynchronization if the data loss happens, and both can be decoded independently [41], [43], [49]. The image can be one or many slices that consists of sequences of Coding Tree Units (CTUs) for both Luma and Chroma components in raster-scan order. The tile was proposed to increase the capability for parallel processing [41], [43]. Tile is a rectangular data structure form that can be independently decoded that can be used as the region of interest (ROI) [111].
- **HEVC supports a large variety of block sizes** using the advanced block partition technique known as Coding-Tree Units CTUs. It is based on the quadtree representation to have blocks of multiple sizes ranging from 64x64 down to 8x8 samples using flexible recursive square or rectangular sub-partitioning mechanisms [112]. This variable coding size is adapted for images contents based on rate-distortion optimization (RDO). Three basic units are defined in HEVC for optimizing the codec performance for various application and devices with a significant increase of implementation complexity [113], those are;

- **Coding Unit (CU)** sets the necessary blocks size for the entire process with multiple transform sizes ranging from 8 × 8 to 64 × 64 pixels, adapted depending on the application. The large CU size provides a significant coding efficiency for higher resolution frames such as 4k and 8k resolutions. The CU can be divided into Prediction Units (PUs) with symmetric or asymmetric partitioning modes; where the Intra-prediction mode can be applied only for the first two symmetric partitions while the inter-prediction mode uses the other six available partitions as illustrated in Figure 3-2 [49], [112].

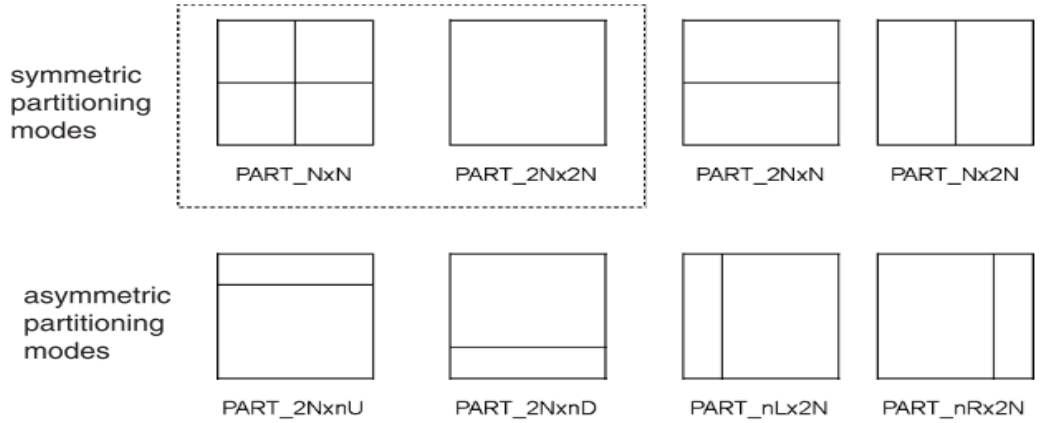


Figure 3-2 HEVC Intra/Inter partitioning modes of a CU to PUs [49], [112].

- **Predicting Unit (PU)** defines the size of possible partitioning for each decomposition level of the quadtree and its size varies depending on the type of prediction mode and the spatial details which creates the texture. The size of PU ranges from 4x4 to 64x64 pixels [41], [112], [114]. The choice of the prediction mode is done at the level of each PU. The set of the possible supported sizes is summarized below according to the type of the prediction: (1) Skip: 2N × 2N, (2) Intra: 2N × 2N, N × N and (3) Inter: 2N × 2N, 2N × N, N × 2N, 2N × U, 2N × D, nL × 2N and nR × 2N with the size 4 × 4 removed.
- **Transform Unit (TU)** sets the size of the transform and quantization processes applied to a prediction unit (PU) in a quad-tree structure. The size of the (TU) can be the same or smaller than CU size [115]. In Intra-prediction mode, (TU) is exclusively square block with multiple-size blocks. for Luma components, it could be 4x4, 8x8, 16x16 or 32x32 with two corresponding chroma block for 4:2:0 color format. In inter prediction mode,

(TU) has rectangular shape to avoid cases where the transform cuts a border of block and produces high frequency coefficients. It could be of sizes as 32 x 8, 8 x 32, 16 x 4, or 4 x 16 with corresponding chroma block [22]. Integer discrete sine transform (IDST) is used for the 4 x 4 Luma intra-prediction residuals instead of using the discrete cosine transform (IDCT). While, 4x4 (IDCT) is available for inter prediction for avoiding mismatch between the encoder and the decoder.

Now, we can specify the partition procedure of HEVC encoding. Firstly, the size of the Coding Tree Unit (CTU) and the depth of partitioning of CU and TUs both are defined. The partition sizes of the three units CU, PU and TU are then recursively determined during coding. Figure 3-3 presents the partition of Large Coding Unit (LCU) of 64X64 samples by a quadtree and by coding tree. CTU is partitioned to many CUs of 8x8 to 32x32.

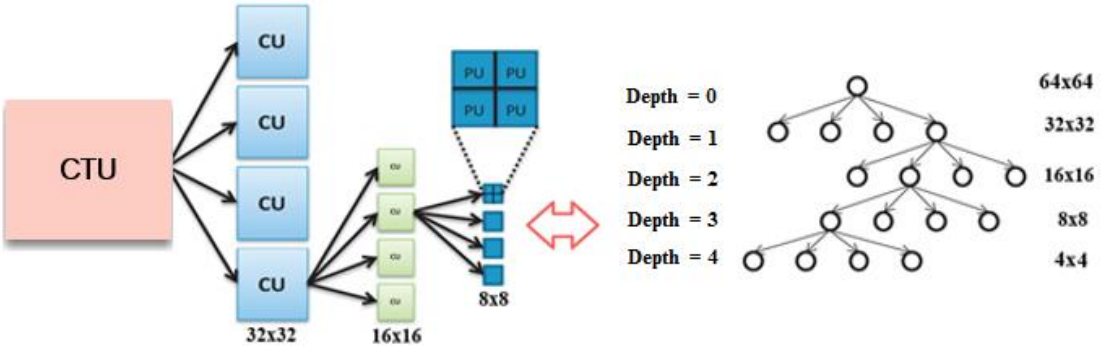


Figure 3-3 Example for the partitioning of a 64x64 coding tree unit (CTU) into coding units (CUs) with different coding depths.

- **HEVC performs the parallel encoding using Wavefront Parallel Processing (WPP) that** allows partitioning each frame slice into rows of CTUs. The CTUs for the first row should be initially processed. While starting CTUs processing for the other rows needs only to wait for encoding the first two CTUs of the previous row.
- **HEVC extends the number of intra prediction modes** to 35 modes instead of 9 modes in H.264/AVC with flexible block size partition ranging from 4x4 up to 64x64 samples, hence limiting the intra prediction residual error. This will be discussed in following section 3.2.
- **HEVC improves the motion vector precision in inter-prediction** mode by estimating specific region movements of particular parts such as rectangular blocks with variable block-size ranging

from 4x4 to 32x32, randomized parts, or even 1⁄4 or 1⁄8 pixel samples motion compensation [16], [41].

- **HEVC standard specifies two-dimensional transforms applied** for inter and intra prediction Luma residuals as; Integer Discrete Sine Transform (IDST) for intra 4x4 TU size and inverse discrete cosine transforms (IDCT) for other available TUs sizes [115], [116]. The integer-to-integer (i2i) approximation of the DCT transform was proposed to transform the prediction residual in HEVC and H.264/AVC lossless compression [117].
- **HEVC applies the uniform reconstruction quantization (URQ)** which use in H.264/AVC for available TU sizes [43].
- **HEVC applies In-loop Filtering;** two filtering stages within the inter-prediction loop; the deblocking filters (DBF) for avoiding the block artifacts before registering the reconstructed block into decoder. Followed by non-linear amplitude mapping filter stage known as sample adaptive offset (SAO) aiming to better amplitude construction for the original signal by using look-up table mapping.
- **HEVC applies Context Adaptive Binary Arithmetic Coding (CABAC)** algorithm to entropy encode all the syntax elements.

HEVC Intra Prediction Coding

Intra-prediction coding in HEVC is an extension of that previously used by H.264/AVC standard. Where both standards are based on block-wise spatial prediction within the same I-frame [41]. Intra prediction exploits in a very efficient way the spatial redundancy inherent in image contents. HEVC implements 35 intra prediction modes to improve the performance including: DC (Mode-1), Planar (Mode-0), and 33 angular (Mode-2…34) modes for all the square block sizes from 4x4 to 64x64 [114]. The luminance component (Y) can be predicted with available Intra-prediction modes; while the two chroma components (U, V) can be predicted with one of the five following modes; Planar, DC, horizontal, vertical and diagonal down right [118]. The Angular prediction mode is performed by extrapolating sample values from the left and upper boundaries neighbors of the reconstructed reference samples depending on 33 directional angles [41], [43]. The Angular prediction is categorized

into two classes: vertical and horizontal angular prediction, where both has negative and positive angles which is more suitable for coding the directional edges [41].

All the pixel samples in same PU are predicted with same angle value. Moreover, each angular mode has specific displacement parameter (***A***) that represents the pixel's displacement expressed in 1/32 fraction of accuracy as presented in Figure 3-4 (a).

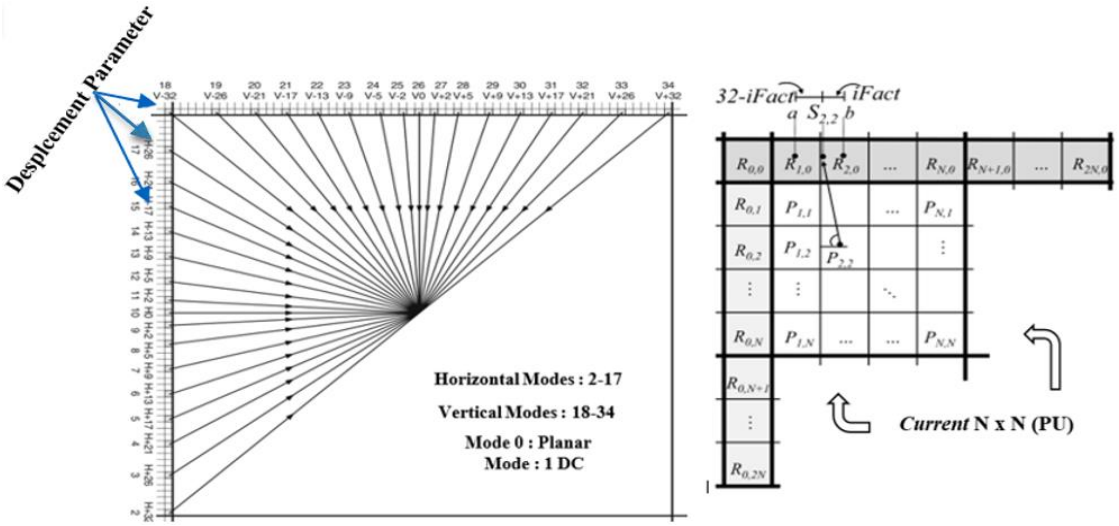


Figure 3-4 (a) HEVC intra-prediction modes (b) Prediction principle for 4x4 PU [119]**.**

These 33 displacement angular parameters increase the possibility for prediction the directional edge structure [41] [120] [110]. The value of the angular displacement parameter become to have the highest value for predicting the diagonal direction edges and become smaller for predict that edges which closed to vertical or horizontal direction. The available 33 Angular displacements represented relative to the mode used are resumed in Table 3-1.

Table 1-1 – Displacement Angle corresponding to Angular prediction Mode [41]**.**

Mode-Horizontal	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Angle	**32**	**26**	**21**	**17**	**13**	**9**	**5**	**2**	**0**	**-2**	**-5**	**-9**	**-13**	**-17**	**-21**	**-26**	**-32**
Vertical Mode	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	
Angle	**-26**	**-21**	**-17**	**-13**	**-9**	**-5**	**-2**	**0**	**2**	**5**	**9**	**13**	**17**	**21**	**26**	**32**	

F

igur

e 3-4 (b) shows the previously decoded boundary samples which are located at upper and left of current prediction unit PU. They are used as reference samples expressed by ($\boldsymbol{R_{0,0}}$, $\boldsymbol{R_{1,0}}$,.... $\boldsymbol{R_{2N,0}}$ Vertical-prediction) and ($\boldsymbol{R_{0,1}}$,$\boldsymbol{R_{0,2}}$... $\boldsymbol{R_{0,2N}}$ Horizontal-prediction). The availability of the neighboring reference samples has to be checked. In two cases it might not be available: (1) when the reference samples lie outside of the picture and do not belong to the same slice as the current block. In this case, a nominal average sample value is used as a replacement for the reference samples depending on the sample bit-depth (8-bit video, this is $= 2^{Bd-1}$ = 128) [41]. (2) When just the left or the top boundary references are missing, in this case these reference samples are filled with the nearest available reference sample value, by generated copying the samples from the closest available reference samples above or on the left respectively. Figure 3-5 illustrates the difference for choosing the reference Index in both positive or negative angle for both vertical and horizontal angular prediction.

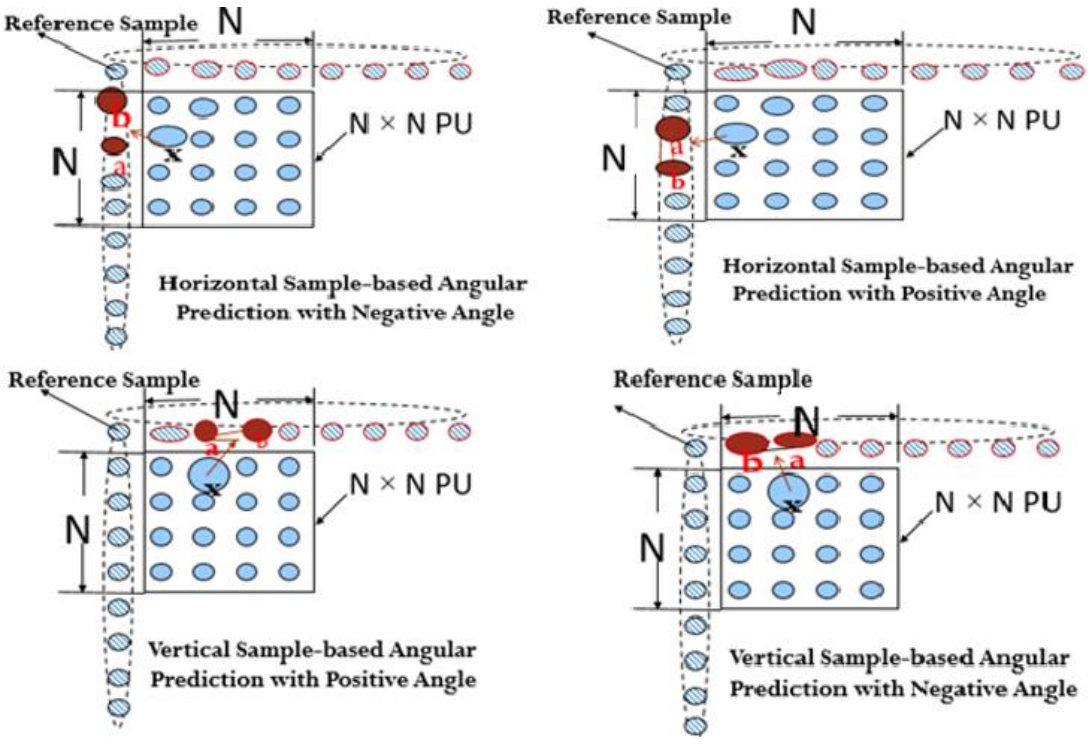


Figure 3-5 The reference sample locations relative to the current sample for Horizontal and Vertical angular intra prediction (with positive and negative prediction angles) respectively (the idea is [121] **)**

The difference between choosing the reference sample index for both positive and negative angle is listed below:

-In positive-angle case; the reference sample (a) is directly indexed by reference picture (Index). While (Index+1) refers to the reference sample (b).

-In negative-angle case; the reference sample (a) is directly indexed by reference picture (Index). While (Index-1) refers to the reference sample (b).

All the reference sample construction conditions are shown in following two figures.

These reference samples used for the prediction may be filtered by 2 or 3 taps smoothing filter. Eq. 3.1 is used to calculate the value of predicted sample P (x, y) by using linear interpolation of reference samples

$$\boldsymbol{P_{x,y}} = ((32 - \text{iFact}) * a + \text{iFact} * b + 16) >> 5 \quad (3.1)$$

Where;

- x, y are the spatial coordinates.
- $\boldsymbol{P_{x,y}}$ is the current predicted sample.
- A is displacement tangent with resolution of 1/32 which ranges from -32 to +32.
- iFact is the distance between two reference sample.
- a and b are the reference samples for current predicted sample which identified by a reference picture Index depends on the intra prediction angle.
- iFact = (x · A) & **31** Horizontal mode & = (y · A) & **31** for vertical mode and the Index = y+ iFact.
- 32 – iFact = (x· A) >>5 Horizontal mode & = (y· A) >>5 for vertical mode and the Index = x+ iFact.
- >> denotes a bit shift operation to the right.
- & denotes logical AND operation.

- **Planar Prediction Mode :**

Planar prediction is a multi-directional prediction process based on interpolating each PU sample from the right column and the bottom line within the current PU. The reference

samples are substituted respectively by the pixels at the top right reference sample ($\boldsymbol{R_{N+1,0}}$) and bottom right reference sample ($\boldsymbol{R_{0,N+1}}$) of the PU as illustrated in Figure 3-6.

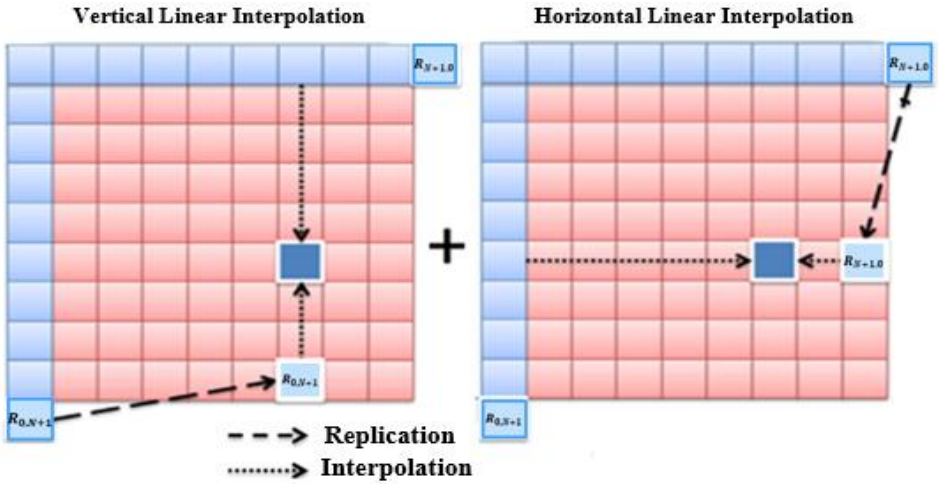


Figure 3-6 Representation of the Planar prediction.

This mode is more suitable for textured areas than classical Intra prediction. It has proven its efficiency to prevent discontinuities along the PU boundaries. The value of predicted sample P (x, y) is obtained by taking the average of vertical and horizontal linear interpolation calculated by the following Eq. 3.4

$$P^{H}_{x,y} = (N-x) * R_{y,0} + x * R_{0,N+1} \quad (3.2)$$

$$P^{V}_{x,y} = (N-y) * R_{0,X} + y * R_{N+1,0} \quad (3.3)$$

$$P_{x,y} = \text{floor}\left(\left(P^{V}_{x,y} + P^{H}_{x,y} + 16 / (\log2(N) + 1)\right)\right) \quad (3.4)$$

- **DC Mode**

This mode is more suitable for smoothly-varying region coding where the Prediction Unit (PU) is filled with the average of all 2N+1 neighboring reference samples as indicated in Eq. 3.5

$$P_{Dc} = \frac{1}{2N}\left(\sum_{x=1}^{N} Rref_{x,0} + \sum_{y=1}^{N} Rref_{0,y}\right) \quad (3.5)$$

Lossless Coding

HEVC has three special modes to perform the lossless coding: lossless mode, I_PCM mode and transform skip mode [41]. (1) Lossless mode can be enabled or disabled by the high-level flag contained in the picture parameter set (PPS). When this mode is active, the codec stages including: transform, quantization, and in-loop filters are skipped. (2) I_PCM mode; can be enabled or disabled by the pcm_flag. It bypasses the following codec stages: transform, quantization, entropy coding and in-loop filtering. In this mode, the video Intra-frame is directly coded with the specified PCM bit depth and transmitted directly to the bit stream like H.264/AVC. (3) Transform skip mode; the transform stage is only skipped for 4x4 Transform Units.

Many several researches have been proposed to improve HEVC lossless compression concerning Intra-prediction coding, transformation and entropy coding [120]. The spatial redundancy was exploited using block-based intra or inter prediction, then the prediction residuals are entropy coded [43], [113], [114], [120], [122]. Differential Pulse Code Modulation (DPCM) is the most used sample-wise prediction method for lossy and lossless compression. The residual differential pulse code modulation (RDPCM) has been proposed for encoding the residual of both vertical and horizontal prediction modes for H.264/AVC [122] and HEVC Version-2 [113]. Sample-based angular intra prediction (SAP) has been proposed in [121]. SAP uses the predicted sample neighbors as the reference sample for the future predicted sample for all angular-prediction modes while authors in [122] proposed a similar idea to improve the H.264/ AVC lossless compression. Different SAP-based encoding where proposed as SAP-HV which implements only for vertical and horizontal modes [123]. SAP-E implements all the angular modes and DC mode [119]. Some researches discussed the invertible transformation for improving lossless compression as: the integer-to-integer (i2i) of the discrete cosine transform (DCT) [117], and integer discrete cosine transform IDST transforms [116].

High Bit Depth Still Picture Coding

From the first time, HEVC v1 designs a specific profile for still picture and synthetic images compression application. HEVC v1 has included three primary profiles; Main 8-bit and Main 10-bit for video coding. Also, Main Still Picture (MSP) Profile with sample format limited to 4:2:0 for 8-bits depth application [43]. MSP profile has proven superior performance for bit-rate saving average ranging from 10% relative to VP9 and more than 44% relative to baseline JPEG [124]. In order to cover a broad range of video requirements, HEVC defines three profiles have introduced by RExt for high bit depth image coding [113]; HEVC Main 4:4:4 16 Still Picture (MSP) profile, and two video coding profiles use both of inter and intra coding; Main-RExt (main_444_16_intra) and High Throughput 4:4:4 16 Intra [125]. In totally, RExt introduces 21 new profile and several additional coding tools for different specific application aiming to reduce the processing expensive [113] In our case, we consider the available HEVC high-throughput profiles allow 8-bits depth up to 16-bits for 4:4:4 16 still-image lossless intra compression, implemented with the used HEVC reference software HM 16.12. The block size has been fixed to 4x4 blocks to have the most excellent analysis size for all available directional edges.

Conclusion

In this chapter, we proposed a brief overview of the state-of-the-art HEVC standard and introduce the new coding tools related to this standard. In particular, we focused on HEVC Intra prediction technique as we will mainly consider All Intra video coding in the following work. We gave a comprehensive technical description of this specific coding tool with details on the intra prediction modes used. The HEVC Lossless coding profiles have been also presented. Indeed, some mechanical applications could require guaranteeing that the reconstructed image after decompression is mathematically identical to the original one in order to avoid that any coding artifacts disturb further mechanical analysis. In the following chapter, we will present the first contribution of our work which concerns the strain field measurement from very high frame rate video sequences.

CHAPTER 4

PERFORMANCE EVALUATION OF STRAIN FIELD MEASUREMENT BY DIGITAL IMAGE CORRELATION USING HEVC COMPRESSED ULTRA-HIGH-SPEED VIDEO SEQUENCES.

Context of the study

Visual information constitutes nowadays one of the most dominant channels for acquiring, processing and communicating information in many sectors including entertainment, medicine, meteorology, transportation systems, or physics [20], [126]. In particular, mechanical engineering generates a huge amount of image media to be processed and stored for further use. For instance, this includes high spatial resolution topographical still images or ultra-high-speed video imagery for material crash analysis [127]. However, it is common in mechanical engineering to store raw visual data without applying any compression at all, or lossless compression only (using solutions like pkzip or tar). But lossless compression leads to limited compression ratios and lossless coded data still require large storage devices (several tens or even hundreds of terabytes). An alternative solution to significantly increase the coding efficiency could be to apply lossy compression algorithms such as JPEG or JPEG2000 for still images, or H.264/AVC or HEVC for video sequences [16], [41]. Unfortunately, compression artifacts introduced by these algorithms not only affect the visual quality of an image but can also distort the features that one computes for subsequent tasks related to image analysis or pattern recognition. Since imaging technologies are widely used to analyze the mechanical properties of materials, considering the image quality is essential. The reason behind is that the central paradigm of materials properties is largely dependent on microstructures which don't accept relevant differences between the original and reconstructed data. Hence, only lossless or nearly lossless compression techniques should be considered. In this chapter we will evaluate the impact of HEVC lossy and lossless compression on characterizing material mechanical response (in terms of strain field). Also, we will test the influence of image compression on other kinds of mechanical response.

Methodology

In order to satisfy the objectives of this study, we considered images previously recorded to characterize material mechanical behavior when subjected to severe loading conditions over a wide range of strain rate. To measure displacement and strain fields over the specimen surface Region of Interest (ROI) throughout loading, a high-speed video camera (up to 25,000 fps) is used. Then, the obtained image sequences are post-processed by Digital Image Correlation (DIC) software to extract the strain fields [128]. We propose to compress these images sequences using HEVC, and to evaluate the impact of HEVC compression on the performance of the subsequent mechanical analysis.

Methods and Materials

We test the influence of image compression on two kinds of mechanical responses (crash and impact). The strain measurement is a key point for evaluating the impact of compression on the recorded DIC image sequence. In this section, we will give a brief overview on the applied image processing techniques, namely DIC software analysis (VIC 2D) and HEVC compression. In Chapter 2, we discussed Digital Image Correlation technique (DIC) as a useful tool for measuring displacement and in-plane strain fields. We analyze the mechanical loading response by Vic-2D after applying two different compression techniques of HEVC (lossless and lossy). Also, we evaluate the efficiency of the proposed compression techniques by considering the compression ratios as well as the quality of the reconstructed video.

4.1.1 High-speed test device

Two kinds of high-speed tests, namely tensile test of a polypropylene and shear Arcan test of a glue, were carried out at room temperature using an Instron 65/20 hydraulic tensile device (i.e. 65 kN load cell sensor, maximum speed 20 m.s-1). For the present tests on polymeric materials, a piezoelectric load cell sensor, calibrated in the range 0-5kN, with a precision of 2.5 N, was fixed on the rigid frame of the device. High-speed camera Photron FASTCAM-APX RS was used.

Images were analyzed using DIC software VIC 2D [127]. The specimen ROI is selected manually and subdivided into several subsets. Subset size and step size (distance between two centers of facets) are selected in order to optimize the signal vs noise ratio. Details are given hereafter.

The idea is based on DIC operation and the main steps are [62]:

1. The random pattern of the reference image is divided into square facets of size 18x18 pixels, each of them being characterized by a unique signature in gray level (figure 4-1).
2. This unique signature allows the tracking of facets by DIC software, using a correlation algorithm. The software compares between the recorded images at a given loading and the reference image.

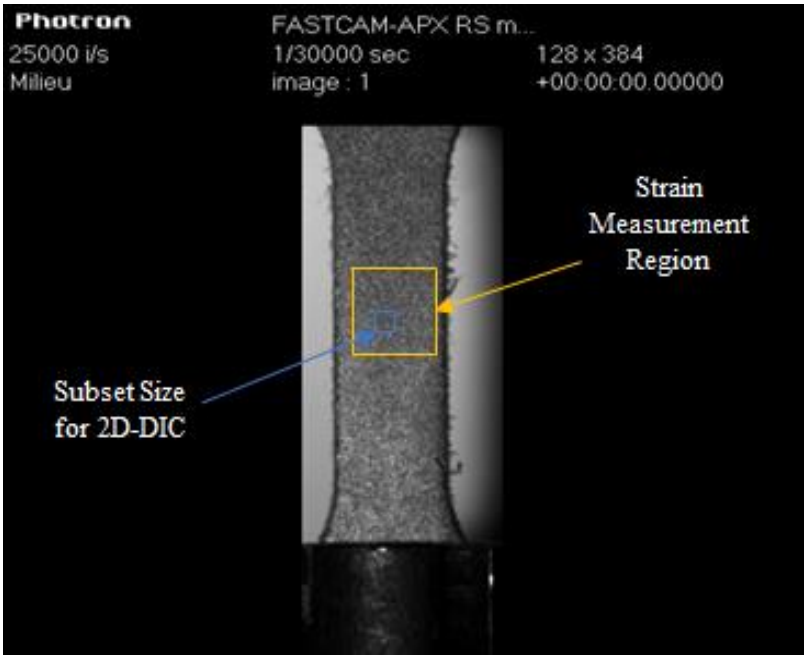


Figure 4-1 Reference image for 2D-DIC specimen measurements with Subset size of 18×18 pixels.

The Local strain measurement means that true strain components are acquired over small sub-surfaces compared to the whole ROI. By following this way, all strain field heterogeneities (strain localization, gradient…) can be known and thus providing an enriched information on material behavior.

4.1.2 HEVC Lossy and Lossless Compression

In this study, we compressed both deformed video sequences of materials under testing captured by ultra-high-speed camera using HEVC standard. The JVT reference software HM version 10.1 was used for HEVC compression configured for Main Profile with a Clean Random Access. The Group of Picture (GOP) size was set to 8 pictures (B- Frames) combined with an Instantaneous Decoding Refresh (IDR) picture (I-Frame). The coding order was set to 0, 8, 4, 2, 1, 3, 6, 5, and 7, while the Reference Frames was equal to 4. The target quantizer is variable, with a Quantization Parameter QP = 0, 5, 12, 17, 20, 22, 25, 27, 32, and 37. The Intra Period was set to (-1) which implies that only the first frame will be coded as Intra. Table 4.1 summarizes the HM reference software encoder configuration.

Table 1-2 HM 10.1 Encoder Parameters

Coding Options	Chosen Parameter
Encoder Version	HM 10.0
Profile	Main
Reference Frames	4
R/D Optimization	Enabled
Motion Estimation	TZ Search
Search Range	64
GOP	8
Hierarchical Encoding	Enabled
Temporal Levels	4
Decoding Refresh Type	1
Intra Period	-1
Deblocking Filter	Enabled
Coding Unit Size/Depth	64/4
Transform Unit Size (Min/Max)	4/32
TransformSkip	Enabled
TransformSkipFast	Enabled
Hadamard ME	Enabled
Asymmetric Motion Partitioning (AMP)	Enabled
Fast Encoding	Enabled
Fast Merge Decision	Enabled
Sample adaptive offset (SAO)	Enabled
Rate Control	Disabled
Internal Bit Depth	8

To turn the encoder into the lossless compression mode, the Lossless Coding configuration parameters are enabled causing the transformation, quantization, and all the in-loop filtering operations to be bypassed [14].

Results

In this Section, we will discuss the obtained results for the impact of HEVC compression on DIC image sequences while computing the in-plane strain fields at different loading speed during two mechanical tests: uniaxial tensile test and Arcan shear test.

4.1.3 Tensile Test of Polypropylene (PP) Specimen

In the present study, DIC technique is applied to measure displacement/strain field during dynamic uniaxial tensile loading of a polypropylene (PP), at a displacement rate of 1m/s at room temperature. Camera frame rate is fixed at 25,000 im/s. 71 images of 512x472 pixels are recorded during the loading, up to specimen failure (test duration of 2.84 μs, nominal axial strain at break of about 6%). It is worth noting that the useful part of the frames, i.e. corresponding to specimen image is of size128x384 pixels while the other parts of the frames are composed of black background or white text, added by the software during image recording (Figure 4-2).

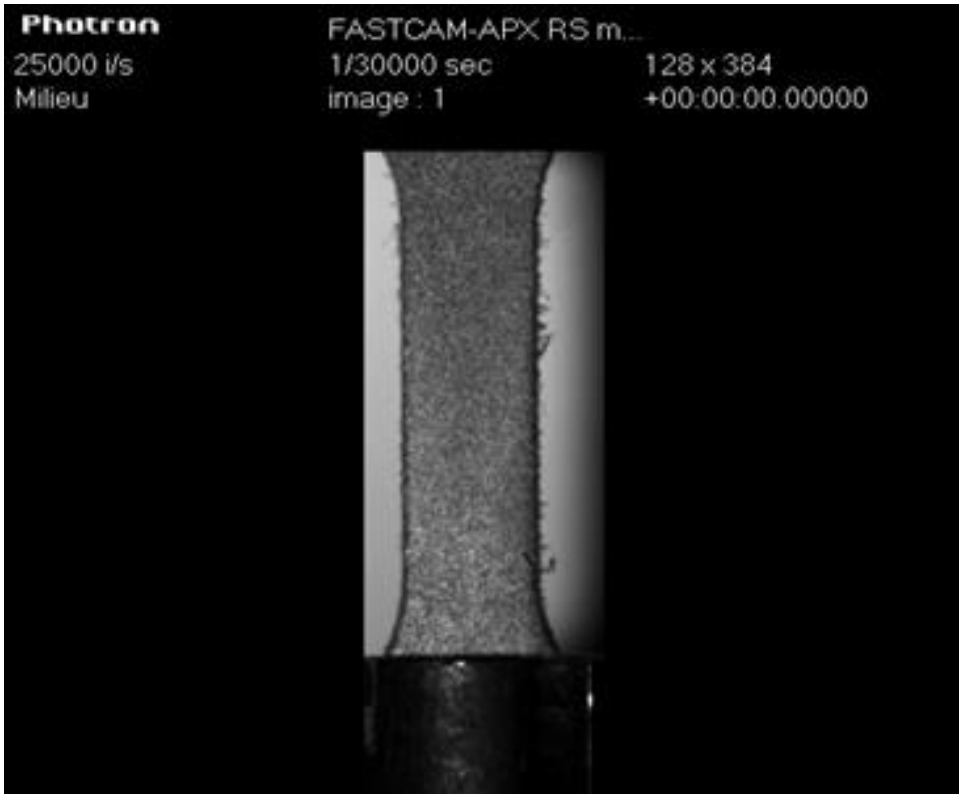


Figure 4-2 First image at undeformed stage of dynamic PP tensile specimen (global size of 512x472 pixels vs. useful part of 128x384 pixels).

In uniaxial tensile tests, the longitudinal and transverse strains are the two main components of the in-plane stress tensor, while shear stress remains very low.

4.1.4 Sikapower Arcan test

In order to test the influence of image compression on other kinds of mechanical response, DIC image sequences recorded during Arcan test at 45° of Sikapower® polymeric glue were also considered (Figure 4-3). The experimental setup for Sikapower Arcan test was similar to the previous setup used in Tensile Test of Polypropylene (PP) Specimen.

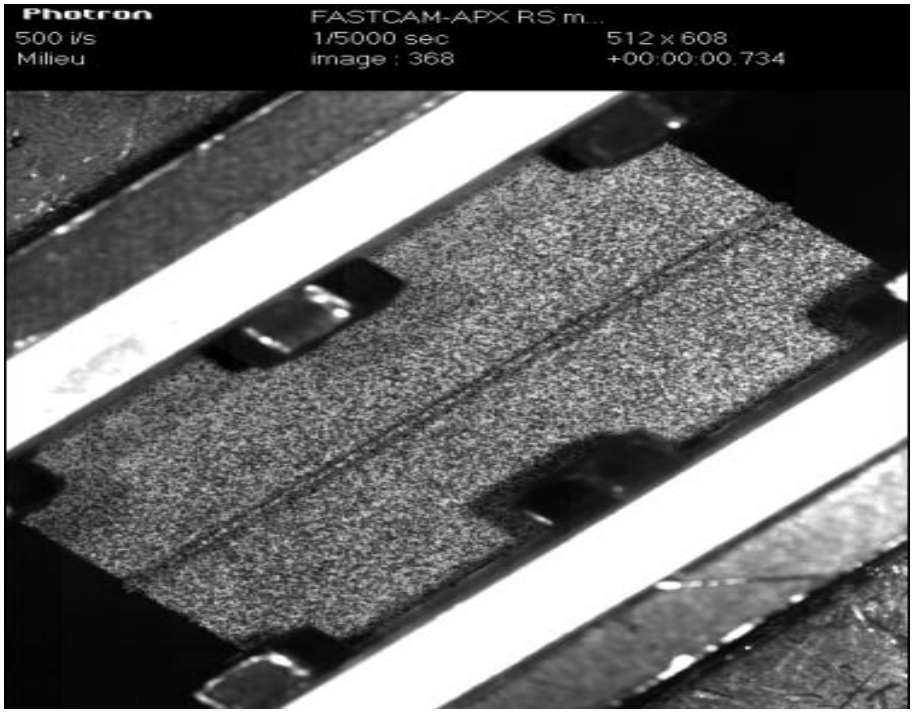


Figure 4-3 First image at undeformed stage of Arcan test at 45° of Sikapower glue joint.

The registered sequence during Sikapower Arcan test is composed of 150 frames of size 512x696 pixels captured at frame rate 500fps

4.1.5 Discussion

These sequences are firstly compressed by the means of HEVC. the Bit-rate as well as the quality of the reconstructed video sequences are both evaluated. Peak Signal-to-Noise Ratio (PSNR) expressed in dB and Structural Similarity Index Measurement (SSIM) are used for video quality evaluation. The SSIM metric varies between 0 (bad quality) and 1(perfect). It is common in broadcast applications to consider that a PSNR value higher than 35dB which corresponds to an excellent video quality. Table 4.2 summarizes the results obtained for the two sequences noted Sequence1 for tensile test of polypropylene (PP) specimen and Sequence2 for Sikapower Arcan test in terms of compression ratio as well as reconstructed video quality. For lossless coding, the achieved compression ratio for both sequences are 12:1 and 3.5:1, respectively. For lossy compression, the video sequences are compressed with a compression ratio varying from 12:1 to 1600:1 for Sequence1, and from 3.5:1 to 800:1 for

Sequence2. These differences are directly related to the spatio-temporal characteristics of video contents.

Table 1-3 HEVC Compression Performances

Test Sequence	PSNR (dB)		SSIM		Compression Ratio	
	Seq 1	Seq 2	Seq 1	Seq 2	Seq 1	Seq 2
Lossless	Inf	Inf	1	1	12.19:1	3.52:1
Lossy_QP0	76	67	1	0.9999	12.15:1	3.64:1
Lossy_QP5	60	54	0.9996	0.9987	17:1	5.6:1
Lossy_QP12	53	47	0.9983	0.9961	41:1	16.6:1
Lossy_QP17	50	45	0.9969	0.9944	77:1	33.2:1
Lossy_QP20	47	43	0.9950	0.9917	148:1	60:1
Lossy_QP22	46	41.3	0.9936	0.9900	226:1	80:1
Lossy_QP25	44	39	0.9908	0.9853	416:1	114:1
Lossy_QP27	43	37.7	0.9884	0.9819	533:1	133:1
Lossy_QP32	39.9	33.4	0.9770	0.9647	1143:1	267:1
Lossy_QP37	37	29.4	0.9562	0.9278	1600:1	800:1

The PSNR decreases to minimal values of 37 dB and 29.4 dB for Sequence1 and Sequence2, respectively. The SSIM values reach the minimal values of 0.9562 and 0.9278 for Sequence1 and Sequence2, respectively. The Rate Distortion (R-D) curves for the two video sequences are shown in Figure 4.4.

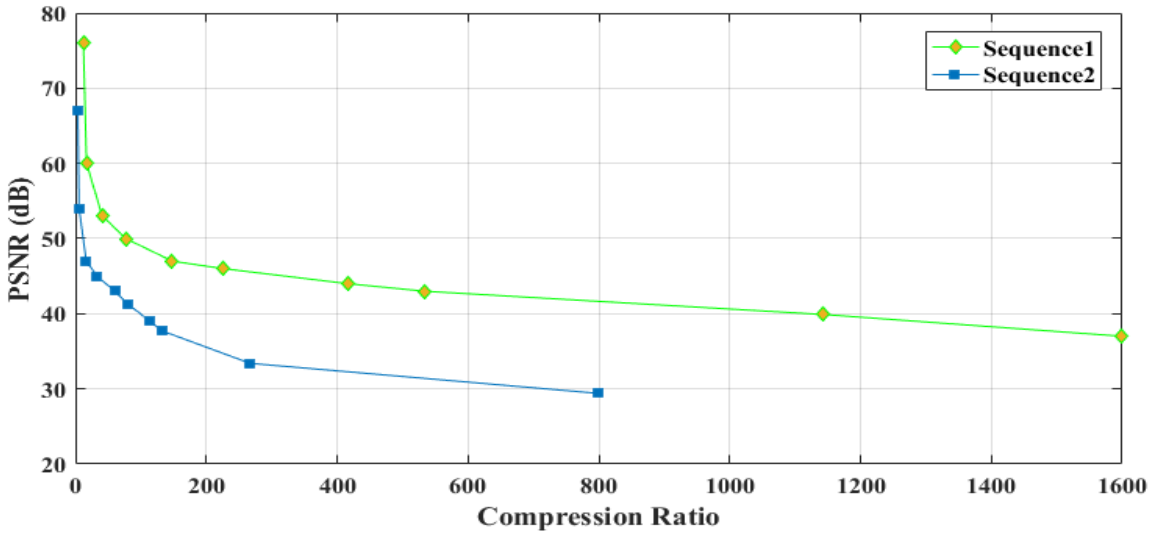


Figure 4-4 R-D curves for the two video sequences.

The average BD-BR saving of Sequence1 relative to Sequence2 is -88.3(%), which indicates the required overhead in bit rate to achieve same PSNR values. The average value of BD-PSNR indicates an increase of 13 dB for Sequence1 at the same bit rate. Therefore, the PSNR and SSIM values correspond to high visual quality levels as illustrated in Figures 4-5 and 4-6 for Quantization Parameter (QP) value =25.

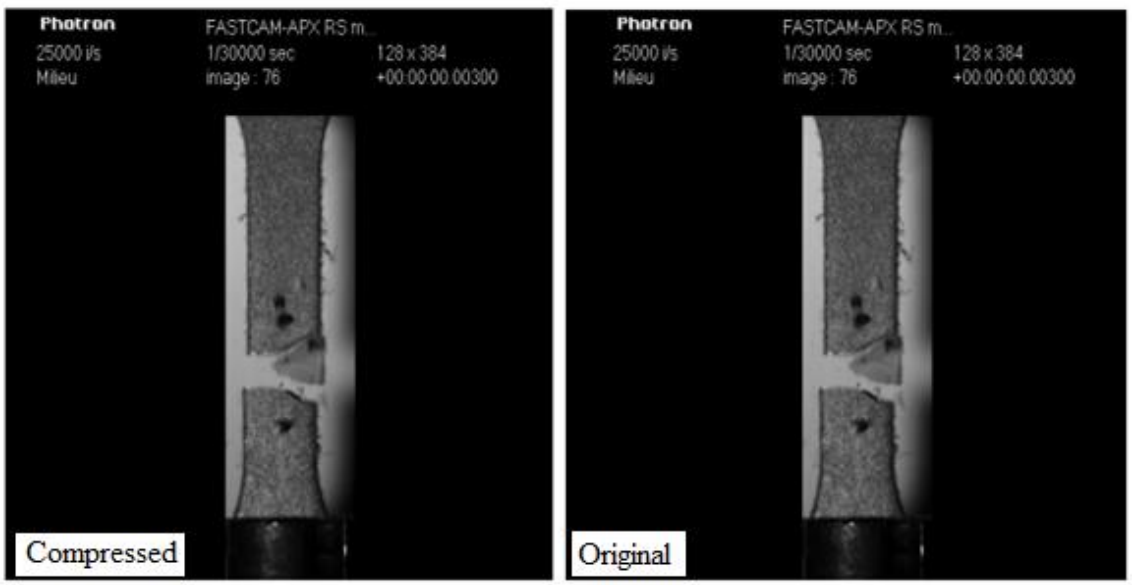


Figure 4-5 Illustration of HEVC high quality performances for compressed sequence1 (QP=25, PSNR =44.4dB and SSIM=0.99) compared with the original sequence1.

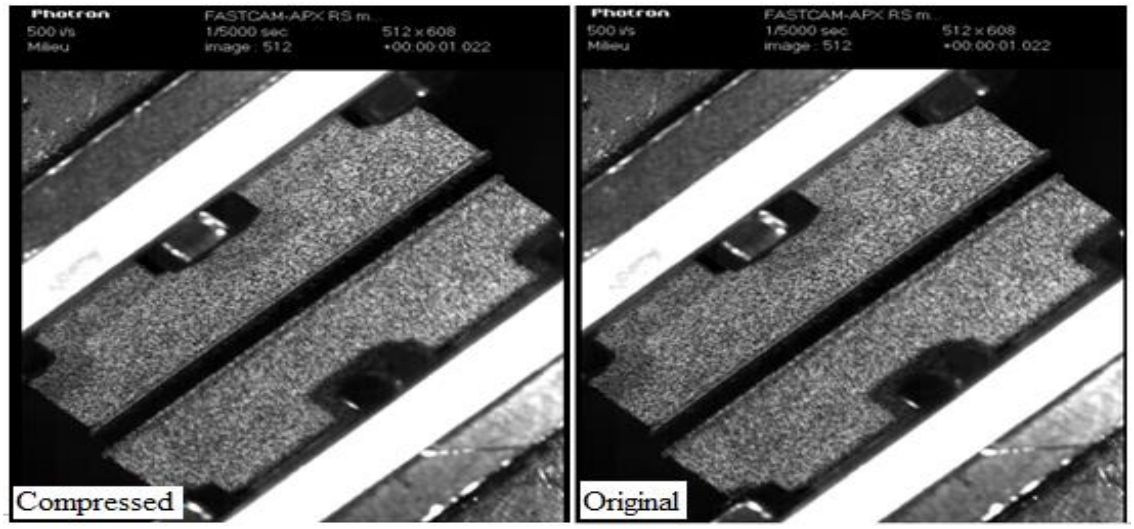


Figure 4-6 Illustration of HEVC high quality performances for compressed sequence2 (QP=25, PSNR =39.2dB and SSIM=0.98) compared with the original sequence2.

Moreover, the variation in SSIM between 0.9853 and 1.0 can be considered as irrelevant. Hence, HEVC compression preserves very well the structural properties of the image contents which are crucial for further mechanical analysis. To illustrate this point, for lossless sequence and all available QP values, true in-plane strain fields were computed using DIC Software (VIC-2D) based on HEVC compressed images for both Sequence 1 (i.e. dynamic tensile test of PP) and Sequence 2 (i.e. Arcan test of glue joint).

These computed In-plane strains are compared to those obtained from initial uncompressed image sequence. It is worth noting that strains computed from Lossless image sequence are identical to reference strains computed from initial uncompressed images for both Sequences 1 and 2. Dealing with Sequence 1, Figures 4.8 to 4.15 show examples of relative gaps ($\Delta\varepsilon_{yy}$) between the computed true axial strain by DIC technique based on a given Lossy sequence (QP=0, 5, 20 and 25 in those Figures), ε_{yy}^{L}, and the true axial strain computed from uncompressed images, ε_{yy}^{ref}, with :

$$\Delta\varepsilon_{yy} = \frac{\varepsilon_{yy}^{L} - \varepsilon_{yy}^{ref}}{\varepsilon_{yy}^{ref}} \qquad (4.1)$$

Relative gaps are computed at each stage of loading, i.e. for each recorded image (time step of 0.04 µs) and on 18 selected centers of DIC facets. Those points are all located at the half width of the ROI but with ordinates (i.e. position along specimen axis) varying by step of 10 pixels, point 1 being the upper point on the ROI (i.e. the closest point to the top jaw) and point 18 being the lowest point on the ROI (i.e. the closest point to the bottom jaw, at 170 pixels from point 1, approximately 20 mm). First, it appears that relative gaps, $\Delta\varepsilon_{yy}$, can be very important in the first stages of loading. Yet, this trend must be analyzed carefully since in the first stage gaps are computed considering very low strain levels, leading to possibly high value of relative gaps even for acceptable variation of strain value. In addition, it must be highlighted that reference axial strain (ε_{yy}^{ref}) are significantly affected by noise of measurement in the early stage of loading, again because of very low strain value. At higher strain level, relative gaps tend to stabilize at significantly lower value. Results show that relative gaps increase when increasing the value of QP. However, they always remain lower than 10% (in absolute value) at all considered specimen locations, except for QP=25 where higher gaps are computed for a few points. It can be noted that gaps tend to increase when the distance to ROI center (near point 9) increases. It can be partially explained by the fact that strain localization appeared at the center of the specimen, that is to say that reference axial strains were higher in specimen center than in other areas of the ROI. Then, for a same gap in strain value, relative gaps are lower in ROI center. The same trends are noticed for the transverse strains. Considering that relative gaps inferior to 10% (in absolute value) are satisfying in terms of accuracy of computed strain data, HEVC compression algorithm can therefore be used with QP value up to 20, providing a very interesting bitrate 0.05bpp on this particular image sequence. In some loading stages, the low quality of the created random pattern on that specimen surface makes VIC 2-D software is not able to track the displacement between the original and the deformed images.

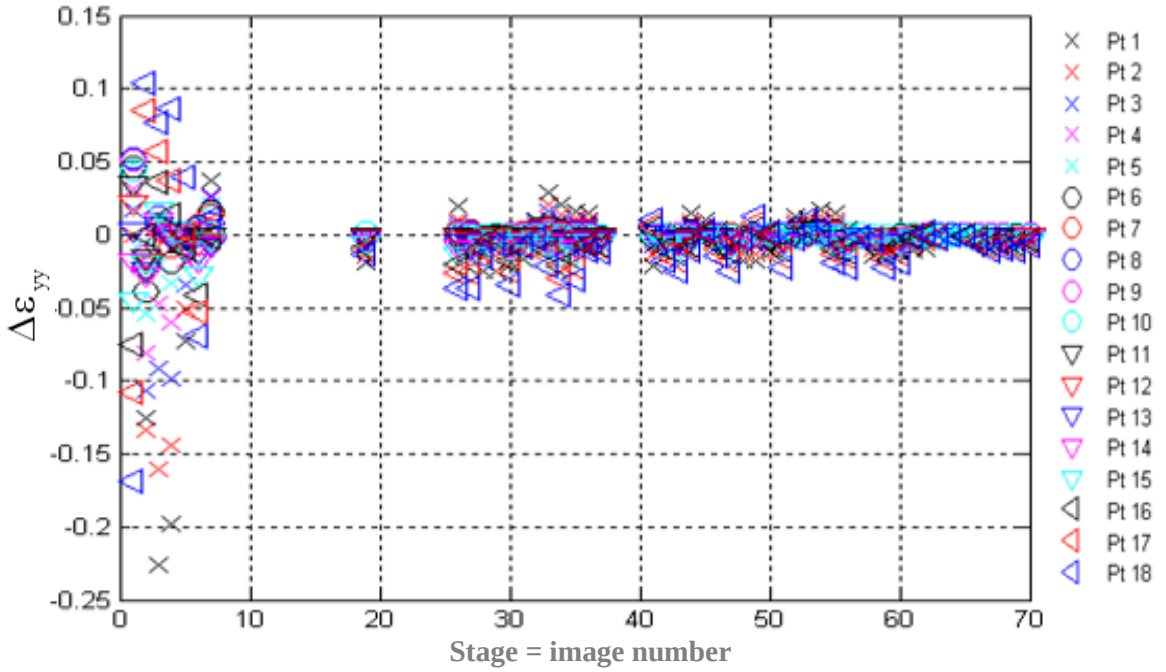


Figure 4-7 Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP0 (All data)

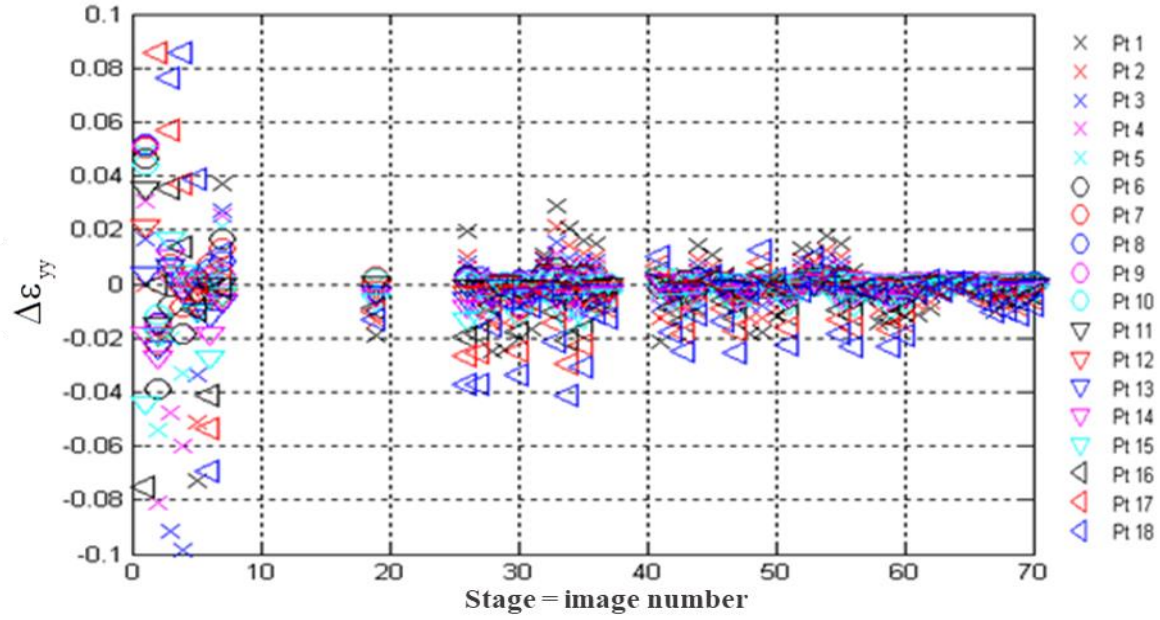


Figure 4-8 Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP0 (Focus on relative gaps between -10% and 10%.

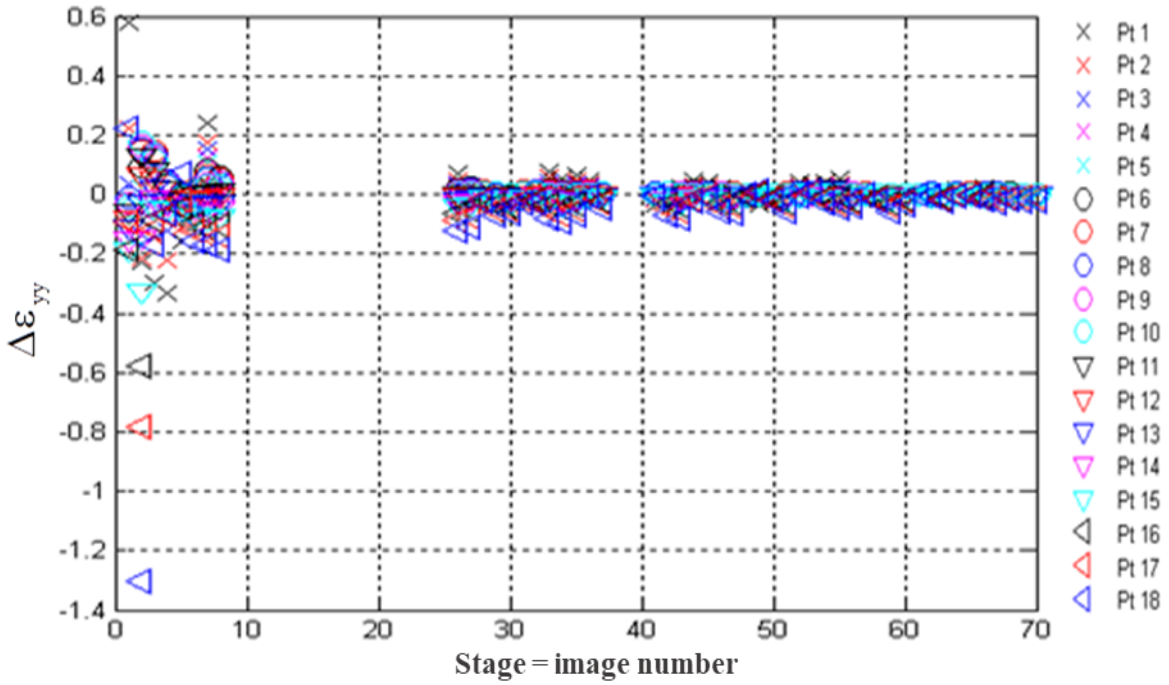


Figure 4-9 Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP5 (All data).

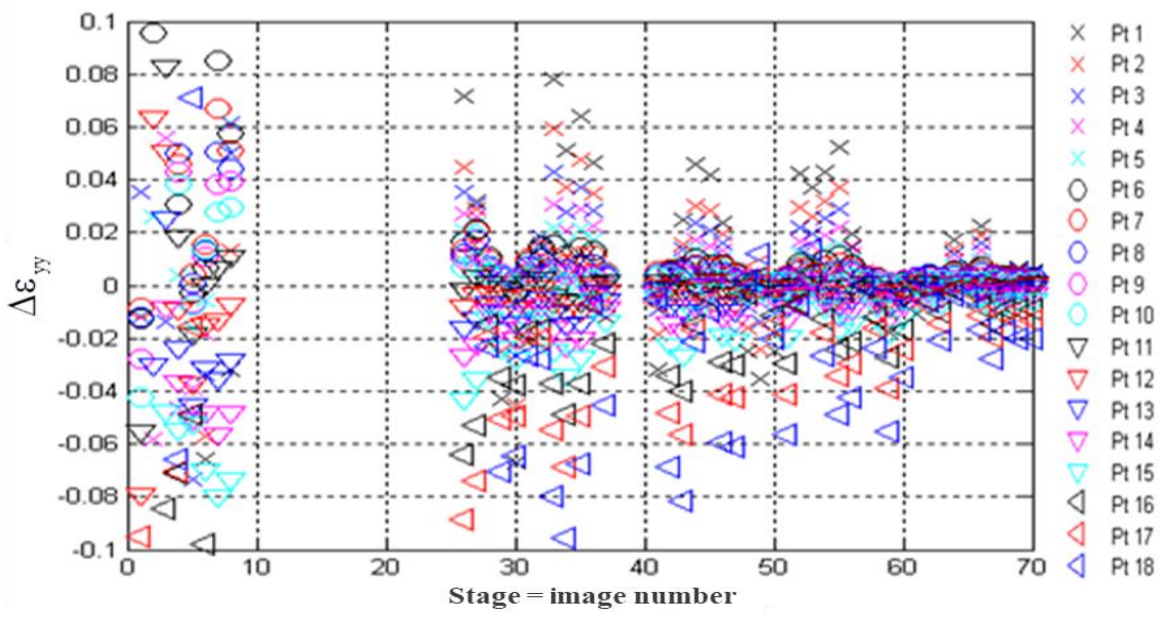


Figure 4-10 Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP5 (Focus on relative gaps between -10% and 10).

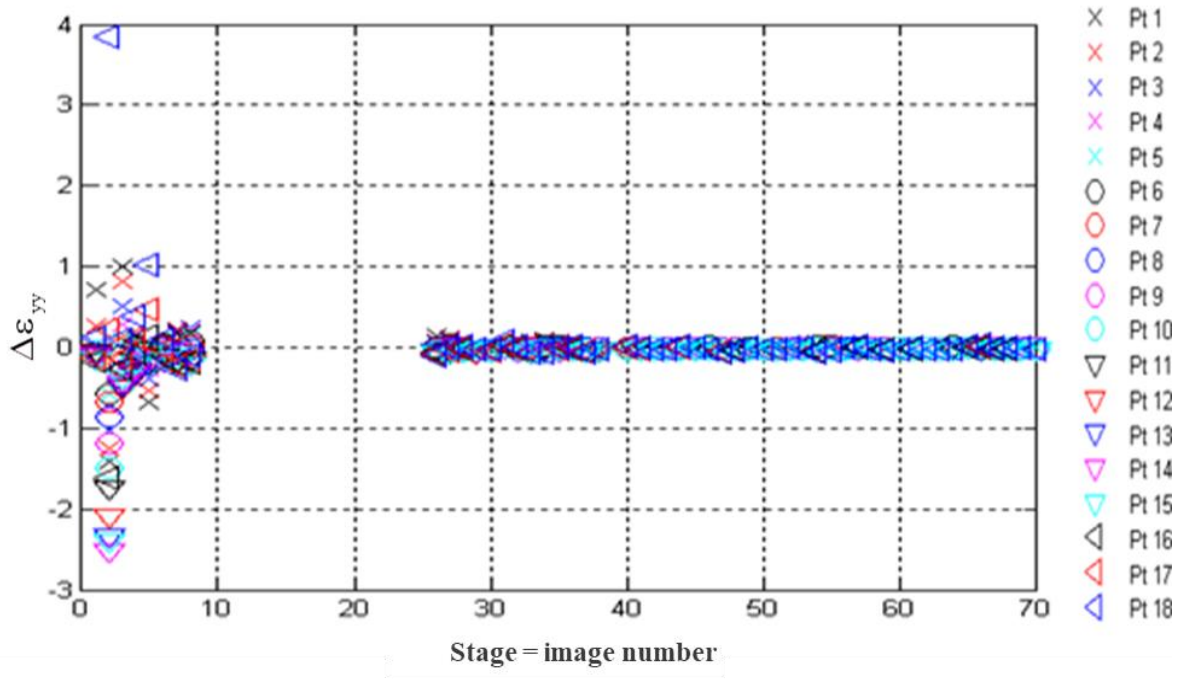


Figure 4-11 PP - Case Lossy - QP20 (All data).

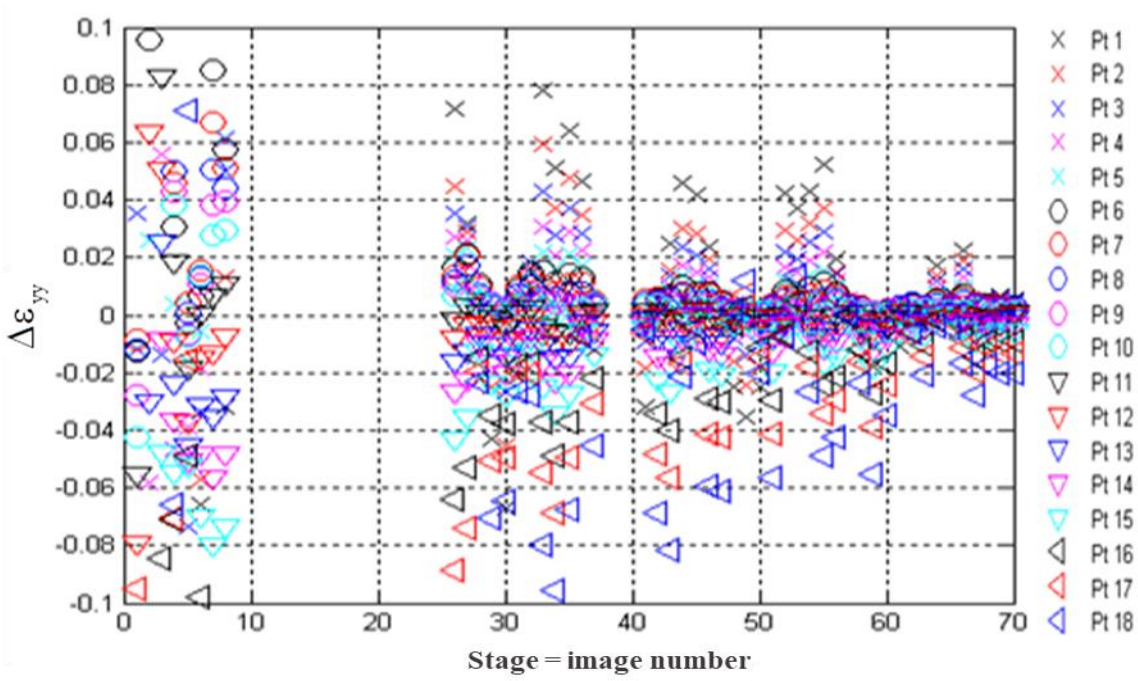


Figure 4-12 Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP20 (Focus on relative gaps between -10% and 10%).

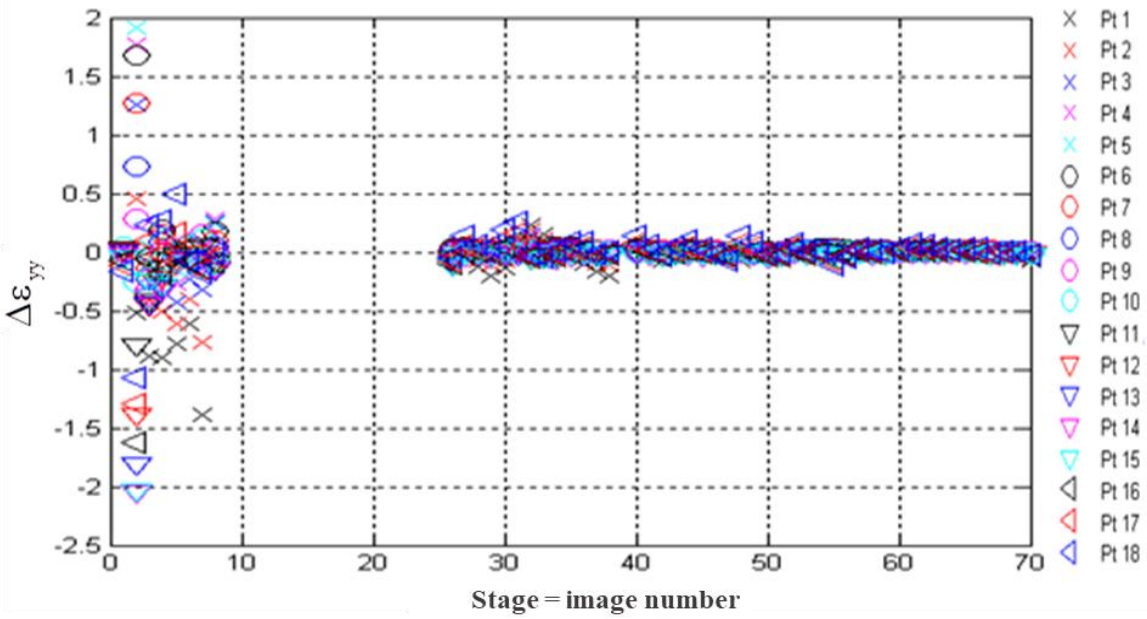


Figure 4-13 Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP25 (All data).

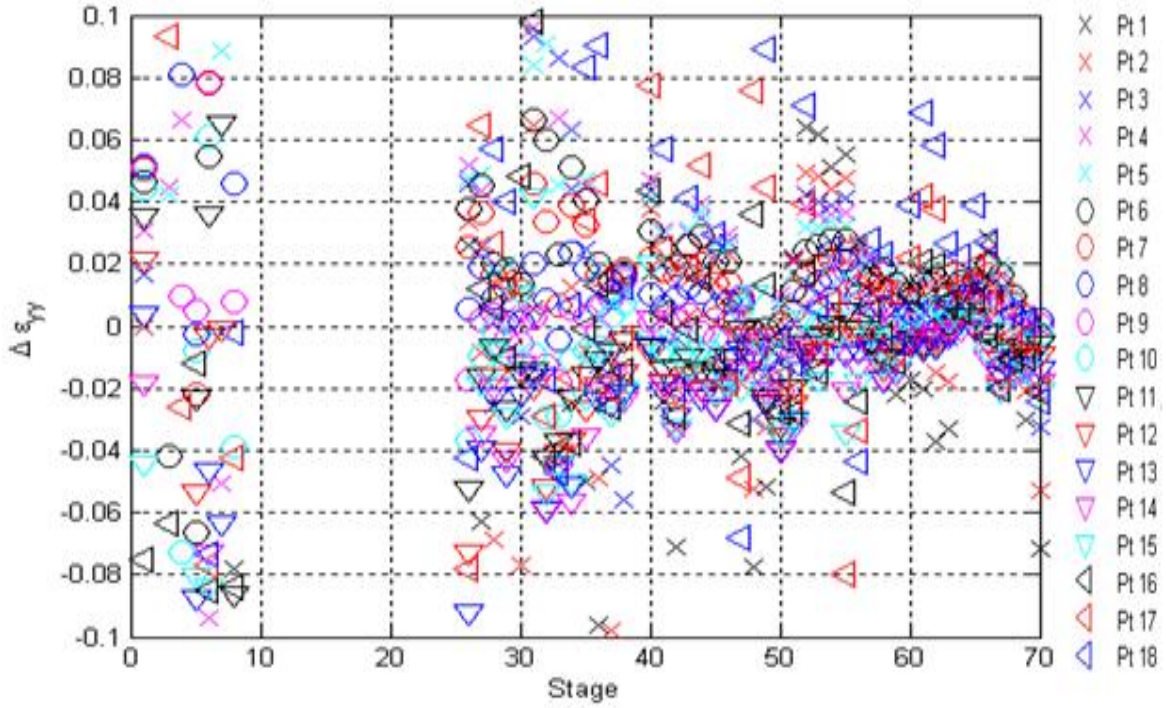


Figure 4-14 Evolution of relative gaps on computed axial strain throughout tensile loading of PP - Case Lossy - QP25 (Focus on relative gaps between -10% and 10%).

The same kind of analysis was done for the Sequence 2, i.e. for images recorded during the Arcan shear test of a glue joint. During this test, the three in-plane strain components develop in a significant way. Computation of relative gaps between axial, transverse and shear strain obtained from compressed images at different QP value and Lossless images. As for Sequence1, gaps are quite important at the beginning of the loading, due to very low strain levels, and then stabilized at satisfying value (i.e. relative gap lower than 10%), even for the highest values of QP (Figure 4-15 and Figure 4-16).

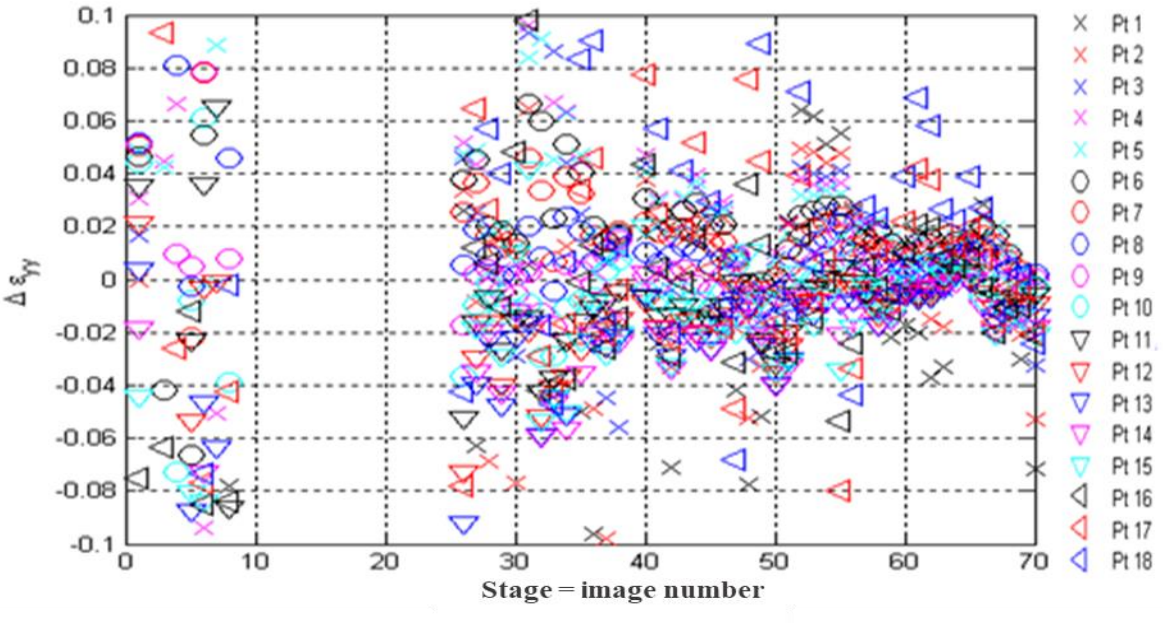


Figure 4-15 Evolution of relative gaps between strains computed from Lossy images and Lossless images of Sequence 2 (Arcan shear test of glue joint), in the ZOI of maximal shear strain (Axial strain).

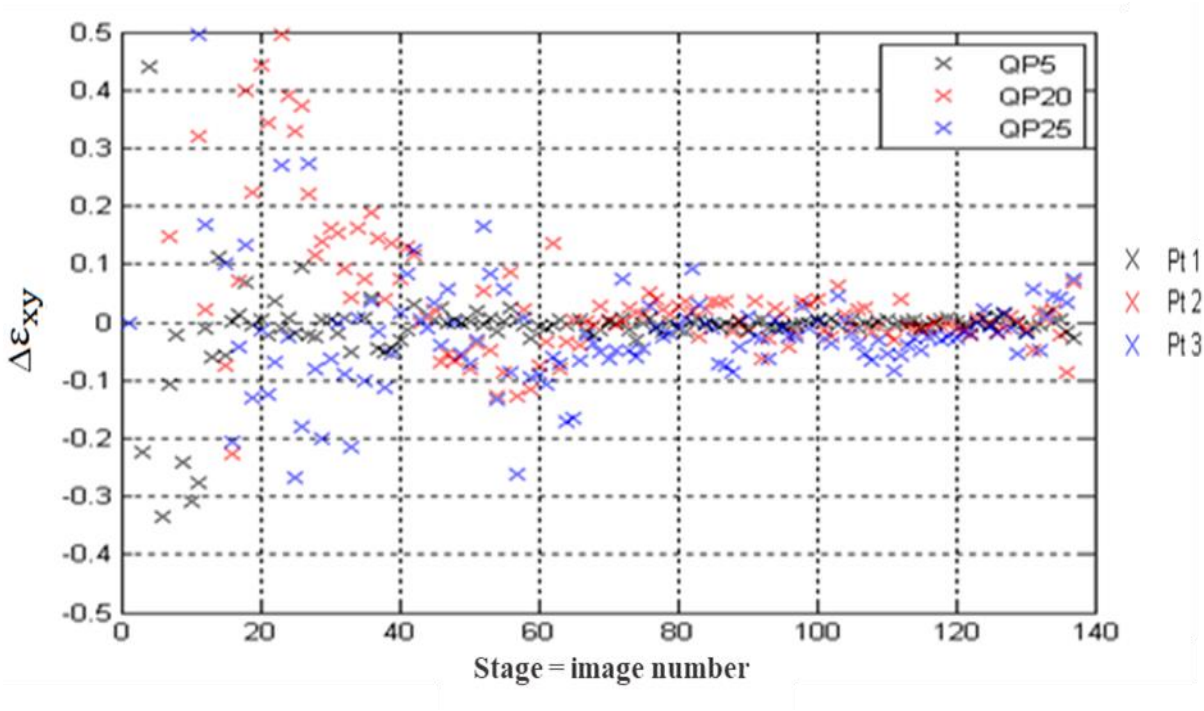


Figure 4-16 Evolution of relative gaps between strains computed from Lossy images and Lossless images of Sequence 2 (Arcan shear test of glue joint), in the ZOI of maximal shear strain (Shear strain).

The comparison of the evolution upon loading of strains computed from sets of images compressed at different QP demonstrates that the compression does not influence the accuracy of results up to QP=20 and only very slightly for QP=25 (Figure 4-17 and Figure 4-18).

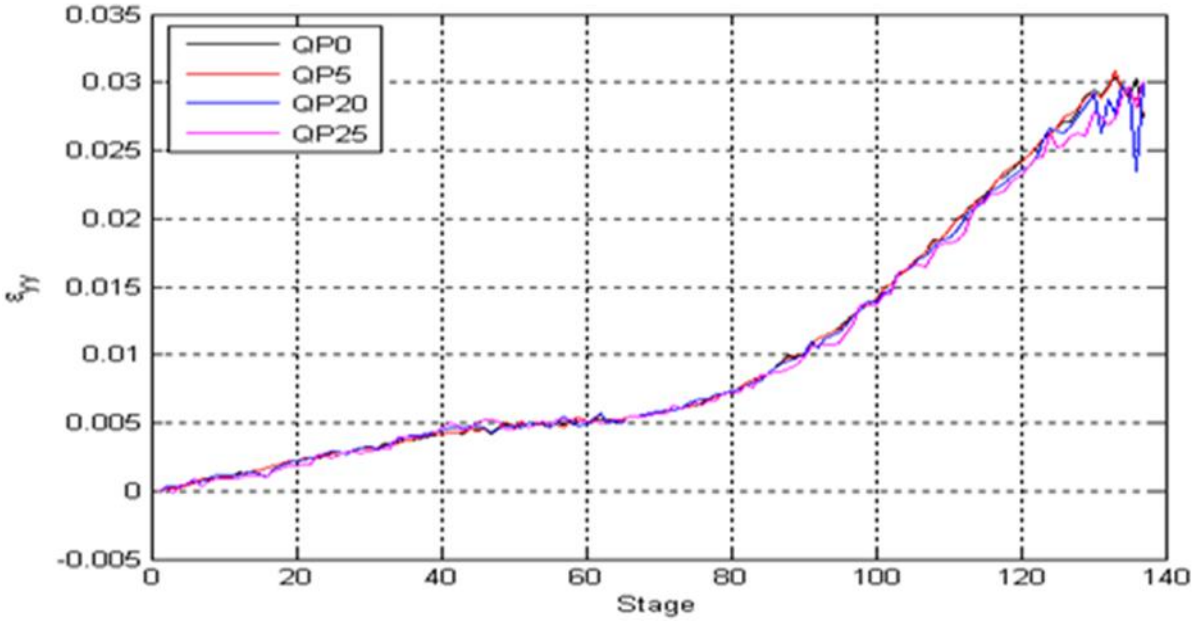


Figure 4-17 Evolution upon loading of strains computed from Lossy images of Sequence 2 (Arcan shear test of glue joint), in the ZOI of maximal shear strain (Axial strain).

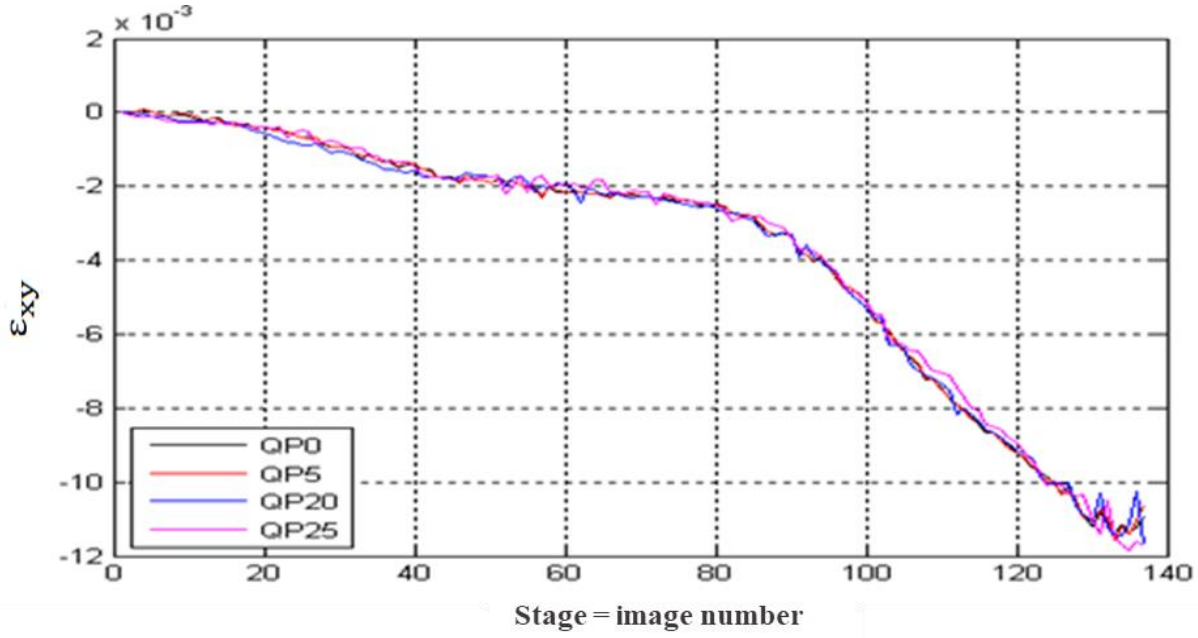


Figure 4-18 Evolution upon loading of strains computed from Lossy images of Sequence 2 (Arcan shear test of glue joint), in the ZOI of maximal shear strain (Shear strain).

Conclusion

In this Chapter, we have evaluated the performances of the state-of-the-art HEVC digital video coding standard on compressing image sequences of material mechanical responses captured by ultra-high-speed camera. We have demonstrated that HEVC provides very high coding efficiency as well as high visual quality. Moreover, further image analysis of mechanical response using DIC software showed that the material response was very well preserved in the moderately compressed sequences (QP ranging from 0 to 20) with an average SSIM nearly equal to 1 for Sequence 1 (i.e. dynamic tensile test of PP) and Sequence 2 (i.e. Arcan test of glue joint).

CHAPTER 5

SVM CLASSIFICATION OF MULTI-SCALE TOPOGRAPHICAL MATERIALS IMAGES IN THE HEVC-COMPRESSED DOMAIN

Context of the study

Materials science engineering is a scientific field that requires the acquisition, processing and analysis of a huge amount of image and video information data [12]. Indeed, various imaging techniques including laser scanning microscopy, X-ray imaging, spectroscopy, high-speed imaging, electron or micro-tomography allow studying the fundamental intrinsic properties of materials with the aim to establish links between the structure, dynamics and functioning of materials [1], [2]. Surface texture or surface morphology is one of the most relevant characteristics of any material surface; it has widely exploited in many machining processes [7]. Typically, the surface texture analysis seeks fundamentally to derive a general efficient mathematical operation from a quantitative texture descriptor that represents the texture with invariance to image transformation: illumination, scale of analysis and rotation. It provides useful information about spatial distribution, used for numerous digital imaging processes including image analysis and classification processes [79]. Surface similarity was proposed for texture analysis in material surface defect detection and surface inspection [13]. The multi-scale surface filtering decomposition techniques have proven their efficiency in roughness functional analysis [15]. Currently, the classification problem becomes more difficult if it is extended to material topography engineering. Indeed, it is very important in this case to have a robust texture descriptor that is invariant to transformation of surface images such as filtering and scale level. As seen in the Section 2.2.1, the surface profile is representing the roughness, the primary form and the waviness with three different frequency ranges [129]. Nevertheless, most of previous surface engineering studies were based on the roughness parameter which is represented by the high frequency (HP) component of the surface profile. More precisely, previous works try to determine which roughness parameter can give better understanding of the surface behavior. However, the material surface imaging analysis is a challenging task because of the

significant changing of the material surface texture, depending on process parameters such as length scales, as well as local physical properties [58], [59]. In addition, most captured image databases for materials science are very often acquired with high spatial resolution, large size and pixel depth, leading to a big amount of data to be stored or transmitted. Hence, it would be of great interest to apply lossy or lossless compression prior to classify or store images, where the decompressed image is used as an input of an image-processing-based material science engineering algorithm [10], [11]. However, it is mandatory that image compression does not affect the structural image properties, which are further exploited during the mechanical analysis of the materials under study. To guarantee such transparency, High-Efficiency Video Coding (HEVC) standard has been retained [24]. Moreover, when considering image compression, one can benefit from relevant compressed-domain information pertaining to the visual content feature extraction techniques in the form of transform coefficients [130], motion vectors [112] block-based segmentation [131] or intra prediction modes recently introduced by the H.264/AVC and HEVC video coding standards as possible image feature descriptors [132]–[134].

In this chapter, we present an original method to perform multi-scale surface classification in the compressed domain. The input image database consists in 13608 multi-scale topographical images. It is widely described in the next section. We show that the proposed solution allows to determine on which surface filtering range and scale length the surface category should be analyzed for classification. Finally, we give classification results on the image database by considering the compression ratios as well as the classification accuracy for each study condition.

5.1 Methodology

In order to evaluate the effects of surface filtering types and the scale of analysis on the performance of six mechanical multi-scale surface classification. Firstly, each surface profile was multi-scale analyzed by using Gaussian Filter analyzing method. The surface topography is decomposed into three multi-scale filtered image types: high-pass HP, low-pass LP, and band-pass BP filtered image data sets. Furthermore, the collected database is lossless

compressed using HEVC, then the compressed-domain Intra Prediction Mode Histogram (IPMH) feature descriptor is extracted. Support Vector Machine SVM is also proposed for strengthening the performance of the system.

5.1.1 Methods and Materials

Digital image compression is a key point for reducing the computational complexity. Where it that compression will simultaneously reduce the bit rate and offer an efficient image feature descriptor. In this section, we first give a brief overview on the collected data base characteristics and IPMH as texture image feature descriptor. SVM modeling process will be also discussed. Finally, the efficiency of the proposed compressed-domain classifier will be evaluated with 13608 multi-scale topographical images by considering the compression ratios as well as the classification accuracy for different study conditions.

5.1.2 Surface Processing

In the present study we have six material categories that come from initial rods of pure aluminum (99.99 %) of 1 meter long and 3mm diameter each. The aluminum rods are sandblasted with 6 different pressures and time of exposures. We used white light interferometer (New View 7300, Zygo) for characterizing and quantifying the surface roughness. The idea of light interferometer is based on using the wave properties of light to generate precisely the 3D topography. It uses a scanning white light interferometry for producing surface raw image and measuring the micro structure of surfaces in three dimensions: it measures the height (Z-axis) over an area in X and Y dimensions. A third-degree polynomial equation is retrieved from surface raw image to suppress the circular aspect of the rod prior to analyses the surface roughness. Then six topographies are extracted from this surface to obtain a 1024x1024 pixels topography map, using a spline interpolation. After that, the topographical map is converted into a grey map with 8 bit-depth. The multi-scale surface filtering decomposition techniques have proven their efficiency in roughness functional analysis [15].

Each topography map is multi-scale analyzed by using Gaussian filter recommended by ISO 11562-1996 and ASME B46.1-1995 standards to discover at which scale extend. each procedure parameter influences the morphology of the surface. This filter was adapted in order to filter the 3D surfaces with a given cut off value. In this study, the low-pass, band-pass and high-pass filter are used as described in [15].

5.1.3 Topographical Materials Texture Image Dataset

The collected mechanical topographic image data set consists of 13608 images with a resolution of 1024x1024 pixels with two internal bit depths: 8 &16 bits, respectively. It represents six mechanical material categories, with the sequence names varying with the following variables written in red line letters to indicate a specific imaging setting as illustrated in Figure-5.1.

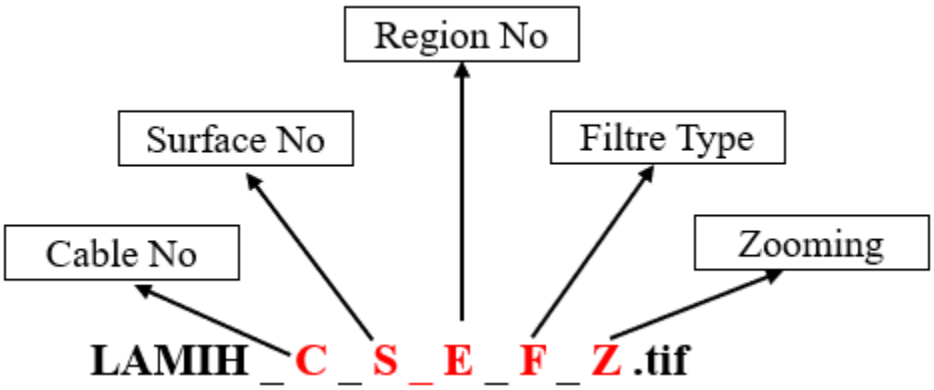


Figure 5-1 Nomenclature used to represent the collected mechanical topographic images

In the present case, the six mechanical categories consist of six surfaces. Each surface includes seven surface regions. Each surface region profile is decomposed into three different types of filtered images: high-pass (HP), low-pass (LP) and band-pass (BP) filtered image. Each filtered image represents the roughness, the primary form, and the waviness of the surface, respectively. Finally, each filtered image type decomposes into 18 different spatial length-scales to result into 4536 images. Figure-5.2 represents MEgABIt mechanical multi-

scale topographical image data bases. It can be noticed that there is a high spatial correlation between any two surfaces topography images from different categories.

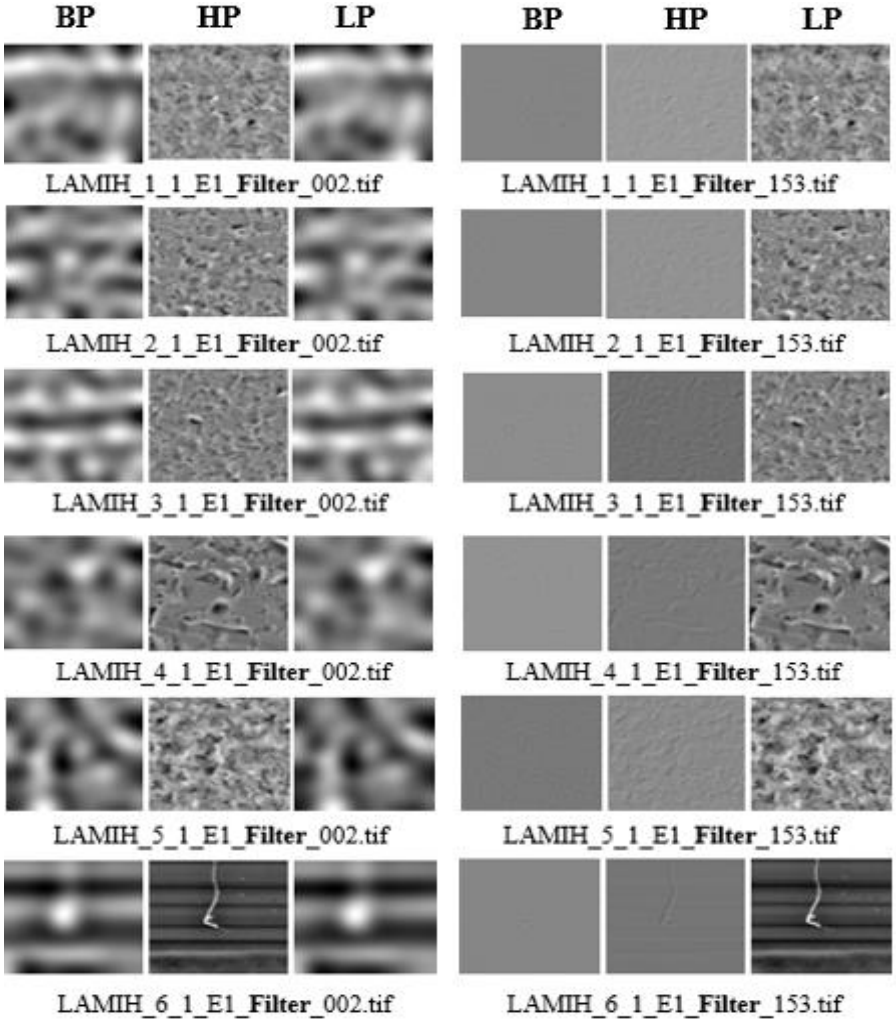


Figure 5-2 Represents one image (Resolution of 1024x1024 pixels) from six mechanical material categories, with two different zooming and three filtered images.

5.1.4 IPHM-Based Classification

As we discussed in Chapter 2, Intra prediction allows exploiting in a very efficient way spatial redundancy inherent in image contents. Figure 5-3 illustrates the effectiveness of HEVC intra prediction process to predict texture contents in images.

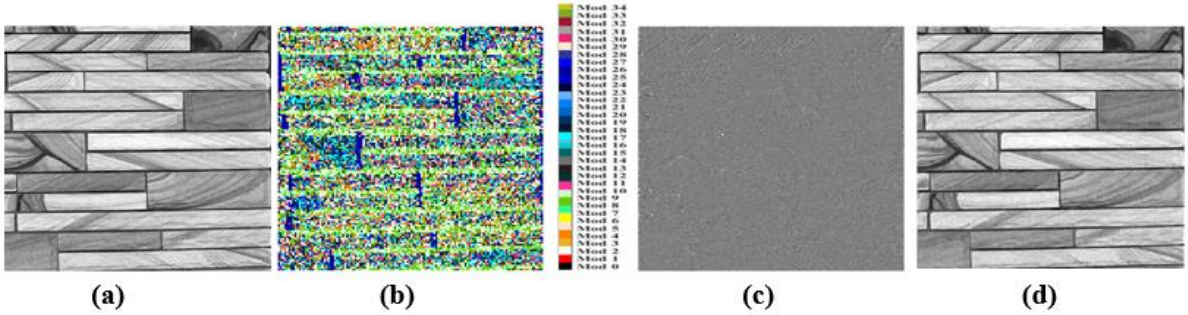


Figure 5-3 Original Image 512x512 (A), selected modes to predict the original image presented with 35 colors (B) Intra Predicted Image (C) and The Residual image (D).

It is clear that the predicted image inherits most of the main characteristics of the original image with a low residual signal. HEVC intra prediction process is very well suited to capture the texture features which represents one of the most important visual descriptors in the field of image classification, pattern recognition and computer vision [77], [78]. Hence, the different intra prediction results should constitute a good candidate for texture feature extraction. Recently, Mehrabi et al [132]–[134] developed a compressed-domain texture feature descriptor based on intra prediction modes computation. The so-called Intra Prediction Modes Histogram (IPHM) descriptor consists in counting the number of blocks predicted by each of the 35 available intra prediction modes. They are calculated directly from the compressed image data without the need to decode the whole image, hence reducing the computational complexity.

The steps to extract the IPMHs are listed below:

• Compressing the entire topographical image databases with HEVC lossless intra prediction coding, by computing the 35 Intra prediction modes for Prediction Units PU of size 4x4 pixels.

• Searching for the best prediction mode, that minimizes the sum of absolute difference (SAD). The selected mode indicates the relation between the pixels inside the Prediction Unit (PU) with the boundary neighbor pixels.

• Counting the frequently utilized prediction modes, to arrange each mode in one histogram bin as follows;

$$H_i' = \{ h_i \quad 0 \leq i \leq 34 \quad \}$$

(5.1)

Where $\mathbf{H_i'}$ is the bin of the histogram for the mode **(i)**.

$\mathbf{h_i}$ indicates the number of blocks in the coded picture which are predicted by mode **(i)**.

The normalized IPMH is generated as:

$$H_i = \frac{H_i'}{X}$$

(5.2)

Where X represents the total number of 4x4 blocks in the image (65536 blocks in the case of a 1024x1024 image).

Finally, the similarity measurement between every two images is based on the intersection between their corresponding normalized IPMH as:

$$\mathrm{Sim}_{a,b} = \sum_{i=0}^{34} \min\left(\left(\mathrm{H_i}, a\ , \mathrm{H_i}, b\right)\right)$$

(5.3)

Firstly, this method was validated in the H.264/AVC compressed domain, then in the HEVC one, using VisTex conventional image databases of natural scenes [133]. Unfortunately, when testing on our complex surface texture image database, the similarity measurement indicates high correlation between many pairs of IPMHs either they were taken from the same category or from different categories. For example, we have used six images from each surface category to evaluate the proposed similarity measurement method on surface categories classification where one image was used for query while the other images were used for testing. The first five retrieved images from each query are ranked in descending order based on the similarity value. This method has classified 36 surface images with poor average accuracy less than 30 % as illustrated in Figure 5-4.

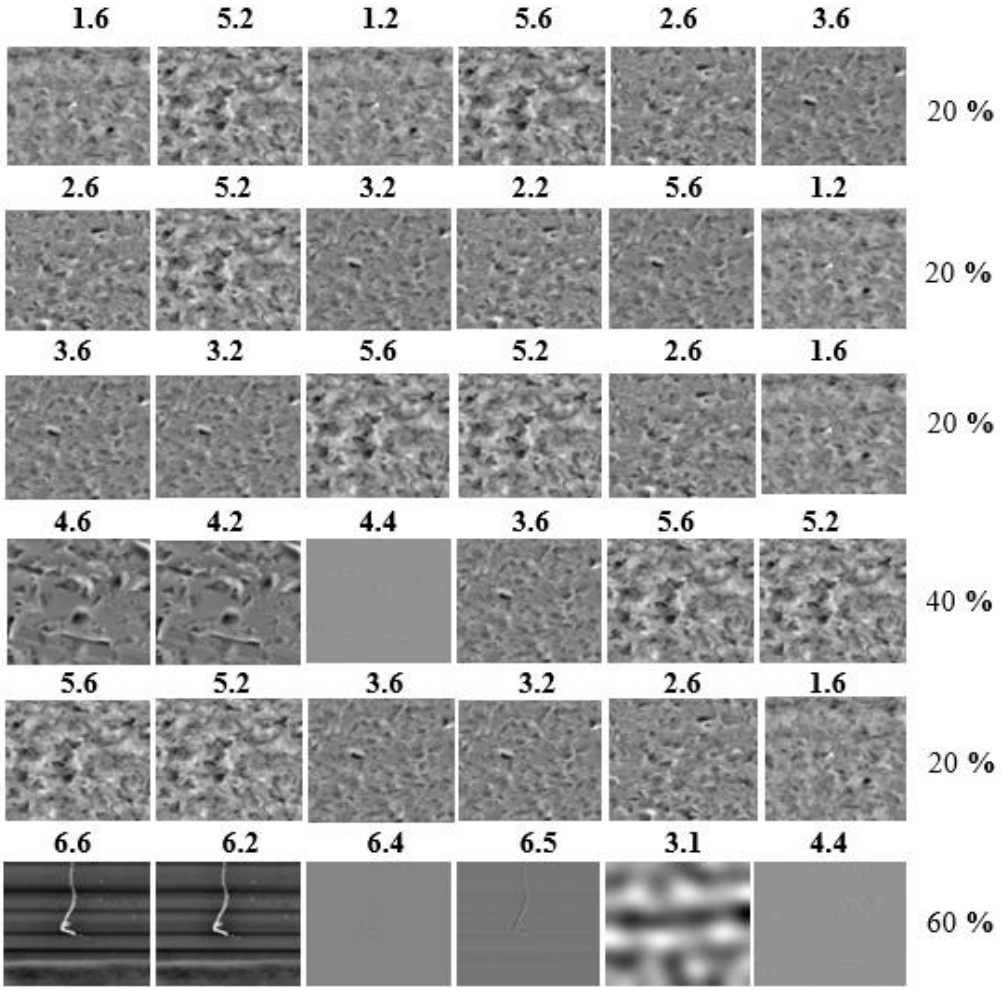


Figure 5-4 First five retrieved images for six images tests (categories 1 to 6) using IPMH, which indicate classification accuracy of 30 %.

From the above Figure 5-4, we can notice poor classification for the first surface image category, which has false prediction in four times and correctly predicted in one-time (image 1.3). These results lead us to develop an original classification algorithm by combining IPMH with Support Vector Machine (SVM) to find the optimal separator between non-linear surface image categories.

5.1.5 HEVC Lossless 4x4 PU Compression

HEVC high-throughput profile is considered for still-image lossless intra compression. This profile supports up to 16-bit depth still image compression. For well characterizing of the texture in localized image area, the Prediction Unit (PU) size is fixed to 4x4 blocks to have the finest analysis size. This profile is implemented by using HEVC reference software HM 16.12. Performances in terms of compression ratio for each surface filtered image types are given in Section 5.4.

5.1.6 SVM Classification

We propose to use SVM to find the optimal separation between these three multi-scales filtered image data sets, to evaluate:

Case-1: the impact of considering the three-filtered image data sets together on the six surfaces categories classification performance.

Case-2: the impact of each filter separately on the six surfaces categories classification performance.

Case-3: the impact of each scale of analysis on the six surfaces categories classification performance.

To perform that, firstly, the dataset is separated into two partitions: a training dataset in order to build the classifier and a testing dataset to evaluate the classifier. Different set sizes are considered in order to evaluate the impact of training data set size on the proposed model performance.

Secondly, for model training, we use a variable number of randomized training datasets in each simulation case. For example, in case-1 where the three dataset images are considered together (4536 images), the training dataset are (454 (10%), 908 (20%), 1362 (30%), 1816 (40%), 2270 (50%), and 2724 (60%)) IPMHs from each surface category respectively, while the rest of the dataset is used for testing. And in the same way for case-2 the training data set when considering the three datasets separately are (76 (10%), 152 (20%), 228 (30%), 304 (40%), 380 (50%), and 456 (60%)) IPMHs from each surface category.

While in case-3, each scale is represented by an image dataset of 42 images each, which is divided into training dataset of 24 images for all available scale of analysis.

Thirdly, to evaluate linear, Poly and RBF (LIBSVM_MODELS) learning algorithms, we perform the 5-k Cross-Validation that has explained in (section 2.3.3); using the training dataset to select the kernel model and to tune the model parameters in each simulation case.

During SVM evaluation, the polynomial function kernel was giving better classification performance in the three cases, with different optimized kernel parameters (C & gamma) for each simulation. Finally, we train the SVM models for case-1 and case-2 with a varied number of randomized training datasets to evaluate the impact of increasing the number of training data set on the classification performance. Figure 5-5 illustrates the procedure for learning and testing the non-linear SVM model, in case of total data set was split into α % for learning ($0 \leq \alpha \leq 1$), and the remaining used for model validation.

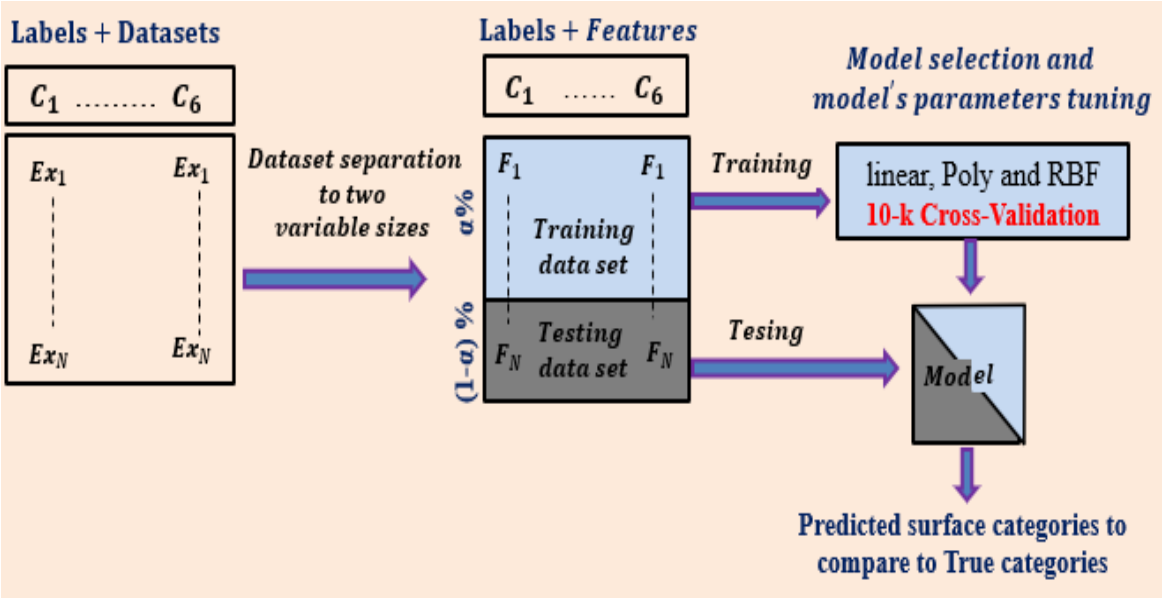


Figure 5-5 Block diagram depicting the procedure for learning and testing the SVM model.

Results

In this Section, we will evaluate the performances of the proposed SVM-based classification algorithm in HEVC compressed domain, using the topographical image

database described in Section 5.3.2. Firstly, we will present the achieved compression ratio for each surface filtered image types. Secondly, we will present the effectiveness of the proposed image texture descriptor to characterize the surface topography with different analyzing conditions. Then, we will present the impact of multi-scale surface filtering types on the model classification performance. Finally, the effect of scale analysis on the model performance will be also evaluated.

5.1.7 The Impact of Surface Topography Filtering Types on Achieved Compression Ratios

In general, the achieved lossless compression ratios depend on image complexity. The compression ratio is high at the lowest scale of analysis, except for high-pass filtered images where there is no difference between compression ratios achieved at any scale value as illustrated in the following Figures 5-6, 5-7 and 5-8.

Figure 5-6 presents HEVC lossless compression ratios for the six multi-scale low-pass filtered surfaces image categories. The average compression ratio is also given.

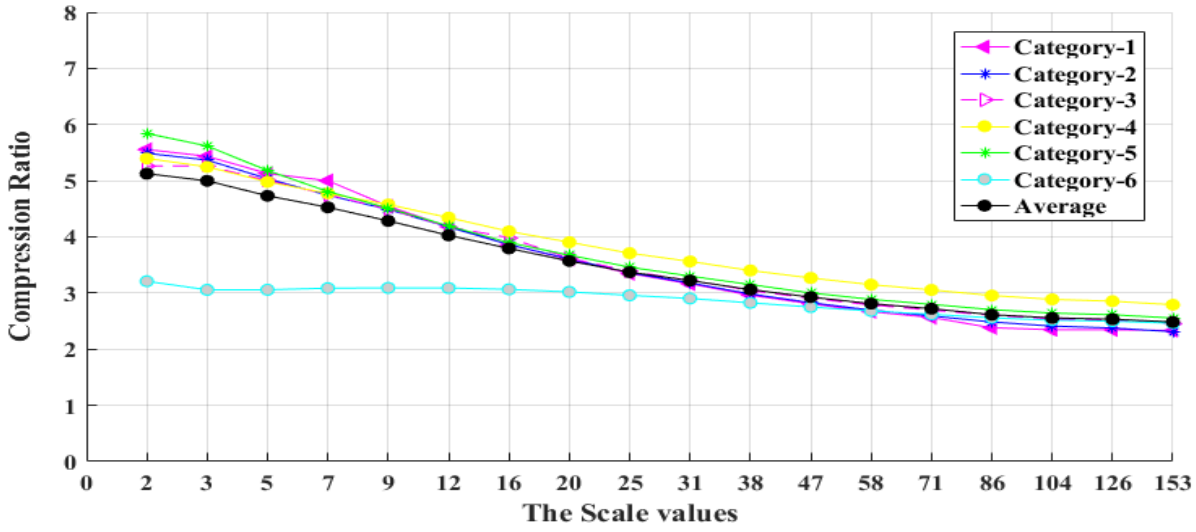


Figure 5-6 Relationship between the scale of analysis and the six surface categories compression performance by using the multi-scale LP-datasets.

The compression ratios vary between 2.3:1 and 5.8:1 depending on the scale value. At the first ten analysis scale values, the compression ratio values are very closed between the six LP multi-scale surface categories except for category-6. Globally, the scale of analysis and the achieved Compression Ratio (CR) are inversely proportional, where CR increases as the scale of analysis decreases.

In Figure 5-7, the compression ratios are very closed for the six band-pass multi-scale surface categories, where the best average CR=5:1 was achieved at the lowest analysis scale, while the highest length-scale was at average CR=2:1.

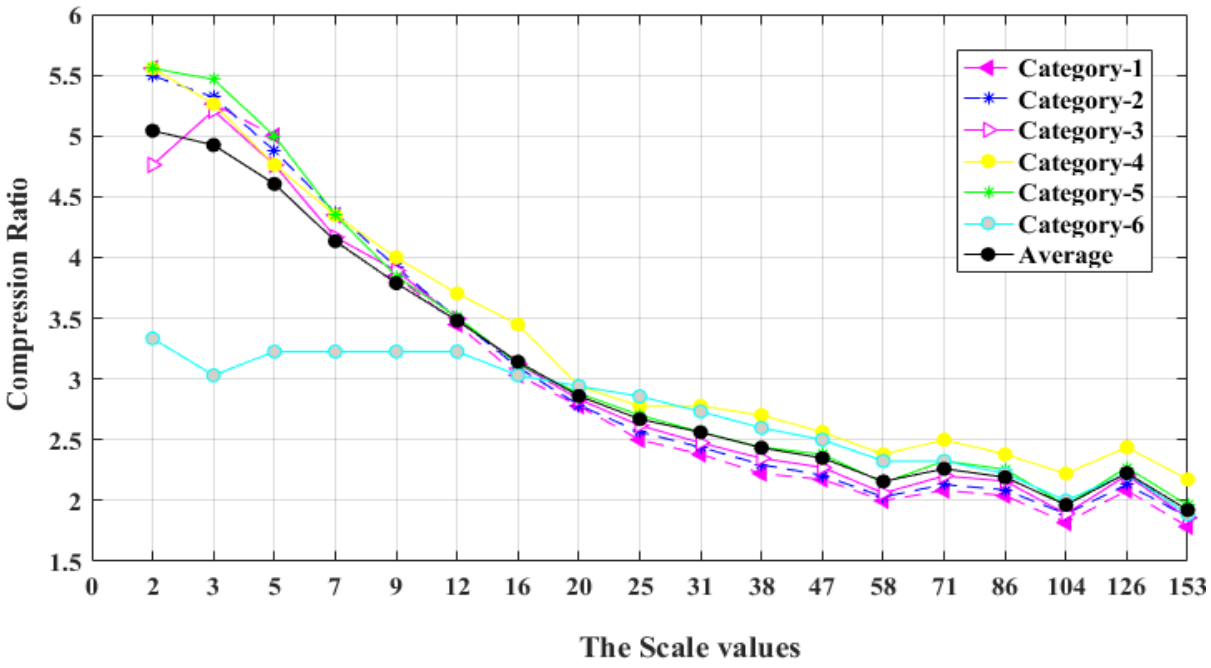


Figure 5-7 Relationship between the scale of analysis and the six surface categories compression performance by using the multi-scale BP-datasets.

In Figure 5-8, it is clear that there is no significant difference between the compression ratio for the six high-pass multi-scale surface categories at different scale of analysis with average CR=2.2:1.

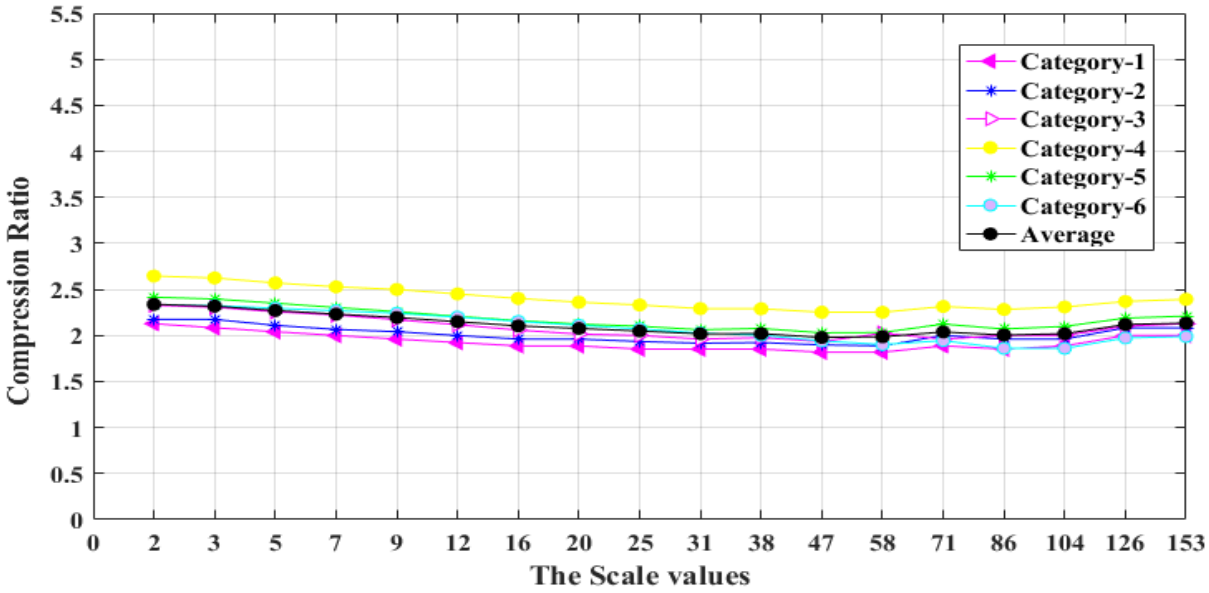


Figure 5-8 Relationship between the scale of analysis and the six surface categories compression performance by using the multi-scale HP-datasets.

5.1.8 Evaluating IPMH As Texture Feature Descriptor

As already mentioned in Section 3.2, the proposed texture feature descriptor is highly related to the specific pattern of the predicted block of pixels. The 33 angular prediction modes can predict all-frequency components for specific predicted 4x4 directional block topography image with a residual signal is nearly null as illustrated in figure 5-9.

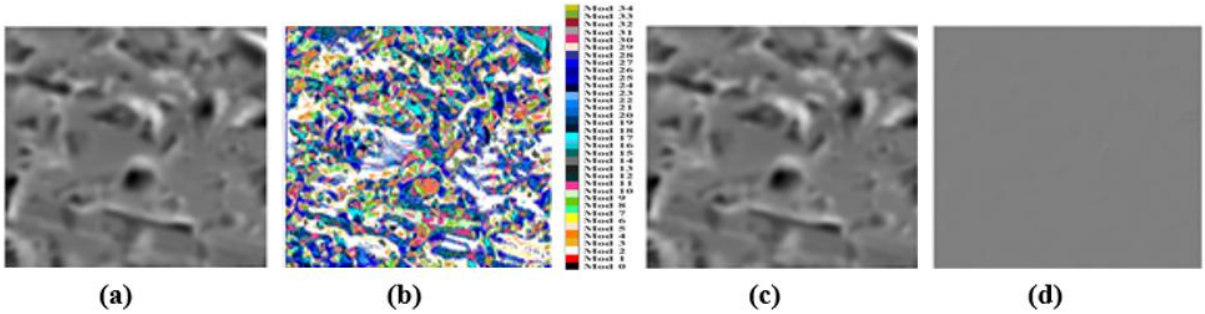


Figure 5-9 Original Image 1024x1024 (A), selected modes to predict the original image presented with 35 colors (B) Intra Predicted Image (C) and The Residual image (D).

In Figure 5-10, the first three sub-figures compare the IPMHs averages for the six categories at three different multi-scale filtered image types: HP, LP and BP filtered image

data set. While, the last sub-figure presents compares the IPMH averages for the three different multi-scales filtered image data sets. The comparison between the average IPMHs was nearly similar for the first five material categories at different prediction modes. While the sixth category, has a small IPMH difference compared to the others.

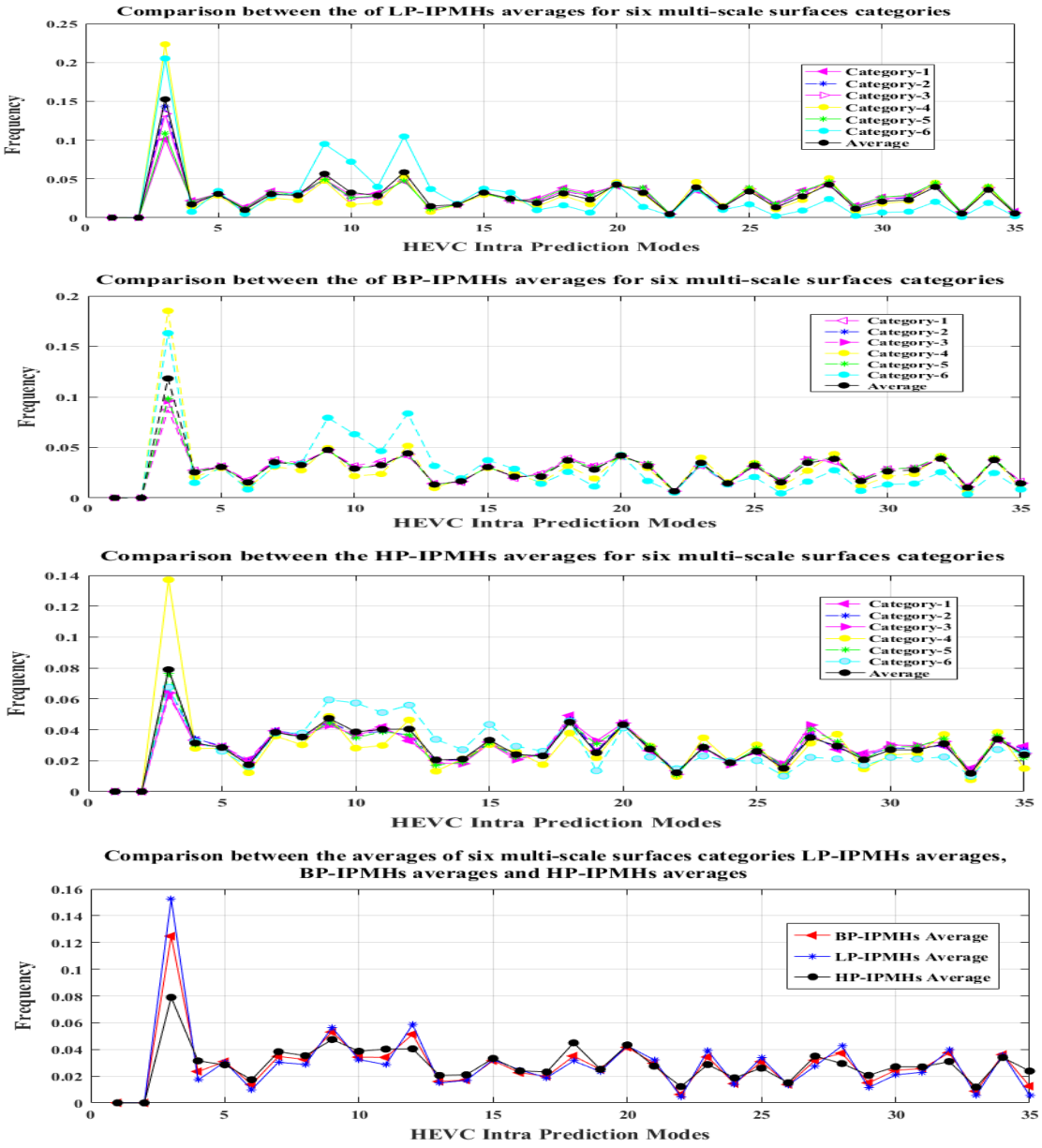


Figure 1-10 Comparison between the IPMHs averages for three different filtered image data sets; LP, BP, and HP data set.

5.1.9 The Impact of Surface Topography Filtering Types on Topographical Images Classification Accuracy

The IPMHs feature descriptors are very efficient for classifying a mix of three multi-scale surface filtered image data sets. The classification accuracy reaches 85 % by using 60 % of total data set as training data set (8170 IPMHs) while using the rest of the dataset (5440 IPMHs) for testing as presented in Figure 5-11.

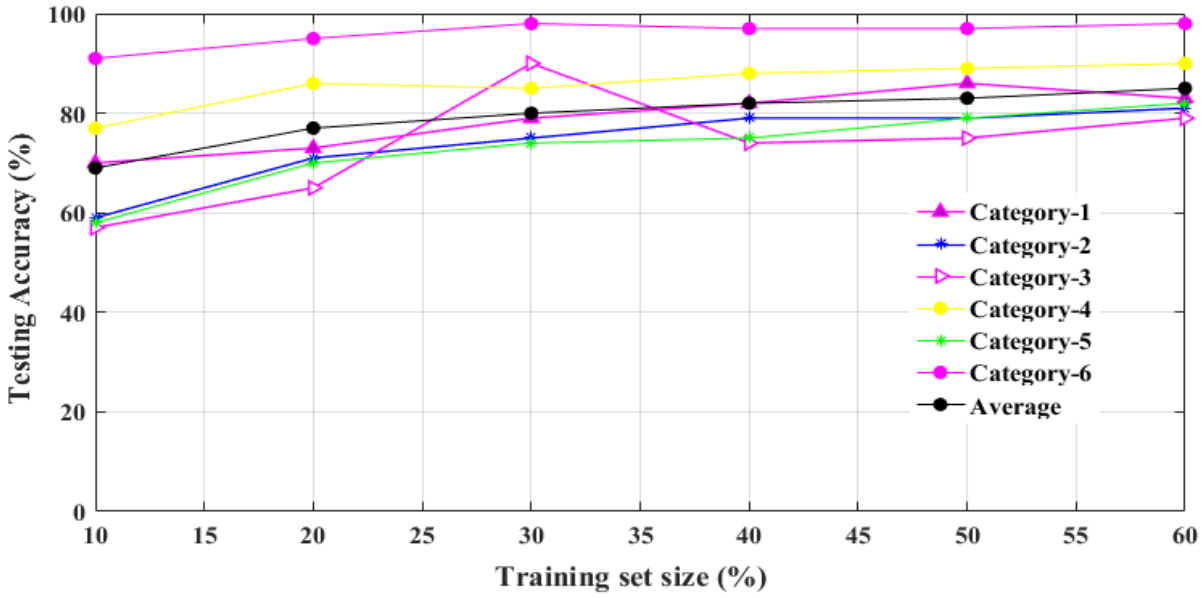


Figure 5-11 The effect of increasing the training set size on the classification accuracy while using mixed multi-scale HP, LP, and BP datasets.

The classification accuracy increases as the number of the training datasets increases. This relation is clearly noticed in Figure 5-11 by plotting the average of the achieved accuracies while classifying the six surfaces categories. The classification accuracy is reported in the following confusion matrix, where the columns and the rows are representing the predicted and the actual classes respectively. The values located at the diagonal of the matrix indicate the exact prediction percentage. Figure 5-12 shows the confusion matrix for the six categories while considering 60 % of the dataset for training. For example, the prediction percentage for category1 is equal to 85%.

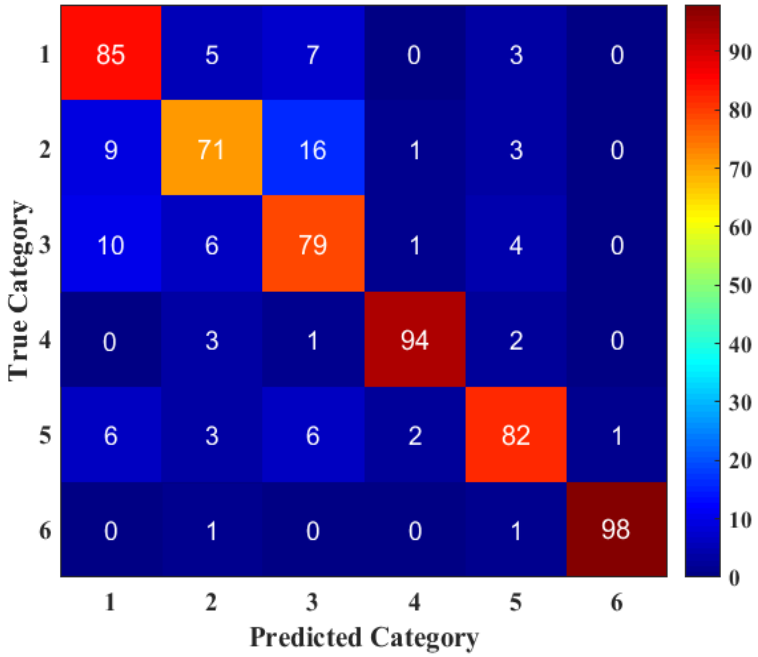


Figure 5-12 The Confusion matrix for classifying the six surface categories (Mixed).

In the case of HP-filtered images dataset, we considered six different percentages (from 10 % to 60 %) of the total dataset (4536 IPMHs) for the training dataset as illustrated in Figure 5-13.

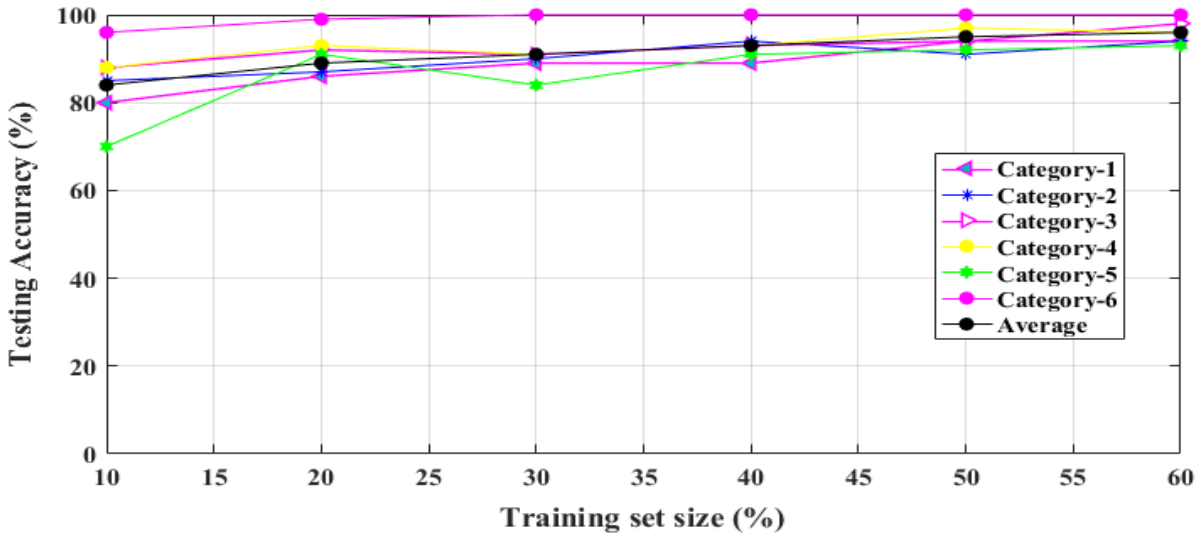


Figure 5-13 The effect of increasing the training set size on the classification accuracy while using HP-datasets.

Figure 5-13 demonstrated the strength of compressed-domain classifier with directly proportional relationship between the size of the training data set and the classification accuracy. Where the classification accuracy can reach 84 % by using just of 10 % from HP-data set (456 IPMHs) for training. Moreover, the classification accuracy improved significantly when considering large size training dataset. For example, the classifier reached 96 % accuracy by using 2724 IPMHs for training.

Figure 5-14 shows the confusion matrix for the six categories for HP filtered image data set. We can notice 10 % average increase in the classification accuracy when considering only HP filtered dataset. For example, the prediction percentage for category1 is equal to 95 %.

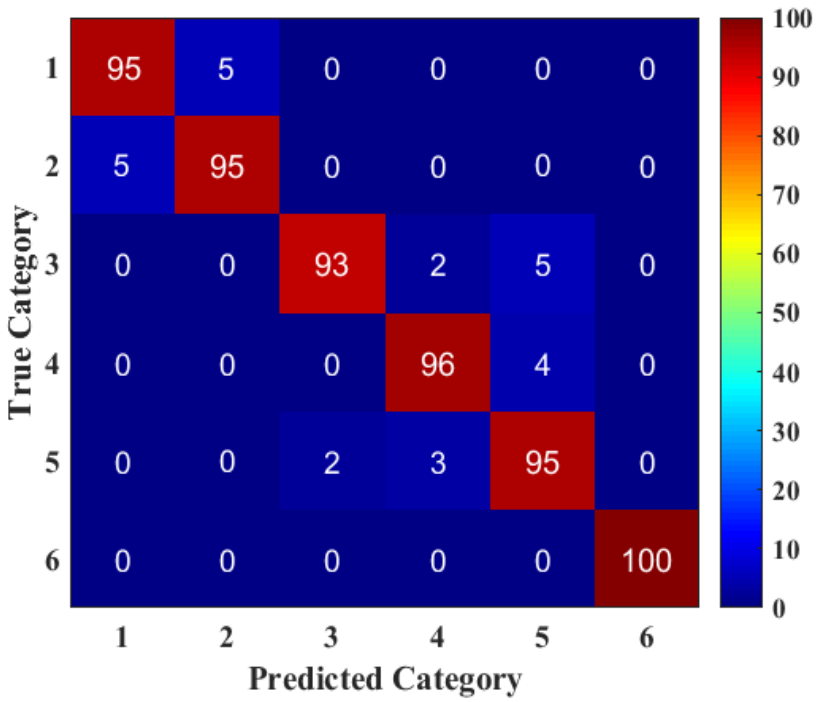


Figure 5-14 Confusion matrix for six surface categories classification by using 60 % of multi-scale HP data set for training.

In the case of LP-filtered images dataset, we considered six different percentage (10, 20 … 60 %) of the total dataset (4536 IPMHs) for the training dataset as illustrated in Figure 5-15. The classification accuracy can reach 87 % by using 60 % of LP-data set (2724 IPMHs) for training.

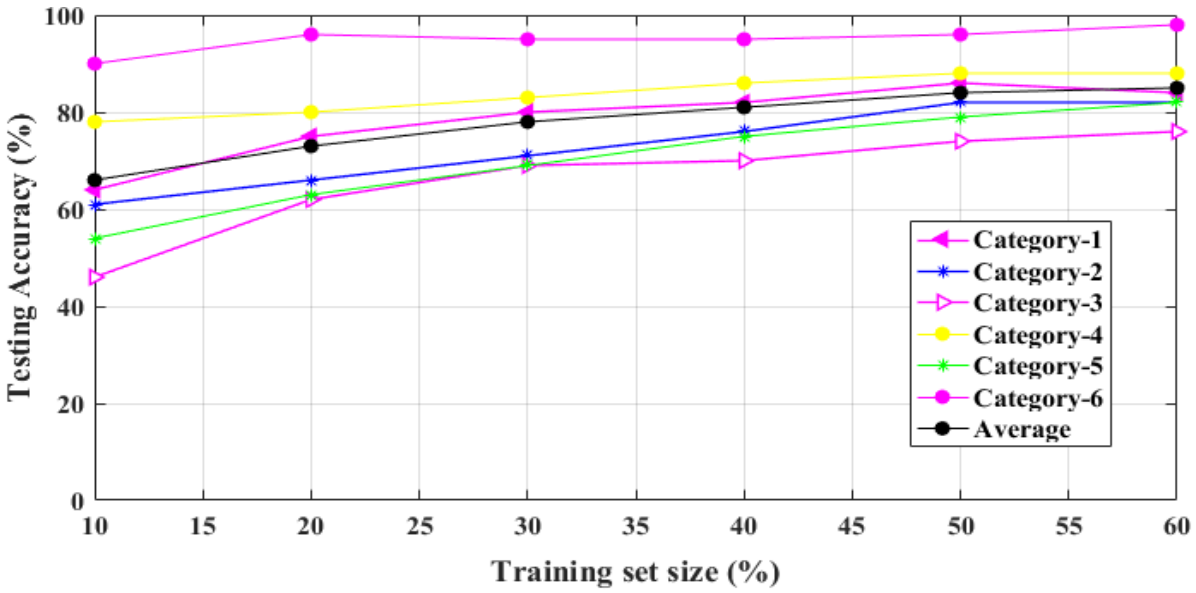


Figure 5-15 The relation between the size of the training data set and the six surface categories classification performance by using the mixed multi-scale LP- datasets.

Figure 5-16 shows the confusion matrix for the six categories for LP filtered image data set. From the figure, we can deduce that the obtained classification accuracy by using LP filtered image dataset is less than that obtained while using HP filtered image dataset.

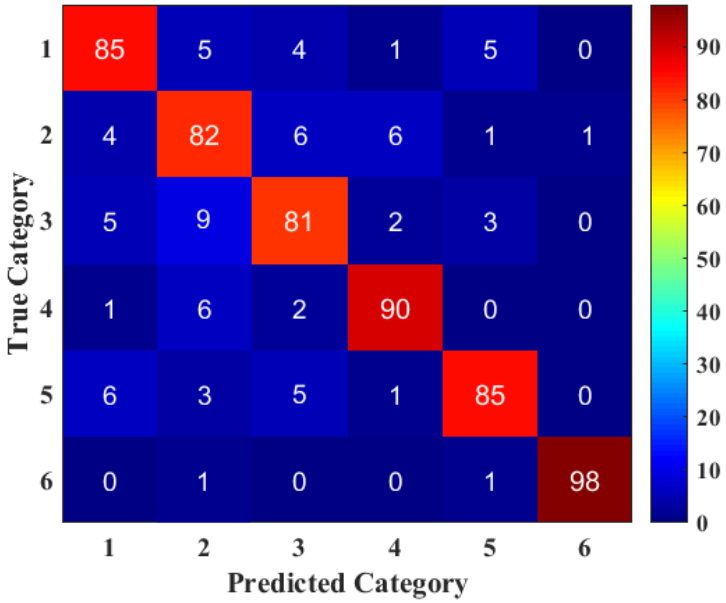


Figure 5-16 Confusion matrix for six surface categories classification by using 60 % of multi-scale LP data set for training.

In the similar way in case of BP-filtered images dataset, six different percentage (from 10% to 60 %) are considered from the total dataset (4536 IPMHs) for the training dataset as shown in Figure 5-17. The classification accuracy can reach 82 % by using 60 % of BP-data set (2724 IPMHs) for training.

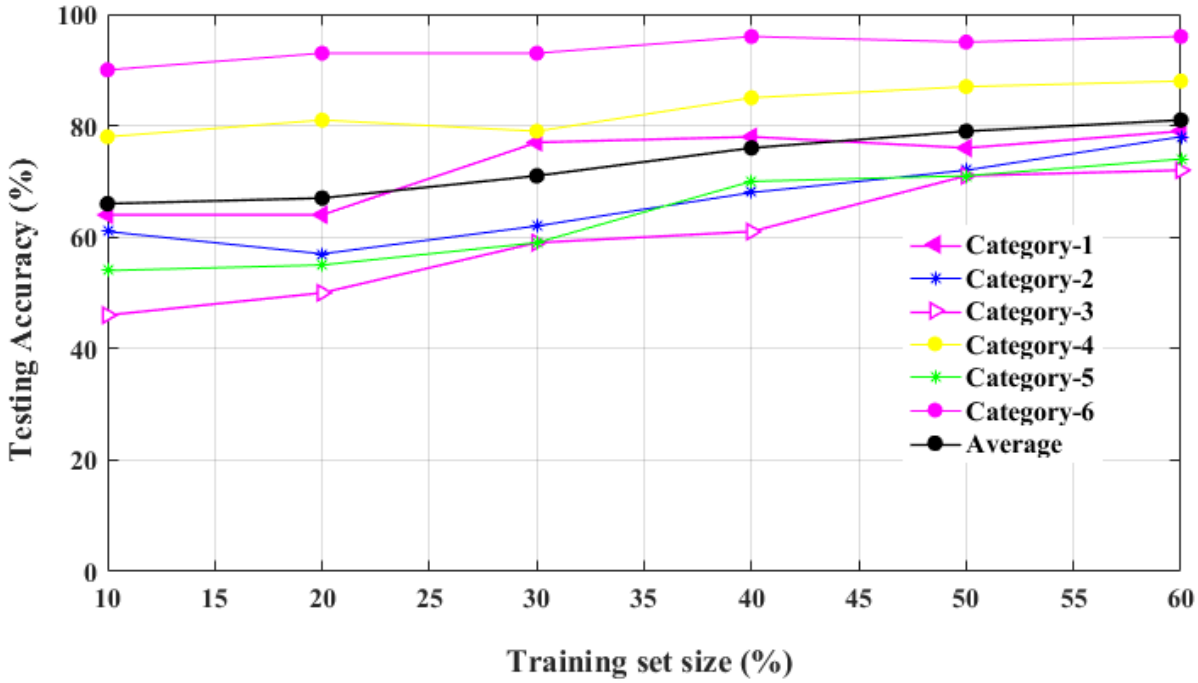


Figure 5-17 The relation between the size of the training data set and the six surface categories classification performance by using the mixed multi-scale BP-datasets.

The following confusion matrix presents the classification accuracy for the six categories by using 60 % of total BP-filtered image data set.

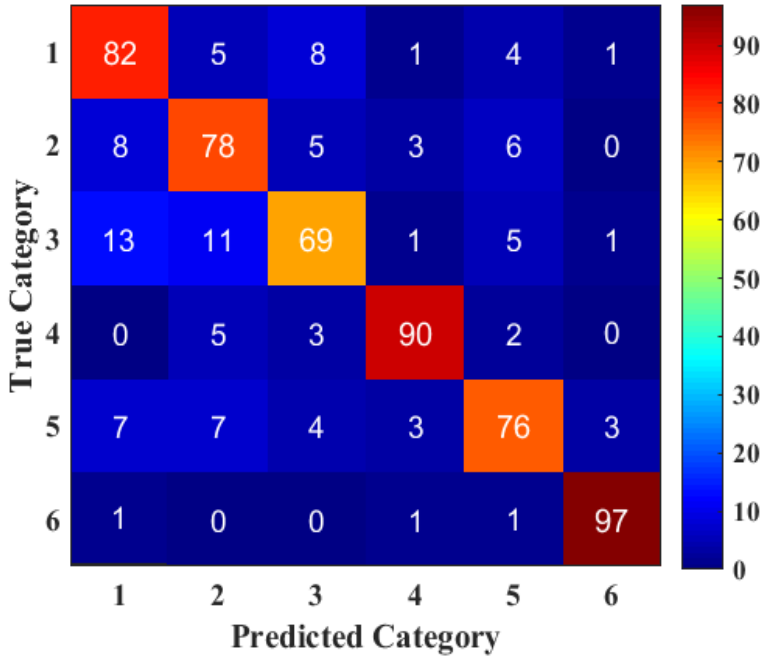


Figure 5-18 Confusion matrix for six surface categories classification by using 60 % of multi-scale BP data set for training.

5.1.10 The Impact of Scale of Analysis on Topographical Images Classification Accuracy

As we previously illustrated, the surface topography profile decomposes into three different filtering methods (Low-pass, Band-pass and High-pass filter) with eighteen different length-scales. In this section, we aim to evaluate the effect of each length-scales on system classification accuracy for four different cases: mixed filtered images dataset, LP-filtered dataset, BP-filtered dataset and HP-filtered dataset. The results have indicated a significant improvement for classification accuracy in the case of the mixed filtered image datasets (LP + BP + HP) at the highest-scale of analysis. The average accuracy reached 88 % by using 60 % (456 IPMHs) of the data sets for training. Where the average accuracy enhanced by only 3 % than in the case of multi scale mixed dataset.

The single scale analysis was more appropriate than multi scale analysis in the case classifying HP, LP and BP data set separately. Figure 5-19 shows the classification accuracy for six LP multi-scale surface categories at different compression ratio and scale of analysis.

In order to increase the classification accuracy, we could use higher scale analysis at lower compression ratios. For example, we can obtain classification accuracy of 86% at scale of analysis = 86 and average compression ratio CR = 2.5:1.

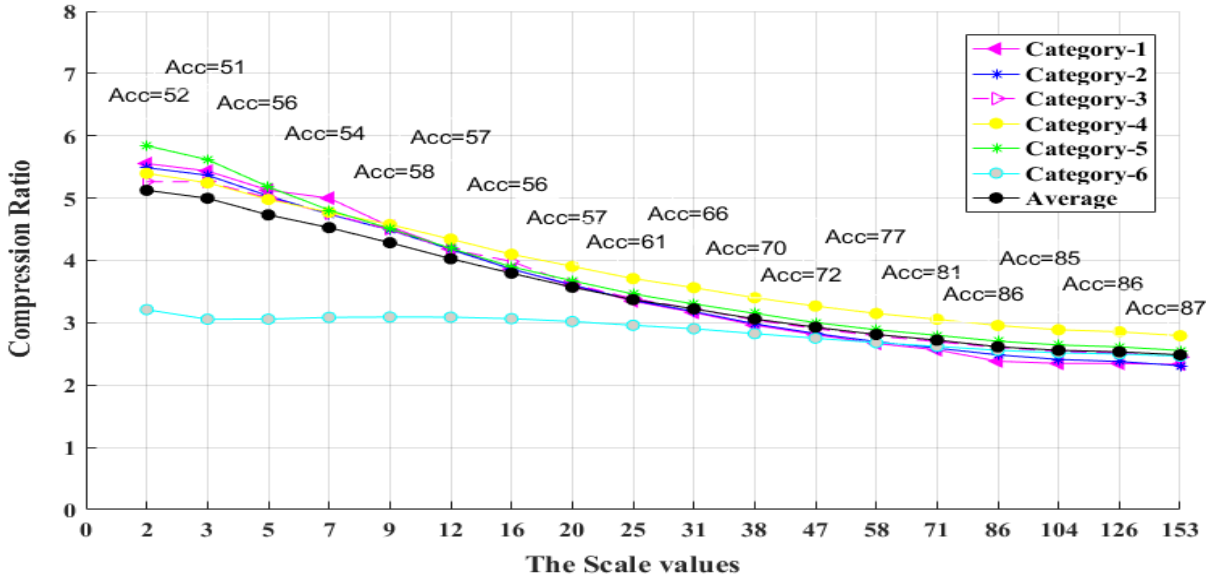


Figure 5-19 The relation between the scale of analysis and the six surface categories compression and classification performance by using the multi-scale LP-datasets

For the six multi-scale band-pass surface categories case, we selected scale of analysis = 71 and average CR =2.16:1 to obtain classification accuracy of 87 % as illustrated in Figure 5-20.

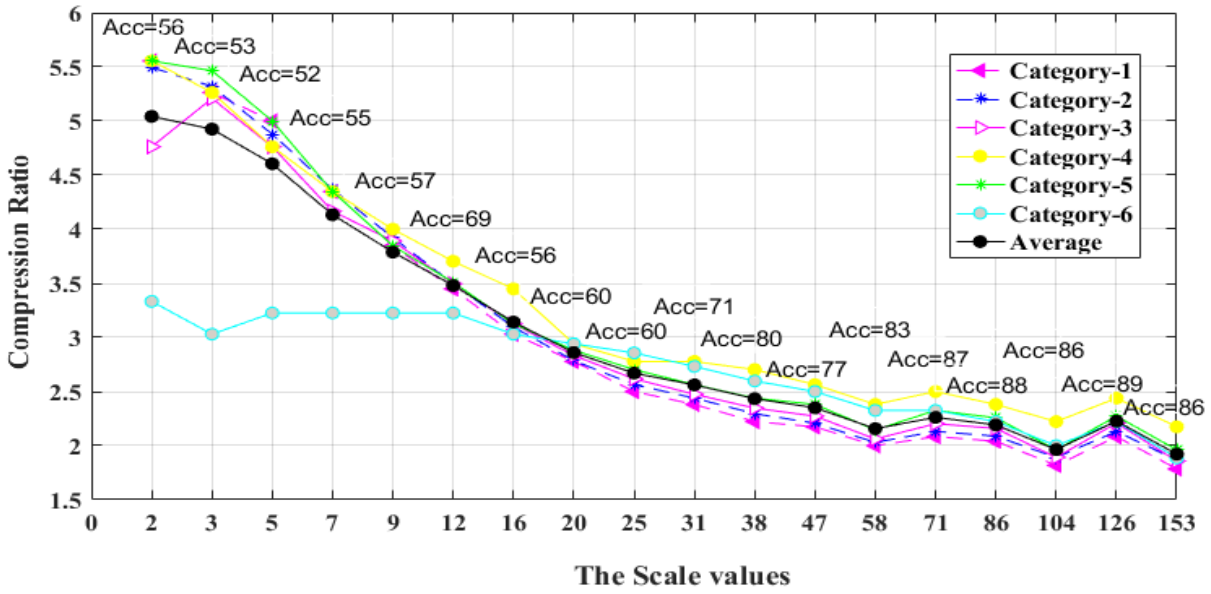


Figure 5-20 The relation between the scale of analysis and the six surface categories compression and classification performance by using the multi-scale BP-datasets.

Finally, the robust performance achieved by using separated scale of HP-datasets, where the scale of analysis has not an impact on the classification accuracy or the compression ratio. In Figure 5-21, there is no significant difference between the compression ratio at the highest and the lowest scale of analysis with compression ratio CR averages = 2.5:1 and 2.2:1 respectively. The best classification accuracy of 93 % was obtained at the highest length-scale 153). Consequently, the six highest-scale high-pass surface categories performances were reported in the confusion matrix shown in Figure 5- 22.

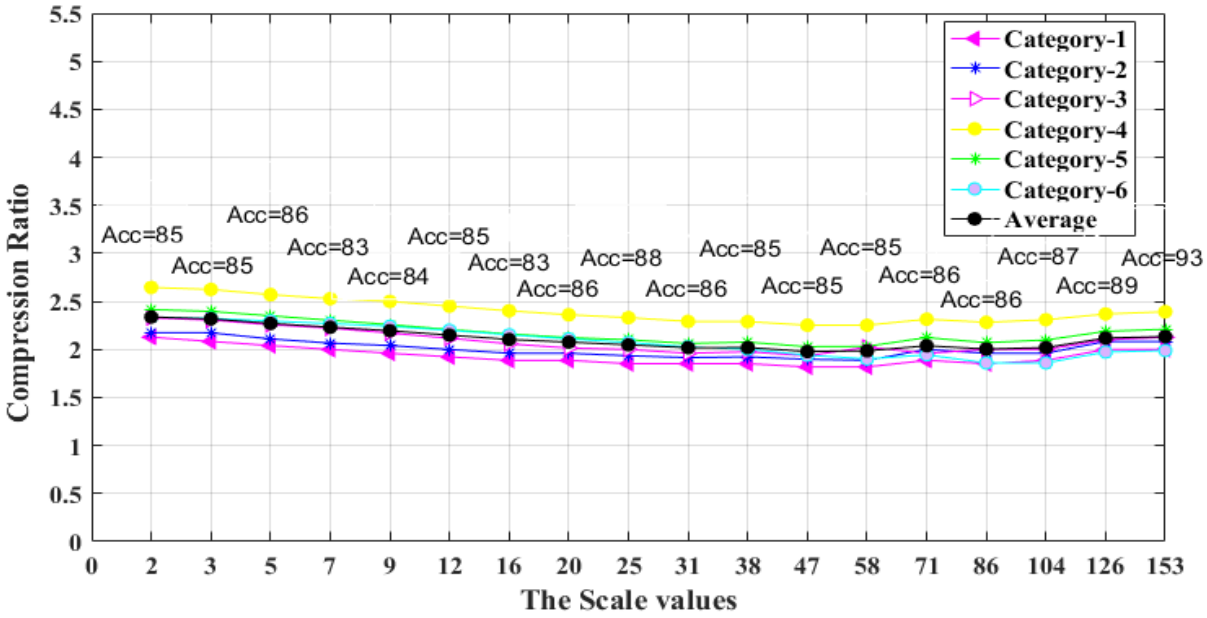


Figure 5-21 The relation between the scale of analysis and the six surface categories compression and classification performance by using the multi-scale HP-datasets.

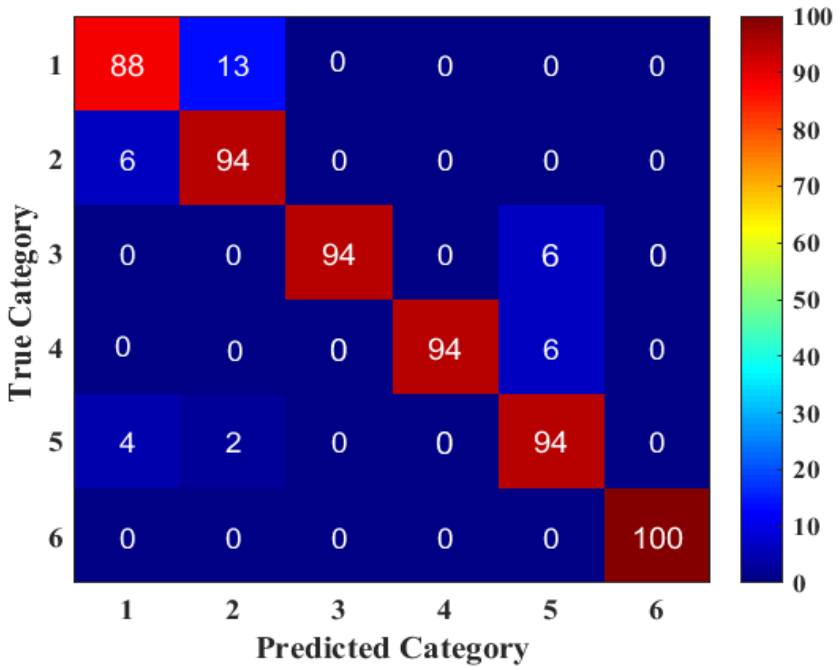


Figure 5-22 Confusion matrix for six surface categories classification by using 60 % of highest-scale HP data set for training.

Conclusion

In this Chapter we evaluated the effects of surface filtering types and scale of analysis on the accuracy of six mechanical multi-scale decomposed surface classification. The surface profile was analyzed by using Gaussian filter multi-scale analyzing technique (HP, LP and BP filters) at all available scales of analysis. We collected three different multi-scale images data sets. The collected datasets were compressed by using HEVC lossless-compression technique. HEVC has guaranteed to preserve the original material parameters. The proposed texture feature descriptor was also extracted from HEVC compressed-domain aiming to reduce the computation complexity. Finally, these extracted feature descriptors were fed into SVM for enhancing the system classification accuracy. The results demonstrated that the robust compressed-domain topographies classifier is either based on the single-scale or multi-scales analyzing methodologies. The high-frequency components (HP-dataset) of the surface profile were more appropriate for characterizing our surface topographies. The best accuracy for the HP image dataset was 96 % and 93 % in the case of multi-scale and single-scale classification respectively.

CHAPTER 6
CONCLUSION AND PERSPECTIVES

In this final chapter, we summarize the achieved contributions of this thesis and we also give research perspectives. The research work presented in this manuscript constitutes a part of the MEgABIt (MEchAnic Big Images Technology) research project supported by UPHF. MEgABIt is an interdisciplinary project which addresses two different areas of expertise: digital images and video compression on one hand, and materials science on the other hand. The MEgABIt project proposes to investigate in the ability to compress large amounts of image data from mechanical instrumentation of large volume deformations. This compression is designed to preserve the maximum of information contained in acquisition systems (high-speed imaging, 3D tomography). Original processing algorithms are developed in the compressed domain in order to make the evaluation of the mechanical parameters possible at the computational level. Also, the classification problem is addressed. This is very challenging because of the high similarity between image pairs in the processed database. In addition to that, storing or transmitting high-resolution images from the MEgABIt database is too expensive without considering the compression techniques.

In this thesis, we use compression as a key point for reducing the computation complexity while preserving intrinsic mechanical parameters. We consider the state-of-the-art HEVC standard for both lossy and lossless compression techniques prior to analysis, classification or storage of image contents, where the decompressed image is used as an input for material image-analysis and classification algorithms.

In **Chapter 2**, we gave a theoretical overview of the background scientific existing approaches in the literature. Section 2.1 introduced the main concepts of digital image and video compression techniques. Section 2.2 illustrated the material surface engineering main processes including surface topography, surface topography measurement, mechanical image deformation analysis, and surface topographical images classification. Finally, Section 2.3 gave the fundamental of Support Vector Machine (SVM).

In **Chapter 3**, we presented an overview of the state-of-the-art HEVC standard. Section 3.1 presented the main new coding tools introduced by the HEVC standard. Section 3.2 gave a comprehensive technical description of the HEVC intra prediction coding technique which was modified concerning to prediction unit sizes down to 4x4 pixels blocks. Finally, the HEVC Lossless coding Profiles are presented as well as its implementation.

In **Chapter 4**, we evaluated the impact of HEVC lossy as well as lossless compression on characterizing material mechanical response from two recorded video sequences corresponding to two different mechanical loading processes: Tensile Test of Polypropylene (PP) Specimen and Sikapower Arcan test. These two sequences were compressed using HEVC Lossy and lossless techniques. Further, these two compressed sequences were reconstructed and post-processed by Digital Image Correlation (DIC) software to extract the mechanical fields. The obtained results demonstrated that HEVC provided very high coding efficiency as well as high reconstructed video quality. The mechanical material response was very well preserved in the compressed sequences at Quantization Parameter ranging from 0 to 20 with an average Structural Similarity or SSIM index nearly equal to one for both Sequence 1 (i.e. dynamic tensile test of PP) and Sequence 2 (i.e. Arcan test of glue joint). We also proposed a HEVC lossy plus lossless coding approach that makes it possible to decompose an image into two separated compressed streams: the first stream is light and allows to display a pre-version of the image with moderate but acceptable quality, while the additional stream makes it possible by adding it to recover the mathematically distortion-free integral image, this which may be necessary for some application contexts.

Finally, we evaluated in **Chapter 5** the impact of both surface filtering type and analysis scale on image classification of six mechanical multi-scale surface categories. Image

classification is performed in the HEVC compressed-domain, where we performed HEVC Lossless coding for reducing the size of the collected database. Texture feature descriptors defined as Intra Prediction Mode Histograms were also extracted from HEVC compressed-domain to reduce the computation complexity. SVM has been retained to distinguish between the highly correlated IPMH pairs either when they were taken from the same category or from different categories. The experiments showed the following results:

1. The Intra Prediction Mode Histograms (IPMHs) feature descriptors were very efficient for characterizing the topographical images with a residual signal nearly equaled to null.
2. The classification accuracy increased as the number of the training datasets increased.
3. The achieved lossless compression ratios up to 6:1 depend on image complexity.
4. The scale of analysis and the achieved compression ratio are inversely proportional, where CR increases as the scale of analysis decreases in cases of: Low and Band-pass filtered image datasets.
5. The results demonstrated that the robust compressed-domain topographies classifier is either based on the single-scale or multi-scales analyzing methodologies.
6. The IPMHs feature descriptors were very efficient for classifying a mix of three multi-scale surface filtered image data sets with accuracy of 85 % and reached 88 % by using highest-scale of analysis.
7. The single scale analysis (highest-scale) was more appropriate than multi scale analysis in the case classifying Low-pass and Band-pass data set separately.
8. The high-frequency components (High-pass-dataset) of the surface profile were more appropriate for characterizing our surface topographies. The best accuracy for the HP image dataset was 96 % and 93 % in the case of multi-scale and single-scale classification respectively.

Further works are numerous; some of them are described herafter:

- Versatile Video Coding (VVC) is already designed as the successor of the current HEVC video coding standard. VVC offers preliminary compression gains of about 40%. Such compression gains are made possible thanks to significantly improved coding tools which

predict closer and closer the real video signal. Hence, VVC should further increase the performances of the HEVC-based solutions proposed in the present work.

- Digital image and video coding tools have been mostly optimized in a HVS point of view. In the present case of the MEGaBIt project, it should be of interest to define a new mechanical-oriented approach which try to optimize the coding process from a mechanical point of view. For instance, motion estimation applied for coding of ultra high-speed videos should be coupled with the DIC process.

REFERENCES

[1] L. Salvo, M. Suéry, A. Marmottant, N. Limodin, and D. Bernard. 3D imaging in material science: Application of X-ray tomography. Comptes Rendus Phys, vol. 11, no. 9–10, pp. 641–649, 2010.

[2] S. C. Mayo, A. W. Stevenson, and S. W. Wilkins. In-Line Phase-Contrast X-ray Imaging and Tomography for Materials Science. Materials (Basel), vol. 5, no. 12, pp. 937–965, 2012.

[3] A. A. G. Bruzzone, H. L. Costa, P. M. Lonardo, and D. A. Lucca. Advances in engineered surfaces for functional performance. CIRP Annals - Manufacturing Technology, vol. 57, no. 2 pp. 750–769, 2008.

[4] M. T. Postek and R. J. Hocken. Instrumentation and metrology for nanotechnology. Report of the National Nanotechnology Initiative Workshop, Gaithersburg, MD, 27–29 January 2004, Available at:
http://www.nano.gov/sites/default/files/pub_resource/nni_instrumentation_metrology_rpt.pdf

[5] National Science Foundation. The future of materials science and materials engineering education. A report from the Workshop on Materials Science and Materials Engineering Education, September 18-19, Arlington, VA (2008).

[6] A Comparison of Scale: Macro , Micro , Nano Learning Module. Southwest Center for Microsystems Education (SCME) NSF ATE Center, 2009.

[7] V. Vakharia, M. B. Kiran, N. J. Dave, and U. Kagathara. Feature extraction and classification of machined component texture images using wavelet and artificial intelligence techniques. 2017 8th Int. Conf. Mech. Aerosp. Eng., pp. 140–144, 2017.

[8] K.J. Stout and L. Blunt. Instruments and Measurement Techniques of Three-Dimensional Surface Topography. Three-Dimensional Surface Topography, pp. 19–94,2000 Elsevier Ltd.

[9] Chaowei Yang, Qunying Huang, Zhenlong Li, Kai Liu & Fei Hu. Big Data and cloud computing: innovation opportunities and challenges, International Journal of Digital Earth, vol. 10, no. 1, 13-53.

[10] E. A. Belyaev, C. Mantel, and S. O. Forchhammer. High bit depth infrared image compression via low bit depth codecs. Infrared Remote Sens. Instrum. XXV, p. 9, 2017.

[11] F. Liu, M. Hernandez-Cabronero, V. Sanchez, M. Marcellin, and A. Bilgin. The Current Role of Image Compression Standards in Medical Imaging," Information, vol. 8, no. 4, p. 131, 2017.

[12] L. Duval, M. Moreaud, C. Couprie, D. Jeulin, H. Talbot and J. Angulo. Image processing for materials characterization: Issues, challenges and opportunities. 2014 IEEE International Conference on Image Processing (ICIP) Paris, pp. 4862-4866, 2014.

[13] Shoyaib, Mohammad et al. “A noise-aware coding scheme for texture classification” Sensors (Basel, Switzerland) vol. 11,8 (2011): 8028-44.

[14] Francisco Barros and all. A DFT-based method for 3D digital image correlation. Procedia Structural Integrity, 2nd International Conference on Structural Integrity, ICSI 2017 Funchal, Madeira, Portugal, vol. 5, pp.1260-1266, 2017.

[15] G. Le Goïc, M. Bigerelle, S. Samper, H. Favrelière, and M. Pillet. Multiscale roughness analysis of engineering surfaces: A comparison of methods for the investigation of functional correlations. Mech. Syst. Signal Process, vol. 66–67, pp. 437–457, 2016.

[16] Iain E. Richardson. The H.264 Advanced Video Compression Standard (2nd Edition) WILEY, 2010.

[17] Cisco - 7 June 2017 – Available at: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html.

[18] DGITAL Tv Europe - 8 February 2017 – Available at: https://www.digitaltveurope.com/2017/02/08/video-to-account-for-three-quarters-of-mobile-traffic-by-2021.

[19] Yun Q. Shi and Huifang Sun. Image and Video Compression for Multimedia Engineering: Fundamentals, Algorithms, and Standards. CRC Press, Inc., Boca Raton, FL, 2008.

[20] Rafael C. Gonzalez and Richard E. Woods. Digital Image Processing (3rd Edition). Prentice Hall, August 2007.

[21] Khalid Sayood. Introduction to Data Compression (Third Edition). Morgan Kaufmann Series in Multimedia Information and Systems, San Francisco, CA, 2005.

[22] Akramullah S. Video Coding Standards. In: Digital Video Concepts, Methods, and Metrics. Apress, Berkeley, CA pp. 30-31 (2014).

[23] Tinku Acharya and Ping-Sing Tsai. JPEG2000 Standard for Image Compression: Concepts, Algorithms and VLSI Architectures. Wiley-Interscience, 2004

[24] T. Eseholi, D. Notta-Cuvier, F. Coudoux, P. Corlay, F. Robache and M. Bigerelle. Performance evaluation of strain field measurement by digital image correlation using HEVC compressed ultra-high speed video sequences. International Symposium on Signal, Image, Video and Communications (ISIVC) Tunis, pp. 142-147, 2016.

[25] Oge Marques, "CH17: image processing and coding. Practical image and video processing using MATLAB, Wiley-IEEE Press, pp. 427-445, 2011.

[26] G. Hudson, A. Léger, B. Niss and I. Sebestyén. JPEG at 25: Still Going Strong. IEEE MultiMedia, vol. 24, no. 2, pp. 96-103, 2017.

[27] M. M. Hasan and K. A. Wahid. Low-Cost Lifting Architecture and Lossless Implementation of Daubechies-8 Wavelets. IEEE Transactions on Circuits and Systems, vol. 65, no. 8, pp. 2515-2523, Aug. 2018.

[28] L. Kau and Y. Lin. Least-Squares-Based Switching Structure for Lossless Image Coding. in IEEE Transactions on Circuits and Systems, vol. 54, no. 7, pp. 1529-1541, July 2007.

[29] J. Uthayakumar, T. Vengattaraman, and P. Dhavachelvan. A survey on data compression techniques :From the perspective of data quality, coding schemes, data type and applications.

Journal of King Saud University - Computer and Information Sciences, 2018, ISSN 1319-1578.

[30] Neeraj Kashyap and Jitesh Modi. JPEG Image Code Format. 140.429.

[31] ITU-T T.87. Information technology – Lossless and near-lossless compression of continuous-tone still images – Baseline (06/98).

[32] Iain E. Richardson. H.264 and MPEG-4 video compression: video coding for next-generation multimedia.Wiley, 2008.

[33] H.R. Wu and K.R. Rao. Digital Video Image Quality and Perceptual Coding. CRC Press November 18, 2005.

[34] Stefan Winkler. Digital Video Quality: Vision Models and Metrics. John Wiley and Sons, 2005.

[35] Z. Wang and A. C. Bovik. Mean squared error: Lot it or leave it? A new look at signal fidelity measures. IEEE Signal Process. Mag., vol. 26, no. 1, pp. 98–117, 2009.

[36] E. Steinbach. Image and Video Compression. 2007.
https://www.tcs.ifi.lmu.de/teaching/ws-2016-17/code/background-image-video-compression

[37] Wei - Yi Wei. An Introduction to Image Compression. National Taiwan University, Taipei, Taiwan, ROC, 2008.

[38] ITU-T T.81. Information technology—Digital compression and coding of continuous-tone still images—Requirements and guidelines (09/92).
https://www.itu.int/rec/T-REC-T.81-199209-I/en

[39] Y. Chen and P. Hao. Integer reversible transformation to make JPEG lossless. Proc. Int. Conf. Signal Processing, pp. 835–838, 2004.

[40] M. Kabir and M. Mondal. Edge-Based and Prediction-Based Transformations for Lossless Image Compression. J. Imaging, vol. 4, no. 5, p. 64, 2018.

[41] G. J. Sullivan and V. Sze. High Eviciency Video Coding (HEVC) Algorithms and Architectures. Springer International Publishing, 2014 https://doi.org/10.1007/978-3-319-06895-4.

[42] Haskell, B. and Puri, A. Mpeg video compression basics. In: The MPEG Representation of Digital Media. Springer, pp. 7–38,. 2012.

[43] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand. Overview of the high efficiency video coding (HEVC) standard. IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 12, pp. 1649–1668, 2012.

[44] P. Yakaiah, C. S. Jakkena, and A. V Paramkusam. Overview of H.265 standards.International Journal of Mechanical Engineering and Technology, vol. 8, no. 8, pp. 48–54, 2017.

[45] S. Vetrivel, M. Gowri, M. S. Sultana, and G. Athisha. An overview of MPEG familly and its applications. Indian Journal of Computer Science and Engineering, vol. 1, no. 4, pp. 240–250.2010.

[46] S. Aramvith and M. T. Sun. MPEG-1 AND MPEG-2 Video Standards. Image and Video Processing handbook (Second edition) Academic press, 1999.

[47] MPEG | The Moving Picture Experts Group website: https://mpeg.chiariglione.org/standards/mpeg-1/video.

[48] T. Sikora. The MPEG-4 video standard verification model. IEEE Transactions on Circuits and Systems for Video Technology, vol.7 n.1, p.19-31, February 1997.

[49] F. Bossen, B. Bross, K. Suhring and D. Flynn. HEVC Complexity and Implementation Analysis. IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1685-1696, Dec. 2012.

[50] T. Jacobs, T. Junge, and L. Pastewka. Quantitative characterization of surface topography using spectral analysis. Surface Topography: Metrology and Properties, vol.5, n.1 2017.

[51] B. Bhushan. Modern Tribology Handbook. (Mechanics and material science series) CRC Press LLC Florida, 2001.

[52] Vorburger, T., Raja, J. Surface Finish Metrology Tutorial (NISTIR 89-4088). U.S. Department of Commerce, National Institute of Standards and Technology, June 1990.

[53] Mitutoyo. Quick Guide to Surface Roughness Measurement. Mitutoyo American Corporation, Bulletin No. 2229.

[54] B. E. Green. Review of surface texture measurement and the associated metrological problems. pp.330-343, PMS* Proceedings, 1967-68.

[55] C. Y. Poon and B. Bhushan. Comparison of surface roughness measurements by stylus profiler. Wear vol. 190, pp. 76–88, 1995.

[56] R. Leach. The measurement of surface texture using stylus instruments. National Physical Laboratory, ISSN 1368-6550 July 2001

[57] M. H. Padgavankar and S. R. Gupta. Big Data Storage and Challenges. Int. J. Comput. Sci. Inf. Technol., vol. 5, no. 2, pp. 2218–2223, 2014.

[58] M. Dalla Costa, M. Bigerelle, and D. Najjar. A new methodology for quantifying the multi-scale similarity of images. Microelectron. Eng., vol. 84, no. 3, pp. 424–430, 2007.

[59] U. Wendt, K. Lange, M. Smid, R. Ray, and K. D. Tönnies. Surface topography quantification by integral and feature-related parameters. Materials Science and Engineering Technology, vol. 33, no. 10, pp. 621–627, WILLEY-VCH 2002.

[60] D. Notta-cuvier, A. Bouzouita, R. Delille, G. Haugou, J. Raquez, and F. Lauro. Design of toughened PLA based material for application in structures subjected to severe loading conditions . Part 1 . Quasi-static and dynamic tensile tests at ambient temperature. Polym. Test., vol. 54, pp. 233–243, 2016.

[61] S. Tung and C. Sui. Application of digital-image-correlation techniques in analysing cracked cylindrical pipes. Sadhana. vol. 35, no. October, pp. 557–567, 2010. https://doi.org/10.1007/s12046-010-0039-4.

[62] J. C. Stinville et al.. Sub-grain scale digital image correlation by electron microscopy for polycrystalline materials during elastic and plastic deformation. Experimental Mechanics, Volume 56, Issue 2, pp 197–216, February 2016.

[63] S. H. Tung and C. H. Sui. Application of digital-image-correlation techniques in analysing cracked cylindrical pipes. Sadhana - Acad. Proc. Eng. Sci., vol. 35, no. 5, pp. 557–567, 2010.

[64] Helin Lu, Chaohong Huang, Cheng Wang, Xiaozhong Wang, Hongyan Fu, and Ziyi Chen. Fast and noninterpolating method for subpixel displacement analysis of digital speckle images

using phase shifts of spatial frequency spectra. Applied Optics Vol. 53, Issue 13, pp. 2806-2814 (2014).

[65] Mark S. Nixon and Alberto S. Aguado. Feature Extraction and Image Processing. Newnes Oxford, pp. 1–30, 2002.

[66] ISO, IEC , 1999 , Information technology - Lossless and near-lossless compression of continuous-tone still images:Baseline , ISO, IES 14495-1 International Standard

[67] Goel R, Kumar V, Srivastava S & Sinha AK. A Review of Feature Extraction Techniques for Image Analysis. International Journal of Advanced Research in Computer and Communication Engineering, vol.6,no.2, pp.153-155, (2017).

[68] R. S. Chora. Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems. International Journal of Biology and Biomedical Engineering vol. 1, no. 1, pp.6-16 (2007).

[69] Y. Mistry, D. T. Ingole, and M. D. Ingole. Content based image retrieval using hybrid features and various distance metric. Journal of Electrical Systems and Information Technology, 2017, http://dx.doi.org/10.1016/j.jesit.2016.12.009

[70] W. K. Pratt. Digital Image Processing (Third Edition). John Wiley & Sons, Inc. (2001) ISBNs: 0-471-37407-5

[71] Y. Mistry, D. T. Ingole, and M. D. Ingole. Content based image retrieval using hybrid features and various distance metric. J. Electr. Syst. Inf. Technol., no. 2016, pp. 1–15, 2017.

[72] X. Wei, S. Lam, and A. Bouzerdoum. Visual descriptors for scene categorization : experimental evaluation. Artif. Intell. Rev., vol. 45, no. 3, pp. 333–368, 2016. https://doi.org/10.1007/s10462-015-9448-4

[73] Y. Lee, H. Ahn, H. Cho, and J. Lee. Object Recognition and Tracking based on Object Feature Extracting. Journal of Internet Services and Information Security (JISIS), volume: 5, number: 3, pp. 48–57, 2015.

[74] C. He, T. Zhuo, X. Su, F. Tu, and D. Chen. Local Topographic Shape Patterns for Texture Description. IEEE Signal Processing Letters vol. 22, no. 7, pp. 871–875, 2015.

[75] E. Hayman, B. Caputo, M. Fritz, and J. Eklundh. On the Significance of Real-World Conditions for Material Classification." In Computer Vision - ECCV 2004, 8th European Conference on Computer Vision. Pro- ceedings, Part IV., pages 253–266, Prague, Czech Republic.

[76] H. Ayad, M. H. Abdulameer, L. E. George, and N. F. Hassan. Descriptor Trends in Texture Classification for Material Recognition Department of Computer Science. · Research Journal of Applied Sciences, Engineering and Technology, vol. 10, no. 10, pp. 1089–1101, 2015.

[77] H. Oulhaj, R. Jennane, A. Amine, M. El, and M. Rziza. Study of the relative magnitude in the wavelet domain for texture characterization. Signal, Image Video Processing vol.12, no.7, pp 1403–1410, 2018.

[78] Y. Dong, J. Feng, L. Liang, L. Zheng, and Q. Wu. Multiscale Sampling Based Texture Image Classification. IEEE Signal Process Letter vol 24, no.5, pp. 614–618, 2017.

[79] L. Liu, P. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen. Local binary features for texture classification: Taxonomy and experimental study. Pattern Recognition, vol. 62, pp. 135–160, 2017.

[80] J. Feng, X. Liu, Y. Dong, L. Liang, and J. Pu. Structural difference histogram representation for texture image classification. IET Image Processing vol.11, no.2, pp. 118–125, 2017.

[81] L. Nanni, M. Paci, and F. Luciano. Texture Descriptors Ensembles Enable Image-Based Classification of Maturation of Human Stem Cell-Derived Retinal Pigmented Epithelium. PLoS ONE Volume 11, Issue 2, pp. 1–29, 2016. DOI:10.1371/journal.pone.0149399.

[82] Matti Pietikainen and Abdenour Hadid. Texture Features in Facial Image Analysis. in Advances in Biometric Person Authentication: International Wokshop on Biometric Recognition Systems, IWBRS 2005, Beijing, China, pp. 22-23, October 2005.

[83] Mohammad Javed. Face Recognition using Principle Gabor Filter. International Conference on Advanced Developments in Engineering and Technology Conference Held in Lord Krishna College of Engineering Ghaziabad, India, vol. 4, no. 2, pp. 176–180, 2014.

[84] M. A. Berbar. Hybrid methods for feature extraction for breast masses classification. Egypt. Informatics J., vol. 19, no. 1, pp. 63–73, 2018.

[85] L. Liu and P. W. Fieguth. Texture Classification from Random Features. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 3, pp. 574 - 586, 2012.

[86] E. Al Daoud. Enhancement of the Face Recognition Using a Modified Fourier-Gabor Filter. International Journal of Advances in Soft Computing and its Applications, Vol. 1, No. 2, ICSRS Publication, 2009.

[87] Zhang, Xin et al. A Study for Texture Feature Extraction of High-Resolution Satellite Images Based on a Direction Measure and Gray Level Co-Occurrence Matrix Fusion Algorithm. Sensors (Basel, Switzerland) vol. 17,7 1474. 22 Jun. 2017.

[88] S.-S. Shai and B.-D. Shai. Understanding Machine Learning:From Theory to Algorithms. Cambridge University Press, 2014.

[89] T. G. Dietterich. Machine learning in ecosystem informatics and sustainability. The 21st International Joint Conference on Artificial Intelligence, Pasadena, California, USA, pp. 8–13. 2009.

[90] M. Praveena and V. Jaiganesh. A Literature Review on Supervised Machine Learning Algorithms and Boosting Process. International Journal of Computer Applications vol. 169, no. 8, pp. 32–35, 2017.

[91] C. Debao. Degree of approximation by superpositions of a sigmoidal function. Approximation Theory and its Applications, vol. 9, no. 3, pp. 17–28, 1993. https://doi.org/10.1007/BF02836480

[92] J. Schmidhuber. Deep learning in neural networks : An overview. Neural Networks, vol. 61, pp. 85–117, 2015.

[93] O. M. Parkhi, A. Vedaldi, and A. Zisserman. Deep Face Recognition. Procedings Br. Mach. Vis. Conf. 2015, no. Section 3, p. 41.1-41.12, 2015.

[94] A. B. Colony. Cancer Classification Based on Support Vector Machine Optimized by Particle Swarm Optimization. Molecules 2017. DOI: 10.3390/molecules22122086.

[95] A. Kampouraki, D. Vassis, P. Belsis, and C. Skourlas. e-Doctor : A Web Based Support Vector Machine for Automatic Medical Diagnosis. The 2nd International Conference on Integrated Information, Procedia - Social and Behavioral Sciences, vol. 73, pp. 467–474, 2013.

[96] M. Malfante, M. D. Mura, J. Métaxian, and J. I. Mars. Machine Learning for Volcano-Seismic Signals. IEEE Signal Processing Magazine, vol. 35 no.2, pp. 20–30, March 2018.

[97] Jaime Ortegon, Rene Ledesma-Alonso, Romeli Barbosa, Javier Vázquez Castillo and Alejandro Castillo Atoche. Material phase classification by means of Support Vector Machines. Computational Materials Science, vol. 148, pp. 336–342 (2018). https://doi.org/10.1016/j.commatsci.2018.02.054

[98] Nello Cristianini and John Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press, 2000.

[99] B.A. Murtagh and M.A. Saunders. Large-scale linearly constrained optimization. Mathematical Programming, vol. 14, pp. 41–72, 1978.

[100] R. J. Vanderbei. LOQO : an interior point code for quadratic programming. Optimization Methods and Software, vol. 11, no. 1-4 pp. 451–484, Taylor & Francis 2008. https://doi.org/10.1080/10556789908805759.

[101] Nurhanna, A. A. and Othman, M. F.. Multi-class Support Vector Machine Application in the Field of Agriculture and Poultry : A Review. Malaysian journal of mathematical sciences (Special Issue: Conference on Agriculture Statistics 2015 (CAS 2015)) vol. 11, pp. 35–52, 2017.

[102] C. Hsu, C. Chang, and C. Lin. A Practical Guide to Support Vector Classification. Department of Computer Science and Information Engineering National Taiwan University Taipei vol. 1, no. 1, pp. 1–12. 2003.

[103] Wei-Lun Chao. A Tutorial on Support Vector Machine. Graduate Institute of Communication Engineering, National Taiwan University Draft version: Dec. 30, 2011.

[104] K. R. Müller, S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf. An introduction to kernel-based learning algorithms. IEEE Trans. Neural Networks, vol. 12, no. 2, pp. 181–201, 2001.

[105] F. Friedrichs and C. Igel. Evolutionary Tuning of Multiple SVM parameters. ESANN'2004 proceedings - European Symposium on Artificial Neural Networks Bruges (Belgium 2004). ISBN 2 930307-04-8, pp. 519-524.

[106] Payam Refaeilzadeh, Lei Tang and Huan Liu. Cross-Validation. Arizona State University Synonyms Rotation, 2008.

[107] Vladimir N. Vapnik. Statistical Learning Theory. John Wiley & Sons, Inc 1998.

[108] Bernhard Sch¨olkopf and Alexander J. Smola. Learning with Kernels. The MIT Press Cambridge, Massachusetts London, England (2001).

[109] C. Chang and C. Lin. LIBSVM : A Library for Support Vector Machines. ACM Trans. Intell. Syst. Technol., vol. 2, pp. 1–39, 2013.

[110] F. Amish and E. Bourennane. Fully pipelined real time hardware solution for High Efficiency Video Coding (HEVC) intra prediction. J. Syst. Archit., vol. 000, pp. 1–15, 2015.

[111] Gary Sullivan, "JCTVC-L_Notes_dB." Meeting report of the 12th meeting of the Joint Collaborative Team on Video Coding (JCT-VC), Geneva, CH, 14–23 Jan. 2013.

[112] F. Rahmani and F. Zargari. Temporal feature vector for video analysis and retrieval in high efficiency video coding compressed domain. Electronics Letters., vol. 54, no. 5, pp. 294–295, 2018.

[113] D. Flynn et al.. Overview of the Range Extensions for the HEVC Standard : Tools , Profiles , and Performance. IEEE Transactions on Circuits and Systems for Video Technology vol. 26, no. 1, pp. 4–19, 2016.
[114] S. R. Alvar and F. Kamisli. Lossless Intra Coding in HEVC with 3-tap Filters. Image Vision and Computing (ICIVC) International Conference, pp. 124 - 128 (2016).
[115] S. Hong, J. H. Kwak, and Y. Lee. Cross residual transform for lossless intra-coding for HEVC. Signal Processing : Image Communication., vol. 28, no. 10, pp. 1335–1341, 2013. https://doi.org/10.1016/j.image.2013.09.004
[116] Fatih Kamisli. Lossless intra coding in HEVC with integer-to-integer DST. 24th European Signal Processing Conference (EUSIPCO) pp. 2440–2444, (2016). DOI: 10.1109/EUSIPCO.2016.7760687.
[117] J. Liang, S. Member, and T. D. Tran. Fast Multiplierless Approximations of the DCT With the Lifting Scheme. IEEE Transactions on Signal Processing vol. 49, no. 12, pp. 3032–3044, 2001.
[118] L. Dong, W. Liu and A. Tabatabai. Improved Chroma Intra Mode Signaling JCTVC-D255. Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 4th Meeting: Daegu, KR, 20-28 January, 2011.
[119] V. Sanchez, F. Aulí-llinàs, J. Bartrina-rapesta, and J. Serra-sagristà. HEVC-based Lossless Compression of Whole Slide Pathology Images. IEEE Global Conference on Signal and Information Processing (GlobalSIP) pp. 297–301, 2014.
[120] V. Sanchez, F. Aulí-llinàs, S. Member, and J. Serra-sagristà. Piecewise Mapping in HEVC Lossless Intra- prediction Coding. IEEE Transactions on Image Processing vol: 25, issue: 9, pp. 4004 - 4017, 2016.
[121] M. Zhou, W. Gao, M. Jiang, and H. Yu. HEVC Lossless Coding and Improvements. Transactions Letters vol. 22, no. 12, pp. 1839–1843, 2012.
[122] Y. Lee, K. Han, and G. J. Sullivan. Improved lossless intra coding for H.264/MPEG-4 AVC. IEEE Transactions on Image Processing vol. 15, no. 9, pp. 2610–2615, 2006.
[123] Gary Sullivan and Jens-Rainer Ohm. JCTVC-M_Notes_d0. Meeting report of the 13th meeting of the Joint Collaborative Team on Video Coding (JCT-VC), Incheon, KR, 18-26 Apr. 2013.
[124] Tung Nguyen ; Detlev Marpe. Objective Performance Evaluation of the HEVC Main Still Picture Profile. IEEE Transactions on Circuits and Systems for Video Technology vol. 25, no. 5, pp. 790–797, 2015.
[125] D. Flynn et al.. Overview of the Range Extensions for the HEVC Standard : Tools , Profiles and Performance. IEEE Transactions on Circuits and Systems for Video Technology Volume: 26 , Issue: 1, pp. 4–19, 2015.
[126] John C. Russ and F. Brent Neal. The Image Processing Handbook (7th ed). CRC Press 2017 ISBN 9781138747494 - CAT# K32564
[127] M. A. S. J. Orteu and H. W. Schreier. Image Correlation for Shape, Motion and Deformation Measurements. Springer 2009.
[128] A. F. Epee, F. Lauro, B. Bennani, and B. Bourel. Cons titutive model for a semi-crystalline polymer under dynamic loading. Int. J. Solids Struct., vol. 48, no. 10, pp. 1590–1599, 2011.
[129] Surface Texture (Surface Roughness, Waviness and Lay) Executive Summary. the American Society of Mechanical Engineers 2001.
[130] M. Mehrabi, F. Zargari, and M. Ghanbari. Compressed domain content based retrieval using H.264 DC-pictures. Multimed. Tools Appl., vol. 60, no. 2, pp. 443–453, 2012.
[131] Farzad Zargari, and Farzaneh Rahmani. Visual Information Retrieval in HEVC Compressed Domain. 2015 23rd Iranian Conference on Electrical Engineering 2015.

[132] M. Yamaghani and F. Zargari. Classification and retrieval of radiology images in H.264/AVC compressed domain. Signal, Image Video Process., vol. 11, no. 3, pp. 573–580, 2017.
[133] F. Zargari, M. Mehrabi, and M. Ghanbari. Compressed Domain Texture Based Visual Information Retrieval Method For I-frame Coded Pictures. IEEE Transactions on Consumer Electronics, vol. 56, no. 2, pp. 728–736, 2010.
[134] F. Rahmani and F. Zargari. Compressed domain visual information retrieval based on I-frames in HEVC. Multimed. Tools Appl., vol. 76, no. 5, pp. 7283–7300, 2017.
[135] Jason Brownlee. Crash Course On Multi-Layer Perceptron Neural Networks. on May 17, 2016.https://machinelearningmastery.com/neural-networks-crash-course.
[136] N. T. Carnevale and M. L. Hines. The NEURON book. Cambridge University Press 2005.
[137] S. Knerr, L. Personnaz and J. Dreyfus. Single-layer learning revisited: A stepwise procedure for building and training a neural network. Neurocomputing: Algorithms, Architectures and Applications, volume F68 of NATO ASI Series, Springer-Verlag, pp. 41-45 (1990).