International Series in
Operations Research & Management Science

Chiwoo Park
Yu Ding

Data Science for Nano Image Analysis

Springer

International Series in Operations Research & Management Science

Volume 308

The book series *International Series in Operations Research and Management Science* encompasses the various areas of operations research and management science. Both theoretical and applied books are included. It describes current advances anywhere in the world that are at the cutting edge of the field. The series is aimed especially at researchers, doctoral students, and sophisticated practitioners. The series features three types of books:

- *Advanced expository* books that extend and unify our understanding of particular areas.
- *Research monographs* that make substantial contributions to knowledge.
- *Handbooks* that define the new state of the art in particular areas. They will be entitled

Recent Advances in (name of the area). Each handbook will be edited by a leading authority in the area who will organize a team of experts on various aspects of the topic to write individual chapters. A handbook may emphasize expository surveys or completely new advances (either research or applications) or a combination of both.

The series emphasizes the following four areas: *Mathematical Programming*: Including linear programming, integer programming, nonlinear programming, interior point methods, game theory, network optimization models, combinatorics, equilibrium programming, complementarity theory, multiobjective optimization, dynamic programming, stochastic programming, complexity theory, etc. *Applied Probability*: Including queuing theory, simulation, renewal theory, Brownian motion and diffusion processes, decision analysis, Markov decision processes, reliability theory, forecasting, other stochastic processes motivated by applications, etc. *Production and Operations Management*: Including inventory theory, production scheduling, capacity planning, facility location, supply chain management, distribution systems, materials requirements planning, just-in-time systems, flexible manufacturing systems, design of production lines, logistical planning, strategic issues, etc. *Applications of Operations Research and Management Science*: Including telecommunications, health care, capital budgeting and finance, marketing, public policy, military operations research, service operations, transportation systems, etc.

More information about this series at http://www.springer.com/series/6161

Chiwoo Park • Yu Ding

Data Science for Nano Image Analysis

Springer

Chiwoo Park
Florida State University
Tallahassee, FL, USA

Yu Ding
Industrial & Systems Engineering
Texas A&M University
College Station, TX, USA

ISSN 0884-8289 ISSN 2214-7934 (electronic)
International Series in Operations Research & Management Science
ISBN 978-3-030-72821-2 ISBN 978-3-030-72822-9 (eBook)
https://doi.org/10.1007/978-3-030-72822-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To our parents:

Chanbum Park and Yongok Lee
Jianzhang Ding and Yujie Liu

and to our families:

Sun Joo Kim, Layna Park and Leon Park
Ying Li and Alexandra Ding

Foreword

The ability of (scanning) transmission electron microscopy (S/TEM) to obtain images on the nanometre to atomic scale has revolutionised our scientific understanding of structure-property relationships in structural, electronic and biomedical materials. More recently, there has been a significant expansion in the use of in-situ/operando gas and liquid stages that moves beyond static samples and/or images and now permits (S)TEM images to be acquired at defined intervals during precisely controlled kinetic reactions. When these in-situ/operando experiments use most advanced CMOS detectors, there is the potential to image dynamic functionality on the scale of individual molecular interactions and for unprecedented fundamental insights into a wide range of next-generation materials to be developed. While it is always tempting for any experimentalist to believe that the imaging challenge is complete when the hardware is designed, installed and commissioned, this book on data science for nano image analysis by Chiwoo Park and Yu Ding shows clearly and explicitly that the hardware is actually just the start of the scientific process. There are tremendous advances in our understanding of dynamics that image analytics can bring to any problem that is being studied in a (S)TEM.

The comprehensive description of the use of data analytical methods for image analysis in this book contains the authors' combined insights from 15 years of pioneering the use of analytics for dynamic imaging in (S)TEM. The contents of the book are structured to take the reader from the basic mathematical definitions of an image and its information content through to the analysis of the main changes taking place in any dynamic sequence: morphology, spacing, temporal evolution and motion/interactions. For students/experimentalists just getting started on the use of analytics, the book follows through the mathematics for a series of examples that are important to imaging both inorganic and organic systems, such as nucleation and growth, coalescence, corrosion and decay. With the current rapid advancement in the use of artificial intelligence (AI) to power scientific discovery, the methods in this book are ideally suited to educate researchers in the benefits of AI for their research and help them with their initial implementation.

This book successfully integrates data science with imaging in a way that allows a new practitioner of the methodology to work through concepts established in each field (data science and imaging) without getting bogged down by differing terminologies and complex mathematical concepts. As such, it is an ideal reference for graduate courses in this area and for scientists who maybe many years from graduate school but now wish to enhance their understanding of new methods for image analysis. This book will be a key reference for me, my research group and our collaborators going forward, and I am grateful for the authors' major accomplishment in completing this text.

Liverpool, UK Nigel Browning
January 2021

Preface

Scientific imaging is one of the crucial steps in biological and materials sciences to see and understand small-scale phenomena that cannot be seen with human eyes. In the earlier years when the resolution of scientific imaging instruments was not very high and the process of imaging was not automated, the volume of the resulting images was manageable. Analyzing the scientific images was by and large manual in nature, yet the manual process was not considered a big burden then. When we first worked on material images nearly 15 years ago, the high-end electron microscope that we could access produced 500 $\times$ 500-pixel images with the resolution of several nanometers per pixel. The time spent from sample preparation to the actual microscope imaging session was over several hours. In a single imaging session, roughly tens of microscopic images were produced. Despite all the complexities in the resulting images and the labor-intensive process of obtaining them, the manual analysis of material images was the norm and default in those days.

Driven by scientific curiosity to see smaller scale objects and fast-changing phenomena and aided with technological prowess and innovations, the improvement in spatial and temporal resolutions of scientific images has seen a dramatic acceleration in the past decade, and the pace of the improvement follows the Moore's Law in the semiconductor industry. Consequently, the data volume and generation rate of scientific images increase at an unprecedented speed. Nowadays the spatial resolution of electron microscopes goes easily below sub-nanometers, resolving every single atom. Millions of such high-resolution images could be produced every second while capturing interesting physical, chemical, and biological phenomena. Manual analysis of such huge volumes of scientific images, easily reaching terabytes per one imaging session, is no longer feasible. If indeed attempted, the manual analysis could take months of dedicated work by a team of experts to obtain a partial analysis. There is a pressing need for automated image analysis of scientific images in order to keep up with the fast pace of data generation.

In spite of its importance, automated material image analysis has progressed relatively slowly, if one compares it to, say, the progress of bio-image analysis. There is a void in terms of research monographs and textbooks dedicated to the topic of automated material image analysis. This void motivated us to write this book covering a broad aspect of nano image analysis problems as well as recent research developments. While the contents of the book are presented in the specific context of nanomaterials, many methods covered herein can in fact be applicable to general material image analysis.

This book describes comprehensively the necessary steps in material image analysis, including mathematical representation of material images, material object detection, separation, and recognition, size and shape analysis, spatial pattern recognition at both local and global scales, time-series modeling of temporally resolved material images, visual tracking of a population of or individual materials for understanding or quantifying their changes, and image super-resolution that enhances the quality of raw images and benefits the downstream analyses.

We intend for this book to be a good reference both for material scientists who are interested in automated image analysis for accelerating their research and for data scientists who are interested in developing machine learning/data science solutions for materials research.

Tallahassee, FL, USA
May 2021

Chiwoo Park

College Station, TX, USA
June 2021

Yu Ding

Acknowledgments

We would like to acknowledge the contribution of our co-authors and collaborators to the research work that forms the backbone of this book: Patricia Abellán, Lawrence Drummy, Ali Esmaieeli, Rolland Faller, Jianhua Huang, Jim Ji, Xiaodong Li, Xin Li, Bani Mallick, Layla Mehdi, Trevor Moser, Lucas Parent, Joseph Patterson, Yanjun Qian, Brian Smith, Anuj Srivastava, Mollie Touve, Garret Vo, David Welch, and Jiaxi Xu.

A special thanks goes to the fellow microscopists: Nigel Browning, James Evans, Nathan Gianneschi, Hong Liang, and Taylor Woehl. The collaborations with them motivated most of the research reported in this book. They provided material images that are hard to get elsewhere and shared their knowledge and experiences in materials image analysis. Our research and contribution would not have been possible without their collaborations.

We would also like to acknowledge the generous support from our sponsors, particularly, the Dynamic Data and Information Processing (DDIP) program (formerly the Dynamic Data-Driven Applications Systems, or the DDDAS program) of the Air Force Office of Scientific Research (AFOSR) and the Cyber-Physical System (CPS) program of the National Science Foundation. We extend a special thanks to Dr. Frederica Darema, former Director of AFOSR, for her appreciation in our work and heartfelt encouragement when we faced difficulties. The countless stimulating discussions with Dr. Darema drove our research to new levels.

The planning for this book started when Camille Price, Series Editor, Springer International Series in Operations Research and Management Science, reached out to us in 2018. We appreciate Camille for her patience and guidance at the book-proposal stage and for meeting with us a couple of times. In the ensuing years during the book-writing process, the Springer team did a fantastic job in managing this project and assisting us in numerous occasions.

Chiwoo is grateful to Florida State University for supporting a year-long sabbatical leave. That time period gave him more freedom and allowed him to concentrate on writing this book. The book's writing would have taken much longer to complete without this support. Thanks to his former students, Xin Li, Garret Vo, and Ali Esmaieeli, for some of the excellent works in this book.

Yu is grateful to his former students and co-authors, Yanjun Qian and Jiaxi Xu, for arranging the image data and computer code for Chaps. 5, 7, 9, and 11. Yu expresses his gratitude to Mike and Sugar Barnes for their generous endowment that supports his work in this important field. Yu is also indebted to his Ph.D. advisor, Dr. Jianjun Shi, at the Georgia Institute of Technology for bringing him to the data science world and for teaching him to be an independent researcher.

Last but not least, Chiwoo would like to express sincere thanks to his wife Sun Joo Kim who has been very supportive of this year-long book-writing project during the difficult time of the global outbreak of COVID-19. Without Sun Joo's sacrifice, Chiwoo could not have been able to concentrate on writing the book. Yu would like to thank his wife, Ying, and their daughter, Alexandra, for their love and support.

Contents

Acronyms

ADMM	Alternating direction multiplier method
AIC	Akaike information criterion
AMISE	Asymptotic mean integrated squared error
ANOVA	Analysis of variance
ARL	Average run length
ARMA	Autoregressive moving average
BIC	Bayesian information criterion
BSP	Binary segmentation process
CSR	Complete spatial randomness
DP	Dirichlet process
EDSR	Enhanced deep-residual networks super-resolution
EM	Expectation maximization
fps	Frames per second
GPU	Graphics processing unit
HPRC	High performance research computing
ID	Index of dispersion
IEEE	Institute of Electrical and Electronics Engineers
i.i.d	Independently, identically distributed
IG	Inverse gamma
LASSO	Least absolute shrinkage and selection operator
LB	Library-based
LP	Linear programming
NLM	Non-local mean
MAP	Maximum a posterior
MCMC	Markov chain Monte Carlo
MCMC-DA	Markov chain Monte Carlo data association
MLE	Maximum likelihood estimation or maximum likelihood estimator
MOTA	Multi-object tracking analysis
MSE	Mean squared error
MWDA	Multi-way minimum cost data association
NN	Nearest neighborhood

NPSD	Normalized particle size distribution
OA	Orientated attachment
OR	Ostwald ripening
PC	Principal component(s)
PCA	Principal component analysis
pdf	Probability density function
PELT	Pruned exact linear time
PSD	Particle size distribution
PSNR	Peak signal-to-noise ratio
RCAN	Residual channel attention networks
ReLU	Rectified linear unit
ScSR	Sparse-coding based super-resolution
SE	Shannon entropy
SEM	Scanning electron microscope
SISR	Single image super-resolution
SK	Skewness
SNR	Signal-to-noise ratio
SSIM	Structural similarity index measure
SPC	Statistical process control
SPM	Scanning probe microscope
SR	Super resolution
SRCNN	Super-resolution convolutional neural network
SRSW	Super-resolution by sparse weight
SQC	Statistical quality control
SSE	Sum of squared errors
STEM	Scanning transmission electron microscope
STM	Scanning tunneling microscope
SVM	Singular value decomposition
TEM	Transmission electron microscope
VDSR	Very deep convolutional neural network for super resolution
UE	Ultimate erosion
UECS	Ultimate erosion for convex set
WBS	Wild binary segmentation

Chapter 1
Introduction

Materials, according to Merriam-Webster online, are "*the elements, constituents, or substances of which something is composed or can be made*." Invention or discovery of new materials has significantly influenced the course of civilization. For example, the discovery of metals prompted the transition from the Stone Age to the Bronze Age, changing drastically the ways to grow crops and enabling people to settle in cities and countries. Various metals and metal processing techniques brought out a series of transitions from the Bronze Age through the Iron Age to the industrial revolution, fundamentally altering human societies. The discovery of electrons and silicon technology led to the development of vacuum tubes, transistors, and semiconductors. More recent materials discoveries include nanomaterials and biomaterials, both of which are regarded as promising materials to shape the future of our lives.

Nanoscience and engineering deal with a specialty material—nanomaterials, defined as materials with nanoscale structures, i.e., either some external dimensions or internal structures are in the nanoscale. By *nanoscale*, we refer to a length scale between one and one hundred nanometers. Figure 1.1 presents a few such examples, of which a nanoparticle is a nanomaterial with all of its external dimensions in the nanoscale, a nano-fiber (nanotudes as hollow nanofibers and nanorods as solid nanofibers) has two of its dimensions in the nanoscale, and a nanoplate has only one of its dimensions in the nanoscale.

The basis of modern materials science and engineering is a trilogy of processing, structure, and properties of materials. The trilogy starts off with studying the structure of a material and its relation to the properties and performance of the material. Understanding of and insights into the structure-property correlation lead to designing the right structure as well as developing new processing or synthesis for making the structure. As such, the desired properties for intended uses can be materialized. Studying the processing-structure-property interrelation appears to be the major paradigm of the modern materials science and engineering (Reid 2018). This major paradigm certainly influences the study of nanoscience and engineering.

C. Park, Y. Ding, *Data Science for Nano Image Analysis*, International Series in Operations Research & Management Science 308,
https://doi.org/10.1007/978-3-030-72822-9_1

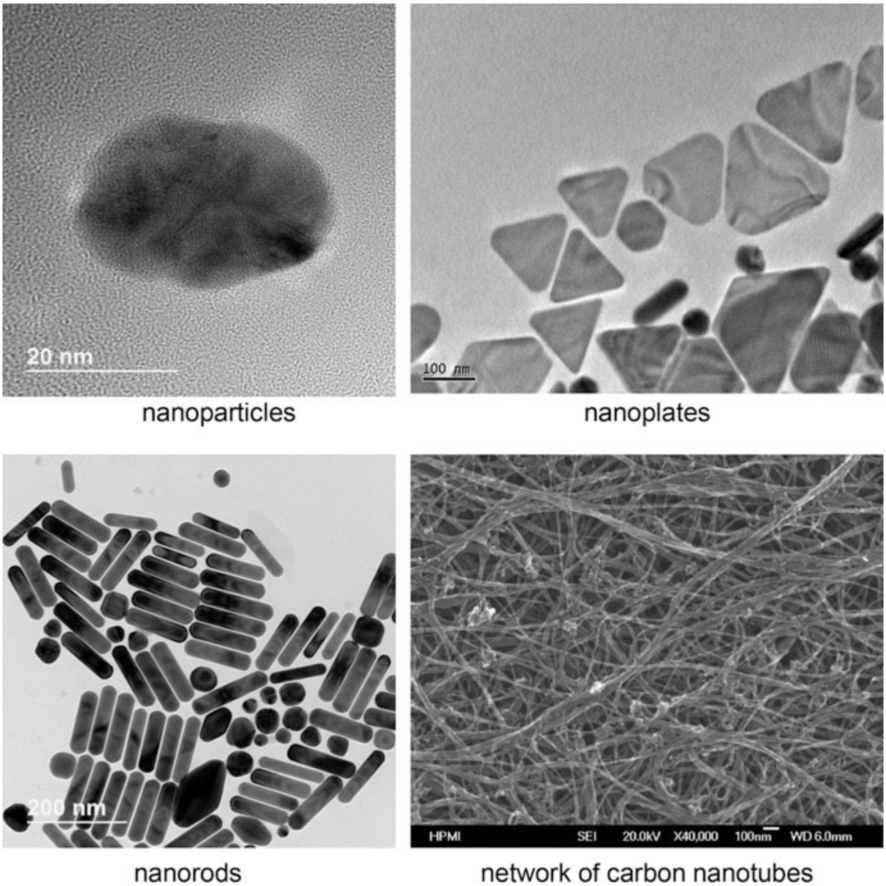

Fig. 1.1 Examples of nanomaterials

In the nanoscale, the structure of a material is more influential to the material's properties or performance, than in traditional scales, making the nanoscale structure an even more crucial factor to be researched and characterized.

While the processing-structure-properties correlation has been studied with a combination of theory, computation, and experiments, metrology and processing techniques play a pivotal role in advancing the capability of materials characterization (Fig. 1.2). There are two major techniques for nanoscale structural characterization—the microscopy imaging techniques and spectroscopy techniques. Both techniques produce imagery data. The microscopy techniques produce real-space images, i.e., images in a spatial domain, while the spectroscopy techniques produce reciprocal-space images, i.e., images in the frequency domain. Theoretically, a reciprocal-space image can be converted to a real-space image by taking the inverse Fourier transform or something equivalent. For many instruments in use,

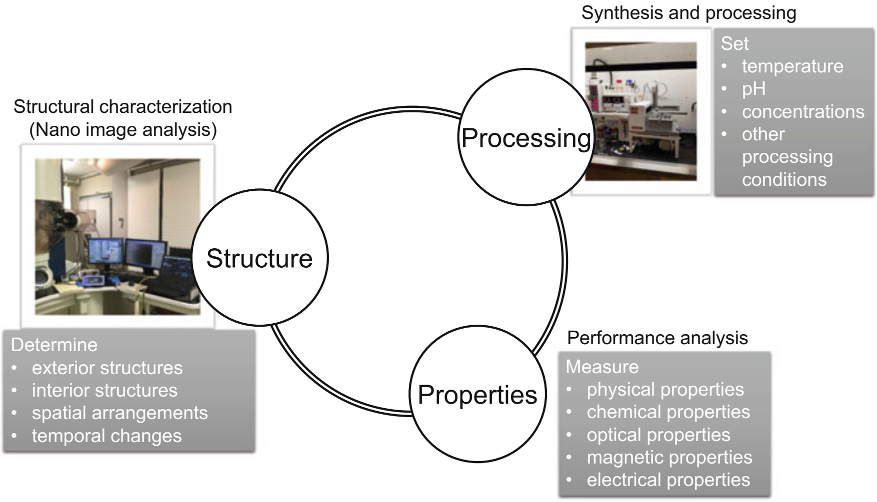

Fig. 1.2 The role of nano image analysis in nanomaterials research, which fulfills principally the characterization of nanoscale structures of materials synthesized under applied processing conditions

however, the reciprocal-space images are taken only in a single dimension, which hampers such inverse conversion. In other words, the reciprocal-space image are mostly in a functional form, rather than in imagery forms.

In this book, the real-space images taken from nanomaterials by microscopes are referred to as nano images. Most of the nano images are understandably two-dimensional, but with recent advances in microscopy, three-dimensional or four-dimensional images are also made available. The three dimensions can be either just three spatial dimensions or two spatial dimensions plus one temporal dimension. Likewise, the four dimensions can be three spatial dimensions plus one temporal dimension, or two spatial dimensions in the real space plus the other two dimensions in the reciprocal space. Chapter 2 will provide a detailed account of the existing microscopes and the corresponding image formats.

Broadly speaking, the purpose of nano image analysis, to be discussed in this book, is to analyze nano images and extract the structural information of nanomaterials from the images. With the processing-structure-property paradigm in mind, it is not difficult to see the significant role that nano image analysis plays, which is a key part in structural characterization and the important enabler for many subsequent actions in materials discovery, design, and synthesis. We want to stress that nano image analysis, like the general image analysis, is basically a data science, considering the rich amount of data, embedding all sorts of complexities, that the nano images bring about. There are certainly specialized tools used for image analysis and processing, but the readers will discover that many modeling tools used in analyzing image data nowadays are also commonly used for other data science problems. In this book, while touching upon basic image processing aspects such as binarization and edge detection, much more attention is given to how to model material objects and structures at a higher level after the initial steps of image preprocessing.

Nano image analysis has a lot in common with generic image analysis. It also has its uniqueness. Nano images are obtained predominantly by electron microscopes, meaning that the physics behind the imaging process is profoundly different from that of visible light microscopes and cameras. The nano objects under observation have their unique characteristics and behaviors, which calls for domain knowledge-guided solutions. One such example is the tailored ultimate erosion procedure that can identify nanoparticles more accurately when the shape convexity of nanoparticles is considered. The nanostructures also presents a range of unique research problems, which may or may not be prevalent in other domains or applications.

Depending on the structures of interest, various kinds of nano image analyses can be performed. Section 1.1 introduces nano image analysis and discusses pertinent research problems by examples. Section 1.2 presents a summary of the topics to be covered in this book and how they are organized. Section 1.3 describes how this book can be used for various groups of intended readers. The last section of this chapter lists the information for online supplementary materials, including datasets and example codes.

1.1 Examples of Nano Image Analysis

The structural features of interest in nano image analysis include individualized features such as morphology (size and shape), facets and interior organization (interior atomic arrangements), as well as ensemble features such as the distribution of the individualized features of a group of nano objects and the local and global spatial arrangements of individual objects (dispersion and network). When a time-resolved material imaging capability is available, the temporal evolution of these structural features for a history of the applied processing conditions is of great interest as well.

In the sequel, we describe a number of nano image analyses for the purpose of delivering concrete examples and guiding the reading of this book. Four examples are presented, as shown in Fig. 1.3, exemplifying the four main categories of nano image analysis. The structural information of interest is highlighted in Fig. 1.3 for each category of the analysis, which is, from top-left to bottom-right, the morphology (size and shape), spatial arrangements (location, dispersion and lattice), distribution changes, and motion and interaction.

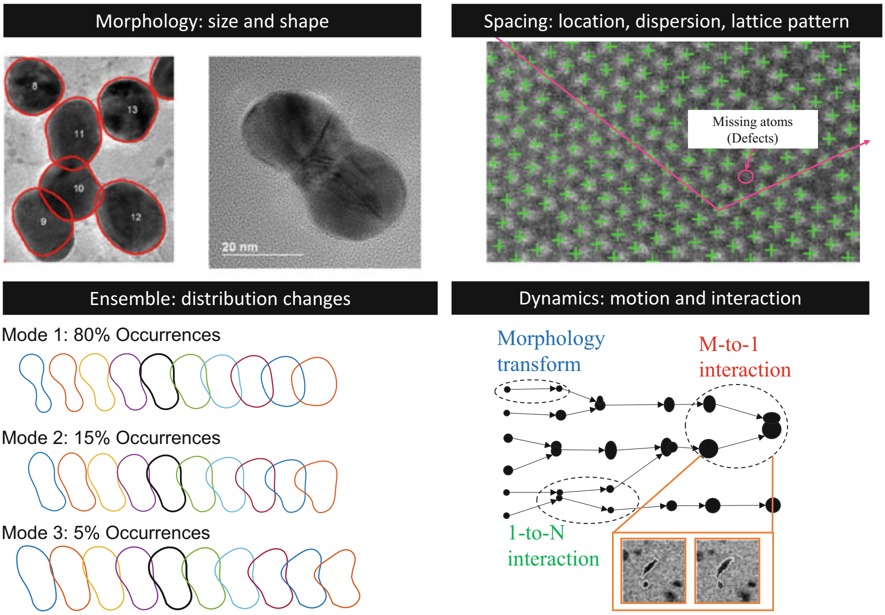

Fig. 1.3 Four main categories of nano image analysis to be discussed in this book

1.1.1 Example 1: Morphology

Our first example is the morphology analysis of nanoparticles. Morphology analysis is to analyze the sizes, shapes and other structural parameters related to the exterior outlines of materials. The morphology analysis of nanoparticles is an essential step in the scientific studies of the morphology-property correlation of the nanoparticle-embedding materials. It is also an integral component when people want to check whether the sizes and shapes of synthesized nanoparticles are within the desirable ranges or design intents, an important question to be addressed for quality monitoring of nanoparticle fabrications.

The input data for morphology analysis is a microscope image of nanoparticles. As illustrated in Figure 1.4, the exterior outlines of nanoparticles shown in the input image are extracted (the top plot). The size and shape information is quantified using the extracted outlines. Then, the size and shape information can be statistically analyzed to provide important statistics such as the probability distributions of sizes, shapes, aspect ratios, and so on. Below summarizes the data analysis sequence:

$$\text{image} \Rightarrow \text{outlines} \Rightarrow \text{sizes, shapes, other structure parameters} \Rightarrow \text{statistics}$$

The step of the analysis encounters technical challenges, including how to identify material objects and their outlines buried under noises, how to recover outline occlusions due to overlaps among materials, how to define and model the

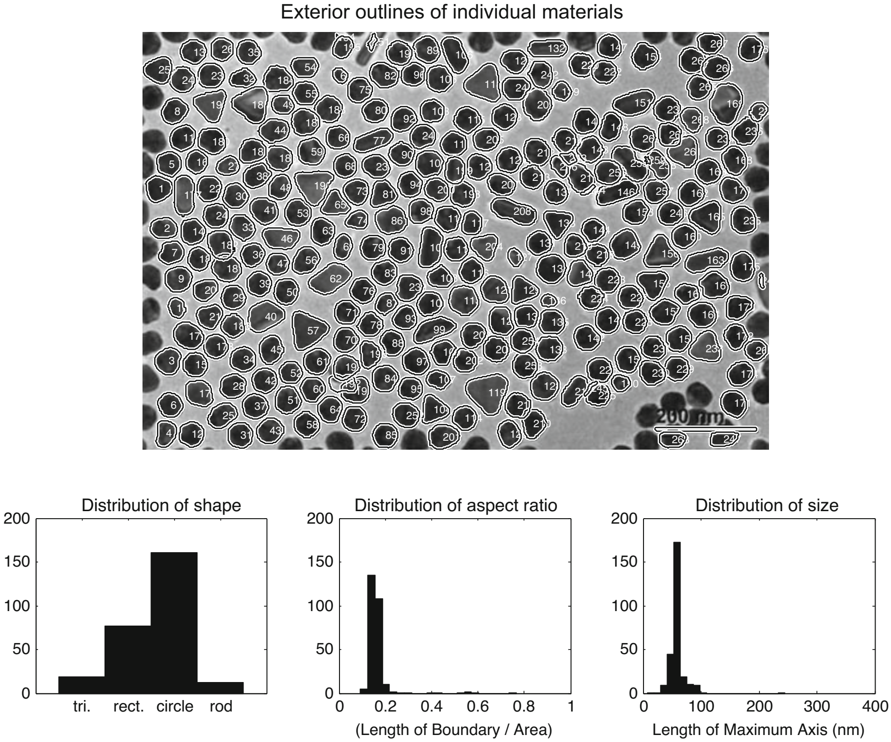

Fig. 1.4 Morphology analysis

sizes and shapes of the outlines mathematically, how to define the probability spaces of the sizes and shapes, how to cluster or group by sizes and shapes and calculate the mean shape and its covariance for a group of materials, and how to perform statistical inferences on a shape space such as model parameter estimation and hypothesis testing. Methodologies addressing these challenges come from the studies of image segmentation in computer vision, statistical shape analysis in statistics, and computational geometry in applied mathematics.

1.1.2 Example 2: Spacing

Our second example is about the interior micro-patterns of materials—specifically, the spatial positioning and arrangements of smaller scale elements within materials. Figure 1.5 presents an example, which is the crystallographic structure of Mo-V-M-O catalysts captured by a high-angle annular dark field scanning transmission electron microscope (HAADF-STEM). The top figure in Fig. 1.5 shows the STEM image taken at an atomic resolution, where each atom is imaged as a white spot over the dark background. Given the input image, many research questions naturally follow, for example, how atoms are spaced, whether the spacing is symmetric (the

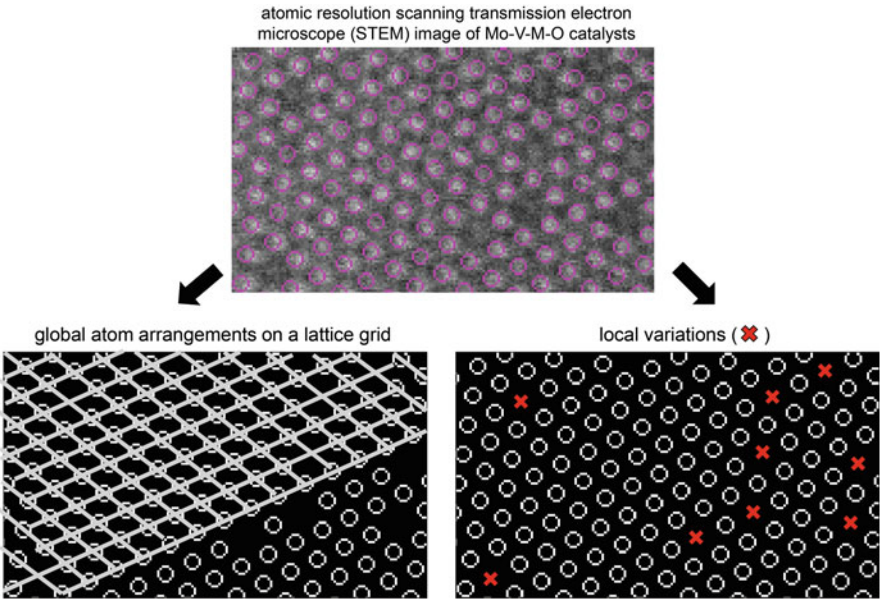

Fig. 1.5 Spatial positioning and arrangements of smaller scale elements within materials

bottom-left figure), and whether there are local variations that break the symmetry (the bottom-right figure). The analysis of such arrangements could yield insights concerning functionalities of materials.

The major steps for this analysis is to first identify all white spots in the image and then analyze the centroid locations of the white spots for identifying possible symmetries or symmetry violations. The first step is well known as the spot detection problem in image processing (Hughes et al. 2010). The spatial symmetry in the second step is mathematically represented as a lattice in geometry, the symmetry group of discrete translational symmetric in $\mathbb{R}^d$, i.e., the d-dimensional space. A point on a lattice in $\mathbb{R}^d$ is represented by a weighted sum of d linear independent vectors in $\mathbb{R}^d$ with integer weights. The basis vectors need to be estimated using the information associated with the detected spots. Not every white spot conforms to the lattice, causing local non-conformity to the global lattice. While many earlier approaches attempt to solve the first and second steps sequentially, more recent attentions are given to integrated approaches, in which spot detection and spatial symmetry analysis are solved together for more robust solutions.

1.1.3 Example 3: Temporal Evolution

While the first two examples use still (snapshot) nano images, the next two examples get into the territory of dynamic imaging. With the time-resolving capability of in situ electron microscopes, the temporal evolution of material structures under an applied processing condition can be imaged in the form of a sequence of images or a video. This capability provides the potential for the structural changes to be correlated with the applied processing condition over time. Depending on whether individual structural changes or the ensemble changes for a group of nanostructures are the topic of interest, different types of analysis can be performed. Our third example, illustrated via Fig. 1.6, is intended to show the need to track the changes for a group of nanostructures in the form of a probability distribution of the structural characteristics.

In the example of Fig. 1.6, a silver nanoparticle growth process is studied. The growth of silver nanoparticles is initiated by an electron beam shot on the growth solution. The growth process is video-imaged by an in situ scanning transmission electron microscope. The top row in Fig. 1.6 presents a few image frames taken

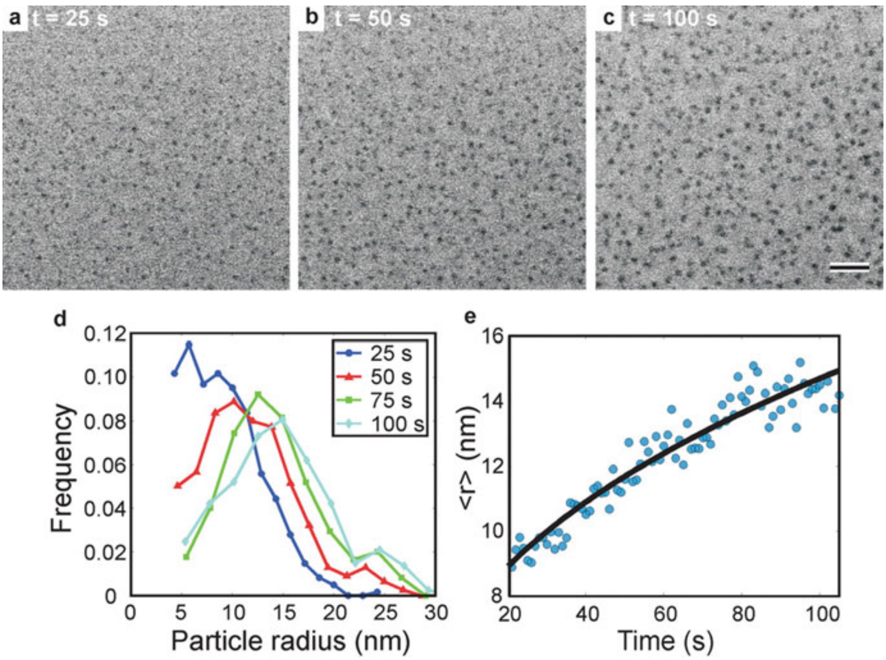

Fig. 1.6 Temporal evolution in particle size distribution. Panels (**a**)–(**c**) show three image snapshots taken from the video of a silver nanoparticle growth process. Panel (**d**) shows the particle size distributions at different times. Part (**e**) shows the trend of the average particle sizes over time. (Reprinted with permission from Woehl et al. 2013)

from the video. The goal of the nano image analysis here is to track the probability distribution of particle sizes over time and understand the particle growth behavior such as the mean growth rate. Figure 1.6d shows the particle size distributions at a few selected time points, whereas Fig. 1.6e shows a polynomial curve fitted to the mean particle size versus time. The exponent of the polynomial quantifies the nanoparticle growth rate. Apparently, a prerequisite for this dynamic image analysis is the morphology analysis of material objects in each image frame.

1.1.4 Example 4: Motions and Interactions

In addition to distribution changes, which are associated with a group of nano objects or structures, individual motions and interactions are also of great interest in materials research. Figure 1.7 shows our fourth example, in which the topic of interest is to understand the orientations of two nanoparticles when they aggregate or merge. Such insights are important to studying how larger nanoparticles are formed and how nanoparticles grow. The in situ scanning transmission electron microscopy, essential to the study in Sect. 1.1.3, is also used for this study. A video sequence of particle aggregation events is captured and then analyzed.

A number of nanoparticle aggregation events are identified and extracted from the video by multi-object tracking algorithms. Figure 1.7a shows two such examples, both of which are aggregations involving two elongated nanoparticles. Figure 1.7b illustrates how the aggregation events are modeled. The orientations of the two elongated objects in the merge are defined as the center coordinates of the merge zone, which is $c_{X,Y}$ in the figure, with respect to the respective coordinate frames of the two objects. The orientations of the two objects are denoted by θ_X and θ_Y, respectively. Figure 1.7c presents the probability distributions describing the orientations. The top-left panel of Fig. 1.7c is the joint distribution of θ_X and θ_Y, fit to the observed data of θ_X and θ_Y, The bottom-left and top-right panels of Fig. 1.7c present the marginal distributions of θ_X and θ_Y, respectively. The bottom-right panel of Fig. 1.7c depicts the mean orientation of θ_X and θ_Y in the fitted distributions. In this example, the bivariate von Mises distribution is used to model the joint distribution. Once the joint distribution is fit, many statistical analyses can be performed; for example, a hypothesis test can be conducted to check whether the two samples of nanoparticle aggregations exhibit the same mean orientations or not.

1.2 How This Book Is Organized

We organize this book into two large parts of nano image analysis: the pre-processing and the main analysis (Fig. 1.8).

The pre-processing analysis is a set of image processing methods that can be performed on any images before the main analysis. This includes image

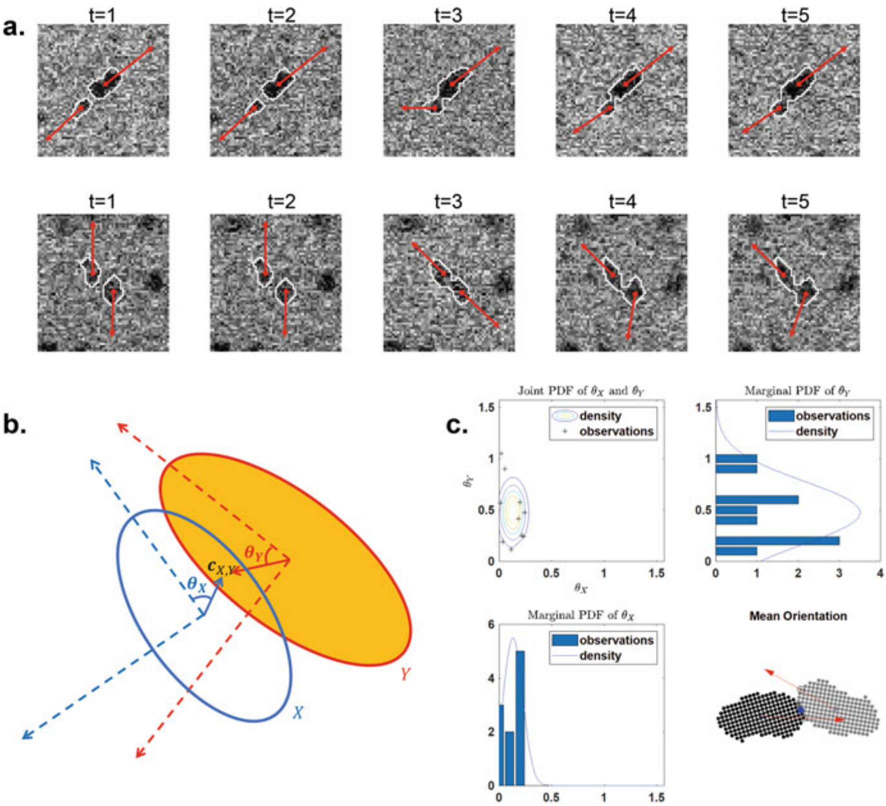

Fig. 1.7 Analysis of particle motions and inter-particle interactions. Panel (**a**) shows two sequences of smaller nanoparticles merged into a larger one. Panel (**b**) presents the diagram illustrating how the orientations of the primary particles involved in the merge are mathematically modeled. Panel (**c**) shows the empirical distributions of the orientations of the two particles observed in multiple merging events, as well as the bivariate angular distribution fitted to the image data

representation, image super-resolution enhancement, and image segmentation. As shown in the top panel of Fig. 1.8, the topics of pre-processing are covered in the following three chapters:

- **Chapter 2** introduces various kinds of nanoscale imaging instruments, i.e., different types of microscopes, and explains the mathematical representations of the resulting material images. Image representations lay the foundation for mathematical formulations and technical solutions in the subsequent nano image analyses.
- **Chapter 3** describes image segmentation, essential for separating the material objects and structures of interest from the noisy background. Image segmentation is the prerequisite to almost all subsequent higher level image analyses.

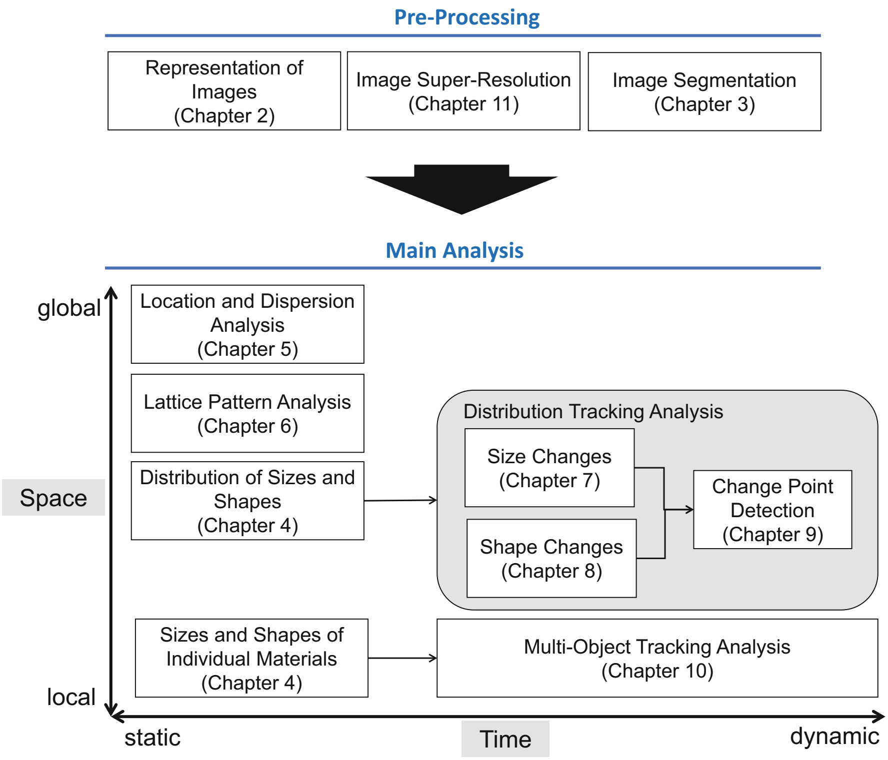

Fig. 1.8 The organization of the topics in the book

- **Chapter 11** discusses an image enhancement approach, known as super-resolution. Image super-resolution is to improve the resolution of raw microscope images by exploiting the relation between a pair of low- and high-resolution images. It is an optional step before image segmentation and the main analysis. But if multi-resolution image data are indeed available, super-resolution actions can definitely help deliver better quality in later analyses.

The organization of main analysis is based on the structural and temporal aspects of nano image analysis, as shown in the bottom panel of Fig. 1.8.

When the topics of static analysis are covered in Chaps. 4 through 6, the focus of the analysis is the static structural features of nanomaterials at the last stage of processing or a certain chosen time point. The structural features of interest include morphology of the exterior outlines of nanomaterials (in **Chap. 4**), spatial arrangements of nanoscale objects in a bulk material (in **Chap. 5**), and spatial arrangements of atoms in the interior of a nanomaterial (in **Chap. 6**).

When the topics of dynamic analysis are covered in Chaps. 7 through 10, the temporal changes of nanostructures are the focus. The modeling and analysis concerning temporal changes are colloquially referred to as *tracking*—individual nanostructures can be tracked or the ensemble changes for a group of nanomaterials can be tracked. When the ensemble analysis is concerned, the state of a group of material objects is described by a probability distribution of the structural features of interest, and the temporal change in the resulting probability distribution is tracked. The ensemble analysis is also known as the distribution tracking, which is covered in **Chaps. 7**, **8**, and **9**. When the individual analysis is concerned, a multi-object tracking approach is employed (in **Chap. 10**). The motion, morphology changes, and interactions of individual nano objects in an image sequence are tracked. Statistical modeling and analysis are performed, using the individual tracking records to reveal profound insights into the nanoscale world.

1.3 Who Should Read This Book

This book is designed for data scientists who are interested in the application of data science to materials research. It covers the topics of statistics, optimization, image processing, and machine learning applied to image and video data captured in materials research. The contents of the book provide data scientists a comprehensive overview in terms of the challenges and state-of-the-art approaches in image data analysis for nanomaterials research. We expect the readers to have at least basic knowledge in statistics, linear algebra, and optimization. The readers do not need any background in materials research, because the book is largely written from the perspectives of a data scientist.

Materials scientists and practitioners who are interested in artificial intelligence and machine learning (AI/ML) could also benefit from this book. In the past few years, we witness a fast increasing presence of AI/ML sessions and workshops in major materials research conferences, such as the Materials Research Society (MRS)'s Spring and Fall Meetings and the Microscopy & Microanalysis (M&M) conference. We also noticed that many students from the materials science departments are taking, or have taken, machine learning courses. With the basic knowledge of statistics and machine learning, materials scientists can branch out to more advanced topics by using this book, which builds upon examples and case studies taken from materials research. Practitioners can run the algorithms and methods through the code and examples accompanying the book and apply the solutions for their own benefit.

1.4 Online Book Materials

Either one of the following book websites has the example codes and datasets used in the book:

https://www.chiwoopark.net/book/dsnia.
https://aml.engr.tamu.edu/book-dsnia.

The contents on both websites are the same. Readers please feel free to access either of them.

References

Hughes J, Fricks J, Hancock W (2010) Likelihood inference for particle location in fluorescence microscopy. The Annals of Applied Statistics 4(2):830–848

Reid R (2018) Inorganic Chemistry. ED Tech - Press, Waltham Abbey Essex, UK

Woehl TJ, Park C, Evans JE, Arslan I, Ristenpart WD, Browning ND (2013) Direct observation of aggregative nanoparticle growth: Kinetic modeling of the size distribution and growth rate. Nano Letters 14(1):373–378

Chapter 2
Image Representation

2.1 Types of Material Images

2D Material Images Most material imaging techniques produce two-dimensional material images over a lateral plane of a material specimen as shown in Fig. 2.1. These material imaging techniques are largely grouped into optics-based microscopes and probe-based microscopes. The optics-based microscopes emit a beam of lights (photons) or charged particles (electrons and ions) to a material specimen of interest and exploit the interaction of the light or charged particle beam with the specimen to produce images. Depending on the beam sources, we have various types of microscopes, such as the optical microscopes (visible light beam), X-ray microscopes (X-ray beams), electron microscopes (electron beams), and ion-beam microscopes (ion beams). The wavelengths of the beam sources influence the resolutions of these microscopes. In the optics-based microscopes, various modes of beam-material interactions such as diffraction, reflection and refraction have been exploited for imaging. For example, in a transmission electron microscope (TEM), an electron beam of uniform current density irradiates one side of a thin material specimen, and the electron current density on the other side of the specimen is imaged through multiple stages of magnification lens (Reimer 2013). A scanning electron microscope (SEM) has a focused beam of electrons probe a raster over a material specimen, and the secondary electrons emitted for each raster spot are then detected (Reimer 2013). The second electron yield is related to the topography of the specimen. The electron probe size is 0.5–2 nm diameter if a field emission gun is used, which determines the spatial resolution of imaging. A scanning transmission electron microscope (STEM) uses a focused electron beam like SEM to probe over a raster of a specimen but measures the transmitted beam intensity like TEM (Pennycook and Nellist 2011).

C. Park, Y. Ding, *Data Science for Nano Image Analysis*, International Series in Operations Research & Management Science 308,
https://doi.org/10.1007/978-3-030-72822-9_2

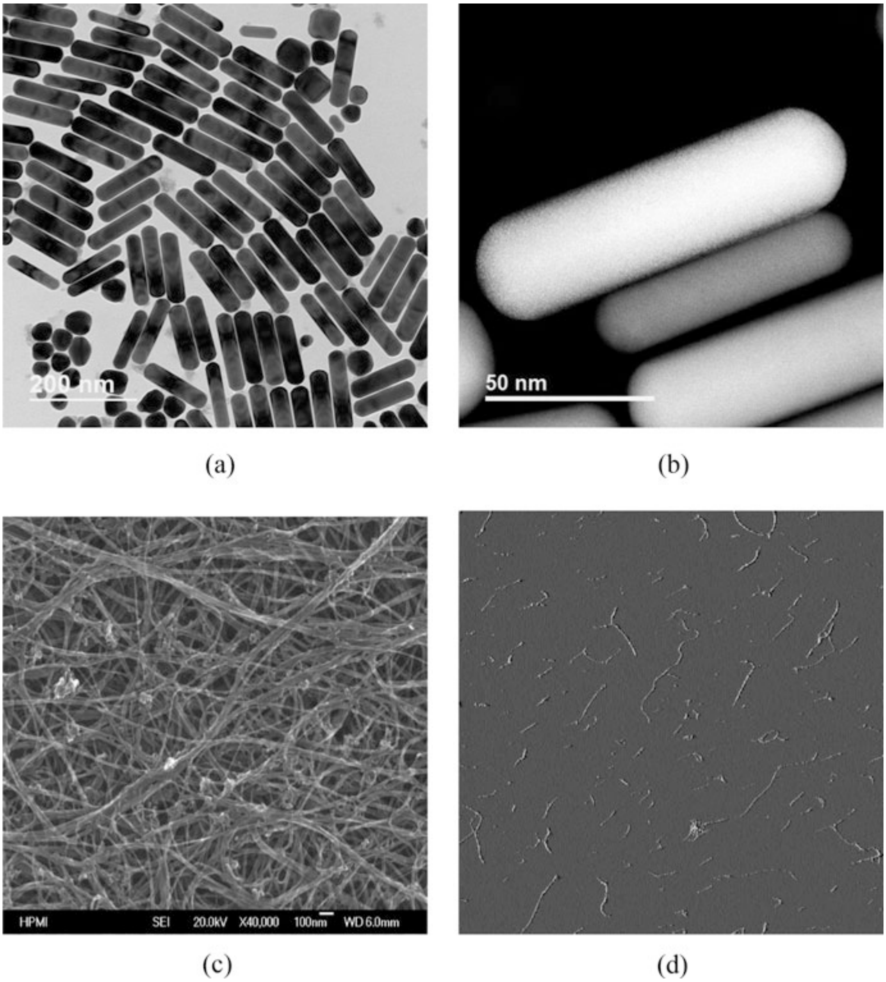

Fig. 2.1 Examples of 2D material images ((**d**) reprinted with permission from Kvam 2008). (**a**) TEM image (Au nanorods). (**b**) STEM image (Au nanorods). (**c**) SEM image (carbon nanotubes). (**d**) AFM image (carbon nanotubes)

The probe-based microscopes use a physical stylus to scan the surface of a material specimen and use the signals generated during the interaction between the stylus and the specimen to produce an image (Meyer et al. 2013). Depending on the types of the interactions exploited, the probe-based microscopes can be subcategorized. In a scanning tunneling microscope (STM), a conductive tip is brought very close to the surface of a specimen, and the tip position is moved over the surface. The tunneling current in the tip and the specimen at each probing location is measured and used to produce images. The tunneling current is a function of

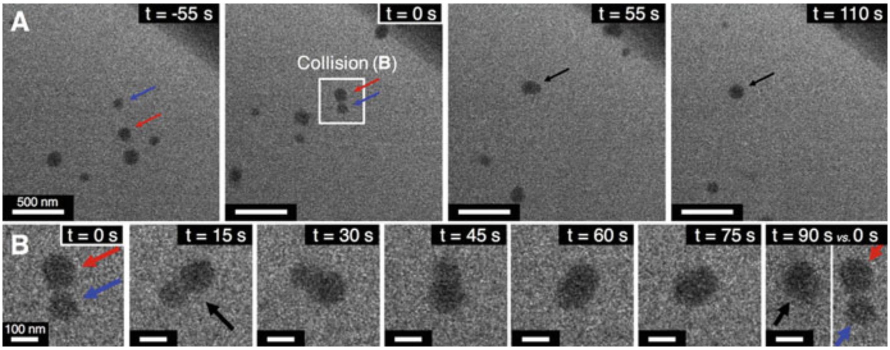

Fig. 2.2 Examples of 2D material video: Micelle fusion process (reprinted with permission from Parent et al. 2017)

the surface height and the local material density of the material specimen (Binnig and Rohrer 2000). Atomic force microscope (AFM, Fig. 2.1d) can work in three modes: contact mode, non-contact mode, and tapping mode (Binnig et al. 1986). In the contact mode, a stylus tip scans the specimen in close contact with the surface, and the deflection of the tip is measured. The tapping mode is similar to the contact mode except that the tip moves up and down, like taping on the surface, to prevent the tip from being stuck to the surface. In the non-contact mode, the tip scans the specimen surface with a constant distance above the surface, and the Van der Waals force between the tip and the surface is detected to construct the topographic image of the surface.

2D Material Videos In situ microscopy (Ross 2007) is the imaging technique to capture the change of a material specimen in response to chemical or physical stimuli in real time. The imaging technique yields a sequence of 2D images recorded over the period of measurements as illustrated in Fig. 2.2. Most of the conventional electron microscopes and some scanning probe microscopes had not been equipped with this real-time imaging capability, mainly because of the nature of their sample preparation process. In order for the conventional electron microscopes to image properly, the sample specimen should be dried first and then placed in a high-vacuum environment. But once dried, the material samples can no longer change or evolve, as most of materials can only change in liquid or gas phases. In situ microscopy uses a special sample holder attached onto the conventional microscopes, which allows the natural phases of a sample to be placed for imaging. For examples, liquid-cell transmission electron microscopes use a liquid cell holding a very thin layer of liquid samples that are sandwiched by two silicon or graphene windows (Ross 2016). Such in situ capabilities were first applied to TEMs. The reason is that SEM rasters over a specimen and acquires its image pixel by pixel. In the earlier days, the raster speed is not fast enough to make the in situ capability meaningful. Unlike the scanning electron microscopes, TEM acquires all its image pixel simultaneously from a single electron beam shot and its imaging process is thus much faster. The scanning speed has since improved greatly, which makes the in situ capability also feasible for the scanning electron microscopes; see the example of in situ STEM (Woehl et al. 2013).

3D Material Images Material samples are inherently 3D. The 2D material images produced by the conventional microscopes are the results of projecting the 3D structures onto a 2D plane perpendicular to the beam irradiation direction or the results of simply scanning the surface of the 3D structures. While these 2D images provide important structural information, they have very little ability in revealing the 3D structures of material specimens. Three-dimensional imaging capabilities have been developed with three major tomography techniques. The first one is the electron tomography, in which a sample is rotated around a single axis, and a series of 2D TEM images are sequentially taken at different rotation angles. The series are analytically combined into a 3D image by a tomography reconstruction method (Mu and Park 2020). The second technique is the X-ray tomography, similar to the electron tomography, except for its use of X-ray photon as the beam source (Salvo et al. 2003). The third technique is the focused ion beam (FIB) tomography, which exploits SEM and a focused ion-beam microscope for achieving the 3D imaging capability. The tomography technique uses an ion beam to peel off the surface of a material specimen little by little via a sputtering process, while the SEM uses an electron beam to scan the new surface once the old surface is peeled off (Holzer and Cantoni 2012). This process creates a series of 2D surface images at different layers of the material specimen, and the 3D structure of the material can be reconstructed by stacking these image series.

4D Material Images Electron tomography produces a static 3D image of material structures: either the state of a solidified material or a snapshot of an evolving material. Four-dimensional electron tomography refers to the imaging technique that creates a series of 3D tomograms representing the dynamic changes of the 3D material structures (Kwon and Zewail 2010). 4D electron tomography first obtains a series of 2D images at a number of projection angles and times using ultrafast electron pulses and then reconstructs the series into 4D tomograms. Another 4D material image data comes from 4D STEM. Unlike the conventional STEM that yields one scattering intensity at each probe location over the raster lines, a 4D STEM yields 2D scattering patterns at each probe location; see an example in Fig. 2.3. The 2D scattering pattern is known as the convergent beam electron diffraction, which contains important local structure information.

The material images so far discussed are mostly available in the form of digital images in two to four dimensions. The digital image data can be modeled through various mathematical abstractions, including a function, tensor, graph and set. The choice of mathematical abstraction of the image data impacts considerably the subsequent image analysis. In the subsequent sections, we describe a few commonly used mathematical abstractions of digital images and the related analysis approaches.

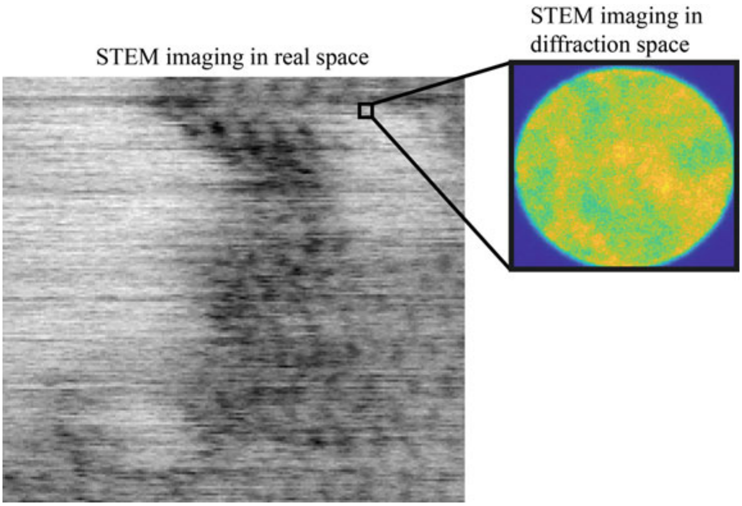

Fig. 2.3 Example 4D material images

2.2 Functional Representation

One can model an image as a function, $f : \mathcal{X} \mapsto \mathcal{Y}$, where $\boldsymbol{x} \in \mathcal{X}$ represents a location of an imaging space $\mathcal{X}$, and $f(\boldsymbol{x}) \in \mathcal{Y}$ represents the image intensity at the location (Sonka et al. 2014). The domain $\mathcal{X}$ represents the imaging space of two to four dimensions, typically rectangular and modeled as the product of respective real intervals, i.e., $\prod_{j=1}^{d}[0, W_j]$, where d is the dimension of the imaging space and $W_j > 0$ specifies the size of the domain for the j-th dimension. The co-domain $\mathcal{Y}$ is often a single-dimensional space of non-negative scalars for a grayscale image or a multi-dimensional space of vector quantities for color or multi-channel images.

For a digital image, both of $\boldsymbol{x}$ and $f(\boldsymbol{x})$ are digitized, implying that both the domain and co-domain are discretized to finite sets. The domain $\mathcal{X}$ is uniformly partitioned into grids of equal sizes, and the set of the grids defines the discrete domain for f. Each cubicle (in 3D) or quadrat (in 2D) of the smallest element in the X-domain partition is called a pixel. The physical dimension of a pixel is naturally known as the pixel size, e.g., a pixel of the size of one nm^2. The pixel size determines the spatial resolution of the image. Similarly, the value of the co-domain $\mathcal{Y}$ is also discretized into a discrete set. The carnality of the co-domain set is called the pixel depth, which determines the resolution of image intensities. For instance, a grayscale image with an 8-bit depth uses 2^8 integer values in the range of [0, 255] to represent the pixel intensity from the lightest to the darkest.

Using the functional representation, some basic image processing can be expressed mathematically as applying either a point-wise transformation or a convolution to f. The point-wise transformation is to apply a function g on the image output f point-wise, denoted as

$$g \circ f(\boldsymbol{x}).$$

For example, the transformation function g can be a thresholding function, i.e., $g(y) = \mathbb{1}_{y \geq \tau}$, where $\mathbb{1}$ is the indicator function and τ is the threshold for binarization. such point-wise transformation converts a grayscale image to a binary image.

A convolution to f is a linear combination of f weighted by a filter function g. Many popular image filtering operations are represented by the convolution of an image $f \colon \mathcal{X} \mapsto \mathcal{Y}$, which is defined as

$$g * f(\boldsymbol{x}) = \int f(\boldsymbol{x}')g(\boldsymbol{x}' - \boldsymbol{x})d\boldsymbol{x}'. \tag{2.1}$$

Naturally, the filter function g defines the weights used in the convolution of f. The support of the filter function is typically symmetric around a zero-origin, so the result of the convolution is often a linear combination of the intensities of the pixels neighboring the evaluation location $\boldsymbol{x}$. This is referred to as a spatial filter, since it is based on image pixels surrounding the evaluation location.

A commonly used spatial filter is the Gaussian filter, which takes the convolution of f with a Gaussian density function,

$$g(\boldsymbol{x}) = (\pi h^2)^{d/2} \exp\{-\boldsymbol{x}^T \boldsymbol{x}/h^2\},$$

where $h > 0$ is the standard deviation of the Gaussian density. With this filter, the value of $g(\boldsymbol{x}' - \boldsymbol{x})$ monotonically decreases as the Euclidean distance from $\boldsymbol{x}'$ to $\boldsymbol{x}$ increases. Therefore, the convolution (2.1) takes the weighted average of the intensities of the pixels neighboring $\boldsymbol{x}$, with more weights given to the pixels closer to the evaluation location $\boldsymbol{x}$, and produces a local estimate of $f(\boldsymbol{x})$. The Gaussian filter is popularly used for image denoising.

Another popular spatial filter is the mean filter

$$g(\boldsymbol{x}) = \frac{1}{\int_{\mathcal{X}} d\boldsymbol{x}}.$$

The filter function is a constant function, with which the convolution (2.1) is just the simple arithmetic average of the neighboring pixel intensities.

There are spatial filters that cannot be represented in the convolution form (2.1), like the rank statistics-based filters such as the median filter and the spatially weighted median filter (Mitra and Sicuranza 2000). These rank-based filters are more robust to outliers, working better on denoising Poisson noises.

There are non-spatial image filters such as the non-local mean filter (Buades et al. 2004), the bilateral filter (Elad 2002), and other variants (Lou et al. 2009), where the filter function does not only depend on the spatial distance $\boldsymbol{x}' - \boldsymbol{x}$ but also depends on $f(\boldsymbol{x}') - f(\boldsymbol{x})$. Therefore, a non-spatial filtering cannot be represented in the convolution form but can be represented in a more general form of

$$g \otimes f(\boldsymbol{x}) = \frac{1}{C_g(\boldsymbol{x})} \int f(\boldsymbol{x}')g(\boldsymbol{x}, \boldsymbol{x}')d\boldsymbol{x}', \tag{2.2}$$

where $g(\boldsymbol{x}, \boldsymbol{x}')$ is a non-spatial filter function, and $C_g(\boldsymbol{x}) = \int g(\boldsymbol{x}, \boldsymbol{x}')d\boldsymbol{x}'$. For example, the non-local mean filter uses the following filtering function,

$$g(\boldsymbol{x}, \boldsymbol{x}') = \exp\{-b(\mu_f(\boldsymbol{x}) - \mu_f(\boldsymbol{x}')^2)\},$$

where $b > 0$ is a filter parameter, and μ_f is the local mean of $f(\boldsymbol{x})$ that can be obtained by applying the mean filter on f. These non-spatial filters are more effective to preserve local intensity jumps than spatial filters that tend to blur out sharp features.

Image reconstruction and estimation are more advanced image processing than the basic operations like image binarization or filtering explained above. The problem of image estimation is to estimate or infer a pixel's intensity based on either noisy pixel measurements or pixel measurements elsewhere. Using the functional representation of image data, the problem can be formulated as a regression analysis. The underlying noise-free image f is linked with its noisy measurements at N grid locations, $\{(\boldsymbol{x}_i, y_i) : i = 1, \ldots, N\}$, through the following model,

$$y_i = f(\boldsymbol{x}_i) + \epsilon_i,$$

where ϵ_i is an independent random noise with zero mean and variance σ_ϵ^2. The image de-noising problem can be formulated as a regularized learning problem, which is to find f that minimizes the sum of a loss function L and a regularization term, i.e.,

$$\underset{f}{\text{minimize}} \sum_{i=1}^{N} L(f(\boldsymbol{x}_i), y_i) + \lambda R(f), \tag{2.3}$$

where $R(f)$ is the regularization term that penalizes the complexity of function f, and the regularization term is added for avoiding the overfitting of f. A popular choice for the loss function is the L_2 norm defined by the squared difference $||f(\boldsymbol{x}_i) - y_i||_2^2$, as used in a least squares regression, and another choice is the L_1 norm defined by the absolute difference $|y_i - f(\boldsymbol{x}_i)|$, as used in robust regression.

The function f is typically represented in either a semi-parametric or a non-parametric form to accommodate complex patterns of images. In the literature, f is commonly represented by a basis expansion with a pre-determined basis function set, i.e.,

$$f(\boldsymbol{x}) = \sum_{k=1}^{K} \beta_k \phi_k(\boldsymbol{x}),$$

where $\phi_k(\boldsymbol{x}) : \mathcal{X} \mapsto \mathcal{Y}$ is the kth basis function, β_k is the respective coefficient that are to be estimated, and K is the number of the basis functions. Popular basis sets for a two-dimensional domain are the Fourier basis or wavelet basis in signal processing (Figueiredo and Nowak 2003) and the spline basis in statistics (Unser 1999). For more than two dimensions, a product basis is popularly used (Verhaeghe et al. 2007). The product basis implies that the basis function for a higher dimensional domain is constructed by taking the product of basis functions defined on low dimensional subspaces. For example, when $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2$ and $\phi_k^{(l)} : \mathcal{X}_l \mapsto \mathcal{Y}$ is the basis function on $\mathcal{X}_l$, the basis function on $\mathcal{X}$ is defined by the product form,

$$\phi_{k,k'}(\boldsymbol{x}) = \phi_k^{(1)}(\boldsymbol{x}_1)\phi_{k'}^{(2)}(\boldsymbol{x}_2)$$

for $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2)$.

With the basis functions defined and chosen, the space of f is now spanned by the basis functions. As such, the function is uniquely decided once the coefficients of $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)^T$ are estimated. The complexity function, $R(f)$, is defined as a function of $\boldsymbol{\beta}$. The popular choices are the L_p norms of $\boldsymbol{\beta}$ for using $p = 1$ or $p = 0$ (although $p = 0$ induces only a pseudo norm). With the orthogonal basis $\{\phi_k\}$ and the L_1 norm as the complexity function, solving the problem (2.3) produces soft-thresholding on an image, while the L_0 norm as the complexity function, the solution produces hard-thresholding with the threshold of λ.

A popular non-parametric representation of f is based on a local linear kernel smoother (Takeda et al. 2007), where the value of f is locally approximated around a neighborhood of $\boldsymbol{x}$ by a simple polynomial function,

$$f(\boldsymbol{x}) \approx \alpha_{\boldsymbol{x}} + \boldsymbol{\gamma}_{\boldsymbol{x}}^T \boldsymbol{x},$$

and with its coefficients $\alpha_{\boldsymbol{x}}$ and $\boldsymbol{\gamma}_{\boldsymbol{x}}$ estimated for each $\boldsymbol{x}$ by

$$(\hat{\alpha}_{\boldsymbol{x}}, \hat{\boldsymbol{\gamma}}_{\boldsymbol{x}}) = \arg\min \sum_{i=1}^{n} K_h(\boldsymbol{x} - \boldsymbol{x}_i)||y_i - \alpha_{\boldsymbol{x}} - \boldsymbol{\gamma}_{\boldsymbol{x}}^T \boldsymbol{x}_i||_2^2,$$

where $K_h(\boldsymbol{x} - \boldsymbol{x}_i)$ is a kernel function with the bandwidth h, and n is the number of pixels in the prescribed neighborhood of $\boldsymbol{x}$. The estimated $\hat{\alpha}_{\boldsymbol{x}}$ becomes the estimate of f at $\boldsymbol{x}$. The kernel function K_h determines the weighting function of the ith data location $\boldsymbol{x}_i$ neighboring $\boldsymbol{x}$, which is typically defined as an isotropic function depending on the distance of $||\boldsymbol{x} - \boldsymbol{x}_i||_2$. The isotropic smoothing gives a consistent estimator of f when f is continuous and smooth around $\boldsymbol{x}$. When f is discontinuous or has directional changes around $\boldsymbol{x}$, locally adaptive kernel or one-sided kernel functions are used to accommodate the anisotropy (Takeda et al. 2008) or the discontinuity (Qiu 2009).

2.3 Matrix Representation

The observations of a function f at grid locations can be represented as a vector, a matrix, or more generally, a tensor. Such representation can be viewed as a discrete version of the functional representation. For example, a 2D grayscale image observed at $m \times n$ grid locations can be represented as an $m \times n$ matrix,

$$\boldsymbol{F} = \begin{bmatrix} f(1,1) & f(1,2) & \cdots & f(1,n) \\ f(2,1) & f(2,2) & \cdots & f(2,n) \\ \vdots & \vdots & \ddots & \vdots \\ f(m,1) & f(m,2) & \cdots & f(m,n) \end{bmatrix},$$

where $f(i, j)$ is the image intensity at the (i, j)th grid location. A c-channel 2D image is then represented by an $m \times n \times c$ tensor, and more generally, a d-dimensional image by a d-dimensional tensor.

With the matrix representation, image processing can be expressed as matrix operations. Unsurprisingly many types of digital image processing are defined as a matrix convolution, which is a discrete version of the convolution of two functions f and w on a two-dimensional domain (x, y). Recall that the convolution of f and w are expressed as

$$w * f(x, y) = \int_{-a}^{a} \int_{-b}^{b} w(s,t) f(x-s, y-t) ds dt,$$

where $[-a, a] \times [-b, b]$ is the support of w. Its discrete version is then

$$f_w(i, j) := w * f(i, j) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(i-s, j-t),$$

where w is called a filter function of the convolution, and $f_w(i, j)$ is the image intensity of the convoluted image at the (i, j)th grid location. Using a matrix representation, the discrete convolution can be written as $\boldsymbol{W} * \boldsymbol{F}$, where $\boldsymbol{W}$ is the matrix representation of the filter function,

$$\boldsymbol{W} = \begin{bmatrix} w(-a,-b) & w(-a,-b+1) & \cdots & w(-a,b) \\ w(-a+1,-b) & w(-a+1,-b+1) & \cdots & w(-a+1,b) \\ \vdots & \vdots & \ddots & \vdots \\ w(a,-b) & w(a,-b+1) & \cdots & w(a,b) \end{bmatrix}.$$

Popular filter matrices and the result of their applications to a material image example are shown in Fig. 2.4.

Singular value decomposition (SVD) is another population matrix operation used in image analysis, which is to decompose an $m \times n$ matrix of rank p as the product of three matrices,

$$\boldsymbol{F} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T,$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are $m \times p$ and $n \times p$ unitary matrices respectively, and $\boldsymbol{D}$ is a $p \times p$ diagonal matrix with non-negative real numbers on the diagonal. When SVD is applied to a matrix $\boldsymbol{F}$ representing an image, many diagonal elements of $\boldsymbol{D}$ are close to zero, for which one can approximate the matrix by its low-rank version,

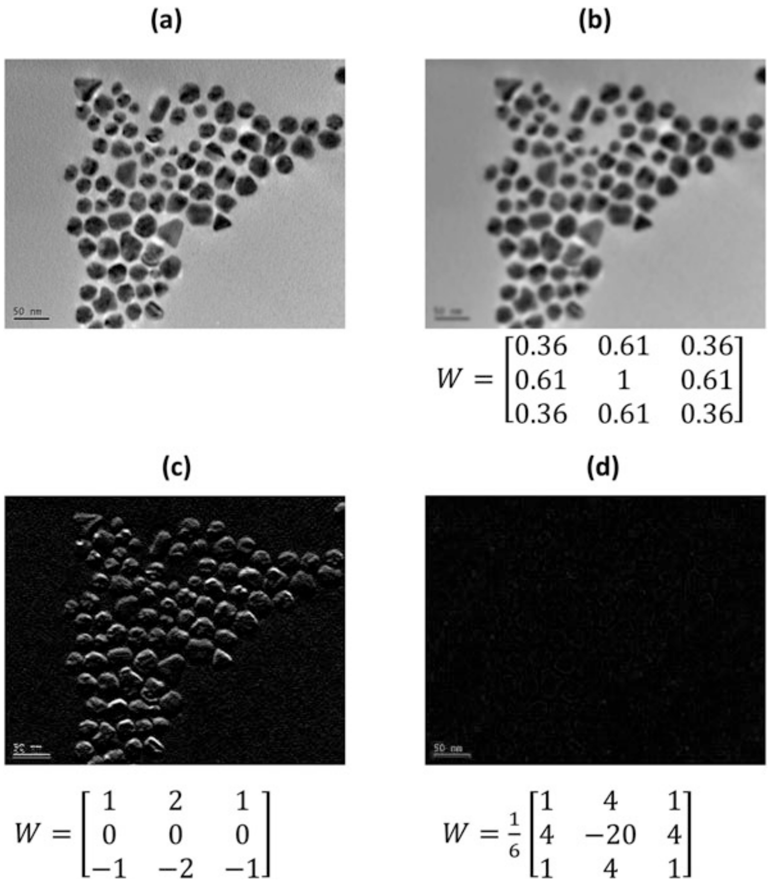

Fig. 2.4 Popular image filters and their exemplary applications. (**a**) Original image. (**b**) Gaussian filter. (**c**) Sobel filter. (**d**) Laplacian filter