Intelligent Systems Reference Library 135

Margarita N. Favorskaya
Lakhmi C. Jain *Editors*

Computer Vision in Control Systems-3

Aerial and Satellite Image Processing

Springer

Intelligent Systems Reference Library

Volume 135

The aim of this series is to publish a Reference Library, including novel advances and developments in all aspects of Intelligent Systems in an easily accessible and well structured form. The series includes reference works, handbooks, compendia, textbooks, well-structured monographs, dictionaries, and encyclopedias. It contains well integrated knowledge and current information in the field of Intelligent Systems. The series covers the theory, applications, and design methods of Intelligent Systems. Virtually all disciplines such as engineering, computer science, avionics, business, e-commerce, environment, healthcare, physics and life science are included. The list of topics spans all the areas of modern intelligent systems such as: Ambient intelligence, Computational intelligence, Social intelligence, Computational neuroscience, Artificial life, Virtual society, Cognitive systems, DNA and immunity-based systems, e-Learning and teaching,Human-centred computing and Machine ethics, Intelligent control, Intelligent data analysis, Knowledge-based paradigms, Knowledge management, Intelligent agents, Intelligent decision making, Intelligent network security, Interactive entertainment, Learning paradigms, Recommender systems, Robotics and Mechatronics including human-machine teaming, Self-organizing and adaptive systems, Soft computing including Neural systems, Fuzzy systems, Evolutionary computing and the Fusion of these paradigms, Perception and Vision, Web intelligence and Multimedia.

More information about this series at http://www.springer.com/series/8578

Margarita N. Favorskaya
Lakhmi C. Jain
Editors

Computer Vision in Control Systems-3

Aerial and Satellite Image Processing

Springer

Editors
Margarita N. Favorskaya
Institute of Informatics and Telecommunications
Reshetnev Siberian State University of Science and Technology
Krasnoyarsk
Russian Federation

Lakhmi C. Jain
Faculty of Education, Science, Technology and Mathematics
University of Canberra
Canberra
Australia

and

Bournemouth University
UK

and

KES International
UK

ISSN 1868-4394 ISSN 1868-4408 (electronic)
Intelligent Systems Reference Library
ISBN 978-3-319-67515-2 ISBN 978-3-319-67516-9 (eBook)
https://doi.org/10.1007/978-3-319-67516-9

Library of Congress Control Number: 2017952191

Printed on acid-free paper

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

Preface

The research book is a continuation of our previous books which are focused on the recent advances in computer vision methodologies and technical solutions using conventional and intelligent paradigms.

- Computer Vision in Control Systems-1, Mathematical Theory, ISRL Series, Volume 73, Springer-Verlag, 2015
- Computer Vision in Control Systems-2, Innovations in Practice, ISRL Series, Volume 75, Springer-Verlag, 2015

The research work presented in the book includes multidimensional image models and processing, vision-based change detection, filtering and texture segmentation, extraction of objects in digital images, road traffic monitoring by UAV, video stabilization, image deblurring, structural verification, matrix transformation and numerical system for computer vision.

The book is directed to the Ph.D. students, professors, researchers and software developers working in the areas of digital video processing and computer vision technologies.

We wish to express our gratitude to the authors and reviewers for their contribution. The assistance provided by Springer-Verlag is acknowledged.

Krasnoyarsk, Russian Federation — Margarita N. Favorskaya
Canberra, Australia — Lakhmi C. Jain

Contents

About the Editors

Dr. Margarita N. Favorskaya is a Professor and Head of Department of Informatics and Computer Techniques at Reshetnev Siberian State University of Science and Technology, Russian Federation.

Professor Favorskaya is a member of KES organization since 2010, the IPC member and the Chair of invited sessions of international conferences. She serves as a reviewer in international journals (Neurocomputing, Knowledge Engineering and Soft Data Paradigms, Pattern Recognition Letters, Engineering Applications of Artificial Intelligence), an associate editor of Intelligent Decision Technologies Journal and Computer and Information Science Journal. She is the author or the co-author of 160 publications and 20 educational manuals in computer science. She co-edited three books for Springer recently. She supervised eight Ph.D. candidates and presently supervising five Ph.D. students.

Her main research interests are digital image and videos processing, remote sensing, pattern recognition, fractal image processing, artificial intelligence, information technologies.

Dr. Lakhmi C. Jain is with the Faculty of Education, Science, Technology and Mathematics at the University of Canberra, Australia, and Bournemouth University, UK. He is a Fellow of the Institution of Engineers Australia.

Professor Jain founded the KES International for providing a professional community the opportunities for publications, knowledge exchange, cooperation and teaming. Involving around 5000 researchers drawn from universities and companies worldwide, KES facilitates international cooperation and generates synergy in teaching and research. KES regularly provides networking opportunities for professional community through one of the largest conferences of its kind in the area of KES www.kesinternational.org.

His interests focus on the artificial intelligence paradigms and their applications in complex systems, security, e-education, e-healthcare, unmanned air vehicles and intelligent agents.

Chapter 1
Theoretical and Practical Solutions in Remote Sensing

Margarita N. Favorskaya and Lakhmi C. Jain

Abstract The chapter presents a brief description of chapters that contribute to theoretical and practical solutions for aerial and satellite images processing and the fields close to this scope. One can find the original investigations in the novel tensor and wave models, new scheme of comparative morphology, warping compensation in video stabilization task, image deblurring based on physical processes of blur impacts, fast and robust core structural verification algorithm for feature extraction in images and videos, among others. Each chapter involves practical implementations and explanations.

Keywords Remote sensing · Multidimensional image processing
Comparative morphology · Digital halftone images · Object extraction
Traffic monitoring · Warping technique · Image deblurring · Feature extraction
Double-sided matrix transformation · Counter noise immunity

1.1 Introduction

Nowadays, computer vision techniques play a significant role in many applications. In spite of pioneer investigations in remote sensing since 1970s, the development of theories and algorithms for this purpose remains the crucial issue yet. The recent

M.N. Favorskaya (✉)
Institute of Informatics and Telecommunications, Reshetnev Siberian State University of Science and Technology, 31, Krasnoyarsky Rabochy Avenue, Krasnoyarsk 660037, Russian Federation
e-mail: favorskaya@sibsau.ru

L.C. Jain
Faculty of Education, Science, Technology and Mathematics, University of Canberra, Canberra, ACT 2601, Australia
e-mail: jainlakhmi@gmail.com

L.C. Jain
Bournemouth University, Poole, UK

M.N. Favorskaya and L.C. Jain (eds.), *Computer Vision in Control Systems-3*,
Intelligent Systems Reference Library 135, https://doi.org/10.1007/978-3-319-67516-9_1

achievements in computer vision permit to have another look at the well-known problems in remote sensing, as well as to increase the speed of computations. This field of investigations is very wide and will require great efforts of the researchers in further development and improvement of classical methods for aerial and satellite image processing.

1.2 Chapters Included in the Book

Chapter 2 contains a detailed description of novel tensor and wave models, as well as the interesting modifications of autoregressive models with multiple roots, in order to synthesize the algorithms of multidimensional image processing and their sequences. The consuming time algorithms on the mathematical level were developed for the applications, which process the images with random nature including TV, multispectral, and radar images and imageries. These mathematical models are based on the random field representation and simulation using the Gibbs, autoregressive, tensor and wave models [1–3]. The random fields are considered on rectangular grids of some dimension and also on the surfaces, such as the cylinder, sphere, and paraboloid that extends the application of the proposed models and algorithms to simulate the planets' surfaces. Another question deals with the anomalies' detection in the images. Four equivalent forms of the optimal decision rules are derived to detect the extensive anomalies. Also the unknown parameters of the affine and curvilinear deformations are determined according to coordinates of the obtained fixed point [4].

Chapter 3 presents a new scheme of Comparative Morphology (CM) as a generalization of the Morphological Image Analysis (MIA) scheme originally proposed by Pyt'ev [5] and further developed in [6, 7]. Such comparative morphologies are based on the mathematical shape theories, which solve the tasks of the image similarity estimation, image matching, and change detection by means of some special morphological models and tools. The proposed CM scheme excludes the MIA ideas of shape and projection and, at the same time, preserves the idea of asymmetrical comparative filtering for the robust similarity estimation and change detection. Comparative filter maps two images, reference and test, as input data and forms the filtered version of test image depending on its relation with reference image [8]. In the suggested terms, a comparative filter is called the morphological filter if it implements a smoothing mapping that preserves any constant (flat) image and preserves the test image equal to the reference image. The experiments estimate the qualitative and quantitative change detection in the real images for different scene types and simulated aerial images from the public benchmark. The obtained results of precision and recall demonstrate that the proposed pipeline is useful for change detection in long-range remote sensing data.

Chapter 4 comprises a development of the combined algorithm for detecting texture regions in noisy digital images that includes two main stages. First, the algorithm recovers the digital halftone (color) images with low signal/noise ratios.

Second, it detects the extended regions with the homogeneous statistical characteristics. The causal multidimensional multi-valued random Markov processes (i.e. the multi-dimensional Markov chain with several states) was used as an approximation of multispectral images [9]. The remote sensing systems can generate the digital images in different spectral bands containing from three (like RGB color images) to hundreds of components, each of which can be considered as the monochrome (grey-scale) image. The color RGB images can be represented as a special case of multispectral images. Big statistical dependence between image elements may exist in the Digital HalfTone Images (DHTI) belonging to different spectral components [10]. Such statistical relationship between the elements within and between elements of the DHTI color components (RG, GB, BR) can be approximated by 3D multi-valued Markov chain. During processing the DHTI with 2^g brightness levels, the problem of memory storage and working with the transition probability matrices $2^g \times 2^g$ in size is appeared. It was suggest to divide the DHTI represented by g bits binary numbers into g bit binary image or bit planes [11]. This allows to reduce the computational resources owing to working with 2×2 transition probability matrices. The quality of segmentation was improved using the evaluation of the statistical characteristics of each local area within and between adjacent bit binary images for different color component [12]. The proposed method allows to segment a noisy image (with signal-to-noise ratio –6 dB) efficiently if the transition probability between elements in the areas does not exceed 0.15. In this case, a segmentation error is less than 8%.

Chapter 5 conducts the investigations for extraction of objects in different types of images like aerial, satellite, and synthetic aperture radar images using the straight edges segments, which are in following grouped in the simple or complex shapes. The problem of feature construction for object description and extraction includes the feature detection and feature description tasks [13]. An advanced edge-based method for automatic detection and localization of straight edge segments was developed that provides an ordered set of straight line segments with their orientations and magnitudes. The algorithm automatically adjusts an appropriate mask size for the slope line filter to give the maximum normalized filter output. It produces a set of segments as affine invariant and scale invariant features for object description. The advanced algorithm performs a reliable detection of local edges and accurately obtains the edge intersection points. The new elements are the global seeking for the most valuable edge direction and automatic adjusting of the mask size for the slope line filter to match with the edge length. The comparative analysis of the noisy image shows that the advanced algorithm is inferior to others in feature detection performance. A hierarchical set of features is developed for object description subject to the proposed feature detector. This set contains four levels of line combination, each level relates to the number of combined line segments. The problem of object selection using the straight line segments extraction and grouping includes several steps, when the following elements are determined sequentially: the straight line segments, crossing points of lines, and closed complex structures [14]. The algorithm can extract not only rectangular or triangular objects or parts of them but other types of objects (roads, polygonal structures). The extended experiments with the noisy images were conducted [15]. Applications to the aerial, satellite, and radar images show a good ability to separate and extract rectangular objects like buildings and other line-segment-rich structures. Most of objects are selected somehow or other and the following problem is how to improve grouping process.

Chapter 6 includes the study of the ground traffic monitoring aided by the Unmanned Aerial Vehicles (UAV) based on applying the on-board computer vision systems. At present, the UAV cannot regard as a direct testifier of a specific road traffic situation, for example, collision of cars [16, 17]. The authors consider a possibility of situation classification based on the observed scene after the situation occurrence including the relative positions of cars, their damages, and position and behaviour of people. For the estimation and forecast of development of the present road situations, it is required to classify the traffic situations. For these purposes, the authors solved the following tasks: the selection of classification type and an alphabet forming of situation class, forming descriptions in the observed scene containing attributes of recognized situations (close to the feature dictionary), and choice of the decision-making algorithms that allows ranging the examined situation to corresponding class. Four classes called as the “Observed Scene”, “Objects of Traffic Accident (TA)”, “External Condition”, and “Additional Objects” were introduced. This description has a hierarchic structure and is divided into classes, subclasses, and divisions of various levels depending on a hierarchic level. In general, various types of descriptions, such as spatial, spatiotemporal, temporal, and causal, may be used [18]. The observed scene description is built using the proposed ontology structure of the TA. For decision making, the production model of knowledge representation and corresponding knowledge base were designed. The functional criteria of the losses, flight safety of the UAV, and reliability of class recognition allow to reconstruct the TA using video sequence obtained from the UAV.

Chapter 7 explores the warping techniques in video stabilization task [19]. All warping techniques are classified as the parametric and non-parametric methods. The parametric methods, such as translation, rotation, procrustes, affine, perspective, bilinear, and polynomial, provide the compact representation and fast computation of a warping but cannot perform well the local distortions. Opposite to them, the non-parametric methods like elastic deformations, thin-plate splines, and Bayesian approach demonstrate a heavy computational load and the presence of local optima. Not all of possible warping techniques are suitable for video stabilization task [20]. Both 2D and 3D stabilization algorithms involve the stabilization of video sequence and reconstruction of missing boundary regions. These stages are implemented differently for 2D and 3D stabilization but in each case the stabilization of video sequence requires the global frame warping, while the reconstruction of missing boundary regions in cropped after stabilization frames demands in the local frame warping. Both types of warping are based on the introduced Structure-From-Layered-Motion (SFLM) model based on the compactness theory of visual objects’ representation. The SFLMs are associated with all foreground moving objects, partly with background moving objects and render the individual motion types of moving objects [21]. The global warping in 2D stabilization is aimed on the compensation of each frame for removal of the unwanted motion, while 3D stabilization requires the generation of a desired 3D camera path and synthesis of the images along 3D path [22]. The reconstruction stages for 2D/3D stabilization are almost similar and based on the special inpainting methods, including the proposed pseudo-panoramic key frame and the multi-layered motion fields. The experiments confirm the quality of the stabilized test video sequences in terms of the content alignment and the frame size preserving.

Chapter 8 provides the investigate study of image deblurring based on physical processes of blur impacts. One can find the detailed overview of the non-blind deconvolution methods including inverse filtering, Wiener filtering, Kalman recursive filtering, minimum Mean Square Error (MSE) method, and various forced iterative methods, and blind deconvolution methods (various types of regularization methods like Tikhonov regularization algorithm, Lucy-Richardson algorithm, Shepp-Logan intuitive regularization method, and Arsenin method of local regularization) for image restoration. The authors proposed the original approach based on the Distortion Functions (DFs) for the small and large linear blur. The corresponding mathematical models as the systems of the linear algebraic equations are built with multiple illustrating examples. A comparison of the restoration results yielded by MatLAB software package using the known deblurring methods shows that application of the proposed model allows to restore the smaller distorted images with low computing complexity. Also the model of non-linear blur was constructed on the basis of the generalized description for processes of video signal distortion. A non-linear vibrational blur can occur, when the unstable position of the camera, the reciprocating motion of the object, or atmospheric turbulence take place [23]. For this case, five systems of linear equations were obtained. All models are based on the accurate background evaluation and perform a promising base for following investigations.

Chapter 9 involves a fast and robust Core Structural Verification Algorithm (CSVA) for feature extraction in images and videos. The proposed algorithm is based on many-to-one matches' exclusion, the improved Hough clustering of keypoint matches, and cluster verification procedure as the modified RANSAC. The feature detection, feature description, descriptors matching, and structural verification are the main stages of keypoint-based methods. These stages are sequentially discussed for many types of known descriptors with their comparative analysis. The input of the CSVA is a set of the nearest neighbour solution for keypoint descriptors (set M). The CSVA takes three steps, such as the primary outlier elimination (set M_{prime}), Hough clustering (set M_{cl}), and verification of the largest cluster(s) relying on mutual relation of the keypoints of different matches (set M_{final}). The ratio of correct/incorrect matches is a key factor for the successful application of geometrical constraints. Three methods, such as the nearest neighbour ratio test, cross-check, and exclusion of ambiguous n-to-1 matches, are considered in details. Verification stage of the CSVA is based on clustering of feature matches in similarity transform parameter space and consists of the Hough clustering and cluster verification [24]. The experiments were conducted using three datasets. Dataset db1 has 20 challenging image pairs of 3D scenes (mostly under strong viewpoint changes). The pairs were matched by different combinations of the SURF and the SIFT detectors and descriptors. Dataset db2 consists of 120 pairs of aerial and cosmic image pairs under the strong viewpoint, season, day-time, and man-made changes. Several manually matched points were used to calculate a ground truth homography H for every image pair. Dataset db3 has more than 300 pairs of images of static 3D scenes (mostly indoor). Few manually labelled corresponding points were used to calculate etalon fundamental matrices E by the least squares. The proposed algorithm uses some specific information (rigidity of objects in a scene), consume low volume memory and only 3 ms in average on a standard Intel i7 processor for verification of 1000 matches. The CSVA has been successfully applied to practical tasks with minor adaptation, such as the matching of 3D indoor scenes, retrieval of images of 3D scenes based on the concept of Bag of Words, and matching of aerial and cosmic photographs with strong appearance changes caused by season, day-time and viewpoint variation [25, 26].

Chapter 10 covers a double-sided matrix transformation that is efficiently used in telecommunication systems. The issues how to increase the interference immunity, transmission speed as well as to provide for the quality of a transmitted image over a telecommunication channel had been appeared since 1980s. Mironovsky and Slaev contributed significantly in this scope, for example [27–29]. The double-sided matrix transformation in respect of its matrices under some serious restrictions (orthogonality, symmetry) is the subject of interest in many researches. In this chapter, a more general case, for which criteria of invariant image existence are introduced, is analyzed. The scale, rotation, and negative (inverse) strip-invariants are investigated and the methods of finding them are proposed. Moreover, the task of arraying the double-sided transformation on the basis of a given set of invariant images is solved. An effective way of finding invariant images at given matrices is based on the use of eigenvectors of matrices. Such productive approach leaded to the novel results in the strip-method useful in cryptography, steganography, and other applications. The multiple examples are met and fully explained during a whole chapter representation.

Chapter 11 is devoted to the numerical systems that may be used in the counters and decoders of digital tools. Nowadays, the classical binary system is the basis of the counters due to their sufficiently simplicity, reliability, and cheapness. However, the counters based upon the binary system lack a natural informational redundancy, which prevents them from finding easily the errors occurring when they work. That is why various methods of enhancing the counters' noise immunity are often proposed. Fibonacci numeral systems with the Fibonacci numbers serving as the weights in the coding words boast such a redundancy to a higher grade than other systems. In addition to their ability to detect errors arising when functioning, the counters based on Fibonacci numeral systems reveal the high computational speed as well. Based on the Fibonacci numeral system, schemes of various noise-proof digital devices such as processors, counters, summers, and even computers have been developed [30]. However, when implementing those devices and circuits, two distinct representation forms: minimal and maximal, were exploited with possible transitions from the minimal form to the maximal one, and vice versa, to do operations over the Fibonacci numbers [31]. In order to complete such transitions, special digital circuits are necessary that realize the operations of catamorphisms and anamorphisms (or “folds” and “unfolds”). These operations made the Fibonacci devices more complicated and slow. This prohibited the thorough use of all potential capabilities of the Fibonacci codes as implemented in efficient Fibonacci counters [32]. The authors investigated the efficiency of the Fibonacci counters handling the minimal presentation form for the Fibonacci numbers, in the part of their information transmission velocity and noise-immunity. Another task is the description of Fibonacci decoders boasting the minimal apparatus costs.

1.3 Conclusions

The chapter provides a briefly description of ten chapters with original mathematical investigations in computer vision techniques applied in remote sensing and close issues. All included chapters involve the long standing original investigations of the authors in novel tensor and wave models and modifications of autoregressive models in order to synthesize the algorithms of multidimensional image processing, new scheme of comparative morphology, which solve the tasks of the image similarity estimation, image matching, and change detection, development of digital halftone (color) images, extraction of straight edges segments with following grouping in the simple or complex shapes, ground traffic monitoring aided by the unmanned aerial vehicles, warping compensation using the proposed structure-from-layered-motion model, image deblurring based on physical processes of blur impacts, core structural verification algorithm for feature extraction, the novel results in the strip-method useful in cryptography and steganography, and the enhanced approach for the counters’ noise immunity in numerical systems based on the Fibonacci numbers. As a rule, each chapter involves large volume of experimental results and justified illustrations.

References

1. Vasil’ev, K.K., Dement’ev, V.E., Andriyanov, N.A.: Application of mixed models for solving the problem on restoring and estimating image parameters. Pattern Recognit. Image Anal. **26** (1), 240–247 (2016)
2. Krasheninnikov, V.R.: Correlation analysis and synthesis of random field wave models. Pattern Recognit. Image Anal. **25**(1), 41–46 (2015)
3. Krasheninnikov, V.R., Kalinov, D.V., Pankratov, Yu.G.: Spiral autoregressive model of a quasi-periodic signal. Pattern Recognit. Image Anal. **8**(1), 211–213 (2001)
4. Krasheninnikov, V.R., Potapov, M.A.: Estimation of parameters of geometric transformation of images by fixed point method. Pattern Recognit. Image Anal. **22**(2), 303–317 (2012)

5. Pyt'ev, Y.P.: Morphological Image Analysis. Pattern Recognit. Image Anal. **3**(1), 19–28 (1993)
6. Vizilter, Y.V., Zheltov, S.Y.: Geometrical correlation and matching of 2D image shapes. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. **I-3**, 191–196 (2012)
7. Vizilter, Y.V., Gorbatsevich, V.S., Rubis, A.Y., Zheltov, S.Y.: (2014) Shape-based image matching using heat kernels and diffusion maps. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. **XL**(3), 357–364
8. Vizilter, Y.V., Rubis, A.Y., Zheltov, S.Y., Vygolov, O.V.: Change detection via morphological comparative filters. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. **III**(3), 279–286 (2016)
9. Petrov, E.P., Medvedeva, E.V.: Nonlinear filtering of statistically connected video sequences based on hidden Markov chains. J. Commun. Technol. Electr. **55**(3), 307–315 (2010)
10. Petrov, E.P., Trubin, I.S., Medvedeva, E.V., Smolskiy, S.M.: Mathematical models of video-sequences of digital half-tone images. In: Atayero, A.A., Sheluhin, O.I. (eds.) Integrated Models for Information Communication System and Networks: Design and Development, pp. 207–241. IGI Global, Hershey (2013)
11. Petrov, E.P., Trubin, I.S., Medvedeva, E.V., Smolskiy, S.M.: Development of nonlinear filtering algorithms of digital half-tone images. In: Atayero, A.A., Sheluhin, O.I. (eds.) Integrated Models for Information Communication System and Networks: Design and Development, pp. 278–304. IGI Global, Hershey (2013)
12. Medvedeva, E.V., Kurbatova, E.E.: Image segmentation based on two-dimensional Markov chains. In: Favorskaya, M.N., Jain, L.C. (eds.) Computer Vision in Control Systems-2. Innovations in Practice, vol. 75, pp. 277–295. Springer International Publishing, Switzerland (2015)
13. Volkov, V., Germer, R., Oneshko, A., Oralov, D.: Object description and extraction by the use of straight line segments in digital images. In: International Conference on Image Processing, Computer Vision and Pattern Recognition (IPCV'2011), pp. 588–594 (2011)
14. Volkov, V., Germer, R., Oneshko, A., Oralov, D.: Object selection by grouping of straight edge segments in digital images. In: International Conference on Image Processing, Computer Vision and Pattern Recognition (IPCV'2013), pp. 321–327 (2013)
15. Volkov, V., Germer, R.: Straight edge segments localization on noisy images. In: International Conference on Image Processing, Computer Vision and Pattern Recognition (IPCV'2010), vol. II, pp. 512–518 (2010)
16. Kim, N., Chervonenkis, M.: Situational control unmanned aerial vehicles for traffic monitoring. Mod. Appl. Sci. **9**(5), 1852–1913 (2015)
17. Kim, N., Bodunkov, N.: Adaptive surveillance algorithms based on the situation analysis. In: Favorskaya, M., Jain, L.C. (eds.) Computer Vision in Control Systems-2, ISRL, vol. 75, pp. 169–200 (2015)
18. Kim, N.: Automated decision making in road traffic monitoring by on-board unmanned aerial vehicle system. Ind. J. Sci. Technol. **8**(S10), 1–6 (2015)
19. Favorskaya, M., Jain, L.C., Buryachenko, V.: Digital video stabilization in static and dynamic scenes. In: Favorskaya, M.N., Jain, L.C. (eds.) Computer Vision in Control Systems-1, ISRL, vol. 73, pp. 261–309. Springer International Publishing, Switzerland (2015)
20. Favorskaya, M., Buryachenko, V.: Fuzzy-based digital video stabilization in static scenes. In: Tsihrintzis, G.A., Virvou, M., Jain, L.C., Howlett, R.J., Watanabe, T. (eds.) Intelligent Interactive Multimedia Systems and Services in Practice, SIST, vol. 36, pp. 63–83. Springer International Publishing, Switzerland (2015)
21. Favorskaya, M., Buryachenko, V.: Fast salient object detection in non-stationary video sequences based on spatial saliency maps. In: De Pietro, G., Gallo, L., Howlett, R.J., Jain, L. C. (eds.) Intelligent Interactive Multimedia Systems and Services, SIST, vol. 55, pp. 121–132. Springer International Publishing, Switzerland (2016)
22. Favorskaya, M., Buryachenko, V., Tomilina, A.: Global motion estimation using saliency maps in non-stationary videos with static scenes. In: De Pietro, G., Gallo, L., Howlett, R.J.,

Jain, L.C. (eds.) Intelligent Interactive Multimedia Systems and Services, SIST, vol. 55, pp. 133–144. Springer International Publishing, Switzerland (2016)
23. Bogoslovskiy, A.V., Zhigulina, I.V., Bogoslovskiy, E.A., Ponomarev, A.V., Vasilyev, V.V.: Linear Blur. Radiotec, Moscow (in Russian) (2015)
24. Malashin, R.: Correlating images of three-dimensional scenes by clusterizing the correlated local attributes, using the Hough transform. J. Opt. Technol. **81**(6), 327–333 (2014)
25. Malashin, R.: Image retrieval with the use of bag of words and structural analysis. J. Phys: Conf. Ser. **735**(1), 12–16 (2016)
26. Malashin, R.: Matching of aerospace photographs with the use of local features. J. Phys.: Conf. Ser. **536**(1), 12–18 (2014)
27. Mironovsky, L., Slaev, V.: Invariants in metrology and technical diagnostics. Meas. Tech. **39** (6), 577–593 (1996)
28. Mironovsky, L., Slaev, V.: The strip method of transforming signals containing redundancy. Meas. Tech. **49**(7), 631–638 (2006)
29. Mironowsky, L., Slaev, V.: Double-sided noise-immune strip transformation and its root images. Meas. Tech. **55**(10), 1120–1127 (2013); Mironovsky, L., Slaev, V.: Implementation of Hadamard matrices for image processing. In: Favorskaya, M.N., Jain, L.C. (eds.) Computer Vision in Control Systems-1, ISRL, vol. 73, pp. 311–349. Springer International Publishing, Switzerland (2015)
30. Stakhov, A.P.: Fibonacci and Golden Proportion Codes as an Alternative to the Binary Numeral System. Part 1. Academic Publishing, Germany (2012)
31. Borisenko, A.A., Kalashnikov, V.V., Protasova, T.A., Kalashnykova, N.I.: A new approach to the classification of positional numeral systems. In: Neves-Silva, R., Tshirintzis, G.A., Uskov, V., Howlett, R.J., Jain, L.C. (eds.) Frontiers in Artificial Intelligence and Applications (FAIA), vol. 262, pp. 444–450 (2014)
32. Borysenko, O.A., Matsenko, S.M., Polkovnikov, S.I.: A noise-proof Fibonacci counter. Trans. Natl. Technol. Univ. Kharkiv **18**, 77–81 (2013) (in Russian)

Chapter 2
Multidimensional Image Models and Processing

Victor Krasheninnikov and Konstantin Vasil'ev

Abstract The problems of developing mathematical models and statistical algorithms for processing of multidimensional images and their sequences are presented in this chapter. Different types of random fields are taken for the basic mathematical image model. This implies two main problems associated with image modeling, namely, model analysis and synthesis. The main attention is paid to the correlation aspect, i.e. evaluation of the correlation function of a random field generated by a given model and, vice versa, development of a model generating a random field with a predetermined correlation function. For this purpose, new models (tensor and wave) and new versions of autoregressive models (with multiple roots) are suggested. The problems of image simulation on the curved surfaces are considered. The suggested models are used to synthesize the algorithms of multidimensional image processing and their sequences. The tensor filtration of imaging sequences and recursive filtration of multidimensional images, as well as the asymptotic characteristics of efficiency of random field filtration on grids of arbitrary dimension are suggested. The problem of object and anomaly detection on the background of interfering images is considered for the images of any dimension, e.g. for multi-zone data. It is shown that four equivalent forms of the optimal decision rule, which reflect various aspects of detection procedure, exist. Potential efficiency of anomaly detection is analyzed. The problems of alignment and estimation of parameters for interframe geometric image transformations are considered for multidimensional image sequences. A tensor procedure of simultaneous filtration of multidimensional image sequence and their interframe displacements are suggested. A method based on a fixed point of a complex geometric image transformation was investigated in order to evaluate large interframe displacements. Options for adaptive image processing algorithms are also discussed in this chapter. In this context, pseudo-gradient procedures are taken as a basis, as they do not require

V. Krasheninnikov (✉) · K. Vasil'ev
Ulyanovsk State Technical University, 32 Severni Venets Street,
Ulyanovsk 432027, Russian Federation
e-mail: kvrulstu@mail.ru

K. Vasil'ev
e-mail: 1948vasiliev@gmail.com

M.N. Favorskaya and L.C. Jain (eds.), *Computer Vision in Control Systems-3*,
Intelligent Systems Reference Library 135, https://doi.org/10.1007/978-3-319-67516-9_2

preliminary evaluation of any characteristics of the processed data. This allows to develop the high-performance algorithms that can be implemented in real-time systems.

Keywords Multidimensional image model · Autoregressive model
Tensor model · Wave model · Curved surface · Processing · Potential efficiency
Prediction · Filtration · Anomaly detection · Recognition · Adaptive algorithm
Pseudo-gradient algorithm

2.1 Introduction

Nowadays, the problems of automatic analysis of images and their sequences are becoming more and more important. This is due to the rapid development of the aerospace Earth monitoring systems, radio and sonar systems with spatial arrays, medical devices for early disease diagnosis, and computer vision systems. For some applications, images can be represented as data files set on a multidimensional index grid changing in discrete time. This allows to describe certain 2D or multidimensional images and their temporal sequences with scalar or vector values, e.g. sequence of TV images, multispectral images, spatial field of wind speeds, and others.

Mathematical models are necessary to formalize image processing problems. Random nature of image values determines the use of probability theory and mathematical statistics methods, i.e. image representation in the form of Random Fields (RFs) [1–5]. This implies the problems of the RF model analysis and synthesis. Composition of information and noise RF is usually used as a model of image observations. This model can also include an additional parameter vector, which allows to consider the peculiarities of image registration, e.g. possible mutual spatial displacement of adjacent frames and various abnormalities (object, signal) [6–9]. The reason for processing the sequences of observations can be development of estimations of an informational RF (prediction, filtration, and interpolation). Another important problem is the estimation of vector parameters, e.g. the image model parameters, spatial displacements, parameters of detected objects, and others.

This chapter presents the results obtained by the authors and their colleagues in modeling and statistical processing of 2D and multidimensional images and their sequences. A number of models [2, 3, 10, 11], including multidimensional ones [7, 12], are known to describe the images and their sequences [13, 14]. Section 2.2 covers the problems of the RF description and simulation. Random field is a set of random variables defined on multidimensional spatial grids with rectangular or more complex cells. For this purpose, the Gibbs, autoregressive, Tensor, and wave models are used [3, 12, 15, 16]. It is difficult to solve synthesis problems using autoregressive models, in particular, even modeling of spatially isotropic RF. However, the use of models with multiple roots of a characteristic equation facilitates to obtain approximately isotropic RF [16–19]. For linear transformations of image frames, it is desirable to use tensor operations that provide the basis for tensor models, in which every recurrent frame is formed from the previous frame and the frame of random variables [16]. The suggested wave models permit to solve the analysis, synthesis, and simulation problems rather efficiently. In these models, the RF is a result of disturbance interaction, which can occur in random places at random times [16, 20]. These models include many well-known models as special cases. Obtaining of the RF with a given type of correlation is achieved by varying a probability distribution of a large-scale disturbance ratio. Section 2.2 also covers different ways to describe the images on arbitrary surfaces. Autoregressive and wave RF models are used for this purpose [16]. For example, the scan of a cylindrical image can be used to represent the speech and other quasi-periodic signals [21], and images on the sphere can be used to describe a planet relief [22].

There are many papers on image filtering. Many methods of algorithm synthesis are developed taking into consideration a type of informative and noise components of observations [6, 15, 23–25]. At the same time, much attention was paid to reduce the computational complexity of filtration procedures, which, in particular, led to the creation of Kalman filters. However, the use of Kalman filtering in image processing causes significant difficulties due to multidimensionality. These difficulties were partially overcome using line-by-line vector filtration, as well as combination of several one-dimensional filters [12]. Section 2.3 introduces a tensor filter of multidimensional image sequences. This filter was developed on the basis of the RF tensor model described in Sect. 2.2. Asymptotic characteristics of efficiency of random field filtration on arbitrary dimension grids are also introduced, as well as a potential accuracy of multidimensional RF filtration.

Many researches are devoted to the problems of object or signal detection and recognition [6, 8, 9, 11, 26, 27]. The problems of synthesis and analysis of optimal decision rules for detecting point and extensive anomalies in the images, including multi-zone data, and their sequences are discussed in Sect. 2.4. Four equivalent forms of the optimal decision rule based on the likelihood ratio are obtained. These four forms are significantly different in computational complexity for large spatial sizes of images [16]. Characteristics of efficiency for anomaly detection in multi-dimensional images are also derived.

An important task of image sequence processing is their alignment, when due to various reasons there are spatial deformations of adjacent frames. Numerous approaches how to solve this problem, for example, search for characteristic points, light flow analysis, morphological analysis, and correlation-extreme methods, are proposed in [7, 10, 11, 14, 28]. In Sect. 2.5, the problems of alignment and estimation of parameters of image interframe geometrical transformations are considered for sequences of 2D and multidimensional images. On the basis of the tensor model of image sequences, a tensor filter for simultaneous filtration of images and their interframe displacement estimation are synthesized [16]. The authors also present a new method to estimate the parameters of image deformation using a fixed point of a complex geometric transformation of two images [16, 29, 30]. This complex transformation consists of the actual deformation plus an additional artificially performed geometric transformation. Unknown parameters of actual deformation are determined according to coordinates of the obtained fixed point.

In practice, the image, noise and observation models are usually only partially known, i.e. there is a priory uncertainty. Thus, the synthesis of adaptive processing algorithms is required. For this purpose, many adaptive algorithms were suggested [31, 32]. Section 2.6 outlines the adaptive procedure options, which are included in the algorithms for solving image prediction and alignment [16]. Special attention is paid to the pseudo-gradient adaptation, on the basis of which highly efficient algorithms with comparatively small computational costs are synthesized. These qualities allow their usage in real-time systems, dealing with large images and their sequences. Section 2.7 concluded the chapter.

2.2 Mathematical Models of Images

The mathematical models suitable for multidimensional processing, such as the random fields, tensor models of random fields, autoregressive models of random fields, wave models of random fields, and random fields on surfaces are discussed in Sects. 2.2.1–2.2.5, respectively.

2.2.1 Random Fields

Nowadays, information systems, including spatial sensor devices and digital computing, are widely used. Therefore, the images with discrete spatial and temporal variables will be primarily considered. Without loss of generality, one may assume that the images are the sets of values arranged on multidimensional rectangular grids with a unit step. Two-dimensional and three-dimensional grids are presented in Fig. 2.1a, b, respectively. In general, an image is a set in n-dimensional grid nodes $\Omega = \{\bar{j} = (j_1, \ldots, j_n) : j_k = \overline{1, M_k},\ k = \overline{1, n}\}$. According to the physical nature, image values may be scalar (e.g. brightness of a monochromatic image), vector (velocity field, color images, displacement field), and more complex (e.g. matrix). If an image value in the node (pixel) $\bar{j}$ is denoted as $x_{\bar{j}}$, then the image is a set of these values on the grid $X = \{x_{\bar{j}} : \bar{j} \in \Omega\}$.

If the data is a time sequence of images, then a sequence can be assumed as a single image, increasing a grid dimension by one unit. For example, the sequence of planar images (Fig. 2.1a) can be regarded as a single three-dimensional image (Fig. 2.1b).

If it is necessary to specify a time variable, let us set it down at the top $X = \left\{x^i_{\bar{j}} : \bar{j} \in \Omega,\ i \in I\right\}$. This image is a set on the direct product $\Omega \times I$ of grids Ω and I, where I is a set of time index values. The cross section $x^i = \left\{x^i_{\bar{j}} : \bar{j} \in \Omega\right\}$, i.e.

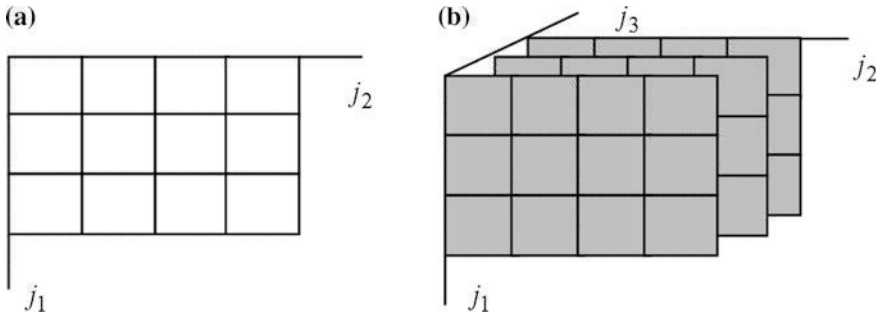

Fig. 2.1 Grids with: **a** two-dimensional image, **b** three-dimensional image

a set of image samples at a fixed value of time index i, is called the ith frame. Each frame is a set on a grid Ω. For example, Fig. 2.1b shows three two-dimensional frames.

Thus, an image can be determined as a function defined on a multi-dimensional grid. The values of the image elements cannot be accurately predicted in advance (otherwise monitoring system would not be necessary). Therefore, it is intrinsic to regard these values as the Random Variables (RVs) using the methods of probability theory and mathematical statistics.

2.2.2 Tensor Models of Random Fields

Consider the RF $X = \left\{ x^i_{\bar{j}} : i \in I, \bar{j} \in \Omega \right\}$ defined on an $(n + 1)$-dimensional grid $\Omega \times I$, where $\Omega = \{j = (j_1, j_2, j_3, \ldots, j_n)\}$ is an n-dimensional $M_1 \times M_2 \times \cdots \times M_n$-grid, $I = \{i : i = 1, 2, 3, \ldots\}$. Index i may be interpreted as a time index, thus the expression $x^i = \left\{ x^i_{\bar{j}} : \bar{j} \in \Omega \right\}$ is called a cross section of a field X as in ith frame.

Let the sequence of frames be described by a stochastic difference Eq. 2.1, where $\left\{ \xi^i_j : i \in I, \bar{j} \in \Omega \right\}$ is an updating standard Gaussian field, $\xi^i = \left\{ \xi^i_j : \bar{j} \in \Omega \right\}$ is the ith field frame, $\varphi^i(x^{i-1}) = \left\{ \varphi^i_j(x^{i-1}) : \bar{j} \in \Omega \right\}$ is $M_1 \times M_2 \times \cdots \times M_n$-matrix function, $\vartheta^i(x^{i-1}) = \left\{ \vartheta^i_{\bar{j}\bar{t}}(x^{i-1}) : \bar{j}, \bar{t} \in \Omega \right\}$ are tensors of rank $2n$ with group indices forming a perturbation component of the ith frame from ξ^i using the rule of tensor multiplication $\vartheta^i \xi^i = \left\{ \vartheta^i_{\bar{j}\bar{t}} \right\} \left\{ \xi^i_{\bar{t}} \right\} = \left\{ \sum \vartheta^i_{\bar{j}\bar{t}} \xi^i_{\bar{t}} \right\}$.

$$x^i = \varphi^i(x^{i-1}) + \vartheta^i(x^{i-1})\xi^i \quad i = 1, 2, \ldots \tag{2.1}$$

Transposing this frame supposes a permutation of its group indices $\vartheta^{iT}_{\bar{j}\bar{\imath}} = \vartheta^{i}_{\bar{\imath}\bar{j}}$. Note that the superscript i indicates a frame number, i.e. time in our case. Thus, index i is considered to be an umbral index and the summation with respect to it is not extended.

Model provided by Eq. 2.1 allows to describe a very broad class of Markov random frame sequences. In particular, a linear model in Eq. 2.1, where tensors P^i and ϑ^i do not depend on x^{i-1}, describes the Gaussian sequence.

$$x^i = P^i x^{i-1} + \vartheta^i \xi^i \tag{2.2}$$

For this field of Covariance Functions (CFs) there exists a multidimensional covariance matrix defined directly

$$V_x(i,j) = M[(x^i - m^i) \times (x^j - m^j)] \quad V_x(i) = V_x(i,i),$$

where $m^i = M[x^i] = \left\{ M[x^i_{\bar{j}}] : \bar{j} \in \Omega \right\}$, and symbol "$\times$" indicates an external matrix multiplication. Thus, $V_x(i)$ and $V^i_\xi = M[\xi^i \times \xi^i]$ are symmetrical $M_1 \times M_2 \times \cdots \times M_n \times M_1 \times M_2 \times \cdots \times M_n$-matrices. For complete definition of a random field with a state Eq. 2.2, it is necessary to set the law of first frame x^0 distribution. It is often a Gaussian distribution with mean m^0 and covariance matrix V^0_x.

Model in Eq. 2.2 with constant tensors $P^i = P$ and $\vartheta^i = \vartheta$ is of particular interest that leads to Eq. 2.3, where (k) determines raising to the kth power.

$$V_x(i, i+k) = P^{(k)} V_x(i,i) = P^{(k)} V_x(i) \tag{2.3}$$

If roots of the characteristic equation $\det(\lambda E - P) = 0$ are less than 1 in modulus, than $P^{(k)} \to 0$ as $k \to \infty$, and from Eq. 2.3 we obtain that $V_x(i, i+k) \to 0$ as $k \to \infty$.

Using z-transformation device, let us write Eq. 2.2 in the form

$$x^i = zPx^i + \vartheta \xi^i$$

or

$$(E - zP)x^i = \vartheta \xi^i.$$

Multiplying the congruence on the left by $(E - zP)^{-1}$, the expression x^i through a perturbing field is obtained.

$$x^i = (E - zP)^{-1}\vartheta\xi^i$$

Hence,

$$x^i(x^i)^T = (E - zP)^{-1}\vartheta\xi^i(\xi^i)^T\vartheta^T(E - z^{-1}P^T)^{-1}$$

that after averaging gives the expression of the tensor spectrum for a stationary field

$$V_x(z) = \sum_{k=-\infty}^{\infty} V_x(0,k)z^k = (E - zP)^{-1}\vartheta V_\xi\vartheta^T(E - z^{-1}P^T)^{-1},$$

which is a Laurent series in powers of z with tensor coefficients. The tensor coefficients can be found from Eq. 2.4, where $C = \{z : |z| = 1\}$ is a unit circumference of a complex plane.

$$V_x(0,k) = \frac{1}{2\pi i}\int_{C_1} (E - zP)^{-1}\vartheta\, V_\xi\vartheta^T(E - z^{-1}P^T)^{-1}z^{k-1}dz \qquad (2.4)$$

It is sufficient to find $V_x = V_x(0,0)$ from Eq. 2.4, other values are obtained from Eq. 2.3 in a view of $V_x(i,i+k) = P^{(k)}V_x$ for a stationary case. To find V_x it is possible to use limit in Eq. 2.3 as $i \to \infty$ instead of integral in Eq. 2.4

$$V_x = PVP^T + \vartheta\, V_\xi\vartheta^T$$

that is a non-singular system of linear equations relative to tensor components V_x.

Now consider a solution to the model synthesis problem Eq. 2.2, i.e. the problem of finding tensors P^i and ϑ^i, when intraframe $V_x(i,i)$ and interframe $V_x(i-1,i)$ covariance tensors are given. From Eq. 2.3 we obtain

$$V_x(i-1,i) = P^iV_x(i-1,i-1)$$

that is a system of linear equations relative to tensor elements P^i. It is obvious that

$$P^i = V_x(i-1,i)V_x^{-1}(i-1,i-1).$$

Chose a perturbing field with covariance $V_\xi^i = E$. Then the following equation is obtained

$$\vartheta^i\vartheta^{iT} = V_x(i,i) - P^iV_x(i-1,i-1)P^{iT}.$$

For example, for tensor components ϑ^i this equation can be solved on the basis of Gram-Shmidt orthogonalization.

Nonlinear stochastic difference equation may be considered as generalization of the considered tensor model (Eq. 2.2). This equation allows to describe very large class of Markov non-Gaussian RF on n-dimensional grids J_t. Here we have $\left\{\xi^t_{\bar{l}}, \bar{l} \in J_t, t \in T\right\}$, a field of independent, generally speaking, non-Gaussian random variables with known probability density function (PDF) $W(\xi^t_{\bar{l}})$, $\varphi^t_{\bar{j}}(x^{t-1}_{\bar{l}})$ and $\vartheta^t_{\bar{j}\bar{l}}(x^{t-1}_{\bar{s}})$ are tensors of ranks n and $2n$ correspondingly, which in a general case nonlinearly depend on the values $\left\{\xi_{\bar{j}}\right\}$ of the (t–1)th frame of a multidimensional RF $\left\{x^{t-1}_{\bar{j}}, t \in T, \bar{j} \in J\right\}$.

2.2.3 *Autoregressive Models of Random Fields*

Tensor representation assumes that for each moment t of discrete time the RF $\left\{x^t_{\bar{j}},\ \bar{j} \in J\right\}$ is formed recurrently on the basis of the previous value $\left\{x^{t-1}_{\bar{j}},\ \bar{j} \in J\right\}$ and updating RF $\left\{\xi^t_{\bar{j}},\ \bar{j} \in J\right\}$ of independent RV. Despite the fact that the calculations are done recurrently while forming each successive frame $\left\{x^t_{\bar{j}},\ \bar{j} \in J\right\}$, it is desirable to conduct linear Eq. 2.1 or nonlinear Eq. 2.2 transformation of all elements $\left\{x^{t-1}_{\bar{j}},\ \bar{j} \in J\right\}$, $\left\{\xi^t_{\bar{j}},\ \bar{j} \in J\right\}$ determined on an n-dimensional spatial grid.

Thus, tensor models make it possible to describe a big class of non-Gaussian and non-homogeneous RF but lead to an overall computational effort during the RF imitation and processing. In this context, there appear questions about the existence of recurrent not only in time but also in space RF representation and the possibility of constructing optimal recurrent algorithms of statistical analysis for such RF.

In their structure, the random fields are much more complex than stochastic processes. First, implementations of the RF are functions of several variables, the theory of which is more complicated than of one variable. Second, the concept of Markov behavior also becomes much more complicated. A random process can develop in course of time. Model from Eq. 2.1 is a mathematical expression of such development. For Markov sequence, the time interval can be broken at any point i for conditionally independent past $\Gamma^- = \left\{x^k : k < i\right\}$ and future $\Gamma^+ = \left\{x^k : k > i\right\}$. However, the RF is defined on an n-dimensional domain Ω. For its geometrical partitioning into two parts Γ^- and Γ^+, at least an (n – 1)-dimensional domain Γ is required. Markov RF suggests that for any set Γ (of a certain class) all RV included in Γ^- are conditionally independent from the RV belonging to Γ^+, when values of Γ are known. It is possible to name Γ^-, Γ, and Γ^+ as past, present, and future only roughly. However, Markov property allows to imagine the RF as one developing in time from Γ^- through Γ to Γ^+, in addition Γ moves along Ω with the course of time. For example, if it is assumed that the lines of a two-dimensional grid Ω represent Γ, then the field is formed line by line.