Roumen Kountchev · Rumen Mironov ·
Ivo Draganov · Roumiana Kountcheva ·
Kazumi Nakamatsu   *Editors*

New Approaches for Multidimensional Signal Processing

Proceedings of International Workshop,
NAMSP 2023

KES International

Springer

Smart Innovation, Systems and Technologies

Volume 385

The Smart Innovation, Systems and Technologies book series encompasses the topics of knowledge, intelligence, innovation and sustainability. The aim of the series is to make available a platform for the publication of books on all aspects of single and multi-disciplinary research on these themes in order to make the latest results available in a readily-accessible form. Volumes on interdisciplinary research combining two or more of these areas is particularly sought.

The series covers systems and paradigms that employ knowledge and intelligence in a broad sense. Its scope is systems having embedded knowledge and intelligence, which may be applied to the solution of world problems in industry, the environment and the community. It also focusses on the knowledge-transfer methodologies and innovation strategies employed to make this happen effectively. The combination of intelligent systems tools and a broad range of applications introduces a need for a synergy of disciplines from science, technology, business and the humanities. The series will include conference proceedings, edited collections, monographs, handbooks, reference books, and other relevant types of book in areas of science and technology where smart systems and technologies can offer innovative solutions.

High quality content is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes will ensure that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

Indexed by SCOPUS, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, Japanese Science and Technology Agency (JST), SCImago, DBLP.

All books published in the series are submitted for consideration in Web of Science.

Roumen Kountchev · Rumen Mironov ·
Ivo Draganov · Roumiana Kountcheva ·
Kazumi Nakamatsu

Editors

New Approaches for Multidimensional Signal Processing

Proceedings of International Workshop, NAMSP 2023

Editors
Roumen Kountchev
Technical University of Sofia
Sofia, Bulgaria

Ivo Draganov
Technical University of Sofia
Sofia, Bulgaria

Kazumi Nakamatsu
University of Hyogo
Kobe, Japan

Rumen Mironov
Technical University of Sofia
Sofia, Bulgaria

Roumiana Kountcheva
TK Engineering
Sofia, Bulgaria

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721,
Singapore

Paper in this product is recyclable.

Preface

This book contains papers presented at the International Workshop "New Approaches for Multidimensional Signal Processing-NAMSP 2023" which was carried out during July 06–08, 2023 at the Technical University of Sofia, Bulgaria. The workshop was a part of the "Days of Science 2023" at the Technical University of Sofia, organized in collaboration of the Research and Development Sector at TU-Sofia, Bulgaria. The workshop was supported by the Bulgarian National Science Fund (BNSF) at the Ministry of Education and Science of Bulgaria. Co-organizers of NAMSP2023 were Interscience Research Network (IRNet) International Academy Communication Center, China; Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Haryana, India, and Interscience Institute of Management and Technology-Bhubaneswar, India. In the workshop scientists participated from China, India, Canada, USA, Hungary, Nigeria and Bulgaria.

The book comprises 25 chapters. The three main topics, on which were focused the accepted papers, are:

(1) Multidimensional Image Tensor Decomposition and Analysis;
(2) Theoretical Approaches in Multidimensional Signal Processing;
(3) Multidisciplinary Applications of Multidimensional Signal Processing in industry, business, medicine, power engineering, ecology, etc.

The aim of the book is to present the latest achievements of the authors in the processing and analysis of multidimensional signals and the related applications, to a wide range of readers: IT specialists, engineers, physicians, Ph.D. students and other specialists.

Sofia, Bulgaria   Roumen Kountchev
Sofia, Bulgaria   Rumen Mironov
Sofia, Bulgaria   Ivo Draganov
Sofia, Bulgaria   Roumiana Kountcheva
Kobe, Japan   Kazumi Nakamatsu
October 2023

The editors express their warmest thanks to the excellent Springer team which made this book possible.

Contents

About the Editors

Prof. Dr. Roumen Kountchev Ph.D., D.Sc. is a professor at the Faculty of Telecommunications, Department of Radio Communications and Video Technologies, Technical University of Sofia, Bulgaria. His areas of interest are digital signal and image processing, image compression, multimedia watermarking, video communications, pattern recognition, and neural networks. He has 360 papers published in magazines and proceedings of conferences; 22 books; 47 book chapters; 21 patents. He had been principle investigator of 38 research projects. At present, he is a member of Euro Mediterranean Academy of Arts and Sciences and President of Bulgarian Association for Pattern Recognition (member of Intern. Association for Pattern Recognition). He has been a plenary speaker at more than 30 international conferences and workshops and edited 28 books published in Springer series.

Prof. Dr. Rumen Mironov received his MSc and PhD in Telecommunications from Technical University of Sofia and MSc in Applied Mathematics and Informatics from Faculty of Applied Mathematics and Informatics. He is currently associated professor in the Department of Radio Communications and Video Technologies, Faculty of Telecommunications, Technical University of Sofia, Bulgaria. His current research focuses on digital signal and image processing, pattern recognition, audio and video communications, information systems, computer graphics, multidimensional signal processing and programming languages. He is a Member of Bulgarian Association of Pattern Recognition (IAPR), Bulgarian Union of Automation and Automation Systems and he is a member of IJBST Journal Group. He participated in more than 20 national and international scientific research projects and he is Author of more than 120 scientific publications in magazines and proceedings of conferences.

Prof. Dr. Ivo Draganov received B.Sc., M.Sc., and Ph.D. in Telecommunications from Technical University of Sofia, Bulgaria, in 2003, 2005, and 2009. Currently, he is an associate professor in the Faculty of Telecommunications and teaches Programming in Assembler, Digital Design, Measurements in Telecommunications, Audio Systems and Speech Coding, Multimedia Systems, Digital Video Broadcasting, and other subjects. His main areas of scientific research are digital image processing,

pattern recognition, multimedia systems, and neural networks. He is a member of the Editorial Board of the Computer and Communications Engineering Journal. During his research activities, he participated in more than 10 national and international projects and published more than 80 papers. He is a member of IEEE.

S.R. Dr. Roumiana Kountcheva is Vice President of TK Engineering. She got her M.Sc. and Ph.D. at the Technical University of Sofia, Bulgaria, and became Senior Researcher (SR) at TIE, in 1993. She had two post-doc trainings in Japan (Fujitsu, 1977, and Fanuc, 1980). She has more than 200 publications (including 28 book chapters and five patents, and presented 22 plenary speeches at international conferences and workshops. She is Member of IRIEM, IDSAI, IJBST Journal Group, and Bulgarian Association for Pattern Recognition. She participated as PI, Co-PI, and Team Member of 47 scientific research projects, from which 14 international. She is Reviewer of WSEAS conferences and journals and edited several books for Springer SIST series. Her scientific areas of interest are in image processing and analysis, image watermarking and data hiding, multidimensional image representation, tensor representation, pyramidal image decompositions, image and video compression, CNC and programmable controllers, etc.

Prof. Dr. Kazumi Nakamatsu received the M.S. Eng. and Dr. Sci. from Shizuoka University, 1976, and Kyushu University, Japan, 1999, respectively. He is Full Professor at School of Human Science and Environment, University of Hyogo, Japan. His research interests encompass various kinds of logic and their applications to Computer Science, especially paraconsistent annotated logic programs and their applications. He is Author of over 150 journal papers, book chapters, and conference papers and edited 14 books published by prominent publishers such as Springer-Verlag. He has chaired various international conferences, workshops, and invited sessions, and he has been a member of numerous international program committees of workshops and conferences in the area of Artificial Intelligence and Computer Science. He has contributed numerous invited lectures at international workshops, conferences, and academic organizations. He also is Recipient of some conference and paper awards. He is Member of Japan AI Society, etc.

Chapter 1
Deep Representation and Analysis of Visual Information, Based on the IDP Decomposition

Lakhmi C. Jain, Roumen K. Kountchev, and Roumiana A. Kountcheva

Abstract We present contemporary methods for image decomposition analysis in the spectrum domain, based on the Inverse Difference Pyramid (IDP) decomposition. The basic IDP implementations in various aspects of visual information processing and analysis are discussed, in the range from 2D images to third-order tensors. Special attention is paid to the main IDP features, which are compared with those of the famous pyramidal decompositions. The basic IDP modifications are presented: the Reduced Branched IDP, which could be implemented on the basis of various 2D orthogonal transforms (WHT, discrete Fourier transform DFT, DCT, KLT, etc.), and the upgrade to the Adaptive IDP, based on neural networks integration. Special approaches are introduced for the IDP-based decomposition for sequences of correlated images, and some important applications in multidimensional image tensor representation are given; the compression of single and groups of correlated multi-spectral, multi-view, and computer tomography images; and the faster object search in large image databases. The experimental results obtained by the approaches based on the IDP decomposition, confirm its efficiency which is very high for some image classes. In the conclusions the analysis results and the trends for future investigations and implementations are explained.

L. C. Jain
KES International, Selby and University of Arad, Arad, Romania

R. K. Kountchev
Technical University of Sofia, Bul. Kl. Ohridsky 8, Sofia, Bulgaria

R. A. Kountcheva (✉)
TK Engineering, Sofia, Bulgaria
e-mail: kountcheva_r@yahoo.com

1.1 Introduction

The widely used methods for hierarchical image decomposition, based on the pyramidal representation in the pixel domain, comprise of two main approaches, the Non-orthogonal Pyramids and the Orthogonal Pyramids [1, 2]:

– Non-orthogonal Pyramids belong the Gaussian/Laplacian Pyramid (GP/LP); the Reduced/Enhanced Laplacian Pyramid, the Reduced-Sum/Reduced-Difference Pyramid, the Hierarchy-Embedded Differential Pyramid, etc.,
– Orthogonal Pyramids are mainly represented by the Sub-Band Decomposition based on filter banks: the Orthonormal Wavelet Pyramid, the Steerable Pyramid (based on directional filter banks), the Curvelets Pyramid, the Ridgelet Pyramid, etc.

The approaches, based on the Laplacian/Gaussian decompositions are well-known and multiple applications already exist for image compression, analysis, machine learning, etc. [1–3]. Specific for these decompositions is, that in the lowest decomposition level the number of components is the largest, and decreases in each consecutive level (Fig. 1.1). The decomposition ends after the last coefficients of the highest level are calculated. Besides, these decompositions require the execution of multiple decimations and interpolations. Unlike this the IDP decomposition starts with a relatively low number of decomposition components and each level contains a larger number, but the decomposition can stop before the last possible level is executed, depending on the needed quality of the restored image.

The Inverse Difference Pyramid [4] is a kind of hierarchical image decomposition, based on pyramidal representation in the spectrum domain. The IDP structure is based on orthogonal deterministic or statistic transforms, such as the Discrete Fourier Transform (DFT), the Discrete Cosine Transform (DCT), the Walsh-Hadamard Transform (WHT), the Complex Hadamard Transform (CHT), the Karhunen–Loeve Transform (KLT), etc. One modification of IDP is the Adaptive Inverse Difference Pyramid (AIDP), which could be based on the following: Back-Propagation Neural Networks (BPNN), Gaussian Radial Basis Function Networks (GRBFN), Self-Organizing Feature Mapping Vector Quantization (SOFM-VQ). Another kind of IDP decomposition is the Non-linear Inverse Difference Pyramid (NL-IDP), based on the Discrete Modified Mellin-Fourier Transform (MMFT).

A simplified structure of the basic IDP decomposition is shown on Fig. 1.2.

The processing of the input (original) 2D image starts at the lowest level. For this, the image is represented by the matrix, [B]. To simplify the calculations, the image is divided into smaller pieces (sub-blocks), each of size $2^n \times 2^n$, and on the figure is shown the processing of one of these blocks only. For the decomposition any 2D orthogonal transform could be used like the Walsh-Hadamard Transform (WHT), the Discrete Fourier Transform (DFT), the Discrete Cosine Transform (DCT), the Karhunen–Loève Transform (KLT) or Principal Component Analysis (PCA), etc. In order to reduce the number of calculated transform coefficients, the "truncated" orthogonal transform is used, i.e., part of the coefficients are not calculated because
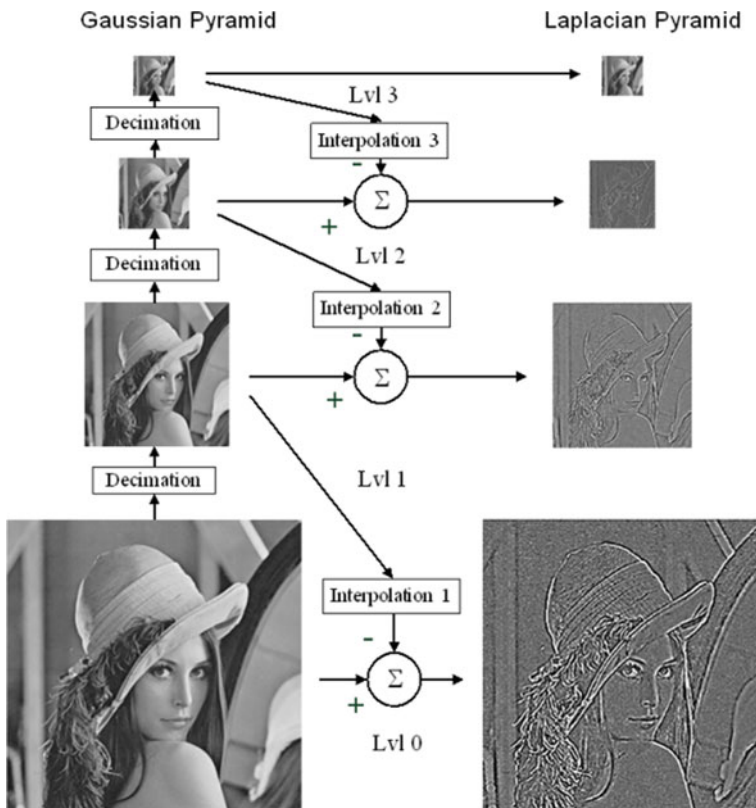
Fig. 1.1 Image decomposition based on Gaussian/Laplacian pyramid

(after analysis) their values are evaluated as too small to influence the quality of the restored image. The corresponding block in the figure is denoted as Truncated Orthogonal Transform, "TOT". The values of the calculated coefficients are saved in the block $[\hat{S}]$ (in the figure are used correspondingly: $[\hat{S}_0]$—for the start level; $[\hat{S}_1]$—for level 1, etc.). From the start level in this example 4 coefficients are retained, $\hat{s}_0(u, v)$ only, as shown in the lower right part of the figure. The processing continues with the preparation of the data for the next decomposition level. For this, on the coefficient s_0 Inverse Orthogonal Transform (the block "IOT") is executed, and the first is obtained, which is the coarse approximation of the processed image block, $[\hat{B}_0]$. This approximation (i.e., the restored image sub-block) is subtracted from the original in "$\sum$" and the difference (error) image is calculated. For this level, the "Error" image matrix is denoted as $[E_0]$. In the second level it is divided into smaller sub-blocks, of size $2^{n-1} \times 2^{n-1}$, and after that processed with TOT again. The values of the calculated coefficients are saved in the corresponding blocks $[\hat{S}_1]$. The calculated 16 coefficients for this level, are saved in the block $\hat{s}_1(u, v)$. The coefficients in the next level are calculated in a similar way and the obtained corresponding 256
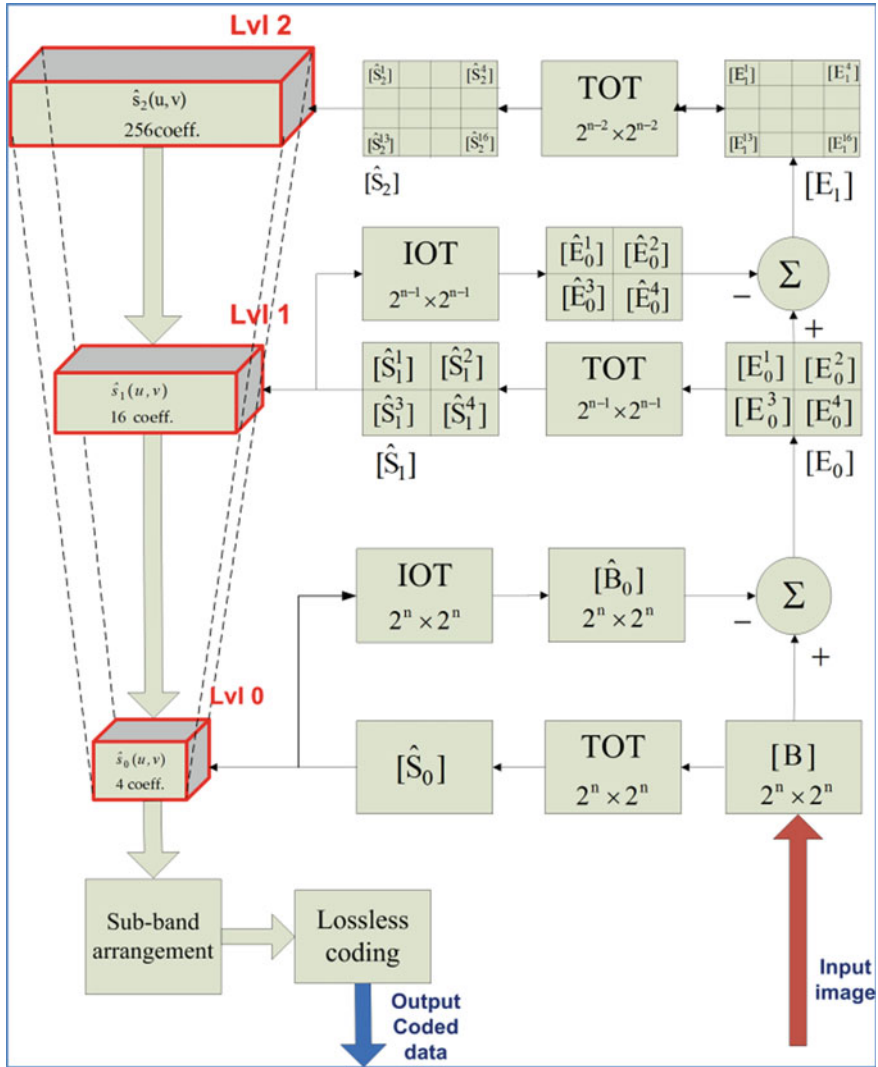
Fig. 1.2 Basic structure of the inverse difference pyramid

coefficients are retained and saved in the block $\hat{s}_2(u, v)$. As it is shown, the number of coefficients is higher in each consecutive decomposition level, and this is why the decomposition is called "Inverse Difference Pyramid". The decomposition can stop in any decomposition level, depending on the achieved quality of the restored image, or on the values calculated for the "difference" image. When the decomposition is stopped, the values of the coefficients, retained in all levels, are arranged in accordance with their spatial frequencies and are then losslessly coded through Adaptive Run-length Encoding (ARLE), followed by a Modified Huffman (MH) coding. The

ARLE and the MH coding were specially developed for the IDP decomposition, in accordance with the specifics of the processed data [5]. The restoration of the decomposed image is executed in reverse order.

The main advantage of IDP is its efficiency when used for the compression of 2D or 3D visual information. Unlike Laplacian/Gaussian pyramids, the IDP decomposition does not use decimations and interpolations, because it is based on some kind of orthogonal transform. The computational cost of the process depends on the number of needed operations, analyzed in previous research of the authors [6]. The authors also proved theoretically the low computational complexity for the tensor representation of multidimensional visual information [7, 8]. One important disadvantage of IDP is, that if all coefficients are retained, the decomposition is over-complete. To avoid this, the properties of the used orthogonal transform, are used. For example, for the WHT it is not necessary to calculate all neighbor coefficients in a group, because for each 4 coefficients is necessary to calculate 3 only (the fourth is calculated by using the values of the remaining three). On the basis of this property, the problem with the decomposition over-completeness, is solved.

1.2 Branched IDP

To achieve a better performance for the processing and analysis of sequences of correlated images, the Branched Inverse Difference Pyramid (BIDP) was developed. This approach is extremely useful for processing sequences of medical images (computer tomography images, magnetic resonance images, etc.), and groups of multispectral or multi-view images. In all these cases, the processed groups of images have high mutual correlation. The BIDP block diagram is shown on Fig. 1.3. In this case, the basic IDP diagram is retained (the red rectangle), but some new relations between images in the processed group, are introduced.

In the case shown on Fig. 1.3, one sequence of $(2N + 1)$ images (representing same object or scene) is processed bound together by similarity, and one of the images is used as a reference.

To select the reference image, various approaches could be used, for example, through PSNR comparison, etc. For video sequences, the middle image in the group is usually the most suitable, and this is the easiest solution. Another approach is based on the analysis of the images histograms for the group: the image, whose histogram is most similar with these of the remaining in the group, is chosen to be used as a reference (R).

The image decomposition starts with this reference image, which is processed with some kind of orthogonal transform, using a limited number (preset) of transform coefficients. After inverse transform of the so calculated coefficients, the coarse approximation of the processed image is obtained. The IDP decomposition then branches out into several decompositions, whose number corresponds to the number of images in the group. The first approximation for all multispectral images is that calculated for the reference image. In case that IDP comprises of 2 levels only, each
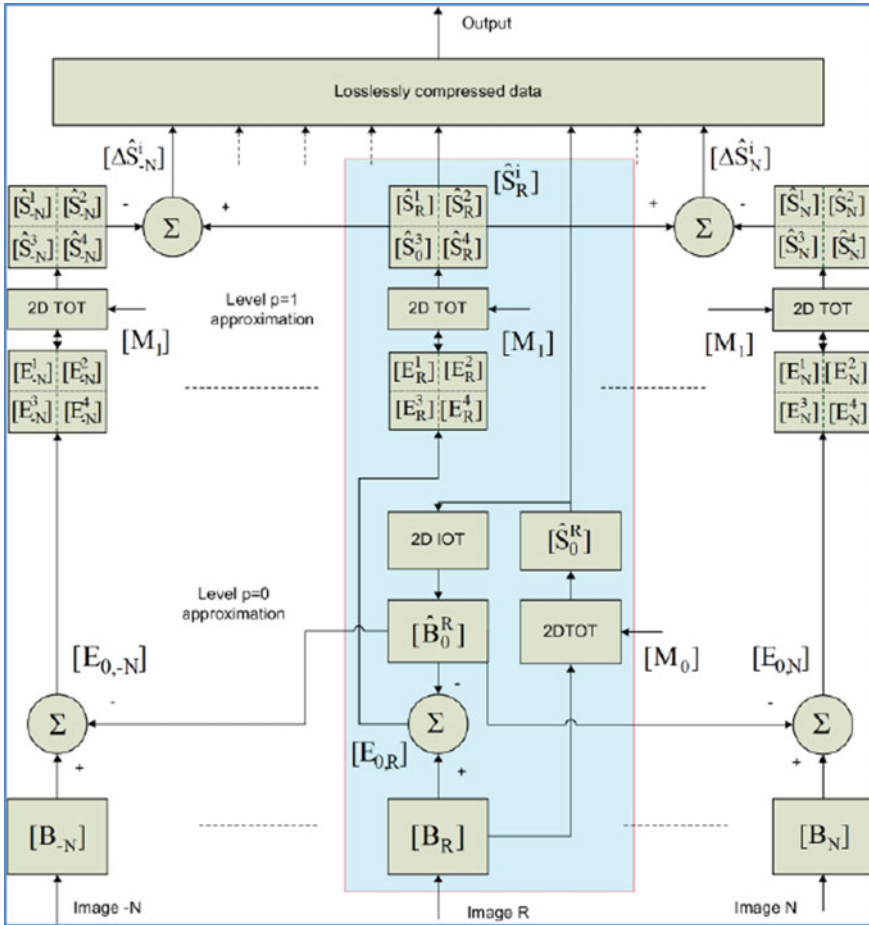
Fig. 1.3 Block diagram of BIDP

branch is built individually in the next level. For IDP of more levels, the reference image approximation could be used also in the second level, etc. The similarity between the processed images, permits the same coarse approximation to be used for the whole group. Depending on the visual contents, the number of images in one group could be different and is set in relation to their mutual correlation. For example, for video sequences, the highest correlation usually exists between each 8–12 sequential frames, and for CT and MRI sequences, longer groups could be used. Due to the result of the processing based on BIDP, high compression and very good visual quality of the restored images are achieved, as confirmed by the experimental results.

1.3 Multi-layer Tensor Decomposition Trough 3D IDP

The common practice in image representation is based on the use of a 2D matrix, where each pixel corresponds to one matrix element. Together with this, many contemporary applications exist, where video sequences and groups of correlated images, obtained from various sources must be stored, analyzed or searched, and for this, the most suitable approach is that they are to be treated as 3D arrays of matrices. Recently tensors, and specially the third-order tensors, are most suitable to represent such sequences. The main obstacle for the wide tensor decomposition implementation in real time applications is the high computational complexity. Tensor decompositions are usually based on deterministic discrete transforms of the kind: Discrete Wavelet Transform (DWT), or the Discrete Cosine Transform (DCT) followed by SVD in the frequency domain [9]. In part of the related publications [10–13], algorithms are proposed for cubical decomposition based on the 3D separable discrete transforms: Discrete Fourier Transform, Discrete Hartley Transform, Discrete Cosine Transform, etc. To reduce the computational complexity, in many cases "fast" algorithms are used, one of which is the 3D Fast Fourier Transform (FFT). Compared to the SVD/PCA-based algorithms, the tensor decomposition based on deterministic orthogonal transforms offers lower energy concentration in the first decomposition components but accelerates the computations. This is why, the tensor decomposition based on orthogonal transforms is reasonable in cases when real-time processing of various multidimensional data is needed.

The approach, presented here, is the hierarchical third-order tensor decomposition based on the 3D-IDP. For this, the tensors are transformed into the 3D-WHT spectrum space. The basic concept is to represent each third-order tensor X of size $N \times N \times N$ through a 3D Reduced IDP (3D-RIDP) [14] of the kind, shown on Fig. 1.4. For this, the tensor X is initially divided into Q sub-tensors X_q for $q = 1, 2, …, Q$, each of size $M \times M \times M$, where $M = \lfloor N/\sqrt[3]{Q} \rfloor$. The value of M is defined in accordance with the condition $M = 2^m$. For the calculation of each sub-tensor X_q of size $N \times N \times N$ ($N = 2^n$) is built the individual n-level 3D-RIDP. In result, the tensor X is transformed into the corresponding spectrum tensor S, which comprises n levels of coefficients. The coefficients in the initial level have the highest energy concentration, while the energy in the next levels decreases quickly. In correspondence with Parseval's theorem, where the total energy of the coefficients of the tensor S is equal to that of the elements of the tensor X, but is redistributed. The main advantages of the method are the lower computational complexity because the only mathematical operations, which are needed, are "additions" and their number is relatively low. Furthermore, the main part of the tensor energy is concentrated in a small number of coefficients from the first pyramid level which permits significant information redundancy reduction, after neglecting the low-energy elements.

The properties of the 3D-RIDP open new possibilities for implementation in various application areas related to processing and analysis of 3D data: sequences of correlated images (video, multi-spectral, multi-view, medical images from various sources), multichannel signals, etc.
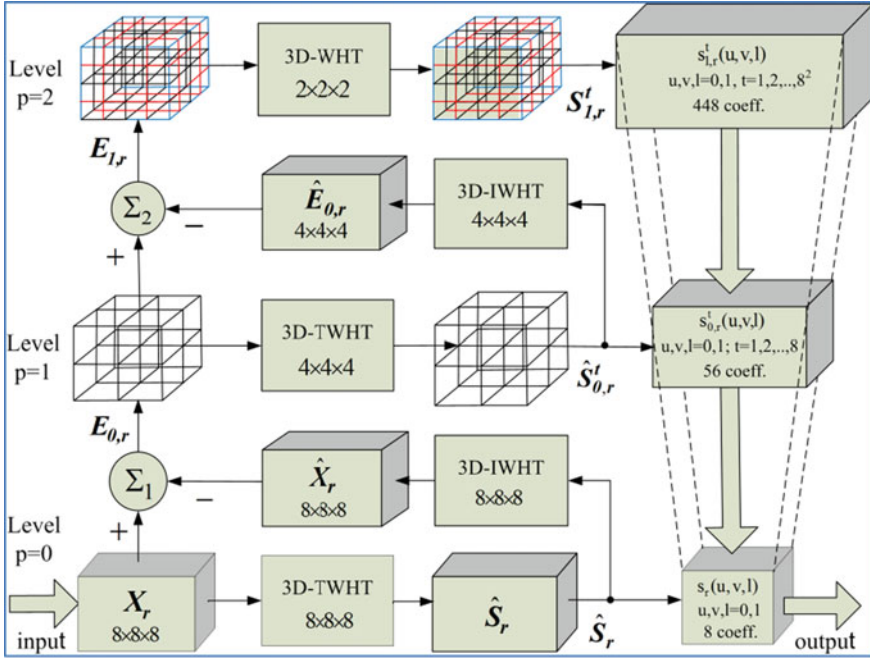
Fig. 1.4 Block diagram of 3D reduced IDP

1.4 Multi-layer Tensor IDP, Based on Hierarchical SVD

To achieve computational complexity reduction, a new non-iterative approach for multi-dimensional tensor representation based on the Multi-layer Tensor Spectrum Pyramid (MLTSP) [15] is proposed, with embedded 3D orthogonal transforms (3D OT) and Hierarchical Tensor SVD (HTSVD) [8]. This approach is illustrated by an example for the representation of a tensor of size $8 \times 8 \times 8$ through a two-layer tensor spectrum pyramid (2LTSP), with embedded 3D Frequency-Ordered Fast Walsh-Hadamard Transform (3D FO-FWHT) [7] and HTSVD for a tensor of size $2 \times 2 \times 2$ (HTSVD$_{2\times2\times2}$) [8]. To explain the multi-layer TSP structure, as an example here is used the two-layer tensor spectrum pyramid (2LTSP), which comprises a coder and a decoder. The structure of the decoder is mirror-symmetrical to that of the coder. Both block diagrams are shown on Fig. 1.5, a, b. The block diagram of the computational graph of the algorithm HTSVD$_{2\times2\times2}$ of two hierarchical levels for the decomposition of the elementary tensor $S_{2\times2\times2}$ of size $2 \times 2 \times 2$, is shown in Fig. 1.6. The decomposition is based on SVD for the matrix $[X]$ of size 2×2, denoted as SVD$_{2\times2}$. The SVD$_{2\times2}$ decomposition is executed through simple relations, of low computational complexity [15]. After the decomposition of $S_{2\times2\times2}$ is finished, the tensors in the resulting sum are arranged following the decrease of the variances for the sub-matrices obtained after the unfolding.
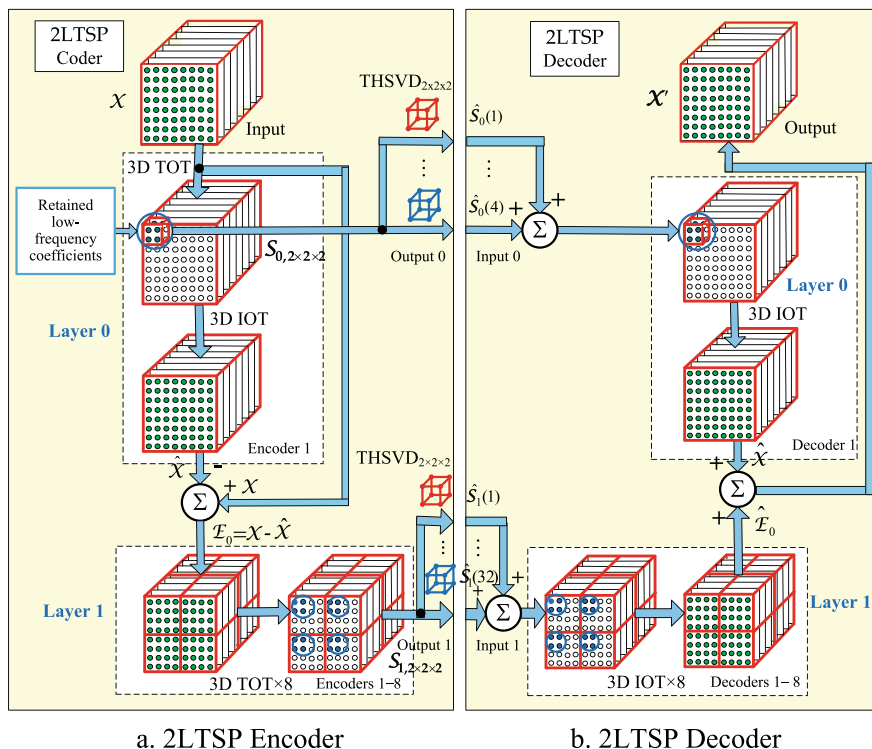
a. 2LTSP Encoder   b. 2LTSP Decoder

Fig. 1.5 2LTSP based on 3D OT and HTSVD$_{2\times2\times2}$ for input tensor X of size $N \times N \times N$, for $N = 8$

The voxels of higher values in the $S_{2\times2\times2}$ decomposition in Fig. 1.6, are colored in red, and these of lower—in blue. The main advantages of the MLTSP algorithm are the low computational complexity, the high flexibility regarding the choice of their parameters, and the high ability for information redundancy reduction in the input tensor.

1.5 3D Adaptive Inverse Difference Pyramid with Convolutional Auto Encoder/Decoder

This adaptive IDP version is based on the use of a Convolutional Auto Encoder/ Decoder (CED) [16]. The two components of the CED neural network are ained by deep learning. These are the Convolutional Coder (CE), is used to compress the input data, and the Convolutional Decoder (CD), which restores the already compressed input data. On Fig. 1.7, the block diagram of one multi-layer CED is shown, of the kind $m^3 \rightarrow n \rightarrow m^3$, i.e., it comprises m^3 input cells, n cells in the hidden layer (the
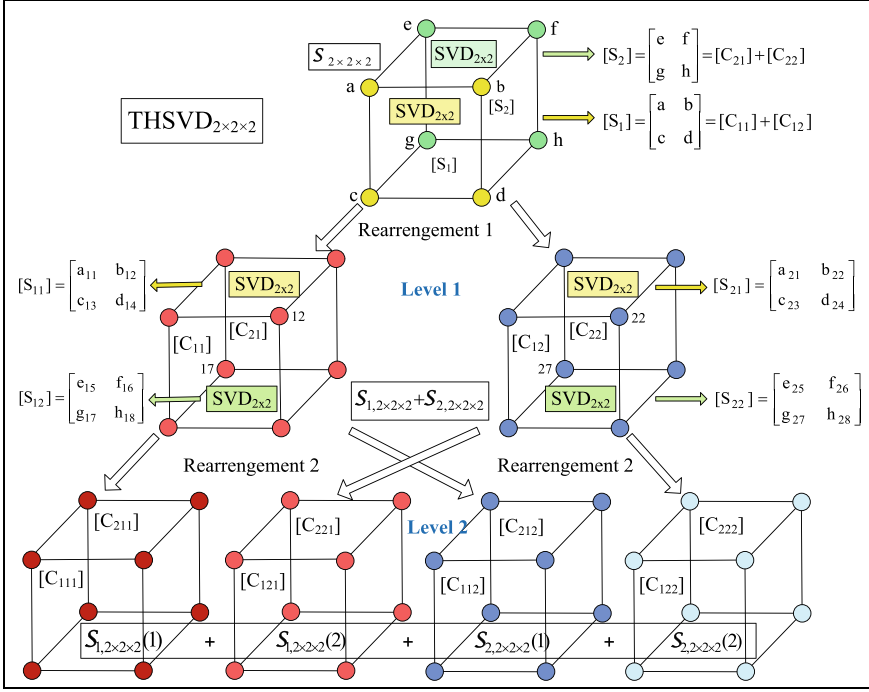
Fig. 1.6 $HTSVD_{2\times2\times2}$ algorithm for the tensor $S_{2\times2\times2}$ decomposition

output of the coder and the input of the decoder, respectively), and m^3 output cells. After the end of the iterative CED training, the values obtained for the output cells approximate these from the input cells, at minimum root mean square error.

Two new approaches are offered here for the compression of a single third-order (cubical) tensor.

The *First approach* is based on the Adaptive Inverse Difference Pyramid (AIDP) structure, combined with CED. On Fig. 1.8 the corresponding three-level block diagram is shown. Here the third-order tensor of size $m \times m \times m$, enters the AIDP input. In the initial (zero) hierarchical AIDP level, the elements of the input tensor are arranged as a sequence of length m^3, following a preselected rule. This sequence defines the m^3-dimensional vector, which enters the auto-encoder CED-0. In this case, the hidden layer contains n cells (for $n << m^3$), while the output layer is of m^3 cells. From them, after inverse rearrangement, the output third-order tensor of size $m \times m \times m$ is restored. After the CED-0 self-training is finished, the so obtained output tensor approximates the input tensor, of size $m \times m \times m$. In the first summator (Σ_1), the approximated tensor is subtracted element-by-element from the input tensor, and as a result the difference tensor is obtained, which corresponds to the approximation error. In the first hierarchical AIDP level, the difference tensor is divided into 8 sub-tensors, each of size $(m/2) \times (m/2) \times (m/2)$. The elements of these sub-tensors are
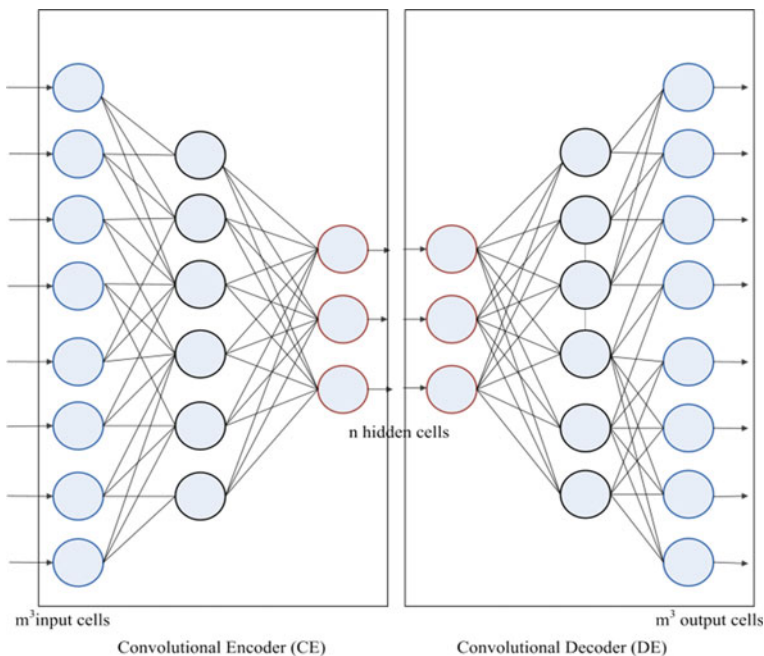
Fig. 1.7 Convolutional E/D (CED)

transformed into eight $(m^3/8)$-dimensional vectors respectively. They enter sequentially the input of CED-1, whose hidden layer contains $(n/2)$ cells, and the output layer—$(m^3/8)$ cells. After the self-training of CED-1 for these sub-tensors is finished, the corresponding 8 sub-tensors of length $(m^3/8)$ are restored, which comprise the approximating difference tensor. In the second summator (Σ_2), the approximated difference tensor is subtracted element-by-element from the first approximation, and in the end the second difference tensor, i.e., the second approximation of the tensor is obtained. In the second hierarchical AIDP level, the difference tensor is divided into 64 sub-tensors, each of size $(m/4) \times (m/4) \times (m/4)$, which are then transformed into the corresponding $(m^3/64)$-dimensional vectors. They enter sequentially the auto-encoder CED-2, whose hidden layer comprises $(n/4)$ cells, and the output layer, $(m^3/64)$ cells.

After the end of the CED-2 self-training, from the so calculated 64 output vectors of length $(m^3/64)$, the corresponding 64 sub-tensors are restored which build the approximated second difference tensor. Each of the restored tensors in the AIDP levels, has a corresponding feedback to the CED-0, CED-1, and CED-2 coders. These connections are used in the self-training process of the auto-encoders. At the output of the AIDP zero level, an n-dimensional vector is obtained, which comprises the cells of the hidden layer of the trained CED-0. At the output of the AIDP first level, the $(4n)$-dimensional vector which is the concatenation of the elements of the 8 vectors is obtained, each of length $(n/2)$, built by the cells of the CED-1 hidden
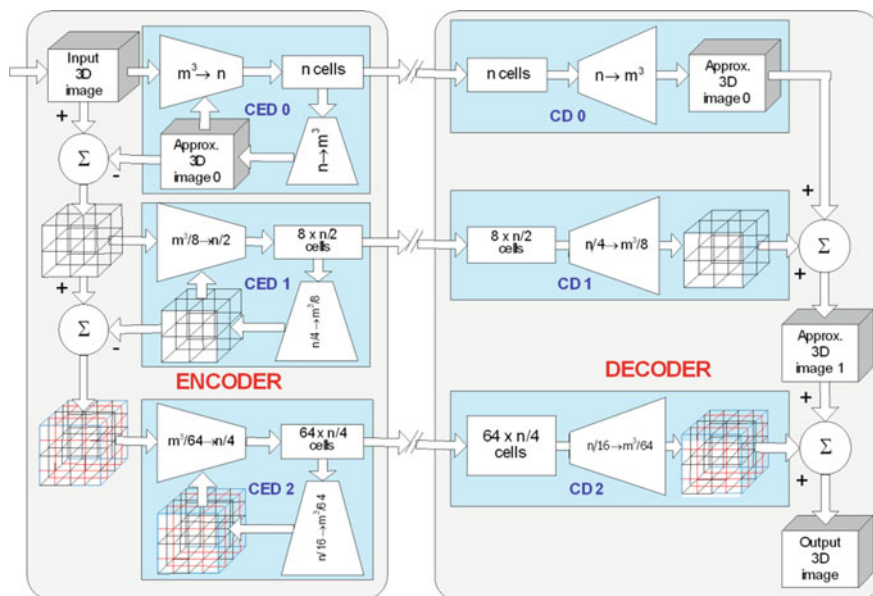
Fig. 1.8 Block diagram of the two-level adaptive IDP with convolutional Encoder/Decoder

layer, for the corresponding input vector. At the output of the AIDP second level, the $(16n)$-dimensional vector is obtained which is the concatenation of the elements of the 64 vectors, (each of length $(n/4)$), built by the cells of the hidden layer of the trained CED-1, for the corresponding input vector. The output n-dimensional vector for the zero AIDP level is the shortest, but it carries the largest information volume for the input tensor of size $m \times m \times m$. The output vectors from the first and second AIDP level, are correspondingly 4 and 16 times longer than these from the zero level, but they carry much less information about the input tensor. This permits significant reduction of the data obtained from the AIDP outputs, without noticeable information loss; i.e., the meaningless information is filtered (neglected). The number of AIDP levels together with the size growth of the input third-order tensor, should be increased.

The *Second approach*, aimed at tensors sequence compression (for example, color RGB images), as illustrated in Fig. 1.9. Here, a tensor decomposition is shown, based on the 2-level 3D Branched IDP (3D BIDP). The levels numbers are $p = 0.1$ [17], in correspondence to the block diagram from Fig. 1.3. Each tensor is of size $M \times N \times 3$; the input sequence is denoted as X_{t-1}, X_t, X_{t+1} for $k = 1$, and in the time moments $t - 1, t, t + 1$, where it contains 3 matrices of size $M \times N$. For the moment, at the input of the trained CED arrives the tensor X, and at the output is obtained the approximated tensor, $\hat{X}$. The approximation accuracy depends both on the CED training, and on the neuron number (n) in the hidden layer. These neurons are the components of the corresponding output n-dimensional vector, s.
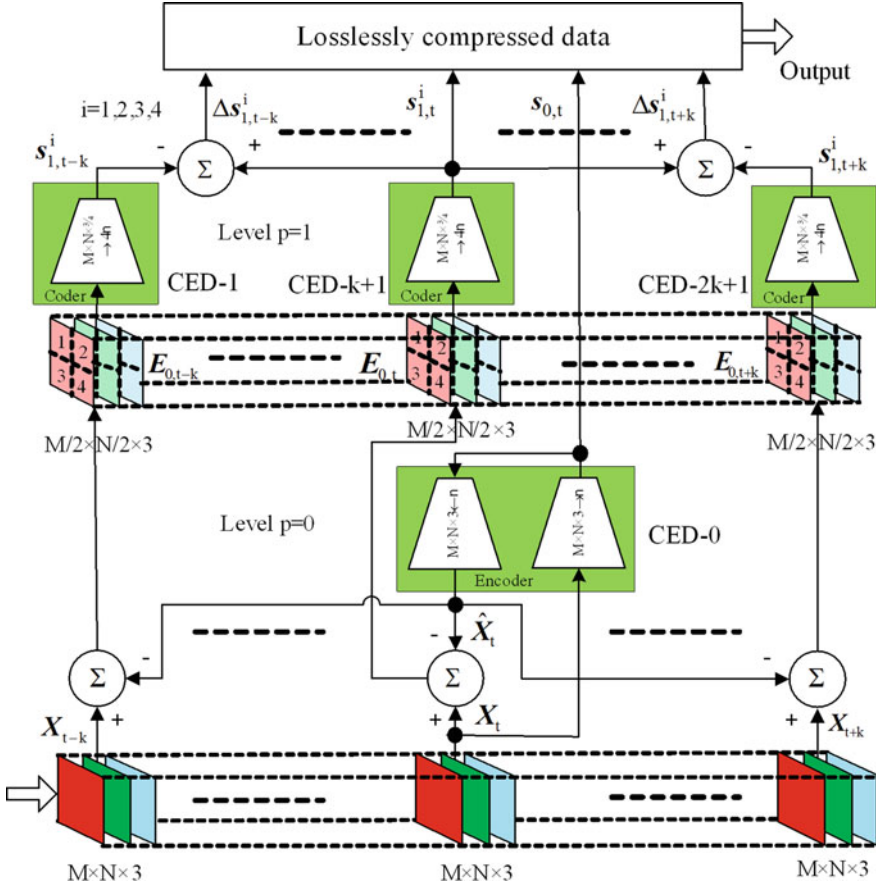
Fig. 1.9 Block diagram of the two-level 3D BIDP based on the convolutional Encoder/Decoder

Due to the result of the 3D BIDP implementation for the tensor sequence X_{t-1}, X_t, X_{t+1}, the output vectors $s_{0,t}$, $\Delta s^i_{1,t-1}$, and $\Delta s^i_{1,t+1}$ for $i = 1, 2, 3, 4$ are obtained of total length $n + 4n + 4n + 4n = 13n$. Due to high correlation existing between the sequential tensors X_{t-1}, X_t, X_{t+1}, the values of the significant part of the components in the difference vectors $\Delta s^i_{1,t+1}$ and $\Delta s^i_{1,t+1}$ are close to zero. In this way, the input tensors are transformed into an output vector of small length which contains many zero values, i.e., the features' space is reduced at minimum computational cost. For the calculations reduction the mutual correlation between the tensors is used, which determines the relation $\hat{X}_t \approx \hat{X}_{t-1} \approx \hat{X}_{t+1}$. In the result, the calculation of tensors $\hat{X}_{t-1}$ and $\hat{X}_{t+1}$ through the corresponding CEDs, is not necessary. In the second level ($p = 1$) of the 3D BIDP, each difference tensor $E_{0,t} = X_t - \hat{X}_t$, $E_{0,t-1} = X_{t-1} - \hat{X}_t$, and $E_{0,t+1} = X_{t+1} - \hat{X}_t$, is divided into four sub-tensors of size $(M/2) \times (N/2) \times 3$, and for each, a corresponding CED is used. The neurons in the hidden layers of all CED in the level $p = 1$ are represented by the vectors $s_{1,t}$, $s_{1,t-1}$, and $s_{1,t+1}$.

Applications

In this part of our work, the experimental results obtained for some of the most important applications are shown: for compression and image content protection, and for efficient object search in large image databases.

1.5.1   Compression of Multidimensional Images

The compression algorithms are developed both for single images, and for groups of correlated images or sequences. The approach is based on the IDP decomposition with Back Propagation Neural Networks (BPNN). In Table 1.1 are shown some comparison results obtained for several widely used test images, when approximately the same quality of the restored images is achieved. The results for the IDP-BPNN decomposition are given for two-level IDP with initial sub-blocks of size 8×8. As it could be noticed, the compression ratio for IDP-BPNN for most of the test images is approximately two times higher, while the quality of the restored images (evaluated by their PSNR) is close or better than this for JPEG 2000.

The use of the BIDP for sequences or groups of correlated images offers similar results, part of which are given in Table 1.2. For the evaluation were used various kinds of medical images. The following abbreviations are used: CT—computer tomography; MG—mammography; NM—nuclear magnetic; US—ultrasound; dcm—Dicom; and jp2—JPEG2000. The "idp" format was specially created for the IDP decomposition [18]. The header contains detailed information about the decomposition structure—the number of levels, the used transform for each level (for example, DCT WHT, etc.), the arrangement of the retained coefficients, and the kind of lossless compression applied on the compressed image data. The lossless compression method is based on adaptive run-length coding, which corresponds to the data statistics. The results show that IDP offers higher compression ratio than Dicom, and is comparable in efficiency with JPEG 2000, but at lower computational cost (the wavelet transform is more complicated than WHT or DCT). Additional advantage of IDP is the sub-block structure, which offers high flexibility in observing medical images, and permits enlargement (on request) of a selected Region Of Interest (ROI). This is an important feature in case of remote diagnostics and medical decision support applications.

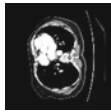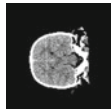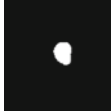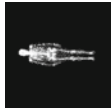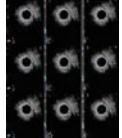### 1.5.2   Content Protection of Visual Information

The IDP-based decomposition permits the insertion of invisible resistant watermark (WM) [19]. The block diagram of a two-level WM insertion algorithm, is shown in Fig. 1.10. In this case, the IDP structure is retained, but the decomposition is based on the Complex Hadamard Transform (CHT). The watermark is prepared in the same

Table 1.1 Comparison results for JPEG2000 and IDP-BPNN

No	Name	8 bpp 512 × 512	JPEG 2 K (Matlab)		IDP-BPNN(8 × 8)	
			CR	PSNR [dB]	CR	PSNR [dB]
1	Boy		25.23	28.98	60.40	29.05
2	Fruits		32.64	32.69	60.29	32.89
3	Tracy		54.13	35.66	59.93	35.32
4	Vase		35.75	27.20	60.18	26.83
5	Clown		31.71	31.47	60.01	31.81
6	Peppers		38.39	30.70	60.23	30.94
7	Text		17.86	18.35	60.22	18.69
8	Lena		32.10	29.20	59.57	29.15

way, as the image (i.e. decomposed into 2 or 3 levels), and in the watermarking process, its coefficients are added to the image decomposition coefficients. To retain the image quality unchanged (i.e., "invisible" watermarking), the WM information is inserted in the phases of selected coefficients, and could be extracted by using a special decoding software. The watermark "depth" depends on the phase rotation

Table 1.2 Compression results for various kinds of medical images

Image size, KB	CTI 512 × 512 14 images	CTI 512 × 512 14 images	CTI 512 × 512 14 images	MGI 1914 × 2294	NMI 1024 × 1024	USI 1020 × 818	USI 1020 × 818
Type: CTI, MGI, NMI, USI							
dcm	545	545	545	4190	2000	1800	1800
jp2	52.2	49.6	3.92	820	103	700	189
idp	50.5	46.9	2.3	801	71	388	134
dcm/jp2	10.44	10.98	139.03	5.11	19.41	3.48	22.16
dcm/idp	10.79	11.62	236.95	5.23	28.17	6.28	31.26

angle, for example, for rotations in the range 0–20°, the PSNR is always higher than 35 dB, i.e., the visual quality of the watermarked test images is retained. In Fig. 1.10, the following notations are used: $w_0(r)$—the decomposed watermark data for level 0; $w_1(r)$—the decomposed watermark data for level 1; $Z_0(r)$—the calculated coefficients for the initial (zero) level, with the inserted watermark; $Z_1(r)$—the calculated coefficients for level 1, with the inserted watermark. For a decomposition of higher number of levels, the structure is retained. The so calculated coefficients are arranged following their spatial frequency, and are losslessly coded. With this, the coding procedure is finished. The decoding is executed in reverse order.

The main advantages of the algorithm [19], are:

– The algorithm is highly resistant against attacks, based on high-frequency filtration (JPEG compression), which is confirmed by the almost constant MSE value for the extracted watermark.
– The algorithm permits insertion of significant amounts of data (the number of inserted bits could be approximately equal to ¾ of the total number of pixels).
– The algorithm is highly resistant against attacks related to image editing of the kind: crop, rotations, etc.
– The algorithm permits to insert different watermarks in each consecutive decomposition level, which is an additional tool to ensure hierarchical access control.
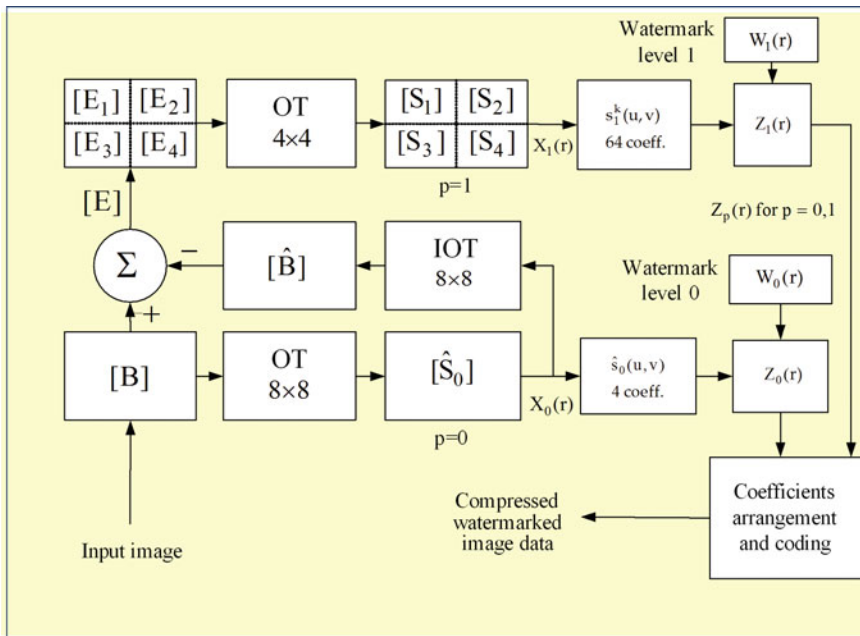

Fig. 1.10 Block diagram for 2-layer WM insertion in one IDP sub-block of size 8 × 8

1.5.3   Object Search in Large Image Databases

Contemporary databases contain huge number of images, video sequences, etc., and sometimes the search needs too much time. The layered structure of the IDP-based decomposition gives significant abilities for the search process enhancement.

For this, the images in the database and the query image are decomposed in a similar way. The retained decomposition coefficients from all decomposition levels, which represent the query image, constitute the "Cognitive 3D IDP model" (Fig. 1.11). The retained coefficients are used for the layer-by-layer comparison and evaluation in the search process.

In case that the search is aimed at a specified group of images, the method permits to select in advance the most suitable group of coefficients (which ensure the highest similarity), so as to enhance the process significantly.

The model is based on the n-level IDP decomposition under Neural Network (NN) control, shown on Fig. 1.12. The accuracy of the 3D model in the IDP level p is defined by the NN in the preceding level $p - 1$, and as a result the minimum
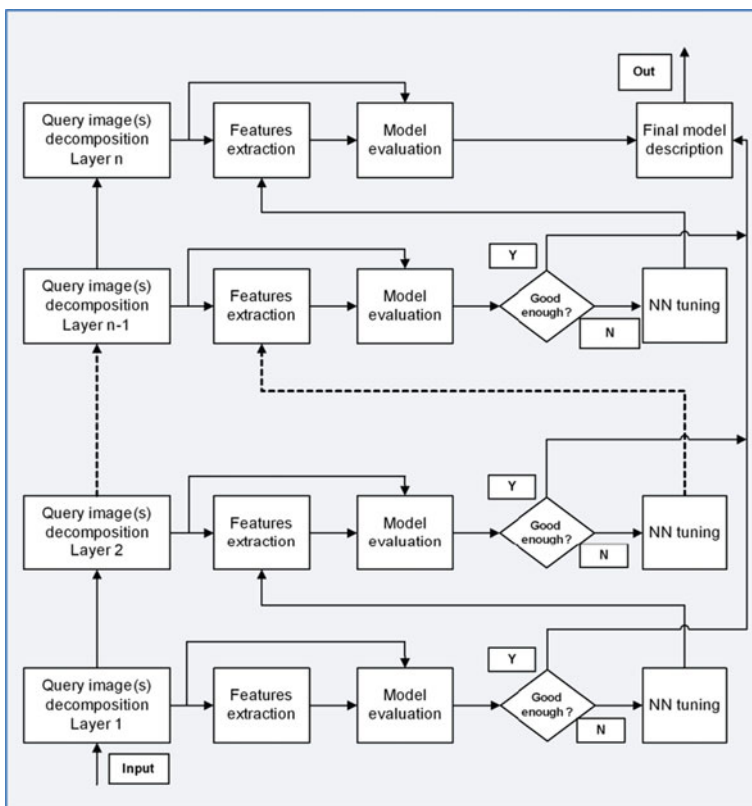

Fig. 1.11 Creation of the cognitive 3D IDP model