ELSEVIER

Computers in Biology and Medicine 37 (2007) 1786–1795

Computers in Biology and Medicine

www.intl.elsevierhealth.com/journals/cobm

Computer-aided grading and quantification of hip osteoarthritis severity employing shape descriptors of radiographic hip joint space

Ioannis Boniatis[a], Dionisis Cavouras[b], Lena Costaridou[a], Ioannis Kalatzis[b], Elias Panagiotopoulos[c], George Panayiotakis[a,*]

[a]*University of Patras, School of Medicine, Department of Medical Physics, 265 00 Patras, Greece*
[b]*Technological Educational Institute of Athens, Department of Medical Instrumentation Technology, 122 10 Athens, Greece*
[c]*University of Patras, School of Medicine, Department of Orthopaedics, 265 00 Patras, Greece*

Received 24 October 2006; received in revised form 17 May 2007; accepted 29 May 2007

Abstract

A computer-based system was designed for the grading and quantification of hip osteoarthritis (OA) severity. Employing an active-contours segmentation model, 64 hip joint space (HJS) images (18 normal, 46 osteoarthritic) were obtained from the digitized radiographs of 32 unilateral and bilateral OA-patients. Shape features, generated from the HJS-images, and a hierarchical decision tree structure was used for the grading of OA. A shape features based regression model quantified the OA-severity. The system accomplished high accuracies in characterizing hips as "Normal" (100%), of "mild/moderate"-OA (93.8%) or "severe"-OA (96.7%). OA-severity values, as expressed by HJS-narrowing, correlated highly ($r = 0.9$, $p < 0.001$) with the values predicted by the regression model. The system may contribute to OA-patient management.

Keywords: Hip; Osteoarthritis; Radiography; Shape; Classification

1. Introduction

The syndrome of osteoarthritis (OA) causes alterations in the tissues of synovial joints, a condition that results in pain, stiffness, deformity as well as limitation of the movement of the affected joint [1].

Plain film radiography is considered as the reference imaging modality for the assessment of the severity of OA [2], despite the fact that Magnetic Resonance Imaging (MRI) is considered as the most promising tool for the investigation of the disease [3]. Radiographic features that signify the presence of OA in a joint include: joint space narrowing (indicating the progressive loss of articular cartilage), subchondral sclerosis, osteophytosis, and cyst formation [4]. For the assessment of the severity of hip OA, several qualitative grading scales have been proposed [5]. According to these scales, a severity grade is, subjectively, assigned to the studied hip joint, while each grade is defined on the basis of joint structural alterations visualized on radiographic images. The Kellgren and Lawrence (KL) grading scale [6] is considered as the golden-standard despite its deficiencies [7]. The specific scoring system evaluates the severity of OA on a five (0–4) grade scale, in which 0 indicates a "normal" joint, while the grades 1–4 are used in order to characterize the osteoarthritic condition as "doubtful", "mild", "moderate", and "severe", respectively [6].

Shape is a visual feature of cardinal importance regarding the description and the recognition of an object within a digital image [8]. In biomedical images, the shape and the size of anatomical structures may provide useful information regarding the physiology or the pathology of the structures [9]. In the field of X-ray musculoskeletal imaging, the shape of lumbar vertebrae has been utilized for the discrimination between normal and defective (having osteophytes) ones by means of a neural network based classification approach [10]. In another study, alterations in the shape of periarticular bone contours of the middle or proximal phalanges, associated to rheumatoid arthritis of the hand, have been assessed by computer-based

* Corresponding author. Tel.: +30 2610 996 113, 999 537; fax: +30 2610 996 113.
E-mail address: panayiot@upatras.gr (G. Panayiotakis).

0010-4825/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiomed.2007.05.005

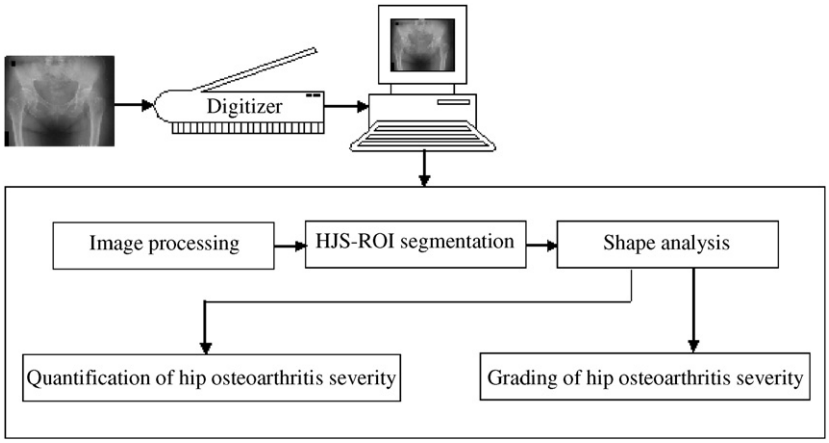

Fig. 1. Diagrammatic representation of the computer-based system for the grading and quantification of hip osteoarthritis severity.

image analysis techniques [11]. Finally, the size of radiographic knee joint space has been used for the automatic discrimination between osteoarthritic and unimpaired knees using a neural network [12].

Referring to hip OA, a characteristic shape alteration associated to the disease is the narrowing of hip joint space (HJS), perceived on radiographic images. The specific radiographic feature has been considered as a defining criterion for epidemiologic studies of the disease [13], whereas the monitoring of HJS-narrowing has been accepted as the most reliable index for the monitoring of the disease progression [4]. The progressive and non-uniform loss of the articular cartilage, indicated by the narrowing of radiographic HJS, is expected to differentiate the morphology of radiographic HJS in osteoarthritic hips. Thus, the utilization of features capable of evaluating the shape and the size of radiographic HJS could establish means for the discrimination between normal and osteoarthritic hips as well as among various grades of hip OA-severity.

Previous studies [13,14] have introduced HJS-width thresholds for characterizing a hip as normal or osteoarthritic. In previous studies performed by our group [15–18], hip joint alterations associated to OA were assessed by means of the radiographic texture of HJS. Additionally, textural information extracted from the region of radiographic HJS has been employed in the design of pattern recognition schemes for the discrimination among OA-severity categories and the quantification of the severity of the disease. In a preliminary study performed by our group, morphological descriptors of radiographic HJS were employed in the design of a probabilistic neural network based classifier, used for the discrimination between: (i) normal and osteoarthritic hips, and (ii) between hips of mild/moderate OA and of severe OA [19].

In the present study, a more robust computer-based system was designed for the grading of hip OA-severity from pelvic radiographs as well as for the quantification of the severity of the disease (see Fig. 1). Within this context: (i) digital image features, capable of evaluating alterations concerning the shape and the size of radiographic HJS due to OA, were generated from the abovementioned anatomical region, (ii) these features were employed in the design of a tree-structure based classification scheme, implemented by suitable combination of pattern recognition algorithms, for characterizing hips either as normal or as of mild/moderate OA or of severe OA, and (iii) a shape features based regression model was introduced for the quantification of the severity of hip OA.

2. Materials and methods

2.1. Patients and radiographs

The sample of the study comprised 64 hips (18 normal, 46 osteoarthritic) corresponding to 32 patients with verified hip OA. Eighteen patients were diagnosed as of unilateral OA, while 14 as of bilateral OA. Diagnosis of hip OA was performed according to the clinical and radiographic American College of Rheumatology criteria [20]. The mean value of patients' ages was 66.7 years (range 49 years–83 years). A pelvic radiograph was available for each patient, while all radiographs were obtained employing a specific radiographic protocol. In particular, each radiograph was taken on an X-ray unit (Siemens, Polydoros 50, Erlangen, Germany) with the following settings: 70–80 kVp, 100 cm focus to film distance, alignment of the X-ray beam 2 cm above the pubic symphysis, use of a fast screen and film cassette 30 cm × 40 cm. Radiographs were digitized at 12 bits (4096 gray levels) with a spatial resolution of 146 ppi (0.17 mm pixel size), using a laser digitizer suitable for medical applications [21]. Digitizer performance was evaluated employing a quality control protocol [22].

The KL grading scale criteria were used for the radiographic assessment of hip OA-severity [6]. In particular, each one of three orthopaedists assigned a KL severity grade to the examined hips, while in order to establish a golden-standard, only those examinations of common consent were retained for the purposes of the present study. Based on the assigned KL grades, the hips of the sample were allocated into one of the following three major OA-severity categories: "normal/doubtfoul" (KL $= 0, 1$), "mild/moderate" (KL $= 2, 3$), and "severe" (KL $= 4$). Within this context, and referring to the unilateral

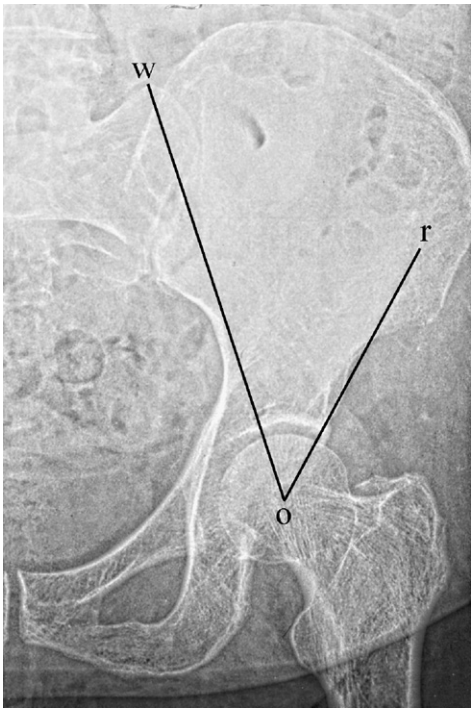

Fig. 2. Determination of the acute angle, defined on patient's standard anatomical landmarks, encompassing the studied hip joint space region of interest. O: centre of femoral head, w: highest point of homolateral sacral wing, r: lateral rim of the acetabulum.

OA-group, 18 hips were assigned to the "normal/doubtful", 9 to the "mild/moderate", and 9 to the "severe" category. The corresponding numbers for the hips of the bilateral group were 0/7/21.

2.2. Active contours segmentation of radiographic HJS

In order to determine the region of radiographic HJS, an active contours (snakes) segmentation model was utilized [23]. In particular, the determination procedure comprised the following steps:

(i) The digitized radiographs were processed by an algorithm implementing the adaptive wavelet transform [24]. This resulted in the contrast enhancement of the radiographic image, while the articular margins of the joint were emphasized.

(ii) On the enhanced radiographs, an acute angle encompassing the weight-bearing portion of the femoral head was formed for each hip (see Fig. 2), employing custom developed software [25,26]. This angle, defined on the basis of patient's anatomical landmarks [27], provided the medial and lateral limits of the studied HJS region of interest (ROI). As it can be observed in Fig. 2, the medial limit of the HJS-ROI was defined by the line joining the centre of femoral head (summit of the angle) and the highest point of the homolateral sacral wing, while the lateral limit was given by the line joining the centre of the femoral head and the lateral rim of the acetabulum.

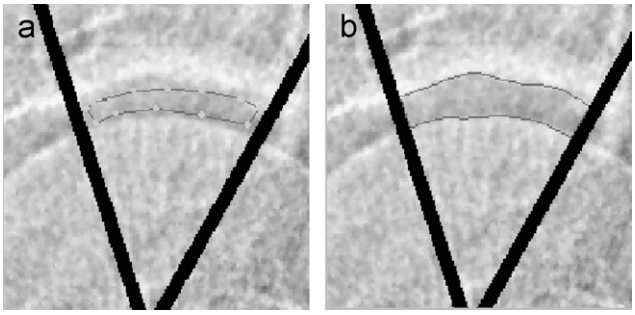

Fig. 3. Active contours based segmentation of the radiographic hip joint space (HJS) region of interest (ROI), corresponding to Fig. 1. The initialization points, were connected by a solid line, which provided an initial rough approximation of the HJS-ROI contour (a). After multiple iterations, the articular margins of radiographic HJS-ROI, indicated by the solid line, were determined by the active contours model (b).

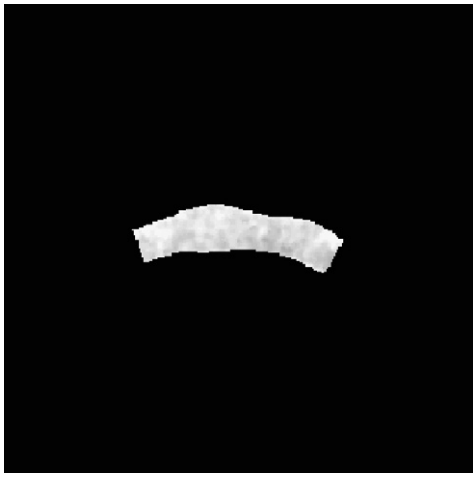

Fig. 4. Segmented hip joint space (HJS) region of interest (ROI), corresponding to Figs. 2, and 3.

(iii) Within this angle, the orthopaedist commenced the segmentation process by providing manually a number of initialization points. As it can be observed in Fig. 3a, these points were placed within the region of radiographic HJS, in the vicinity along the articular margins of the hip joint. The initialization points were connected automatically by a line, which provided an initial rough approximation of the contour of the HJS-ROI (see Fig. 3a). After the specification of the parameters of the snakes model (elasticity, rigidity, number of iterations) the articular margins of the HJS-ROI were determined by the segmentation algorithm (see Fig. 3b). An example of segmented HJS-ROI, corresponding to Figs. 2 and 3 is presented in Fig. 4.

On each enhanced pelvic radiograph, two ROIs were segmented corresponding to patient's both HJSs, according to the previously described procedure. It has to be mentioned that in cases characterized by complete loss of the HJS due to narrowing, the shape analysis could not be performed, so an acceptance threshold was set (number of pixels corresponding to the ROI greater than 200).

2.3. Shape analysis of radiographic HJS

For the needs of the present study, the following features [8,28,29] were generated from the segmented HJS-ROIs (see Fig. 3b), in order to quantify shape and size aspects of radiographic HJS:

- f_1 area, the number of pixels comprising the ROI
- f_2 perimeter, the number of pixels in the boundary of the HJS-ROI
- f_3 minimum upper distance, the minimum distance between the centre of mass (centroid) and the upper boundary (roof of the acetabulum) of the HJS-ROI
- f_4 minimum lower distance, the minimum distance between the centroid and the lower boundary (upper margin of the femoral head) of the HJS-ROI
- f_5 maximum medial distance, the maximum distance between the centroid and the medial boundary of the HJS-ROI
- f_6 maximum lateral distance, the maximum distance between the centroid and the lateral boundary of the HJS-ROI
- f_7 major axis length, the length of the major axis of the ellipse that has the same normalized second order spatial central moments as the HJS-ROI
- f_8 minor axis length, the length of the minor axis of the ellipse that has the same normalized second order spatial central moments as the HJS-ROI
- f_9 eccentricity, the eccentricity of the ellipse that has the same normalized second order spatial central moments as the HJS-ROI. The specific feature is defined as the ratio of the distance between the foci of the ellipse and its major axis length
- f_{10} orientation, the angle (in degrees) between the x-axis and the major axis of the ellipse that has the same normalized second order spatial central moments as the HJS-ROI
- f_{11} convex area, the area of the convex hull corresponding to the HJS-ROI. For an object within a digital image the convex hull is defined as the minimal convex shape that entirely bounds the object
- f_{12} equivalent diameter, the diameter of a circle with the same area as the HJS-ROI
- f_{13} extent, the ratio of the area of the HJS-ROI to the area of the Bounding box corresponding to the HJS-ROI. The Bounding box of an object is defined as the smallest rectangle (oriented along the rows of the image) containing the object.

All the abovementioned areas and distances were expressed in pixels.

2.4. Design of the grading classification system

For the computer-based grading of hip OA-severity, a classification system was implemented as a two-level hierarchical decision tree (see Fig. 5). The first level of the system was designed so as to discriminate between normal and osteoarthritic hips. At the second level, the hips that had correctly been characterized as osteoarthritic at the first level were further classified as of mild/moderate or of severe OA. As it can be observed in Fig. 5, the classification task concerning the first level of the tree structure was performed by an ensemble of 3 individual classifiers. In particular, the k-nearest neighbor (k-NN), the Bayes and the quadratic-least squares minimum distance (Quadratic-LSMD) classifiers were combined according to the majority vote (MV) principle, in order to improve the classification accuracy of the system. According to the MV combination rule, an unknown pattern is assigned to a specific class ω_k if the majority of individual classifiers, implementing the combined classification scheme, assigns the pattern to ω_k [30].

At the second level, the quadratic-LSMD classifier was used for the further discrimination of osteoarthritic hips. At both levels, the classifiers were designed employing the shape features that were generated from the region of radiographic HJS.

2.4.1. k-nearest neighbor classifier

The k-nearest neighbor (k-NN) classifier implements an effective non-parametric classification approach that does not require the prior knowledge of patterns distributions. The specific classification algorithm opines regarding the assignment of an unknown pattern to a class relying on a distance function between patterns. In particular, an unknown pattern is assigned the label of the class in which the most nearest training patterns (neighbors) belong [31]. For the needs of the present study, the Euclidean distance was used as the metric of proximity between patterns [9]. Considering the p-dimensional feature space, the Euclidean distance d_E between the unknown pattern vector $\mathbf{x}$ and a training pattern vector $\mathbf{y}$ is given by

$$d_E = \sqrt{\sum_{i=1}^{p} (x_i - y_i)^2}. \quad (1)$$

2.4.2. Quadratic-least squares minimum distance classifier

The Quadratic-LSMD classifier discriminates among classes by implementing quadratic decision boundaries. The specific classification algorithm is a modified version of the least squares minimum distance (LSMD) classifier, in the sense that employs a quadratic equation within the least squares method. In the quadratic approach, for a pattern vector $\mathbf{x}$ of p-dimensionality, a pattern vector $\mathbf{y}$ of augmented dimensionality q is formed, through a one-to-one transformation. Due to the latter, the quadratic decision functions involving the elements of pattern $\mathbf{x}$, can be written in the form of linear equations employing the elements of the augmented pattern $\mathbf{y}$. Thus, the classifier finally opines similarly to the LSMD classifier. Accordingly, the augmented pattern $\mathbf{y}$ is mapped from the augmented feature space into a decision space wherein the patterns of a class are clustered around a pre-selected point. The transformation that implements the mapping from the feature space to the decision space is chosen so as the overall mean-square mapping error is minimized (least squares). An unknown pattern is assigned to a class if it is closest (minimum distance) to the predefined point of the decision space, corresponding to the class [32].

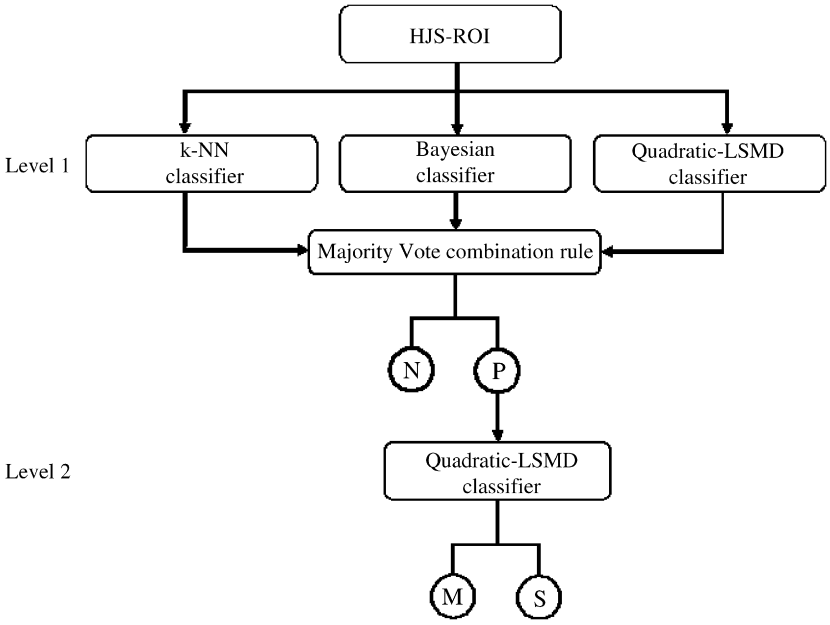

Fig. 5. Two level hierarchical decision tree structure for grading osteoarthritis severity employing shape features of the hip joint space (HJS) region of interest (ROI). Level 1: discrimination between Normal (N) and osteoarthritic (P) hips by the Majority Vote based combined classification scheme. Level 2: discrimination between hips of mild/moderate (M) osteoarthritis and of severe (S) osteoarthritis employing the quadratic-least squares minimum distance (LSMD) classifier.

2.4.3. Bayes classifier

The Bayes classifier implements a statistical pattern recognition approach, based on Bayes formula from probability theory. The classifier is designed so as to minimize the average risk (probability of misclassification) associated to a discrimination task. For a classification problem concerning the assignment of an unknown pattern $\mathbf{x}$ into one of N classes $\omega_1, \omega_2, \ldots, \omega_N$, the Bayes classifier decides by utilizing functions of the form

$$d_k(\mathbf{x}) = p(\mathbf{x}|\omega_k) \cdot P(\omega_k), \quad k = 1, 2, \ldots, N. \tag{2}$$

In Eq. (2), the term $p(\mathbf{x}|\omega_k)$ stands for the class-conditional probability density function, which describes the distribution of pattern vectors within the class ω_k. On the other hand, $P(\omega_k)$ represents the *a priori probability* concerning the occurrence of class ω_k. The pattern $\mathbf{x}$ is assigned to the class ω_k associated with the highest value of the decision function $d_k(\mathbf{x})$. Assuming: (i) probability density functions $p(\mathbf{x}|\omega_k)$ of Gaussian form, and (ii) equal probabilities $P(\omega_k)$, the decision function of the classifier can be written in the form of Eq. (3):

$$d_k(\mathbf{x}) = \ln P(\omega_k) - \tfrac{1}{2} \cdot \ln |\mathbf{C}_k| - \tfrac{1}{2} \times [(\mathbf{x} - \mathbf{m}_k)^{\mathrm{T}} \cdot \mathbf{C}_k^{-1} \cdot (\mathbf{x} - \mathbf{m}_k)], \quad k = 1, 2, \ldots N, \tag{3}$$

where $\mathbf{C}_k$ and $\mathbf{m}_k$ represent the covariance matrix and the mean vector of class ω_k while T indicates transposition [9].

2.4.4. Feature selection and evaluation of system performance

In order to determine the feature combination providing the highest classification accuracy with the minimum number of features ("optimum" or "best" feature combination) the exhaustive search procedure was followed in conjunction with the leave one out (LOO) performance evaluation method [9]. In particular, the generated features were exhaustively combined with each other (i.e. combinations of two, three, etc. features) in order to form a pattern vector. For every feature combination, each one of the classifiers was designed employing all the patterns of the sample, but one ("leave one out"). This pattern was considered as an unknown one and was used in order to determine the committed classification error. By this way, the performance of the classifier was evaluated employing patterns that were not used for its design. The classification performance was expressed in terms of overall accuracy, specificity and sensitivity [9]. The above described procedures were followed at each level of the decision tree-structure.

In order to safeguard against variations in the dynamic range of the generated features, a fact that could result in inaccurate classification scores, the features were normalized to zero mean and unit standard deviation [9], according to

$$f_{i\_\mathrm{norm}} = \frac{f_i - \mu}{\sigma}, \tag{4}$$

where f_{i_norm} is the normalized value of the f_i shape feature, while μ and σ are the mean value and standard deviation, respectively, of feature f_i over all HJS-ROIs.

2.5. Quantification of hip osteoarthritis severity

In an effort to provide a quantitative estimation of the severity of the disease, a regression model was introduced for the specific purpose. In particular, for each patient of the unilateral

OA group:

(a) OA_severity was expressed as the difference between the area of radiographic HJS of the osteoarthritic (HJS_Area_{path}) and contralateral normal (HJS_Area_{norm}) hip, according to Eq. (5) [15]:

$$OA\_severity(\%) = \frac{HJS\_Area_{norm} - HJS\_Area_{path}}{HJS\_Area_{norm}} \times 100. \quad (5)$$

In this context, OA_severity takes values between 0 and 100, expressing the percentage reduction of the HJS-area due to OA. Considering that: (i) the monitoring of HJS-narrowing is regarded as the most reliable index for monitoring OA-progression [4], (ii) the unaffected hip in patients with unilateral OA can be used as reference for normal HJS-area [27], and (iii) in normal individuals, right and left hips do not differ significantly [33], the HJS-area difference may be employed for evaluating the cartilage loss due to OA, and thus for assessing OA-severity.

(b) Utilizing multiple linear regression analysis [34], a multivariate model of shape features was generated for the quantitative assessment of OA using the "Matlab Statistics Toolbox" (The MathWorks, Inc., Natick, USA). Following multiple trials for different combinations of the generated features, the Equivalent Diameter and Maximum Medial Distance of the pathological HJS-ROIs were found to give best fit and they were employed as the independent (predictor) variables of the model, while the OA_severity was used as the dependent variable. The regression equation describing the model was of the form

$$\begin{aligned} OA\_severity(\%) = & -1.9 \cdot \text{Equivalent Diameter} \\ & + 0.3 \cdot \text{Maximum Medial Distance} \\ & + 119.6. \end{aligned} \quad (6)$$

In order to test the quality of the regression OA-severity model a "Leave-6-Out" model was used. In particular, data for six patients were reserved for verification purposes and evaluated by the model developed using the remaining 12 patients.

2.6. Statistical analysis

The student's t-test was used in order to investigate the existence of statistically significant differences ($p < 0.05$) between normal and osteoarthritic hips for the shape feature values. Multiple linear regression analysis [34] was utilized for the generation of the multivariate model concerning the quantitative assessment of OA. The existence of statistically significant differences between the OA_severity values predicted by the regression Eq. (6) and the "Leave-6-Out" model was investigated by means of the student's paired t-test. Pearson's correlation coefficient was used in order to assess the correlation between the measured and the predicted values. The student's t-test was also employed so that to verify whether Eq. (6) predicted OA_severity values for hips of mild/moderate and of severe OA differed significantly. The coefficient of variation (CV) [35] was used in order to assess the reproducibility of the HJS-ROI determination process. In particular, each of the orthopaedists segmented the HJS-ROIs twice, according to the previously described procedure. A time interval of about a month was intervened between the two determinations, while the evaluation scores were employed for the calculation of the CV. High degree of reproducibility is indicated by low values of the CV coefficient, and vice versa. Student's paired t-test was used in order to investigate whether shape features extracted from the two determinations differed significantly ($p < 0.05$). All statistical processing was performed utilizing the "Matlab Statistics Toolbox".

3. Results and discussion

This study proposes a computer-based classification system for the grading and quantification of hip OA-severity from radiographic images. The classification system, which is implemented in the form of a two-level hierarchical decision tree structure, utilizes shape properties of the radiographic HJS in order to characterize a hip as normal, of mild/moderate OA or of severe OA.

The degenerative process of OA causes the progressive and non-uniform loss of articular cartilage. In a radiographic image this loss is indicated by the narrowing of HJS, which induces alterations in the morphology of the specific anatomical region. Thus, the shape and the size of radiographic HJS in osteoarthritic hips, are expected to differ in comparison to normal ones. This differentiation was verified by the results of statistical analysis of the present study, which revealed the existence of statistically significant differences ($p < 0.001$) between normal and osteoarthritic hips for the shape feature values. So far, the well-known HJS-width and/or HJS-area parameters have been employed in previous studies for the assessment of differentiations concerning the size of HJS in the osteoarthritic condition [13,14,27,36]. In the present study, the alterations of the radiographic HJS associated to OA were evaluated by means of shape descriptors, which provide quantitative information concerning not only the size, but also the shape of the specific anatomical region. The segmentation of radiographic HJS-ROI was found to be reproducible. Regarding the intra-observer reproducibility, the CV was found equal to 2.3%, on average, indicating the reliability of the segmentation process. Inter-observer reproducibility was also high, since the corresponding value for the CV was 2.9%, on average. In addition, the feature values that were generated from the twice-determined HJS-ROIs were found not to differ significantly ($p > 0.05$).

At the first level of the hierarchical decision tree structure, the overall classification accuracy achieved by the system was 98.4%, since 63 out of 64 hips were correctly classified. As it can be observed from Table 1, all the normal hips were properly characterized, resulting in the highest possible specificity accuracy (100%). On the other hand, only one osteoarthritic hip was misclassified, corresponding to a sensitivity accuracy of 97.8%. It has to be mentioned that the specific hip was of "mild/moderate" OA, a fact that may justify the specific classification error.

The abovementioned classification scores were derived from the suitable combination of the k-NN, Bayes, and quadratic-LSMD classification algorithms. In particular, the outputs cor-

Table 1
Truth table demonstrating classification results regarding the discrimination between normal and osteoarthritic hips at the first level of the hierarchical decision tree

Hip characterization	Normal	Osteoarthritic	Accuracy (%)
Normal	18	0	100
Osteoarthritic	1	45	97.8
Overall accuracy (%)			98.4

Table 2
Classification performances of the individual classifiers implementing the combined classification scheme at the first level of the hierarchical decision tree

Classifier	Overall accuracy (%)	Specificity (%)	Sensitivity (%)
k-NN	95.3	88.9	97.8
Bayes	93.8	100	91.3
Quadratic-LSMD	93.8	88.9	95.6

k-NN is the k-nearest neighbour, LSMD the least squares minimum distance.

responding to the best classification performance of each of the individual classifiers were combined according to the "majority vote" principle [30]. Table 2 tabulates the best classification scores for each of the individual classifiers. As it can be observed, the k-NN algorithm accomplished the highest accuracy in discriminating between normal and osteoarthritic hips. In particular, 61 out of 64 hips (95.3% overall accuracy) were assigned to the correct categories, for the optimum feature combination comprising the features [Area, Perimeter, Maximum Lateral Distance]. Sixteen out of 18 normal hips were properly classified (88.9% specificity accuracy), while only one osteoarthritic hip was incorrectly assigned to the normal category (97.8% sensitivity accuracy). After multiple trials, the number of neighbours in the k-NN was determined to be equal to 3 ($k = 3$). The Bayes classifier discriminated successfully 60 out of 64 hips, accomplishing an overall accuracy of 93.8%. The specificity accuracy was 100%, since all the normal hips were correctly classified, while the sensitivity accuracy was 91.3% (proper classification for 42 out of 46 osteoarthritic hips). The optimum feature combination provided the aforementioned scores comprised the features [Area, Major Axis Length, Extent, Perimeter]. Finally, the quadratic-LSMD classifier achieved the same overall accuracy as the Bayesian algorithm (93.8%) for the three-dimensional optimum feature combination [Eccentricity, Equivalent Diameter, Minimum Upper Distance]. The specificity accuracy accomplished was 88.9%, since 16 out of 18 normal hips were properly classified. Referring to the osteoarthritic hips, all but 2 were correctly characterized, resulting in a sensitivity accuracy of 95.6%. As it can be concluded, the combined classification scheme improved the overall classification accuracy regarding the discrimination between normal and osteoarthritic hips. This can be justified by taking into consideration the fact that each of the individual classifiers, implementing the ensemble structure, commits classification errors in different regions (sub-spaces) of the entire

Table 3
Truth table demonstrating classification results regarding the discrimination between hips of mild/moderate osteoarthritis and of severe osteoarthritis at the second level of the hierarchical decision tree

Osteoarthritis severity category	Mild/moderate	Severe	Accuracy (%)
Mild/moderate	15	0	100
Severe	1	29	96.7
Overall accuracy (%)			97.8

Table 4
Truth table demonstrating classification results for the hierarchical decision tree structure

Osteoarthritis severity category	Normal	Mild/moderate	Severe	Accuracy (%)
Normal	18	0	0	100
Mild/moderate	1	15	0	93.8
Severe	0	1	29	96.7
Overall accuracy				96.9

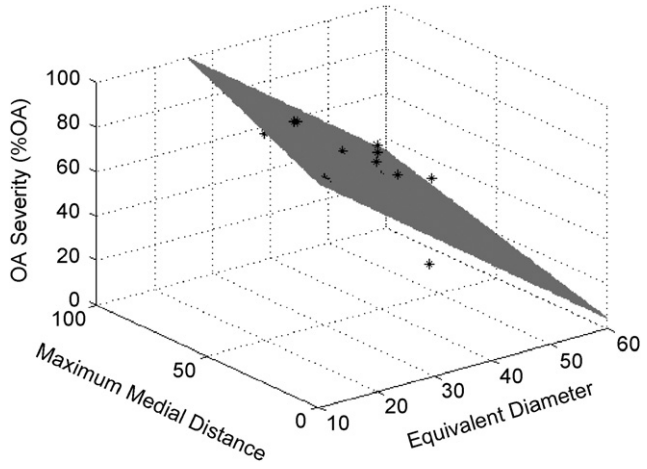

Fig. 6. Three-dimensional scatter diagram and regression surface of the multiple regression analysis model for quantitative estimation of osteoarthritis (OA) severity.

feature space. This complementary information was utilized by the MV-based ensemble, which accomplished a superior classification performance for the specific discrimination task, in comparison to the individual classifiers.

At the second level of the hierarchical tree, the hips that had been correctly classified as osteoarthritic at the first level, were further characterized as of mild/moderate or of severe OA. The quadratic-LSMD classifier performed the specific discrimination task accomplishing an overall classification accuracy of 97.8%. As it can be observed from Table 3, all the hips of mild/moderate OA were assigned to the correct category, resulting in the highest possible accuracy (100%). Referring to the hips of Severe OA, only one hip was misclassified providing an accuracy of 96.7%. The optimum feature combination employed for the specific classification task comprised 5 features: [Area, Convex Area, Extent, Perimeter, Maximum Lateral Distance].

Table 5
Measured and predicted values of OA_severity

Patient	Measured OA_severity value (%)	Shape feature equivalent diameter[a]	Shape feature maximum medial distance[a]	Predicted OA_severity value[b] (%)	Predicted OA_severity value by "Leave-6-Out" model (%)
1	84.2	24.7	62.7	91.5	97.7
2	92.7	16.1	13.2	93.0	92.8
3	92.0	15.6	11.2	93.3	92.9
4	84.8	25.5	63.6	90.2	96.4
5	82.3	31.0	83.6	85.8	93.7
6	67.3	41.3	54.9	57.6	55.7
7	82.2	31.8	67.7	79.5	76.4
8	73.1	29.7	47.4	77.4	74.3
9	86.8	29.5	60.3	81.6	78.5
10	51.8	44.2	61.1	54.0	50.9
11	66.0	38.5	63.6	65.5	62.4
12	77.9	33.3	34.5	66.7	63.6
13	53.3	40.7	58.7	59.9	63.0
14	21.6	59.0	73.2	29.5	38.8
15	60.1	46.8	69.1	51.4	55.5
16	58.9	42.6	49.6	53.5	59.4
17	18.7	46.2	44.5	45.2	53.9
18	46.1	52.3	59.2	38.0	46.8

OA is osteoarthritis.
[a]Values in pixels.
[b]Equation of multiple linear regression model: OA_severity (%) $= -1.9 \cdot$ Equivalent Diameter $+ 0.3 \cdot$ Maximum Medial Distance $+ 119.6$.

Summarizing, the overall classification accuracies accomplished by the hierarchical tree structure for normal hips, hips of "mild/moderate", and of "severe" OA were 100%, 93.8%, and 96.7%, respectively (see Table 4).

Besides the grading of hip OA, the shape features were also employed in the design of a multivariate regression model for the quantification of the severity of the disease. The latter was expressed as the percentage of HJS-area difference between the osteoarthritic and contralateral-normal HJS of patients with unilateral OA (see Eq. (5)). According to the results of multiple regression analysis, the model employed as independent variables the shape features Equivalent Diameter and Maximum Medial Distance (see Eq. (6)). The graphical representation of the relationship between the variables involved in Eq. (6) is shown in Fig. 6. The measured OA_severity values, expressed by Eq. (5) and the corresponding ones predicted by the model (see Eq. (6)) are presented in Table 5. Statistical analysis revealed that the measured and the predicted values did not differ significantly ($p > 0.05$), while the correlation between them was strong and significant ($r = 0.9$, $p < 0.001$). In order to assess the reliability of the quantitative estimation, a "Leave-6-Out" model was used, which predicted the OA_severity values displayed in the last column of Table 5. According to the results of statistical analysis, the OA_severity values, predicted by the Leave-6-Out model, and the corresponding ones, predicted by the model of Eq. (6), did not differ significantly, a finding that may imply the reliability of the suggested approach and its potential use in a clinical environment.

Since the proposed regression model involves shape features generated from pathological HJS, the osteoarthritic hips of the sample were used in order to obtain average OA_severity values for the "mild/moderate" and "severe" categories. Regarding the mild/moderate category, the mean value (±SD) of OA_severity was 45.7% (±15.4%), while the corresponding numbers for the "severe" class were 68.5% (±20.9%). Statistical analysis demonstrated the existence of significant differences ($p < 0.001$) between the two osteoarthritic categories for the predicted OA_severity values. As it can be observed, the hips of severe OA provided a significantly higher mean value of OA_severity in comparison to the hips of mild/moderate category. This finding complies with clinical data concerning the progressive deterioration of the joint structure due to the degenerative action of the disease [4–6]. Finally, and taking into consideration that the quantification of the severity of the disease is based on shape features generated from osteoarthritic hips, the proposed regression model could be of value to orthopaedists in monitoring OA-progression on the same individual, facilitating the management of osteoarthritic patients.

4. Conclusion

In conclusion, structural alteration of the hip joint, associated to OA, can be assessed reliably by computational shape descriptors of radiographic HJS. The proposed two-level hierarchical decision tree discriminated efficiently normal from osteoarthritis hips, while graded reliably the severity of OA in osteoarthritic hips. In addition, the cartilage degeneration due to OA and the induced size and shape alterations of radiographic HJS were utilized for the design of a regression model, which quantified reliably the OA-severity in osteoarthritic hips. The proposed system could positively contribute to the management

of osteoarthritic patients. In particular, the suggested computer-based approach could enhance the reliability of OA-severity assessment, safeguarding against subjective judgement variations associated to the use of qualitative scales (objectification of severity scales in OA). Furthermore, the quantitative estimation of OA-severity provided by the system could facilitate the comparison of OA-progression in follow-up studies, i.e. therapeutic response to a potential intervention aiming to the retardation of the disease progression. Finally, the system could be used as a diagnosis decision support tool to un-experienced physicians and/or as a 'second opinion' tool to experienced ones.

5. Summary

The radiographic assessment of hip osteoarthritis (OA) severity mainly relies on the employment of qualitative scales. Among them, the Kellgren and Lawrence (KL) grading scale is considered as the golden-standard for epidemiological studies of the disease. Within the context of the qualitative assessment of OA, a severity grade is, subjectively, assigned to the studied hip joint, while the severity grades are defined on the basis of joint structural alterations visualized on radiographic images. Among the characteristic radiographic features of hip OA, the narrowing of radiographic hip joint space (HJS) has been accepted as a defining criterion of the disease, while it is considered as the most reliable index of the disease progression.

In the present study, a computer-based system was designed for the grading and quantification of hip OA-severity, employing computational shape descriptors of radiographic HJS. The sample of the study comprised 64 hips (18 normal–46 osteoarthritic), corresponding to 32 patients with verified unilateral or bilateral OA. For each of the patients a pelvic radiograph was available. Hips were allocated by orthopaedists into 3 OA-severity categories, formed in accordance with the KL scale: "normal-doubtful", "mild-moderate", and "severe". Patients' pelvic radiographs were digitized, while their radiographic contrast was enhanced by the adaptive wavelet transform. On each enhanced radiograph the articular margins of the hip joint were determined by an active contours (snakes) segmentation model. Accordingly, 64 regions of interest (ROIs), corresponding to patients' both HJSs were segmented, while a number of shape features were generated from the HJS-ROIs. These features were employed in the implementation of a two-level hierarchical decision tree, designed so as to assign a hip into one of the 3 OA-severity categories. Additionally, a shape features based regression model was introduced for the quantification of hip OA-severity. The latter was expressed as the percentage difference between the area of radiographic HJS of the osteoarthritic and contralateral-normal hip of unilateral osteoarthritic patients.

Statistical analysis revealed the existence of statistically significant differences ($p < 0.001$) between normal and osteoarthritic hips for the generated features. The latter were sensitive enough to capture HJS shape alterations, and as a result, the classification scheme accomplished high accuracies in characterizing hips as "normal" (100%), of "mild/moderate"-OA (93.8%) or "severe"-OA (96.7%). The OA-severity values, as expressed by HJS-narrowing, correlated highly ($r = 0.9$, $p < 0.001$) with the predicted values by the regression model. In addition, OA-severity values, predicted by the model, for patients of severe OA were significantly higher in comparison to those predicted for patients of mild/moderate OA.

Taking into consideration the results of the present study, the proposed system may contribute to the management of osteoarthritic patients by objectifying the severity scales in OA, enhancing the comparability of grading of OA in progression, and acting as a diagnosis decision support tool to un-experienced physicians and/or as a 'second opinion tool' to experienced ones.

Acknowledgments

The first author was supported by a grant by the State Scholarship Foundation (SSF), Greece. The authors thank the staff of the Departments of Orthopaedics and Radiology for their contribution to this work.

References

[1] R.A. Stockwell, Cartilage failure in osteoarthritis: relevance of normal structure and function, A review, Clin. Anat. 4 (1990) 161–191.
[2] P. Garnero, P.D. Delmas, Biomarkers in osteoarthritis, Curr. Opin. Rheumatol. 15 (2003) 641–646.
[3] C.G. Peterfy, Imaging of the disease process, Curr. Opin. Rheumatol. 14 (2002) 590–596.
[4] R.D. Altman, J.F. Fries, D.A. Bloch, J. Carstens, T. Derek Cooke, H. Genant, P. Gofton, H. Groth, D.J. McShane, W.A. Murphy, J.T. Sharp, P. Spitz, C.A. Williams, F. Wolfe, Radiographic assessment of progression in osteoarthritis, Arthritis Rheum. 30 (1987) 1214–1225.
[5] Y. Sun, K.P. Günther, H. Brenner, Reliability of radiographic grading of osteoarthritis of the hip and knee, Scand. J. Rheumatol. 26 (1997) 155–165.
[6] J.H. Kellgren, J.S. Lawrence, Radiological assessment of osteoarthrosis, Ann. Rheum. Dis. 16 (1957) 494–501.
[7] T.D. Spector, C. Cooper, Radiographic assessment of osteoarthritis in population studies: whither Kellgren and Lawrence?, Osteoarthritis Cartilage 1 (1993) 203–206.
[8] W.K. Pratt, Digital image processing, third ed., Wiley, New York, Chichester, Weinheim, 2001 pp. 588–612.
[9] S. Theodoridis, K. Koutroumbas, Pattern Recognition, second ed., Elsevier Academic Press, Amsterdam, 2003 pp. 269–319, 13–54, 385–395, 163–205.
[10] M. Cherukuri, R.J. Stanley, R. Long, S. Antani, G. Thoma, Anterior osteophyte discrimination in lumbar vertebrae using size-invariant features, Comput. Med. Imag. Grap. 28 (2004) 99–108.
[11] M.A. Browne, P.A. Gaydecki, R.F. Gough, D.M. Grennan, S.I. Khalil, H. Mamtora, Radiographic image analysis in the study of bone morphology, Clin. Phys. Physiol. Meas. 8 (1987) 105–121.
[12] T.L. Mengko, R.G. Wachjudi, A.B. Suksmono, D. Danudirdjo, Automated detection of unimpaired joint space for knee osteoarthritis assessment, in: Proceedings of 7th International Workshop on Enterprise Networking and Computing in Healthcare Industry, HEALTHCOM 2005, 23–25 June 2005, pp. 400–403.
[13] P. Croft, C. Cooper, C. Wickham, D. Coggon, Defining osteoarthritis of the hip for epidemiologic studies, Am. J. Epidemiol. 132 (1990) 514–522.
[14] T. Ingvarsson, G. Hägglund, H. Lindberg, L.S. Lohmander, Assessment of primary osteoarthritis: comparison of radiographic methods using colon radiographs, Ann. Rheum. Dis. 59 (2000) 650–653.

[15] I. Boniatis, L. Costaridou, D. Cavouras, E. Panagiotopoulos, G. Panayiotakis, Quantitative assessment of hip osteoarthritis based on image texture analysis, Br. J. Radiol. 79 (2006) 232–238.
[16] I. Boniatis, L. Costaridou, D. Cavouras, I. Kalatzis, E. Panagiotopoulos, G. Panayiotakis, Assessing hip osteoarthritis severity utilizing a Probabilistic Neural Network based classification scheme, Med. Eng. Phys., 29 (2007) 227–237.
[17] I. Boniatis, L. Costaridou, D. Cavouras, I. Kalatzis, E. Panagiotopoulos, G. Panayiotakis, Osteoarthritis severity of the hip by computer-aided grading of radiographic images, Med. Biol. Eng. Comput., 44 (2006) 793–803.
[18] I. Boniatis, L. Costaridou, D. Cavouras, E. Panagiotopoulos, G. Panayiotakis, A computer-based image analysis method for assessing the severity of hip joint osteoarthritis, Nucl. Instrum. Meth. A, 569 (2006) 610–613.
[19] I. Boniatis, L. Costaridou, D. Cavouras, E. Panagiotopoulos, G. Panayiotakis, A morphological descriptors based pattern recognition system for the characterization of hip osteoarthritis severity from x-ray images, in: Book of Abstracts of IMAGING 2006, 27–30 June 2006, Stockholm, Sweden, p. 68.
[20] R. Altman, G. Alarcón, D. Appelrouth, et al., The American college of rheumatology criteria for the classification and reporting of osteoarthritis of the hip, Arthritis Rheum. 34 (1991) 505–514.
[21] Lumiscan 75, system specifications. Lumisys Inc. 1998, ⟨http://www.lumisys.com/support/manuals.html⟩.
[22] E.P. Efstathopoulos, L. Costaridou, O. Kocsis, G. Panayiotakis, A protocol-based evaluation of medical image digitizers, Br. J. Radiol. 74 (2001) 841–846.
[23] C. Xu, J.L. Prince, Snakes, Shapes, and Gradient Vector Flow, IEEE Trans. Image Process. 7 (1998) 359–369.
[24] P. Sakellaropoulos, L. Costaridou, G. Panayiotakis, A wavelet based spatially adaptive method for mammographic contrast enhancement, Phys. Med. Biol. 48 (2003) 787–803.
[25] P. Sakellaropoulos, L. Costaridou, G. Panayiotakis, An image visualisation tool in mammography, Med. Inform. Internet Med. 24 (1999) 53–73.
[26] P. Sakellaropoulos, L. Costaridou, G. Panayiotakis, Using component technologies for web-based wavelet enhanced mammographic image visualization, Med. Inform. Internet Med. 25 (2000) 171–181.
[27] T. Conrozier, A.M. Tron, J.C. Balblanc, et al., Measurement of the hip joint space using computerized image analysis, Rev. Rhum. Engl. Ed. 60 (1993) 105–111.
[28] L.G. Shapiro, G.C. Stockman, Computer vision, Prentice-Hall, Upper Saddle River, New Jersey, 2001 pp. 51–91.
[29] J.C. Russ, The image processing handbook, third ed., Co-published by CRC Press LLC, Boca Raton, Florida and Springer, Heidelberg, 1999 pp. 509–574.
[30] L.I. Kuncheva, Combining Pattern Classifiers: Methods and Algorithms, Wiley, Hoboken, New Jersey, 2004 pp. 111–149.
[31] B. Sankur, Y.P. Kahya, E.Ç. Güler, T. Engin, Comparison of AR-based algorithms for respiratory sounds classification, Comput. Biol. Med. 24 (1994) 67–76.
[32] N. Ahmed, K.R. Rao, Orthogonal Transforms for Digital Signal Processing, Springer, Berlin, Heidelberg, New York, 1975 pp. 225–258.
[33] B. Goker, A. Sancak, M. Arac, S. Shott, J.A. Block, The radiographic joint space width in clinically normal hips: effects of age, gender and physical parameters, Osteoarthritis Cartilage 11 (2003) 328–334.
[34] M.J. Campbell, D. Machin, Medical statistics, second ed., Wiley, Chichester, 1996 pp. 87–104.
[35] G. van Belle, L.D. Fisher, P.J. Heagerty, T. Lumley, Biostatistics. A methodology for the health sciences, second ed., Wiley-Interscience, New Jersey, 2004, pp. 25–60
[36] C.L. Gordon, C. Wu, C.G. Peterfy, J. Duryea, C. Klifa, H.K. Genant, Automated measurement of radiographic hip joint space width, Med. Phys. 28 (2001) 267–277.

Ioannis Boniatis was born in Rhodes, Greece, in 1977. He received his B.Sc. (1999) in Medical Instrumentation Technology from the Technological Educational Institute of Athens, Greece and his M.Sc. (2004) in Electronics and Information Processing from the University of Patras, Greece. He is currently a Ph.D. candidate in the Department of Medical Physics, University of Patras. His research interests include medical image processing and analysis and pattern recognition.

Dionisis Cavouras was born in Kalamata, Greece, in 1951. He received his B.Sc. (1974) in Electronic Engineering, and the M.Sc. (1976) and Ph.D. (1981) in Systems Engineering from the City University, London, UK. He was a research assistant at the Department of Nuclear Medicine, Guy's Hospital, London, between 1976 and 1981, he worked as a Research Fellow at the Department of Computed Tomography, Hellenic Air-force Hospital, Athens, Greece, between 1984 and 1991 and since then he is a professor of medical imaging processing at the Department of Medical Instrumentation Technology in Technological Educational Institute of Athens, and Director of the Laboratory of Medical Image and Signal Processing. His research interests include medical image processing, image analysis, pattern recognition, medical statistics, and medical physics. He has published numerous technical and medical papers as well as conference proceedings.

Lena Costaridou received a B.Sc. in Physics from the Department of Physics of the University of Patras, Greece, a M.Sc. degree in Medical Engineering from the Department of Electrical Engineering and Applied Sciences of the George Washington University, Washington, DC, and in 1997 a Ph.D. degree in Medical Physics from the University of Patras, Greece. She is an assistant professor in the Department of Medical Physics, School of Medicine, University of Patras. Her research interests include medical image processing and analysis applications, especially mammographic image analysis, and evaluation of medical imaging systems and techniques. She is the author or co-author of 40 articles in international peer-reviewed journals and more than 70 international conference papers, as well as the editor of a book on Medical Image Analysis Methods.

Ioannis Kalatzis completed his B.Sc. in Physics from the University of Athens, Greece in 1987 and his Ph.D. in Medical Physics from the University of Athens in 2000. He is currently a research fellow at the Laboratory of Medical Image and Signal Processing in the Department of Medical Instrumentation Technology of the Technological Educational Institute of Athens. His research interests include medical image and signal processing and analysis, with focus in the field of pattern recognition.

Elias Panagiotopoulos was born in Gortynia, Greece, in 1953. He received his M.D. (1978) and his Ph.D. (1988) from the Medical School, University of Athens, Greece. He is a professor of Orthopaedics in the School of Medicine, University of Patras, Greece. His research interests include degenerative alterations of the joints, musculoskeletal imaging and bioengineering. He has published numerous medical papers as well as conference proceedings.

George Panayiotakis was born in Chania, Greece, in 1956. He received his B.Sc. in Physics (1979) and his Ph.D. in Medical Physics (1986) from the University of Patras, Greece. He is a professor in the Department of Medical Physics, School of Medicine, University of Patras, Greece and head of the Medical Radiation Physics Unit at the University Hospital of Patras. He currently is the head of the Steering Committee of the Interdepartmental Program of Postgraduate Studies in Medical Physics and member of the Steering Committee of the Interdepartmental Program of Postgraduate Studies in Electronics and Information Processing, both offered by the University of Patras. He is a member of the Greek federal committee for Medical Physics licensing and consultant of the National Food and Drug Administration, Greece. He is a member of the overseas advisor committee of The British Journal of Radiology and reviewer in various international journals. He has published over 130 papers in international peer-reviewed journals and over 300 papers in international conferences. His research interests focus on medical imaging and include medical image processing and analysis, medical radiation physics, medical image detectors, mammography and simulation using Monte Carlo techniques.