Towards applying content-based image retrieval in the clinical routine

Marcelo Costa Oliveira\textsuperscript{a,}\textsuperscript{*}, Walfredo Cirne\textsuperscript{a,1}, Paulo M. de Azevedo Marques\textsuperscript{b}

\textsuperscript{a}Laboratório de Sistemas Distribuídos, Departamento de Sistemas e Computação, Universidade Federal de Campina Grande, Campina Grande, Brazil
\textsuperscript{b}Centro de Ciências das Imagens e Física Médica, Faculdade de Medicina, Universidade de São Paulo, Brazil

Received 16 December 2005; received in revised form 26 April 2006; accepted 25 June 2006
Available online 10 August 2006

Abstract

Content-based image retrieval (CBIR) has been one of the most vivid research areas in the field of computer vision, and substantial progress has been made over the last years. As such, many have argued for the use of CBIR to support medical imaging diagnosis. However, the sheer volume of data produced in radiology centers has precluded the use of CBIR in the daily routine of hospitals and clinics. This paper aims to change this status quo. We here present a solution that applies Computational Grids to significantly speed up the CBIR procedure, while preserving the security of data in the clinical routine. This solution combines texture attributes and registration algorithms that together are capable of retrieving images with greater-than-90\% precision, yet running in a few minutes over the Grid, making it usable in the clinical routine.

Keywords: Content-based image retrieval; Texture attributes; Image registration; Grid Computing

1. Introduction

The volume of data produced in hospitals and medical centers is increasing fast. The annual production of the large radiology centers is about ten Terabytes per year [24]. This situation exists due to the ease with which the data of the patients are obtained and stored, resulting principally from the reduction of the cost of the computing equipments and image devices during the last years. However, more and more medical images represent huge quantities of data that need to be safely stored, automatically processed and must be indexed in an intelligent way, because they are fundamental pieces in the clinical diagnosis [23,21].

The increasing use of computers to aid the diagnosis (CAD) produced a fast development of computing algorithms applied to medicine in the last decades. The objective of CAD is to improve the accuracy of the diagnosis, and the consistency of the interpretation of the radiological image [2]. However, some CAD tools that show great results are not used in the clinical routine because they have high computational cost [13].

The difficulties in applying these CAD algorithms in the clinical routine and the limitations that still exist related to the storing, processing, searching and retrieving of images in large databases has been motivating companies and research institutions to find new solutions to solve these problems [3].

The Grid Computing (GC) technology represents the most recent and promising tool in distributed computing. GC is the integration of computers distributed geographically, making it possible to create a virtual computing application to solve problems related to the storing and access of mass data and to the processing of applications with high computing costs.

The shared processing capacity of GC allows the study of large data quantities using images processing algorithms that require high computing power. The use of GC also allows that health regional networks combine resources into one large distributed image database, making it possible for the medical communities to share data as well as medical applications and knowledge, and therefore making possible a greater interaction among the medical centers [18].

However, although very promising, we do not know of the use of GC to support diagnosis as part of the clinical routine of a hospital or clinic. The use of GC to support CAD seems to be yet restricted to medical research. The goal of this work is to analyze and introduce an accessible cost methodology based on the existing techniques, requirements and technologies in
Grid Computing. This work used Texture Analysis to select the largest volume of images in the database that need to be submitted to Grid Computing. This capability avoids large data transference and the bandwidth bottleneck. In fact, Texture Analysis was capable of reducing the number of images that need to be processed using costly registration algorithms.

In this work, we also show the higher efficiency of registration algorithms in the retrieval of medical images. Moreover, we present the effective use of Grid Computing to run the images registration algorithm applied to the CBIR in the clinical routine.

2. Content-based medical images retrieval

Among the various CAD techniques, Content-Based Medical Image Retrieval Applications (CBIR) are a major beneficiary of the Grid Computing technology due to their characteristics and necessities: high processing and great need of large storage [24].

The objective of CBIR in radiology information systems is to give the right information to the specialist in the appropriate time, in order to improve the quality and efficiency of the diagnoses. In the process of clinical decisions, CBIR offers great benefits, being able to retrieve images in databases of the same form, anatomical region and pathology [25].

The use of image registration techniques applied to CBIR obtained considerable success in the medical images processing community due to the capacity of offering easy comparison between two images. This technique seeks for a space transformation able to map points of an image in the corresponding points of another image. The image registration can be done by means of a 2D rigid coordinate transformation, which combines rotation, translation and scale, in seeking for the maximum matching between a reference image and a target image. The rigid transformation is based on the squared error minimization between structures’ outlines using similarity measure algorithms between the intensity of two images. The non-rigid transformation techniques are out of the scope of this paper and are described in [31,15].

However, to make a comparison between a reference image and a large medical image database using only one computer demands much more time than is reasonable for computer-aided diagnosis. Grid Computing makes feasible the use of the images registration technique applied to the CBIR, employing parallelism to reduce the time used to process the algorithm, so making this technique a promising tool to apply in the clinical routine [13].

In [24], there is presented a prototype to make a CBIR using medical image registration by means of Grid Computing resources. The application consists of a hybrid technique involving metadata and rigid registration algorithms. From a reference image, the application makes a selection based only on the metadata of the image and then these selected images are processed using the images’ registration in the Grid. However, the authors says that the methodology applied to the selection resulted in a large volume of data to be sent and processed in the Grid, and that, in the clinical routine, this volume of data would be a lot bigger. Besides, the information contained in the medical images metadata can have a high error rate, there being related cases of up to 16% error [11].

The development of a CAD application must offer an effective and fast answer to the user. Therefore, solutions to optimize the volume of data that will be distributed and processed with assurance in the Grid are necessary.

3. Necessary requirements to the development of CAD applications using the Grid

Since the first project able to identify the potential of the Grid in biomedicine [23], a growing number of papers have shown the importance of the use of Grid Computing in the Life Sciences [21,23,26,20]. However, the state of art of the use of Grid Computing in CAD is still in prototype or for research only.

A great difficulty in introducing a CAD application using Grid Computing resources in the clinical routine is mainly related to the integrity and safety in the manipulation of the medical data in the Grid. We can also mention the complexity and little maturity of current Grid technology, bandwidth network problems that slow data transfers down, lack of a friendly and intuitive interface for the specialist (in the majority of cases laical in computing) and the development of a consistent application with the technologies already established of existing medical data management and medical images.

The main concern related to the distribution of medical data in the Grid is privacy. The major goal for any application that manipulates medical information is the respect for the privacy of the patient. The information and the images of the patients are confidential and should only be accessible by the medical team involved and the patient himself. This way, a medical Grid open to an institution or to a hospital federation must have a restricted access control and must ensure safe transference and safe data storing. The absence of a safe data integration is one of the greatest challenges to be solved in Grid Computing applications in medicine [13].

Data transfers are one of the main components in the total running time of CAD applications in the Grid. The Grid access to the database images involves concurrent access and the transference of a great volume of data. The solution proposed in [22] was to pre-replicate the data in the sites of the Grid, which are connected via a LAN. However, the storing of the distributed data in the Grid introduces data control issues that are more complex than in closed application. In order to assure this functionality, the applications should consider the safety requirements, as already pointed out in this paper.

The solution used in [26] to optimize the Magnetoencephalography data transference in a WAN was the use of a 1 GB/s network. However, this is a high cost solution and often not applicable to small clinics and public hospitals.

A procedure involving a CAD tool normally involves not only one algorithm, but a set of algorithms that use techniques coming from two knowledge areas: computing vision (which involves image processing to enhance, window/level, zoom,
pan, segmentation and extract attributes) and artificial intelligence (which includes methods to select attributes and pattern recognition) [2]. Medical applications also require an intuitive and easy graphic interface. Therefore, CAD applications that use the Grid should have all these requirements, besides using a middleware able to give transparent yet secure access to Grid data and resources.

The technologies and specifications of medical data management were strongly established in the last 20 years following the ACR/NEMA standard [30]. ACR/NEMA established the DICOM (“Digital Imaging and Communication in Medicine”) standard for communication and storage of medical images and related information. The standardization by means of DICOM was fundamental to the development and deployment of the Picture Archiving and Communication System (PACS). PACS has the function of managing the storing of DICOM images in hospitals and clinics.

Related to PACS, hospitals need a Hospital Information System (HIS) and a Radiology Information System (RIS) to correctly relate information and examinations of the patient with his respective digital images, and clinical information obtained from the patient record [9]. Faced with this situation, to be introduced in the clinical routine, medical applications must be integrated to the PACS/DICOM system. Therefore, the Grid Computing technology applied to the CAD will only be adopted if it is integrated to the PACS and RIS. In [18], some integration prototypes between the PACS and the Grid Computing architecture applied to the distributed storing of medical images are shown. On the other hand, the projects [24, 22] related to the application of the Grid to CAD do not address the need for integration with the PACS and RIS.

In this work, we also have shown the higher efficiency of registration algorithms in the retrieval of medical images. Moreover, we present the effective use of Grid computing to make viable the running of the image registration algorithm applied to the CBIR in the clinical routine.

4. Materials and methods

The application was developed in the operational system GNU/Linux Debian using Java 1.5, the image registration algorithm being based on the sum squared difference (SSD), using the implementation of InsightToolkit [14]. For its assessment, a heterogeneous image base was used, with 2400 magnetic resonance images of different anatomic regions, sequences and acquisition plans and gray levels varying between 4096 and 65,536.

4.1. Application description

The application has a graphic interface that requires the specialist to authenticate himself. After authentication, the specialist obtains permission to seek and visualize the examination of interest in the DICOM standard that belongs to PACS. The developed PACS is still a prototype; however, it already stores and makes available images in DICOM standard. The overall application is depicted in Fig. 1.

The PACS was constructed utilizing PostGreSQL-8.1-3 and utilized the Hibernate 3.0 technologies to insert and retrieve images. Each image of PACS has an associated characteristic vector (see Section 4.2) that is obtained when the images are inserted into the PACS. As a result, the PACS serves as an image catalog, with all necessary information to locate the medical image (patient-id, study-id and image-id), and a path to find the image files by means of the Network File System.

The application also allows the specialist to select a stored examination in his computer. In order for the specialist to define a better reference image, the application offers him the basic medical image processing tools such as zoom, pan and window/level.

This application was separated in two modules. In the first module, the texture analysis algorithm selects the images with the smallest Euclidean distance (Section 4.3) between the characteristic vectors extracted from the reference image and the characteristic vector of images that are in the database. This module begins after one reference image has been chosen. The algorithm will select the 1000 most similar images that are in the database to be used in the second module. After completing the first module, the application offers the user an interface with the classification of the most similar images (Fig. 2).

In the second module, the image registration algorithm based on the sum of squared difference (SSD) (Section 4.4) is run in parallel in Grid using the reference image and the selected images (target images) are retained on catalog. The
application is divided into 20 independent tasks, each having 50 target images and the reference image. In order to optimize the application, the images are compacted before being sent to the Grid and the SSD algorithm libraries are stored in the Grid machines. After receiving the similarity values of all 1,000 images from the Grid, the application offers the user an interface with the classification of the most similar images (Fig. 3).
Facing the classified images, the specialist has access to the patient electronic record and can compare his diagnosis with others already made. In order to ensure patient privacy, the patient name and patient identification are hidden from the specialist. We are currently adding more functionality to this module, which is a part of the GridVida project. This project will construct an environment based in Grid Computing to give support to an application to integrate the patient electronic record systems of the Brazilian Public Health System.

4.2. Texture Analysis

The Grey Level Co-occurrence Matrix is a technique to extract information from second-order texture. Studies using Texture Analysis (AT) to classify images are already well established [5,19,8,1]. The co-occurrence method offers a second-order solution that produces the texture characteristics. Given an image, the co-occurrence method obtains from an image the occurrence probability \( C(i, j | \Delta x \Delta y) \) of a pixel pair with intensity \( i, j \) and spacing between the pixels of \( \Delta x \) and \( \Delta y \) in the dimensions \( x \) and \( y \) respectively.

The second-order histogram statistics are applied to the co-occurrence matrix producing the texture attributes. The texture attributes used in this project were suggested by [12].

Energy: \( \sum_{i,j} C(i, j)^2 \); \hspace{1cm} (1)

Entropy: \(- \sum_{i,j} C(i, j) \log C(i, j)\); \hspace{1cm} (2)

Inverse difference moment: \( \sum_{i,j} \frac{1}{1 + (i - j)^2} C(i, j) \); \hspace{1cm} (3)

Shade: \( \sum_{i,j} (i + j - \mu_x - \mu_y)^2 C(i, j) \); \hspace{1cm} (4)

Inertia: \( \sum_{i,j} (i - j)^2 C(i, j) \); \hspace{1cm} (5)

Promenance: \( \sum_{i,j} (i + j - \mu_x - \mu_y)^4 C(i, j) \); \hspace{1cm} (6)

Correlation: \(- \sum_{i,j} \frac{(i - \mu_x)(j - \mu_y)}{\sqrt{\sigma_x \sigma_y}} C(i, j) \); \hspace{1cm} (7)

Variance: \( \sum_{i,j} (i - \mu)^2 C(i, j) \). \hspace{1cm} (8)

In our work, all of the PACS images have an associated 32 dimension characteristic vector. The characteristic vector was obtained by means of the calculation of the eight co-occurrence matrices in orientations 0°, 45°, 90° and 135°. We also adopted the distance between the pixels equal to one.

4.3. Distance measure between the characteristic vectors

The images select is performed by means of the smallest values obtained by Euclidean distance, defined in (9), between the characteristic vectors of the reference image and the PACS images. We have utilized Euclidean distance because of its very simple implementation.

\[ d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \] \hspace{1cm} (9)

where \( x \) and \( y \) are the characteristic vectors of the reference and target image respectively, and \( n \) is equal to the size of the characteristic vectors, that is, 32 in our work.

4.4. Image registration

In this study, we have used image registration based on the similarity measure. This technique calculates the sum of squared difference (SSD) between the intensities of two images. The algorithm determines in an iterative way transformations that optimize the similarity measure of the voxel intensities between two images. Every iteration makes transformations on the reference image by means of transformation estimates so as to make it more similar to the target image, and then generates a new estimate of the similarity [27]. Mathematically, the sum of squared difference (SSD) is defined by:

\[ S(p|F, M, T) = \frac{1}{N} \sum_{i} [F(x_i - M(T(x_i, p)))|^2 \] \hspace{1cm} (10)

where \( F \) is the reference image intensity function, \( M \) is the target image intensity function, \( T \) is the rigid space transformation function, \( x_i \) are the pixels of the reference image considered by the metric (SSD), \( p \) is the transformation parameter used to map the points between the reference and target images and \( N \) is the total number of pixels of the reference image [31].

The rigid transformation \( (T) \) used in this application was affine transformation. It is one of the most popular transformations used for image registration and the main advantage comes from the fact that it is represented as a linear transformation. Affine transformation is composed of rotation, scaling, translation and shearing [14].

4.5. Grid middleware

The Grid Computing used in this application was OurGrid (Fig. 4) [6,28]. OurGrid is a cooperative and free-to-join Grid, where sites share their idle computing resources and, when necessary, receive idle resources of other sites. OurGrid assumes that the parallel applications run on it are Bag-of-Tasks (BoT), in other words, that tasks are independent from each other. Besides, OurGrid ensures priority to the local user over the local resources. This characteristic ensures that the SSD algorithm will be run no worse than not using the Grid, using only local resources.

In order to protect the data and the local resources OurGrid uses SWAN (Sandboxing Without A Name). SWAN is an application based on the Xen virtual machine that isolates the users Grid applications in a sandbox; this way it does not allow Grid tasks to access the network or local data. Only the necessary basic resources to run the remote tasks are available to the Grid tasks [6].
The user interaction with OurGrid is done with MyGrid. MyGrid is a broker that manages the applications and offers a set of abstractions that hide the Grid heterogeneity from the user. MyGrid uses task replication in order to obtain a better performance without relying on Grid or application information [7].

MyGrid allows the data to be stored in the Grid machines or removed when the running of the application is over. Then, in order to assure the safety of the patient data, the images are not left in the Grid machines. However, in order to optimize the time of application processing, the libraries of the SSD algorithm were stored in the Grid machines.

In order to give further easiness and dynamism in the application to the specialist, MyGrid API was connected to the application. After the running of all the tasks, MyGrid communicates the end of the running to the application. Then, the application begins the reading of the received files with similarity values. Besides the similarity values, the files have identification about the patient and examination code and the number of the image, which allows an easy association between the similarity value and the image organized into the PACS catalog.

The application orders in a list the most similar images according to the SSD values. The resulting images are shown as thumbnails, which allows the specialist to select the target images that he wants to visualize in the main window of the application. The application will allow, in the future, that after the user has selected a target image, he also has access to the related information in the RIS.

5. Results

The results of the CAD developed in this project were obtained thorough the selection of two anatomic regions, knee and head, belonging to PACS. The knee images are weighted in T1 with sagittal cut plan and the head images are weighted in T2 and axial cut plan. PACS has one examination with 20 images of the knee in sagittal plan and one examination with 40 images of the head in the axial plan. The experiments were repeated three times, choosing different cuts for the described examinations. The images were classified as right when the application returned images of the same plan and sequence of acquisition of the reference images.

The selection of the images allowed the classification of the most similar images according to the texture attributes. The average processing time of the first module (Texture Analysis) was 1.5 min, obtained by the calculation of the Euclidean distance between the characteristic vector of each database image and the characteristic vector of the reference image. The algorithms were processed in the local computer using a Pentium 4 processor at 2.8 GHZ and 1 GB RAM.

In order to assess the results obtained, we use Precision and Recall, which are parameters normally used to assess Content-Based Image Retrieval systems and information retrieval. Recall indicates the proportion of relevant images in the database that were retrieved, out of all relevant images available. Precision, on the other hand, is the proportion of the retrieved images which are relevant to the specialist [4].

Fig. 5 presents the averages obtained by the running of the first module of the Recall vs. Precision curves of the Euclidean distance between the reference image characteristic vectors in relation to the database. This result allows us to assess the efficacy of the CBIR using texture in the classification of the most similar images by the second module. Although the average precision obtained in the first experiments is 0.54 (sagittal knee) and 0.4 (axial head), it was sufficient to select the images that were submitted to the second module.

In the second module of the application, the images were processed by the sum of squared difference (SSD) algorithm using Grid Computing processing. The CBIR using the SSD technique produced satisfactory precision in both of the studied cases, 0.95 (sagittal knee) and 0.92 (axial head), according to the Recall vs. Precision curves average between the reference
images and the images selected by the first module (Fig. 6). Fig. 7 illustrates the classification of the most similar images after the running of the application. Due to space problems, only the nine most similar images are illustrated.

Grid Computing was able to greatly reduce the high processing time of the SSD algorithm. The processing time of the SSD algorithm applied to the experiments reduced the processing time to 116.97 (sagittal knee) and 95.15 (axial head) min, compared to the processing time obtained in one machine. The experiments used 50 Grid processors and a 100 MB/s LAN (Fig. 8). The application total time, considering the time to calculate the Euclidean distance and the time to execute the SSD algorithm, was 5.02 min. We divided the application into 20 tasks consisting of 50 images each. The images were compacted before being sent to the Grid and the average size of the files with 50 images was 4 MB.

The graphic interface allowed the specialist to have an easy interaction with the application, because all the image processing techniques and Grid complexity were made invisible to the specialist. The thumbnail images sample allowed the specialist to have an easy navigation over the algorithm result. The use of the MyGrid API in the application allowed absolute transparency in relation to the complexity of the Grid and removed the necessity of the specialist to use MyGrid.

The control of passwords in the beginning of the application made it possible that only specialists registered had access to PACS. OurGrid made privacy possible when medical images were distributed through SWAN. This solution permitted that an application did not have access to the data of other applications, or vice versa. Besides, after the running of the task in the Grid machines the images were removed.

MyGrid optimized the application executing time, because it allowed the SSD algorithm libraries (10 MB) to be stored in the Grid machines, thus avoiding large data traffic in the Grid.

Fig. 7. The ranking of the nine most similar images retrieved using the sagittal knee image as reference.
The number of digital medical images produced is in constant growth. Large radiology centers produce more than ten thousand digital images daily. Videos and images produced by cardiology and endoscopy are already being integrated to the Picture Archiving and Communication System (PACS), increasing even more the amount of data stored [9]. The management, privacy and processing of this large quantity of data are not simple. Faced with these challenges, companies and research centers seek for new applications that are able to store and process large volumes of data.

The Grid Computing (GC) technology has emerged as a promising tool in the processing and storing of huge volumes of data. The possibilities for exploiting GC technologies in medicine are numerous. For example, [18] shows the use of the Grid to make a backup of medical images in different PACS federated in a Grid. However, ensuring security and confidentiality is, of course, a crucial issue within the field of a Health Grid [29].

This project shows a mixed approach of CBIR techniques to classify similar images from different anatomic regions and orientations using the high processing power of GC. In order to make feasible the use of Grid Computing technology in the clinical routine, the application has adopted safety requirements, is user friendly, and interoperates with image management technologies (PACS) and medical images standards (DICOM). Although the developed PACS is only a prototype, its use allowed us to assess the application. The integration of the application with the Radiology Information System (RIS) is still being developed.

The application dealt with the CBIR techniques based on texture classification and on the similarity measure. The texture classification shows a good approximation to human visual perception and it has been used in many systems to aid the clinical diagnosis [16].

The precision average value of the texture classification was efficient in acting as an initial filter to the second CBIR algorithm, despite the precision average being relatively low, 0.54 (sagittal knee) and 0.40 (axial head). Therefore, the texture classification optimized the application without decreasing the precision, because the final match depends on the second, much more costly CBIR algorithm. Moreover, the texture selecting decreased the total time of the application processing by limiting the volume of data that are distributed in the Grid network to only a thousand images.

A possible solution to increase the efficiency of this selection would be the development of methods to detect movement artifacts, since texture information can be lost when rotation, translation and scale are included in the characteristic space [10].

The use of the sum of squared difference of similarity measure algorithm (SSD) as the second more precise and costly CBIR algorithm promoted a precision average value quite satisfactory, 0.95 (sagittal knee) and 0.92 (axial head). The algorithm was able to retrieve similar images from different anatomic regions and orientations. This ability of the SSD algorithm is due to the capacity of making comparisons based on the image data. Most of the articles in the literature are restricted to a certain anatomic region, acquirement plane or diagnosis procedures because they only use characteristic vectors [17]. However, the computing cost of the similarity measure when executed in only one computer is intractable to computer-aided diagnosis. The Grid has made possible the use of the similarity measure technique due to its ability of processing tasks in parallel in the different computers that form the Grid.

In order to improve the precision of the application we are developing a similarity measure algorithm based on the crossed correlation (CC). The crossed correlation algorithm has a higher computing cost than the SSD technique [24]; however, it will allow the application to make a CBIR between the different forms of images, which is a limitation of the SSD algorithm. Another limitation of the SSD technique is the high sensibility to small quantities of pixels that have large differences of intensity between two images, like, for example, in the cases of contrast injection [31].

OurGrid made integrity and safety possible to the medical data by means of the removal of the images in the Grid machines and through the use of SWAN technology. However, OurGrid still does not allow CBIR to be done in PACS of a distributed health network. Therefore, it is necessary the development of a service associated to the OurGrid peers in order to make viable the distributed CBIR. This solution will allow the health regional networks to use a unique large distributed database, allowing the medical communities to share resources, data, and more importantly, medical expertise.

We are also developing an algorithm to calculate the ideal granularity of the application dynamically based on the behavior of OurGrid. This algorithm will allow us automatically to define the ideal granularity of the application, so we expect that the total time of the application will be decreased.

7. Conclusion

This application showed a solution to the deployment of a Content-Based Medical Image Retrieval application in the
clinical routine using the computing power of OurGrid, a Grid Computing technology. The application allows the specialist, through an intuitive and friendly graphic interface, to use the benefits of CBIR optimized by Grid. OurGrid ensured the application the patient safety and privacy requirements pointed out in the literature. The application also allowed the specialist to seek for images in PACS and to use basic image processing tools. All these characteristics formed a friendly, powerful, low cost tool in computer-aided diagnosis.

However, in order to introduce this solution in the clinical routine, it is necessary to have integration with the Radiology Information System (RIS). Besides, although acceptable, the total running time of the application is still relatively high for the clinical routine. Therefore, the application could still benefit from faster algorithms (without compromising the accuracy) and larger number of machines in the Grid.

References


Marcelo Costa Oliveira is Ph.D. student candidate at the Faculdade de Medicina, Universidade de São Paulo (USP), Brazil. He is a bachelor in computer science and received his master’s degree in physics applied to medicine from Universidade de São Paulo (USP). He has been working as research assistant at the Laboratório de Sistemas Distribuídos da Universidade Federal de Campina Grande (UFCG), Brazil. His current research interests include Grid computing and medical image processing.

Walfredo Cirne is a member of the faculty at the Computer Science Department of the Universidade Federal de Campina Grande, in Brazil. Dr. Cirne holds a Ph.D. from the University of California San Diego, in the USA. He had previously worked on Machine Learning and Computer Networks, and, since 1997, his research focuses on Grid Computing. Currently, Dr. Cirne leads OurGrid, a project developed in cooperation with Hewlett Packard that aims to provide an open, free-to-join grid solution for bag-of-tasks applications.

Paulo Mazzoncini de Azevedo Marques was born in Ribeirão Preto, SP, in 1964. He is a full-time Associate Professor of Medical Physics and Biomedical Informatics at the Internal Medicine Department, University of São Paulo, School of Medicine in Ribeirão Preto (FMRP-USP) — Brazil. He received his B.Sc. and M.Sc. degrees in Electrical Engineering and his Ph.D. in Applied Physics from the University of São Paulo in 1986, 1990 and 1994 respectively. He had a 6 month post-doctoral experience (2001) at the University of Chicago (Chicago, IL, USA), under the supervision of Prof. Kunio Doi, where he mastered several techniques for pattern recognition in medical images focused on computer-aided diagnosis (CAD) and content-based image retrieval (CBIR). His main subject interest areas are CAD, CBIR and PACS (Picture Archive and Communication Systems).