

# A Computer-Aided Diagnostic System using a Global Data Grid Repository for the Evaluation of Ultrasound Carotid Images

Marco Antonio Gutierrez<sup>(1)</sup>, Silvia Helena Gelas Lage<sup>(1)</sup>, Jasper Lee<sup>(2)</sup>, Zheng Zhou<sup>(2)</sup>

1: Informatics Division and Critical Care Unit, Heart Institute (InCor),

University of São Paulo Medical School, São Paulo, Brazil

2: Image Processing & Informatics (IPI) Laboratory, Department of Radiology,

University of Southern California, California, USA

marco.gutierrez@incor.usp.br

## Abstract

*A computer-aided diagnostic (CAD) method of calculating lumen and wall thickness of carotid vessels is presented. The CAD is able to measure the geometry of the lumen and plaque surfaces in ultrasound carotid images using a least-square fitting of the active contours obtained automatically from the vessels border. To evaluate the approach, ultrasound image sequences from 30 patients were submitted to the procedure. The images were stored on an international data grid repository that consists of three international sites: IPI Laboratory at University of Southern California, USA; Heart Institute at University Sao Paulo, Brazil, and Hong Kong Polytechnic University, Hong Kong. The three chosen sites are connected with high speed international networks including the Internet2, and the Brazilian National Research and Education Network (RNP2). The Data Grid was used to store, backup, and share the ultrasound images and analysis results, which provided a large-scale and a virtual data system.*

## 1. Introduction

The use of digital imaging technology has become a key component in both basic research and clinical practice. Although advances in data acquisition technology have improved the resolution and speed at which we can collect image data, most researchers have access to a limited image repository mainly due to lack of efficient software for managing, manipulating, and sharing large volumes of data [1]. In general, the image collection in research and clinical studies intrinsically creates a distributed database, as it involves many hospitals and/or centers in different locations. In addition, the amount of data generated by some imaging procedures can be so large that it is not efficient to concentrate them in a single computing center. As an example, a single study to evaluate ultrasound images of the carotid for the assessment of

the degree of atherosclerotic disease has an average size of 64 Mbytes (250 images, 16 sec of acquisition time). This quantity linearly increases with time and a full transfer over the network from the acquisition site to a central repository would be large enough to saturate the available commodity connections.

Despite of the problems involving the manipulation and sharing of large medical image volumes, current image acquisition equipments do not have the software or computational power for the state-of-the-art data post-processing, typically available at computer-aided diagnostic (CAD) systems. Typically, on-line computation is oriented towards image acquisition and reconstruction. As a result, computationally demanding tasks are performed off-line, on desktop workstations after the completion of an exam. This has the adverse consequence of delaying the interpretation and communication of radiological results and makes adoption of advanced imaging techniques inefficient for radiologists.

On the other hand, making the whole image database available to authorized users, regardless of the data distribution, would provide several advantages. For example, a CAD system could be tested and trained on a much larger data set, with an improvement of its performance in terms of both sensitivity and specificity. The CAD system could be used as a real time selector of images, with a remarkable reduction of the delay between image acquisition and diagnosis. Moreover, data associated to the images, or metadata, would be available to select the proper input for epidemiology studies or for the training of new radiologists.

In this work, we present a CAD system method to segment carotid vessels in a large scale data grid repository. The operator selects a region-of-interest (ROI) in a series of carotid images obtained from B-mode ultrasound. This set of images is convolved with the corresponding partial derivatives of the Gaussian filter. The filter response is used to compute a 2D gradient magnitude image in order to refine the

vessel's boundaries. Using an active contour technique the geometry of the lumen and the plaque surface are determined automatically. The near wall media-adventitia (NWMA), far wall media-adventitia (FWMA) and far wall lumen-intima (FWLI) borders are obtained by a least-square fitting of the active contours result. The distance between NWMA and FWLI (vessel diameter) and between FWLI and FWMA (far wall intima-media thickness) are obtained for all images and the mean value is computed during systole and diastole. To evaluate the approach, ultrasound image sequences from 30 patients were submitted to the procedure. The images were stored on an international data grid repository that consists of three international sites: Image Processing and Informatics (IPI) Laboratory at University of Southern California, USA; InCor (Heart Institute) at Sao Paulo, Brazil, and Hong Kong Polytechnic University, Hong Kong. The three chosen sites are connected with high speed international networks including the Internet2, and the Brazilian National Research and Education Network (RNP2). The Data Grid enables the three sites to share the image data and to analyze the results to improve the clinical research outcome.

## 2. Methods

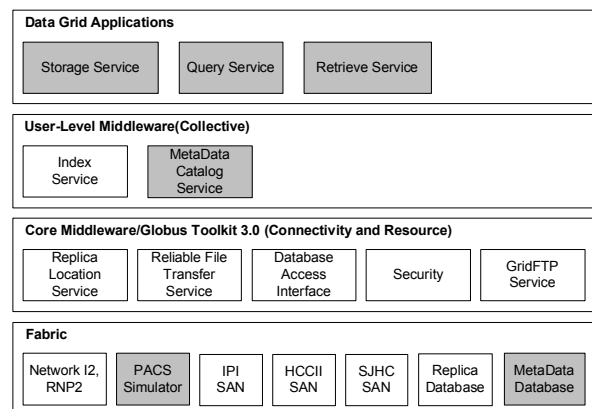
### 2.1. Grid technologies

Grid technologies provide mechanisms for sharing and coordinating the use of diverse resources and thus enable the creation, from geographically and organizationally distributed components, of virtual computing systems that are sufficiently integrated to deliver desired quality of service. These technologies include a high-performance hardware and software infrastructure to provide: a) security solutions that support management of credentials and policies when computations span multiple institutions; b) resource management protocols and services that support secure remote access to computing and data resources and the co-allocation of multiple resources; c) information query protocols and services that provide configuration and status information about resources, organizations, and services; d) data management services that locate and transport datasets between storage systems and applications.

Several large-scale Data Grids, such as TeraGrid (<http://www.teragrid.org/>) and Data Replication for LIGO (<http://www.globus.org>) have been established for fast movement of large amount of data among multiple research institutes. A Data Grid [2] specifically for clinical image backup and disaster recovery has been developed previously at IPI (Image

Processing & Informatics Laboratory) using the Globus Toolkit 4 (GT4) [3]. This Data Grid was designed to utilize the strengths of grid technology along with PACS (Picture Archiving and Communication Systems)/DICOM (Digital Imaging and Communications in Medicine) technology for storing and distributing clinical images [4,5]. In particular, some PACS/DICOM resources are embedded within the five layer grid architecture. These include Storage services, Query services, Retrieve services, which are integrated with the DICOM standard protocols in addition to the use of other Data Grid Services. The five layer architecture of the Data Grid embedded with DICOM technology is shown in Figure 1, which illustrates some of the basic components already developed at IPI, such as the Metadata Catalog Service. The three services in Data Grid Applications layer are not the standard Grid Services defined in the GT4. These services include Storage services, Query services, Retrieve services, which are integrated with the DICOM standard protocols in addition to the use of other Data Grid Services.

The experience and knowledge learned from the Data Grid for clinical image recovery has been utilized to design the Data Grid architecture for CAD assessment of carotid wall thickness.



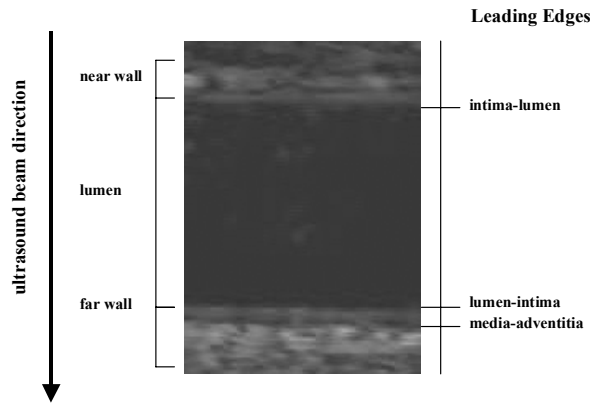
**Figure 1.** Five-layer architecture and the contents of the Data Grid for clinical image backup: The combination of 2nd and 3rd layers of connectivity and resource, respectively, are sometimes referred to as the Core Middleware. Gray-shaded boxes represent components developed or modified at the IPI Lab.

### 2.2. Carotid wall thickness in ultrasound images

Non-invasive ultrasonic B-mode imaging is an increasingly important method for studying progress

and regress of atherosclerotic lesions in the carotid artery. Figure 2 shows a representative B-mode ultrasound image of the carotid artery and a schematic illustration of the relevant leading edges of echo responses. Previous studies [6,7] have shown that the leading edges can be mapped to the following interfaces: near-wall intima-lumen, far-wall lumen-intima and far-wall media-adventitia. The lumen diameter (LD) is defined as the distance between the media-adventitia interface of the near-wall and the lumen-intima interface of the far-wall. The far-wall intima-media thickness (IMT) is defined as the distance between the far-wall lumen-intima and the far-wall media-adventitia interfaces.

The complete geometry of the carotid vessels can be manually determined using B-mode ultrasound. However, this approach is a time consuming procedure and based on subjective operator assessment that inevitably results in inter and intra-observer variability. Efforts have been made to make the measurement less operator dependent by introducing automated image analysis procedures.



**Figure 2.** Interfaces between carotid tissue layers obtained from B-mode ultrasound.

### 2.3. Artery boundary enhancement

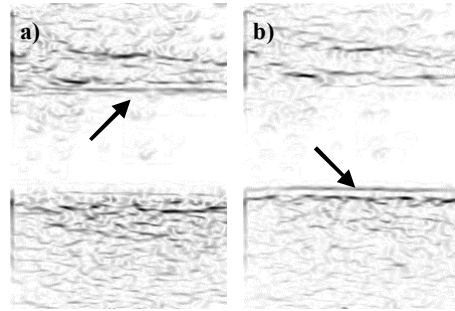
To enhance border detection accuracy, a multiscale border identification was implemented using filters in the form of scaled convolution operators [8]. The scale space of an image is constructed through convolution of the image with a two-dimensional (2D) Gaussian density kernel with zero mean and standard deviation:

$$G(\vec{x}, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}^D} e^{-\frac{\|\vec{x}\|^2}{2\sigma^2}} \quad (1)$$

where  $D$  denotes the dimension of the input domain. A blurred replica of the original image is obtained by

convolution with  $G(\vec{x}; \sigma)$  for a specific  $\sigma$ . The stack of images as a function of increasing scale parameter  $\sigma$  is coined a linear scale space. Hence, as  $\sigma$  increases the detailed object structures vanish, while gross structures persist.

Based on these features a scaled artery image is used to identify the approximated position of the near and far walls. Two complementary images, based on the gradient value in y-direction are obtained: one that enhances pixel values transitions from high to low echoes, such as edges encountered in near wall tissue interfaces, and other that enhances pixel values transitions from low to high echoes (such as edges encountered in far wall tissue interfaces). Figure 3 shows the boundary enhancement of the near and far wall.



**Figure 3.** Boundary enhancement of the near wall a) and far wall b)

The contour of each wall is modeled following the Geometrically Deformed Model proposed by Lobregt and Viergever [9]. In this model, a set of vertices which are connected by straight line segments or edges forms the basic contour structure (Figure 6). The position of a vertex  $V_i$  is represented by a vector  $p_i$ , and the edge between  $V_i$  and  $V_{i+1}$  by a vector  $d_i$ . The contour deformation is caused by a combination of forces which act on the vertices. The resulting acceleration in vertex  $V_i$  is denoted by a vector  $a_i$ .

The contour local curvature at a vertex  $V_i$  is defined as the difference between the directions of the two edge segments that join at that location:

$$\hat{c}_i = \hat{d}_i - \hat{d}_{i-1} \quad (2)$$

The local tangential unit vector  $\hat{t}_i$  is defined as the normalized sum of the unit vectors of two joining edge segments:

$$\hat{t}_i = \frac{\hat{d}_i + \hat{d}_{i-1}}{\|\hat{d}_i + \hat{d}_{i-1}\|} \quad (3)$$

The local radial direction at a vertex  $V_i$  is obtained from  $\hat{t}_i$  by a rotation over  $\pi/2$  radians:

$$\hat{r}_i = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \hat{t}_i \quad (4)$$

## 2.5. Dynamic force formulation

In the model definition, the dynamic in each vertex  $V_i$  must satisfy the Newton's second law,

$$\begin{aligned} F_i &= F_{\text{int},i} + F_{\text{ext},i} + F_{\text{damp},i} \\ F_i &= \mu_i a_i \end{aligned} \quad (5)$$

where  $\mu_i$  is a coefficient that has a mass unit,  $F_{\text{damp},i}$ ,  $F_{\text{int},i}$  and  $F_{\text{ext},i}$  are the damping (or viscous), the internal and the external forces, respectively.

The internal force can be estimated from the local contour curvature along the local r-axis,

$$F_{\text{int},i} = (c_i \cdot \hat{r}_i) \quad (6)$$

The external force acting in each vertex can be approximated by some image feature. In this paper we used the information obtained from the local image gradient as the external force.

The damping force is proportional to the velocity of the vertex and points in opposite direction:

$$F_{\text{damp},i} \approx -k \cdot v_i \quad (7)$$

The total force  $F_i$  acting on a vertex is a weighted combination of damping, internal and external

$$F_i = w_{\text{int}} F_{\text{int},i} + w_{\text{ext}} F_{\text{ext},i} + w_{\text{damp}} F_{\text{damp},i} \quad (8)$$

where  $w_{\text{int}}$ ,  $w_{\text{ext}}$  and  $w_{\text{damp}}$  are the weighting factors.

The deformation process over the contour is implemented as a numerical time integration process in which the complete state of the contour is calculated at a sequence of discrete positions in time [16,17]. A set of state equations controls the deformation process in terms of position, velocity and acceleration of each vertex on the contour:

$$\begin{cases} p_i(t + \Delta t) = p_i(t) + v_i(t) \\ v_i(t + \Delta t) = v_i(t) + a_i(t) \cdot \Delta t \\ a_i(t + \Delta t) = \frac{1}{m_i} F_i(t + \Delta t) \end{cases} \quad (9)$$

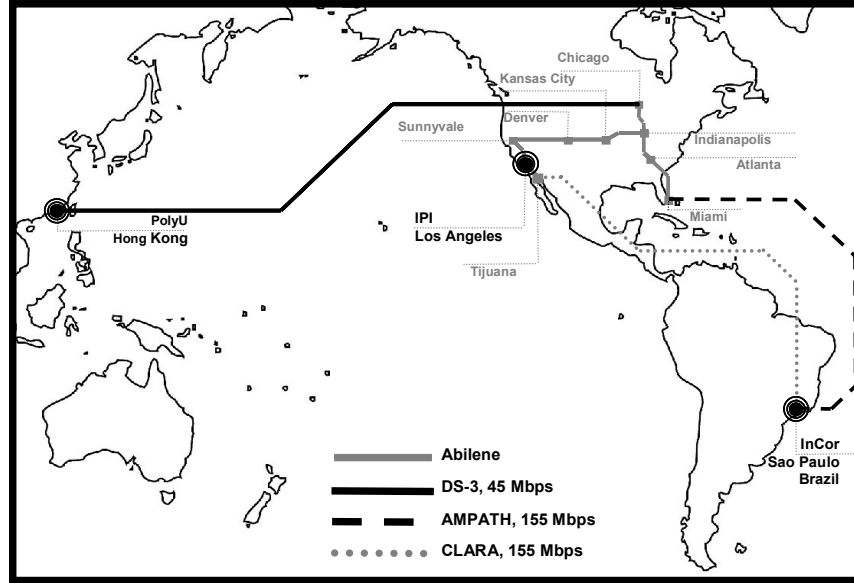
where  $p_i(t + \Delta t)$ ,  $v_i(t + \Delta t)$  and  $a_i(t + \Delta t)$  define the position, velocity and acceleration, respectively, of the vertex in a incremental time  $\Delta t$ . The vertex mass,  $m_i$ , is setting constant for all vertices and the resulting force,  $F_i$ , is calculated using equation (8).

## 3. Results

### 3.1. Data collection

A sequence of B-mode ultrasound images of the carotid artery was acquired with spatial resolution of 456 x 576 pixels, 8 bits/pixel, 15 frames /sec and acquisition time of 16 sec. This setup gives an average data size of 64 Mbytes/exam (250 images). The Data Grid was used to store, backup, and to share the ultrasound images and to analyze the results, which provided a large-scale and a virtual data system. The exams were stored on an international data grid repository that consists of three international sites: Image Processing and Informatics (IPI) Laboratory at University of Southern California, USA; InCor (Heart Institute) at Sao Paulo, Brazil, and Hong Kong Polytechnic University, Hong Kong. The three chosen sites are connected with high speed international networks including the Internet2, and the Brazilian National Research and Education Network (RNP2).

The Data Grid enables the three sites to share the image data and to analyze the results to improve the clinical research outcome (Figure 4).



**Figure 4.** Topology of the International Internet-2 Connectivity Between the Three Sites Linking IPI/USC in North America, InCor in South America, and PolyU in Asia. Note that the Routing Path from InCor to PolyU utilizes AMPATH instead of CLARA which is the routing path between IPI and InCor.

### 3.2. Evaluation procedure

After the retrieval of an exam previously stored in the Data Grid, the operator selects a region-of-interest (ROI) in a series of carotid images. This set of images is convolved with the corresponding partial derivatives of the Gaussian filter. The filter response is used to compute a 2D gradient magnitude image in order to refine the vessel's boundaries. Using an active contour technique the geometry of the lumen and the plaque surface are determined automatically. The near wall media-adventitia (NWMA), far wall media-adventitia (FWMA) and far wall lumen-intima (FWLI) borders are obtained by a least-square fitting of the active contours result. The distance between NWMA and FWLI (vessel diameter) and between FWLI and FWMA (far wall intima-media thickness) are obtained for all images and the mean value is computed during systole and diastole.

In this evaluation a total of 180 images from 30 patients (3 images in diastole and 3 images in systole for each patient) were analyzed, all of which included manually defined interfaces for reference. The minimum and maximum artery diameters were measured for each patient using the manual and the automatic procedure. However, for clinical purposes, some interactive tools for manual tracing were incorporated to the model to correct remaining detection errors in regions with poor image quality. In order to study the variability between the automatic

and manual definition of artery boundaries, the pooled mean,  $\bar{\mu}$ , and the standard deviation,  $\sigma$ , for the difference between automated and manual measurements of lumen diameter were computed. The coefficient of variation,  $CV$ , was calculated according to:

$$CV = \left( \frac{\sigma}{\bar{\mu}\sqrt{2}} \times 100 \right) \% \quad (10)$$

The strength of the relationship between automated and manual methods is indicated by the correlation,  $R_{a,m}$ , between the two measurements:

$$R_{a,m} = \frac{Cov_{a,m}}{\sigma_a \cdot \sigma_m} \quad (11)$$

where  $Cov_{a,m}$  is the covariance between the automated and manual and  $\sigma_a$  and  $\sigma_m$  are the standard deviation of automated and manual measurements, respectively.

The results obtained for the means ( $\mu_{a,m}$ ) and standard deviations ( $\sigma_{a,m}$ ) for the differences between the automatic and manual methods and for the parameters  $CV_{a,m}$  and  $Corr_{a,m}$  are summarized in Table 1.

**Table 1.** Lumen diameter (LD) and Intima-Media Thickness (IMT) measured using automatic and manual methods (n=30). The difference,  $\Delta$ , the coefficient of variability,  $CV$  and the correlation,  $Corr_{a,m}$ , between both measurements are also presented

	<b>Automatic</b>	<b>Manual</b>	<b>Difference <math>\Delta</math></b>	<b>Variability</b>	<b>Correlation</b>
	$\mu_a \pm \sigma_a$ (mm)	$\mu_m \pm \sigma_m$ (mm)	$\mu_{a,m} \pm \sigma_{a,m}$ (mm)	$CV$ (%)	$Corr_{a,m}$
<b>Lumen Diameter</b> (diastole)	7.85 $\pm$ 1.01	7.78 $\pm$ 1.01	0.13 $\pm$ 0.09	0.83	0.99
<b>Lumen Diameter</b> (systole)	6.81 $\pm$ 1.06	6.77 $\pm$ 1.05	0.12 $\pm$ 0.10	1.00	0.99
<b>Intima-Media</b> <b>Thicknes</b>	0.72 $\pm$ 0.14	0.63 $\pm$ 0.12	0.09 $\pm$ 0.06	6.16	0.90

## 4. Conclusions

Measurements of lumen diameter (LD) and intima-media thickness (IMT) of carotid and femoral arteries from B-mode ultrasound are defined as the average distance of interfaces between vessel tissue layers. In order to determine the interface location a manual tracing is commonly used. However, this approach is a time consuming procedure and based on subjective operator assessment. We have proposed a method that uses the active contour technique where the external forces are proportional to the local image gradient obtained from a multiscale analysis. The automated measurements, when compared to those obtained by manual tracing, are equally accurate and the coefficients of variability between both methods are below 1.0% for Lumen Diameter and 6.5% for Intima-Media thickness measurements.

In this paper, we present a Data Grid testbed for evaluation of ultrasound carotid images. Three international sites were used for store, query and retrieve of medical images. The laboratory prototype of this DICOM compliance Data Grid has been completed. Currently, the three International sites have been connected, and we are implementing the Data Grid at the Hong Kong Polytechnic University, and the Heart Institute at Sao Paulo, Brazil.

Once the Data Grid testbed for research and clinical study is established, we can incorporate the multiple trials databases into the Data Grid in the future. The multiple trials databases are equally important as the image data. Such a Data Grid can provide three benefits to the trials databases: 1) fault-tolerance, 2) data and result sharing using for instance the DICOM Structured Reporting, and 3) dynamic creation and modification of data model to support any new trial or change of trials.

## 5. Acknowledgements

This work has been supported by Grant No. 2436 of the National Network for Education and Research (RNP), Scientific and Technology Ministry, Grant No. 300499/2005-1 of the National Council for Scientific and Technological Development (CNPq) and the Zerbini Foundation.

## 6. References

- [1] R. Bellotti, P. Cerello, S. Tangaro, V. Bevilacqua, M. Castellano. Distributed medical image analysis on a Grid infrastructure. *Future Generation Computer Systems*, 23:475-484, 2007.
- [2] B.J. Liu, M.Z. Zhou and J. Documet, Utilizing data grid architecture for the backup and recovery of clinical image data, *Computerized Medical Imaging and Graphics* 29 (2005), 95-102.
- [3] I. Foster, Globus Toolkit Version 4: Software for Service-Oriented Systems, *IFIP International Conference on Network and Parallel Computing*, Springer-Verlag LNCS 3779, 2005, 2-13.
- [4] H.K. Huang, PACS and Imaging Informatics: *Basic Principles and Applications*, Wiley & Sons, NJ, 2004.
- [5] B.J. Liu, F. Cao, M.Z. Zhou, G. Mogel and L. Documet, Trends in PACS Image Storage and Archive, *Computerized Medical Imaging and Graphics* 27 (2003), 165-174.
- [6] C. Liguori, A. Paolillo, A. Pietrosanto. An automatic measurement system for the evaluation of carotid intima-media thickness. *IEEE Transaction on Instrumentation and Measurement*, 50(6):1684-1685, 2001.
- [7] A. Zahalka, A. Fenster. An automated segmentation method for three-dimensional carotid ultrasound images. *Physics in Medicine and Biology*, 46:1321-1342, 2001.
- [8] J.J. Koenderink. The structure of images. *Biological Cybernetics* 1984;50:363-370.
- [9] S. Lobregt, M. Viergever. A discrete dynamic contour model. *IEEE Trans. Medical Imaging*, Vol.14, No. 1, pp.12-24, 1995.