

# Caracterización de Imágenes Médicas Basadas en la Transformada de *Wavelet* para Indexación en Estructuras de Datos Multi-dimensionales

César A.B. Castañón

Agma J.M. Traina,

ICMC–Instituto de Ciências Matemáticas e de Computação - Universidade de São Paulo - São Carlos  
Av. do Trabalhador São-Carlense, 400 - Centro - Cx. Postal 668 São Carlos - São Paulo - Brasil  
{cesar, agma}@icmc.usp.br

## Abstract

This paper proposes a characterization of medical images by a *wavelet* transform approach. Typically, a medical radiology consists of gray level variations in highly localized regions of the image. Currently, it is not possible to extract these regions by fully automatic image segmentation techniques. To address this problem we consider the generation of feature vectors based on *wavelet* transform. Afterwards, the extracted features create a vector that characterizes the image. The system extracts the relevant image features, computes the Euclidean distance of the query image to all image indices in the database, and retrieves the  $n$  most similar images. Our approach is based on the application of Daubechies4 *wavelets* over the global characteristics of the image. The focus of this paper is to assess the utility of the *wavelet* transform on medical image characterization and their feasibility for an appropriate indexing into a multi-dimensional data structure.

**Keywords:** *Wavelet*, Content Based Image Retrieval, Feature Vector, Image Data Base.

## Resumen

En el presente artículo se propone una técnica de caracterización de imágenes médicas aplicando la transformada de *wavelet*. Una de las propiedades de las imágenes médicas, es la de presentar cambios en tonos de gris dentro de regiones localizadas de la imagen. Hasta ahora no se tiene una técnica adecuada que haga posible extraer estas regiones mediante un completo procesamiento automático de las imágenes. Una estrategia para abordar este problema consiste en el uso de filtros de convolución de *wavelet* a fin de identificar numéricamente las características más resaltantes de la imagen, estos valores extraídos generarán un vector de características que se va a constituir en la identificación de la imagen. En la presente propuesta se genera este vector de características basado en la transformada de *wavelet*, se calcula la distancia entre una imagen de consulta y las que se encuentran en el banco de datos, recuperándose las  $n$  imágenes más similares. El enfoque presentado está basado en la aplicación de los filtros de las *wavelets* de Daubechies4 sobre las características globales de la imagen, utilizando como función de similaridad la distancia Euclidiana. El objetivo del presente trabajo es destacar la utilidad de la transformada de *wavelet* en la caracterización de imágenes médicas y su posterior aprovechamiento en la indexación dentro de una estructura de datos multi-dimensional.

**Palabras claves:** *Wavelet*, Recuperación de Imágenes Basadas en Contenido, Vector de Características, Base de Datos de Imágenes.

## 1 Introducción

Los sistemas de recuperación de imágenes basadas en el contenido (*Content-based image retrieval - CBIR*) se refieren a la capacidad de recuperar imágenes considerando el contenido de las mismas. Dada una imagen de consulta, el usuario gustaría que el sistema CBIR recupere las  $n$  imágenes más similares a aquella dada en la consulta, dentro de un banco de imágenes, basado en algún criterio definido. Si almacenamos en la base de datos, el total de las imágenes, y se efectúan las comparaciones de similitud entre todas ellas, este proceso nos consumirá gran tiempo de procesamiento. Por lo que como alternativa, para evitar este oneroso proceso, se opta por la generación de una representación alternativa de la imagen, lo cual es comúnmente conocido como firma o *vector de características*.

Un vector de características es una representación numérica de una imagen, o de una porción de ella, sobre sus aspectos más representativos. Este vector de características es un vector  $n$ -dimensional que contiene elementos cuyos valores describen en forma sucinta el contenido de una imagen. Esta nueva representación de la imagen puede ser almacenada en una base de datos y así permitir una rápida recuperación de imágenes [4]. Existen distintas técnicas que facilitan la generación de este vector representativo que tratan con características de color, textura y forma de los objetos contenidos en la imagen. Una alternativa para facilitar este proceso es utilizar técnicas de transformación de dominio. La principal ventaja de una transformación es su efectiva caracterización de las propiedades locales de la imagen, siendo la transformada de *wavelets* una de las que mejor realizan esta tarea, aprovechando técnicas de multi-escala que implica la descomposición jerárquica de una señal (imagen) a través de la aplicación sucesiva de filtros de transformación.

Hace pocos años atrás que investigadores de las áreas de matemática aplicada y de procesamiento de señales desarrollaron una técnica de transformación de señales, denominada de *wavelet* para la representación multi-escala y análisis de señales [2]. Las transformadas de *wavelets*, a diferencia de la tradicional técnica de Fourier, localizan la información en el espacio de tiempo-frecuencia, en forma particular, tienen una capacidad para intercambiar una resolución por otra, lo cual hace de ellas especialmente indicadas para el análisis de señales en distintas bandas de frecuencia, donde cada resolución tiene su correspondiente escala [2]. Estas transformadas fueron aplicadas con mucho éxito en aplicaciones como compresión de imágenes, mejoramiento, análisis, clasificación y recuperación [9] [1] [3]. Una de las privilegiadas áreas de aplicación donde estas propiedades fueron encontradas como relevantes es la área médica [8] [7].

En este artículo se describe una estrategia CBIR para base de datos de imágenes médicas a fin de extraer características basadas en la transformación de su dominio a través de filtros de *wavelets*. Aspecto importante de la presente investigación está en como conseguir caracterizar las imágenes de forma que el proceso de recuperación pueda retornar las imágenes visualmente más similares en el dominio de interés.

El artículo está organizado en cinco secciones. La sección 2 realiza una breve introducción a los conceptos y definiciones de las transformadas de *wavelets*. La sección 3 describe la técnica utilizada para la caracterización e recuperación por similitud de imágenes. Los resultados y la evaluación respectiva son presentadas en la sección 4. Finalmente, las conclusiones y recomendaciones para trabajos futuros son propuestos en la sección 5.

## 2 Transformadas de Wavelet

Las transformadas de *wavelet* proporcionan una representación tiempo-frecuencia en forma simultánea, es decir, la pregunta de “qué frecuencias están presentes en la señal?” es reemplazada por “qué frecuencias están presentes en la señal y cuando (o donde) se presentan?”. En el caso de imágenes la señal es discreta y bidimensional <sup>1</sup> siendo que las regiones donde se presentan fluctuaciones altas entre los pixels, implica la presencia de mudanza de regiones o variación en la textura de la imagen. El conocer donde ocurren estas fluctuaciones será de gran importancia en la caracterización, puesto que estas fluctuaciones se presentaran de manera homogénea en imágenes que presenten alguna similitud entre ellas. A través de la aplicación de las transformadas de *wavelets* estas fluctuaciones son adecuadamente caracterizadas.

Para un mejor entendimiento y fácil implementación algorítmica de las transformada de *wavelet* Mallat propone un abordaje a través del análisis de multi-resolución [5]. El análisis de multi-resolución permite estudiar una señal o una función en diferentes escalas. Se considera el significado de “escala” de la misma forma como es usado en términos no matemáticos. Una analogía muy útil es considerar la observación de un objeto a diferentes distancias,

---

<sup>1</sup> cada fila de pixels conforma una señal discreta unidimensional, cada pixel representa un valor discreto entre 0 a 255 en el caso de imágenes de tonos de gris, el total de filas conforman una señal bidimensional, donde cada pixel representa un valor de dos componentes (filas, columna)

visto desde una gran distancia se percibe una representación a gran escala del objeto, en la cual es distinguible solo la información global del objeto (uno no se encuentra en la capacidad de alcanzar a ver los detalles del objeto). Así como cuando uno tiene una mayor aproximación, el objeto es visto a menor escala, y mayores detalles son percibidos. De la misma forma, una señal puede ser analizada en diferentes escalas. Escalas mayores tratan con la información global, que implica la representación de la información a bajo nivel. Escalas menores otorgan una más detallada y menos compacta representación de la señal.

El generar una escala mayor implica crear dos nuevos sub-espacios (descomposición) <sup>2</sup>. El primer sub-espacio es el que corresponde al de baja frecuencia, que es una versión reducida (comprimida) del original; y el segundo sub-espacio corresponde al de alta frecuencia, que contiene los detalles del primero. Una de las razones por las que esta técnica es muy atractiva para el procesamiento de señales es el hecho de que esta descomposición de la señal en dos nuevos sub-espacios se obtiene mediante la aplicación de dos funciones (filtros), función escala (filtro de pasa baja) y función *wavelet* (filtro pasa alta).

La función escala, o función  $\phi(t)$ , representa la señal mediante la traslación de la misma a través de la señal; esto presupone que debe existir una secuencia de números  $\{h\} = \{h_0, h_1, \dots\}$  tal que la siguiente relación se cumple :

$$\phi(t) = \sum_n h_n \phi_1(t - \frac{n}{2}) \quad (1)$$

De la misma manera el sub-espacio de detalles es generado mediante la dilatación y la traslación de una función  $\psi(t)$ , implicando la existencia de  $\{g\} = \{g_0, g_1, \dots\}$  para cumplir la siguiente relación :

$$\psi(t) = \sum_n g_n \phi_1(t - n) \quad (2)$$

La fig. 1 muestra algunas funciones escala  $\phi(t)$  y su correspondiente función wavelet  $\psi(t)$  las cuales son de extensión finita y con área distinta de cero, como característica de este tipo de funciones. Dentro de la literatura, la función  $\phi(t)$  es conocida como *wavelet padre* y la función  $\psi(t)$  como *wavelet madre*, siendo que la función padre determina la tendencia de la señal, y la madre caracteriza los detalles de la misma durante el proceso de descomposición. En la fig. 2 puede apreciarse (a) la función escala  $\phi(t)$ , (b) representación de una señal análoga y (c), (d) las señales de aproximación.

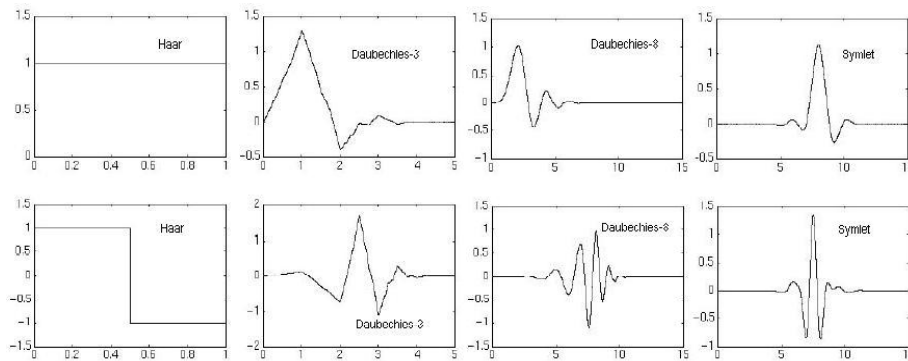


Figure 1: Diseño de algunas *wavelets* analizadoras. Primera columna: *wavelets* padre  $\phi(t)$ . Segunda fila : *wavelets* madre,  $\psi(t)$ . [9].

Una transformada de *wavelet* puede ser aplicada a una una señal multi-dimensional simplemente mediante la aplicación de una transformación 1D a través de todas las dimensiones en forma sucesiva. Una aplicación común de las *wavelets* para señales 2D es en el procesamiento de imágenes. Por ejemplo, la fig. 3 muestra (a) una imagen (b) su primer nivel de descomposición *wavelet* y (c) su segundo nivel de descomposición. Verifícase que en el sub-espacio <sup>3</sup> *LL* se encuentra la versión reducida de la imagen original, en el primer nivel de descomposición; los sub-espacios *HL*, *HH*, *LH* representan la información de detalle en diferentes direcciones (horizontal, diagonal y vertical respectivamente). De la misma forma los sub-espacios *LLLL*, *LLHL*, *LLHH*, *LLLH* son los correspondientes en la siguiente escala de descomposición.

<sup>2</sup>en el caso de generar una escala menor precisaremos de dos sub-espacios para obtener uno nuevo (composición)

<sup>3</sup>denotados tambien como sub-bandas

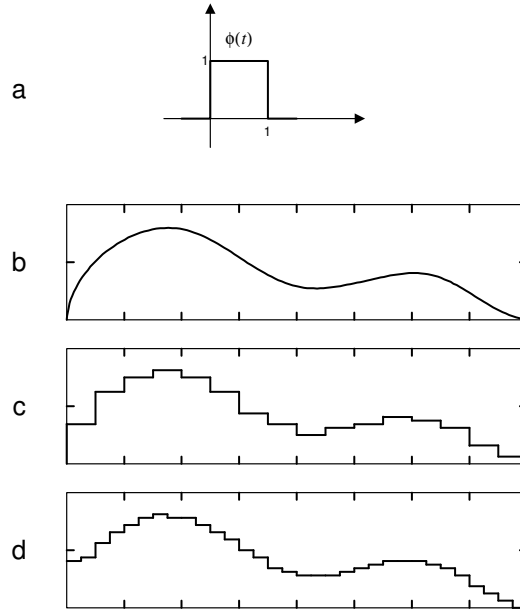


Figure 2: (a) Función escala de Haar. (b) Ejemplo de una señal analógica. (c), (d) Versión de aproximación en escalas 0 y 1 de la señal en (b) usando la función de escala en (a).

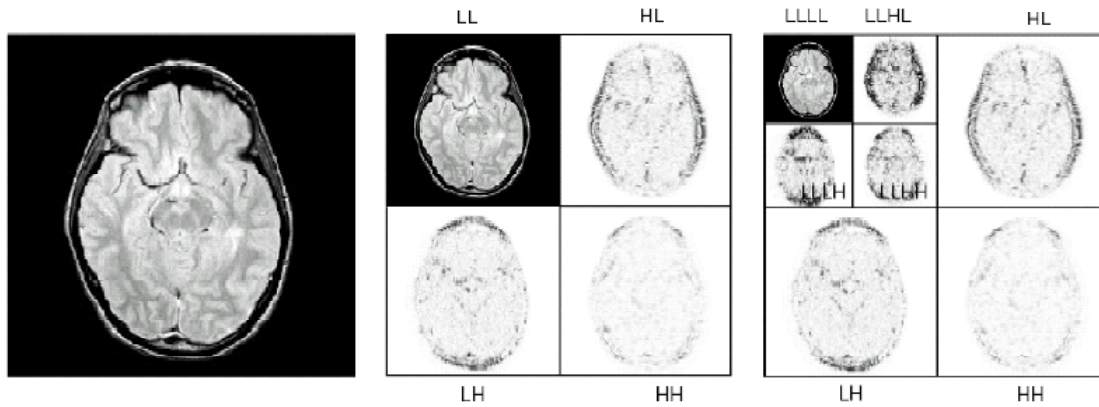


Figure 3: Descomposición *Wavelet* 2D. (a) Imagen original. (b) Primer nivel de descomposición (c) segundo nivel de descomposición.

### 3 Caracterización de Imágenes

Para caracterizar cada imagen, la estrategia seguida aplica el proceso de multi-resolución basado en los filtros de *wavelets*, seguidamente los vectores de característica son calculados considerando las características globales en cada sub-banda de *wavelet*. La presente sección muestra los filtros usados para el proceso de generación de los vectores de característica y la técnica usada para la caracterización de la imagen.

#### 3.1 Filtros de *Wavelet*

La transformada de *wavelet* tiene varias propiedades que la hacen muy atractiva para construir una representación natural de señales e imágenes. Estas propiedades incluyen localización, multi-resolución y compresión [1]. En los experimentos desarrollados, fue usado la base de filtros de *wavelet* correspondientes a la Daubechies4. Para analizar los detalles en una escala determinada, se define una base ortonormal  $\psi_j(x)$  con propiedades similares a aquellas las de  $\phi_j(x)$  descritas anteriormente. Las funciones  $\phi_j(x)$  y  $\psi_j(x)$  son denominadas de *wavelet padre* y *wavelet madre* respectivamente. Estas son las funciones *wavelet* prototipo requeridas para el análisis de *wavelet* [9].

Debido a que las funciones base de *wavelet* otorgan derivaciones continuas, durante este proceso se puede apreciar

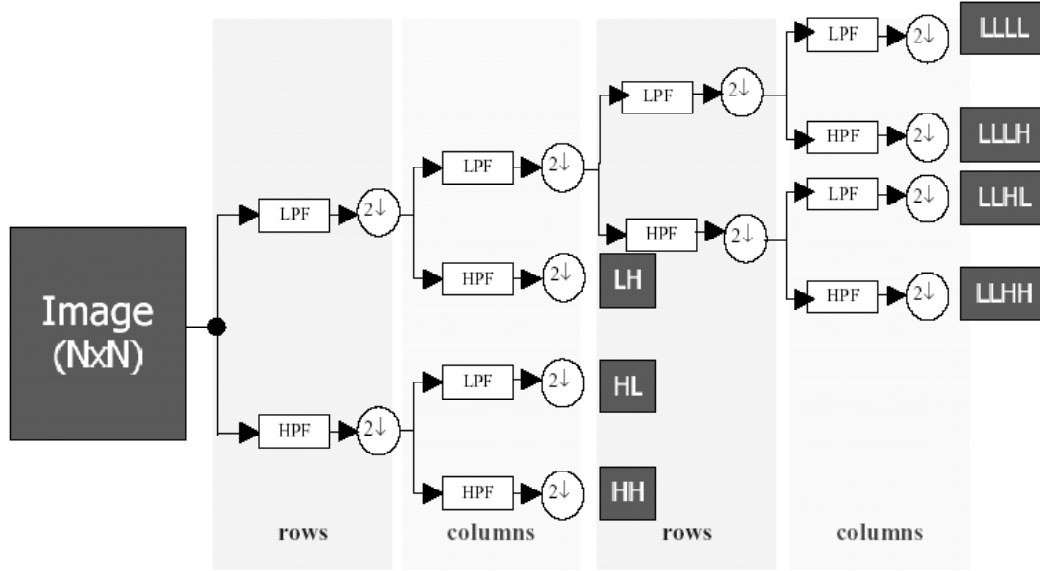


Figure 4: Descomposición jerárquica a través de los filtros de pasa baja (*Low Pass Filter-LPF*) y pasa alta (*High Pass Filter-HPF*) [1].

que las funciones *wavelet madre* son usadas para caracterizar los detalles en una señal, y la información de tendencia es almacenada en los coeficientes obtenidos mediante la *wavelet padre*. La representación de la *wavelet* de Daubechies de una función es una combinación lineal de las funciones base de *wavelet*.

La transformada de *wavelet* Daubechies4 es usualmente implementada numéricamente mediante filtros de espejo cuadrático [6] [2]. El análisis de multi-resolución de la tendencia y fluctuación de una función es implementada mediante convoluciones con un filtro pasa-baja (Eq. 3) y un filtro pasa-alta (Eq. 4) que son versiones de la misma *wavelet* definidas mediante las secuencias :

$$h = \left\{ \frac{(\sqrt{3} + 1)}{4\sqrt{2}}, \frac{\sqrt{3}(\sqrt{3} + 1)}{4\sqrt{2}}, \frac{(\sqrt{3} - 1)}{4\sqrt{2}}, \frac{\sqrt{3}(\sqrt{3} - 1)}{4\sqrt{2}} \right\} \quad (3)$$

$$g = \left\{ -\frac{\sqrt{3}(\sqrt{3} - 1)}{4\sqrt{2}}, -\frac{(\sqrt{3} - 1)}{4\sqrt{2}}, \frac{\sqrt{3}(\sqrt{3} + 1)}{4\sqrt{2}}, -\frac{(\sqrt{3} + 1)}{4\sqrt{2}} \right\} \quad (4)$$

Donde  $\{h\}$  es la secuencia de escalas y  $\{g\}$  se refiere a la secuencia de *wavelet*, descritas en la sección 2.

### 3.2 Generación de Vectores de Características

Para la generación de vectores de características se aplicó tres niveles de descomposición *wavelet*, generando diez sub-espacios (nueve sub-espacio de alta frecuencia y uno de baja frecuencia). El sub-espacio de baja frecuencia corresponde a la versión comprimida de la imagen original. En la implementación seguida, los parámetros específicos extraídos de cada sub-espacio de alta frecuencia fueron *media*, *energía*, *entropía*, *desviación estandard*, generando nueve valores por cada tipo de característica. La fig. 5 muestra la generación del vector de características, donde  $F_n$  corresponde a los valores característica de cada uno de los respectivos sub-espacios.

En el presente artículo, se presenta la caracterización de imágenes basado en la generación de vectores de característica, para de esta forma orientar la indexación de estos espacios de características a través de alguna estructura de datos multi-dimensional.

Como fue descrito en 3.1, los algoritmos de indexación aplican la transformada de *wavelet* rápida con la versión especial del conjunto de *wavelets* de Daubechies4. Los coeficientes de la transformada de *wavelet* en las bandas de baja frecuencia normalmente representan configuraciones de objetos dentro de las imágenes y la banda de alta frecuencia representa la textura y la variación local de color.

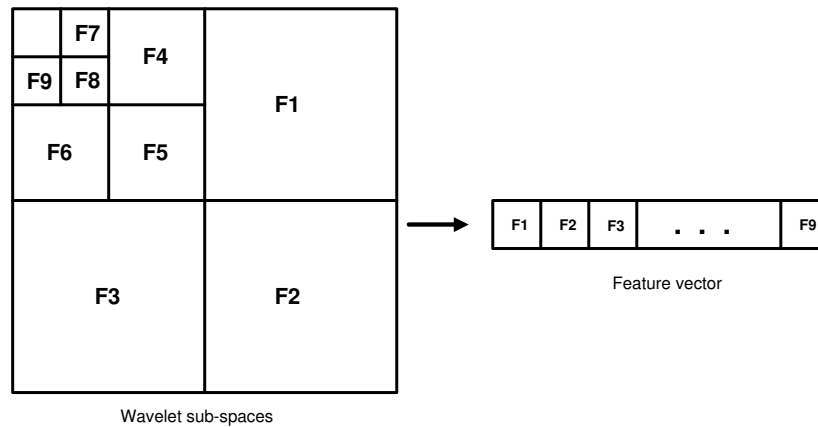


Figure 5: Generación del vector de características a partir de los sub-espacios de *wavelet*.

La búsqueda está basada en la similitud semántica que presentan las imágenes, por lo que imágenes del mismo dominio semántico serán casadas. Por ejemplo, dada una imagen de cerebro, otras imágenes de cerebro serán recuperadas y mostradas como respuesta. Para recuperar las imágenes relacionadas, la búsqueda es realizada en dos pasos, en el primer paso se genera el vector de características basado en la desviación estándar, media, energía y entropía para cada sub-banda. En el segundo paso es calculada la distancia Euclideana entre cada una de las imágenes del banco de datos y la imagen de consulta. Así las distancias son ordenadas y las primeras “*n*” distancias más cortas son recuperadas.

## 4 Experimentos y Resultados

La real prueba de un sistema CBIR es si efectivamente es utilizado por los interesados en la práctica, en nuestro caso si las consultas satisfacen el requerimiento del médico. Para demostrar la factibilidad de los vectores generados, fue desarrollado un pequeño prototipo a fin de visualizar los resultados. El conjunto de imágenes corresponde a grupos de exámenes de tomografía donde imágenes consecutivas presentan cambios muy pequeños.

La primera prueba efectuada fue desarrollada separando ocho grupos de seis imágenes cada una. Luego de realizadas las consultas, el sistema mostró que una imagen de consulta es recuperada junto a las otras imágenes que son parte del mismo grupo, la fig. 6 presenta una consulta con las imágenes separadas en grupos.

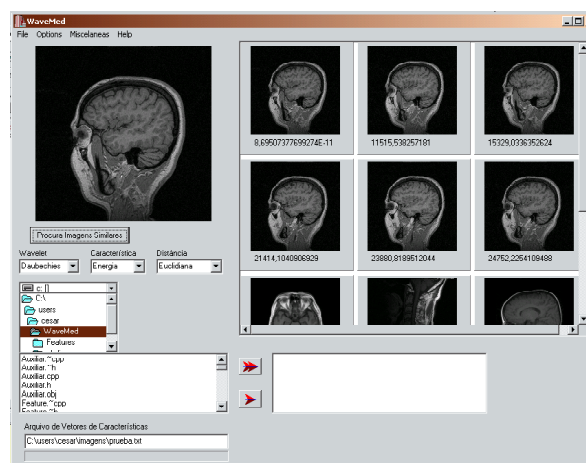


Figure 6: Resultado de la consulta, donde las seis primeras imágenes corresponden al mismo grupo de la imagen de consulta.

La segunda prueba se realizó sobre un conjunto de 1200 imágenes médicas correspondientes a diferentes grupos semánticos. La respuesta para cada consulta satisface el resultado esperado, excepto cuando la imagen de consulta no

es parte del dominio.

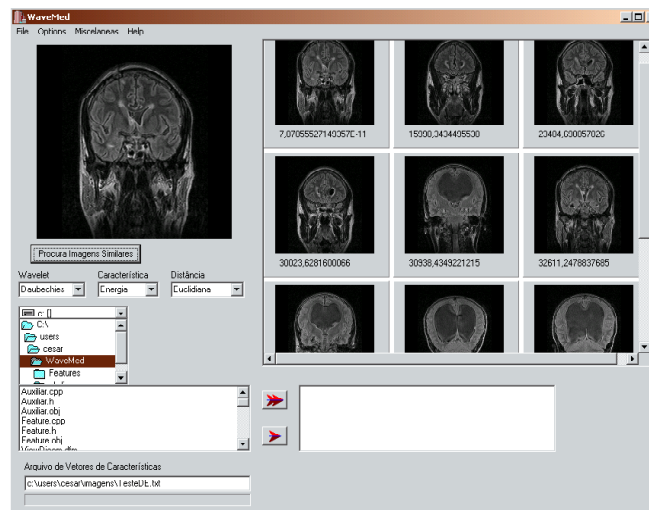


Figure 7: Resultado de una consulta dentro del conjunto de 1200 imágenes.

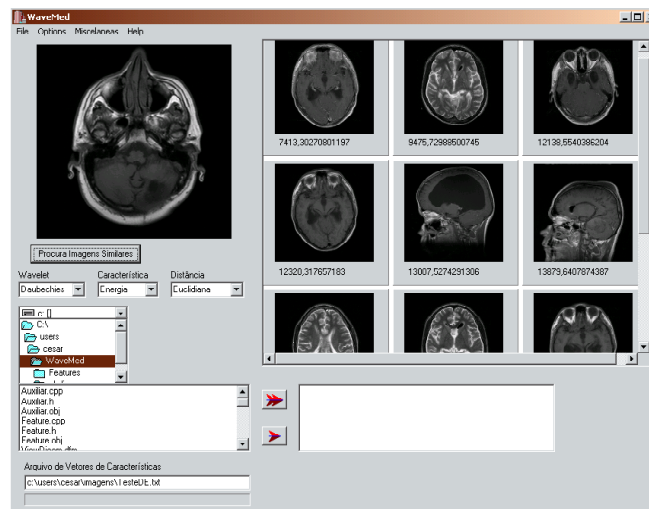


Figure 8: Resultado de una consulta dentro del conjunto de 1200 imágenes donde la imagen de consulta no pertenece al dominio.

#### 4.1 Características de Desempeño

El sistema fue implementado en una PC Pentium III 866MHz con RAM de 128MB, sistema operativo Windows2000. Para calcular los vectores de características de 1000 imágenes de tamaño 256 x 256 en nuestra base de datos de imágenes médicas se requirió aproximadamente de 18 minutos; en promedio, fueron necesarios 1,12 segundos para aplicar la transformada de *wavelet* y calcular las características de todas las regiones.

El proceso de casamiento de imágenes similares es rápido. La diferencia entre un conjunto de 100 imágenes y otro de 1000 imágenes es de escasos milisegundos. El proceso que toma el mayor tiempo de CPU corresponde al cálculo y ordenamiento de distancias. Debido a que en el presente trabajo no es usado alguna metodología de indexación multi-dimensional, el tiempo de casamiento crece linealmente a medida que el número de características por vector y la cantidad de imágenes a procesar también crezca.

## 5 Conclusiones y Trabajo Futuro

En el presente artículo fue presentado el uso de las transformadas de *wavelet* para caracterizar imágenes médicas con el objeto de luego poder recuperarlas en base al contenido de las mismas. Los vectores de características fueron generados sobre las bandas de alta frecuencia generados a partir de la descomposición en tres niveles. Es interesante resaltar que estos sub-espacios generados presentan la propiedad de ortogonalidad entre ellos, lo cual permite la existencia de características no correlacionadas. Los experimentos del sistema implementado mostraron que la performance de recuperación fue aceptable. Puede afirmarse que el presente trabajo muestra la potencialidad de las *wavelets* para caracterizar imágenes médicas.

El sistema puede ser mejorado pudiendo ser más preciso mediante la adición de características que permitan las propiedades de textura y forma. Estos serán los siguientes pasos en el desarrollo de trabajos futuros.

## Agradecimientos

Este trabajo está auspiciado por Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ)- Ministério de Ciência e Tecnologia - Governo do Brasil. También agradecer al Grupo de Base de Datos e Imagenes (GBDI) de la Universidad de São Paulo en São Carlos.

## References

- [1] Elif Albuz, Erturk Kocalar, and Ashfaq A. Khokhar. Scalable Color Image Indexing and Retrieval using Vector Wavelets. *IEEE Transaction on Knowledge and Data Engineering*, 13(5):851–861, September/October 2001.
- [2] Ingrid Daubechies. *Ten Lectures on Wavelets*, volume 61. CBMS-NSF Regional Conference Series in Applied Mathematics, Rutgers University and ATandT Bell Laboratories, 1992.
- [3] Charles E. Jacobs, Adam Finkelstein, and David H. Salesin. Fast Multiresolution Image Querying. *Computer Graphics*, 29(Annual Conference Series):277–286, 1995.
- [4] Murray H. Loew. *Feature Extraction*, chapter 5. SPIE, Bellingham, WA, m.sonka and j. michael fitzpatrick edition, 2000.
- [5] Stephane G. Mallat. *Multiresolution Representation and Wavelets*. Ph.d. thesis, University of Pennsylvania, 1988.
- [6] Yves Meyer. *Wavelets : Algorithms and Applications*. Society for Industrial and Applied Mathematics, Philadelphia, first edition, 1993.
- [7] Michael Unser and Akram Aldroubi. A Review of Wavelets in Biomedical Applications. *Proceeding of the IEEE*, 84(4):626–638, April 1996.
- [8] James Ze Wang. *Semantics-Sensitive Integrated Matching for Picture Libraries and Biomedical Image Databases*. Phd. dissertation, Departament of Biomedical Informatics of Stanford University, 2000.
- [9] James Ze Wang. Methodological Review - Wavelets and Imaging Informatics : A Review of the Literature. *Journal of Biomedical Informatics*, pages 129–141, July 2001. <http://www.idealibrary.com>.