

# IMAGE REPRESENTATION USING PHOTONS

Andreea-Mihaela Pintilie<sup>1</sup>

Mihai Zaharescu<sup>2</sup>

Ion Bucur<sup>3</sup>

## ABSTRACT

*Genetic brain maps can help physicians to discover patterns of brain structure and how it changes in disease or reacts to medication. In order to generate a brain map, an image model is needed. A much discussed subject, nowadays, is the subject of improving images obtained from devices like MRI, PET, CT etc. In medical area there is a need to improve image segmentation and image resolution; images might be blurred or might contain noise due to the patient's movement during the process of acquiring them. Imaging studies of the human brain at active medical institutions today routinely accumulate more than 5 terabytes of clinical data per year. The present paper concentrates on neurological field: brain imaging and on genetic field based on the results of the brain imaging. This paper proposes a new medical image format representation, using photons, a tree-like structure in order to improve the inefficiency problem on large medical datasets and an algorithm for eliminating noise from images.*

**KEYWORDS:** Image representation, medical image acquisition, image formats, raster images, vector images, photon images, quad trees, binarization, segmentation, edge detection

## 1. INTRODUCTION

Image representation is important when trying to analyze different images. Nowadays medical image is a very important component when trying to establish a diagnostic, when planning an evaluating surgical and radiotherapeutic / chemotherapeutic treatments. There are different registration methods for medical images, based on the purpose of the investigation the medicinal doctor is conducting for a patient. In the following section we provide a short classification of the registration methods in medical imaging area. [1]

One of the first criterions is the dimensionality. The main division is based on the spatial - time dimensions. We can register two 3D/3D images - obtained from two tomographic data sets with no time dimension or 2D/2D images - obtained as separate slices from tomographic data, or even a 2D/3D registration used for the alignment of spatial data to

---

1 Engineer, Department of Computer Science and Engineering, Faculty of Automatic Control and Computers Science, University "Politehnica" of Bucharest, Splaiul Independenței 313, Bucharest, 060042, Romania, andreea.pintilie@cti.pub.ro

2 Engineer, Department of Computer Science and Engineering, Faculty of Automatic Control and Computers Science, University "Politehnica" of Bucharest, Splaiul Independenței 313, Bucharest, 060042, Romania, mihai.zaharescu@cti.pub.ro

3 Associate Professor PhD Eng., Department of Computer Science and Engineering, Faculty of Automatic Control and Computers Science, University "Politehnica" of Bucharest, Splaiul Independenței 313, Bucharest, 060042, Romania, ion.bucur@cs.pub.ro

projective data, where the first image can be a CT one, while the second one a X-ray. Time series of images are important when trying to monitor a tumor growth or the effectiveness of a treatment as a pre or post surgical monitoring of healing in different time intervals (short/medium/long).

Registration can furthermore be divided into extrinsic registration or intrinsic methods. When acquiring an extrinsic image usually artificial, foreign, invasive objects are used; less invasive objects or even non-invasive markers can be used but with lower accuracy. As an example, before a neurosurgery operation, a stereotactic frame can be used for localization and guidance.

Intrinsic methods rely mainly on patient generated image content only. Landmarks for example can be locatable points of the morphology of the visible anatomy or geometrical points providing geometric property such as corners, local curvature extremes etc. Usually, the set of identified points is compared to the original image content and they are used to measure distance. Another intrinsic method is segmentation: same anatomical structures are extracted from two different images and used for the alignment procedure. One image can be elastically deformed in order to be compared with the second image: can be zoomed in or out. The voxel property registration method is different from the other two intrinsic methods described above because the method is applied on the image gray values.

Ironically, another registration method is the non-image based one. In this type of registration, two devices are calibrated to each other and that the patient must stay motionless between the first and the second image acquisition.

Another classification of the registration task is based on the modalities that are involved. In monomodal registration - the images belong to the same modality, while in multimodal the images are provided from two different modalities. In monomodal method only one image is involved and the other one is the patient himself or a model. As an example, when trying to analyze the myocardial dysfunction in a patient two images can be acquired: under rest and under stress condition. In multimodal method an image obtained from a PET device can be registered to an MR image in order to relate an area of dysfunction to anatomy or the registration of an MR brain image to a defined model of gross brain structures.

When a single patient provides the images involved, the method is referred as intrasubject registration, when accomplished using images from distinct patients (or a model and a patient) this is known as intersubject registration; when one image is constructed from a database, an ontology or a knowledge base and the other one is provided by a single patient, the method is called atlas registration.

All these image registration methods require different transformations:

- rigid: when only translation and rotations are allowed
- affine: if the transformation maps the parallel lines onto parallel lines
- projective: transformation maps lines on lines
- curved: transformation maps lines to curves
- global: if the transformation is applied to the entire image

- local: if subsection of the image have their transformations defined

In general rigid and affine transformations are used as global transformations and curved transformations are local. Affine transformations are used in rigid body movements, when the scale factors applied on the image are suspected to be incorrect or unknown: as an example, during an MR image acquisition of the lungs: the patient might breath - which will modify the lungs volume and will provide a distorted image. The projective transformation type can be used in 2D/3D image registration as a constrained elastic transformation when a fully elastic transformation behaves inadequately.

## 2. MOTIVATION

In medical image acquisition it is important to have a history in order to establish the rate of progress / regress of a disease. Therefore, it is important to be able to map or to compare different images containing the same anatomical structure.

The image file format is defined as the standardized means of organizing and storing digital images. Image files are composed of either pixel or vector (geometric) data that are rasterized to pixels in graphic display (such as monitor, print and other devices). The pixels that constitute an image are ordered in a grid; each pixel has magnitudes of brightness (expressed by gray-scale value) and color. The file size increases directly with the number of pixels composing an image and the color depth of the image: the greater the number of columns and rows, the greater the image resolution will be, hence the size. The size also increases when the pixels color depth increases: an 8 bit pixel will store 256 colors, while a 24-bit pixel will store 16 million colors. The compression method associated with the file format is also important when trying to establish the image size. The images are classified into three families of graphics: raster, vector and metafile (which combine raster and vector information).

From the common image file formats we mention **JPEG** (compression by eliminating non-visible frequencies, using the cosine transform, and storing color with fewer bits than luminance) **GIF** (format limited to an 8-bit palette (256 colors); they are suitable for storing graphics with relatively few colors like shapes, diagrams and not for natural images) **TIF** (saves 8 bits or 16 bits/color for a 24 and 48-bit total) **PNG** (provides simple detection of common transmission errors as well as full file integrity checking. The format can support up to 16 million colors in a lossless format).

## 3. RAW (RAW IMAGE FORMAT)

The RAW formats use a lossless or nearly lossless compression and produce full size processed images from the same cameras. ISO 12234-2 represents a standard for the raw image format, but most cameras are not standardized or documented. Raw images are also known as digital negatives in film photography. After the acquisition, the image is processed by a raw converter to "positive" file format for further manipulation, which often encodes the image in a device-dependent colorspace. Many medical devices use RAW data format such as Metafile.

Metafile represents a generic term for a file format that can store multiple types of data: raster, vector and even type data. The common use is to provide support for the OS computer graphics.

### **DICOM (Digital Imaging and Communications in Medicine)**

DICOM represents a standard for storing, handling, transmitting and printing medical imaging. Besides being an image format standard, it also represents a network communication protocol. The protocol uses TCP/IP for communication between systems. The DICOM data format contains a structure which allows maintaining not only image information, but as well as other patient information and medical identification. A DICOM data object consists of a number of attributes, including items such as name, ID, etc and also one special attribute containing the image pixel data. A single DICOM object can only contain one attribute containing pixel data. Pixel data can be compressed using a variety of standards, including JPEG, JPEG Lossless, and JPEG 2000.

When acquiring the images the time and the spatial dimensions are important. As a consequence algorithms that map a 3D image onto a 2D image must be used. Hence, images must support transformations. We decided to provide a new image format that can support different geometrical transformation and compression and that will maintain its quality when geometrical transformations are applied.

## **4. OBJECTIVES**

This paper is mainly concerned with finding an image format that can preserve the image quality event after different geometric transformations are applied on the image. A good example would be a patient who is having multiple RMI, representing different stages in the tumor's evolution. In order to provide a probability based on the scan images for the malignity of the tumor it is necessary to analyze the images, but analyzing images that have different dimensions, different intensities or that might even contain the object of interest in different positions can be a very difficult task.

A map reduce algorithm will be used in order to improve the algorithm computational speed. In order to extract different objects from the image, a canny edge detection algorithm will be used.

Another interesting aspect is extracting the tumor from a Pet or RMI image. The main drawback would be sensitivity to false edges, but usually the intensities of the tumor cells are higher than the ones of the normal brain cells and they can be better preserved after considering the luminosity of the pixels around them.

## 5. IMAGE REPRESENTATIONS

### Image representation based on spatial relationships

In article “Similarity Searching for Chest CT Images Based on Object Features and Spatial Relation Maps”[2] is presented an object base image retrieval system for CT images. The paper proposes an image segmentation method which combines the anatomical knowledge of the chest and the well-known watershed segmentation algorithm. The purpose of segmentation is to identify the mediastinum and the two lung lobes in a chest CT image. The proposed solution has a first step the construction of ARGs (attributed relational graphs) which are used to describe the features of segmented objects.

Traditional text-based image retrieval system retrieves images by text keywords. The keywords used cannot reliable describe the features of the images, such as shapes and textures.

Several methods have been proposed to approach the CBIR content-based image retrieval problem. To archiving and retrieving images by content, the images need to be: analyzed, indexed and stored to database. Analysis of images can be divided into several steps, like image segmentation, feature extraction[14][15], etc. The contents of a medical image are complex, therefore automatic image segmentation cannot be easily achieved. Features of images can be color, shape and texture of image content [3]. Because most medical images are represented in gray levels, shape and texture features are more often used in describing medical images. Moreover, the spatial relationships among different objects were proposed to further handle multiple objects in an image [4]. In order to achieve a good representation of separated objects in an image, different strategies should be considered for different kinds of images, especially in processing of analyzing medical images acquired from different devices and/or from different parts of the body.

The first step is the preprocessing module. The CT image files store the pixel-values as CT numbers. The CT number represents the relative X-ray absorption properties of different organs or tissues with respect to water [5]. Most physiological organisms have positive CT numbers. The air has a negative CT number of -1000. In a typical chest CT image, air occupies much of the space in the lung lobes, hence the property to separate lung lobes (and air) from other parts in the images.

The second step is the Modified watershed segmentation algorithm. The method is based on the geographic phenomenon of water flooding up in area with hills and valleys: the water is filling up from the lowest to the highest valleys. The borders between two neighboring valleys are detected when the water in them merge. The watershed segmentation algorithm proposed is the following:

1. The input pixel values are sorted. The purpose is to access the pixels in the same gray level efficiently.
2. Different regions are labeled when the flooding step is processed.
3. The image is segmented in small regions.
4. A region merging step is proceeded to acquire meaningful objects out of enormous small segments.

The article proposes 4 merging steps:

1. Merging of the air part with the background: merge regions whose mean pixel values are less than threshold  $T$  and are contact with the image background.
2. Merging of regions with similar CT numbers: merge two neighboring regions if the difference in CT number is less than a threshold of 40.
3. Merging of lung lobes: merge regions surrounded by the labels as lung lobes.
4. Merging of mediastinum: merge regions positioned between the sternum, spinal cord, and two lung lobes.

After the four merging steps, the three objects (mediastinum, left and right lung lobes) in the chest are identified. After the segmentation step, the objects can be labeled. A spatial-relationship model was used to describe the object features and the relations among different objects. Three properties are used to describe the spatial relationships between two objects:

1.  $D$  (Distance): the minimum distance between the two objects
2.  $P$  (Relative Position): the angle between the connection of the two mass centers and the horizontal line.
3.  $R$  (Ratio): the ratio of the areas of the two objects

### Vector Image representation

Vector Image representation is important because of the vector-based graphical content properties: of being editable and scalable. The main trend in vector image representation is developing scalable (resolution-independent) representations of full-color raster images. In a usual path the image plane is decomposed in triangular patches with potentially curved boundaries, and the color signals over the image plane are handled as height fields. The image representation is based on piecewise smooth subdivision surfaces. The image is modeled with discontinuity curves for providing the continuity conditions useful in vectorization and subsequent vector editing operations.

Vectorization techniques:

- **Triangulation:** In representations based on triangulation, each curvilinear feature is approximated by multiple short line segments. This form of representation is not resolution independent because the difference between a smoothly curved feature and a polyline with only  $C0$  continuity at the vertexes become obvious when zoomed in. A technique to overcome the problem is by fitting subdivision curves to patch boundaries that have  $C2$  continuity.
- **Parametric patches:** When grids or Bezier patches are involved, it is necessary to use techniques that are based on higher-order parametric functions. An example of a vectorization technique based on optimized gradient meshes is manually aligning mesh boundaries with salient image features. The main drawback is the fact that mesh placements can be time-consuming. Meshes might introduce color discontinuities that can be approximated using degenerate quads, but the real challenge is to align color discontinuities with the image features. [7]
- **PDE solutions:** The PDE solution is a mesh-free image representation. The technique is based on diffusion curves [6] and it relies on curves with color and blur attributes as boundary conditions of a diffusion process. The color variations

of a vector image are represented by the final solution of the diffusion process. The main limitation is that the curves are not coupled together; therefore it is difficult to edit the image using region-based color or shaping operations.

### **Quadtree representation**

Quad trees are mainly used for image representations because they reduce the storage space of images and the time required for image manipulations. The quad tree structure is efficient to store 2D images. Maintaining a quad tree as an image representation makes it easier to cluster images based on their characteristics such as color, shape, semantics or texture or even by their history. To represent an image by a quadtree the image is recursively split in four disjoint quadrants or squares having the same size. The root node represents the initial quadrant containing the whole image. Different criteria can be used to define the homogeneity property: pixel color, same texture etc. If the image is not homogeneous (according to the criterion established above), the quadtree root has for descendant nodes which represent the four first levels of the image: northwestern, northeastern, southwestern, and southeastern. A node is a leaf when its corresponding image is homogeneous. Generally the quad tree is unbalanced.

Each quad tree node has a key, or a location code/quad code. The quad tree node and the corresponding image have the same key.

The main operations applied on quad trees are reading and modifying an image quad tree, comparing two quad trees and determine the Q-similarity distance. When trying to modify an image quad tree, one can use the complement operation: each leaf node is accessed and has its value changed. As an example, assuming that the image has a total of 256 different intensity levels to represent a pixel  $p$ , the value of  $p$  will be changed with  $(255 - p)$ . In the case of a binary image the black and white values are interchanged. When comparing two quad trees, nodes with the same identifier are actually compared. The basic operations that can be applied on two quad trees are: union, intersection, comparison and difference). When trying to compare a gray-scale or colored image, a similarity measure based on image features is necessary. The node values might be considered as similar if the similarity distance between those two nodes is under a given threshold. The Q similarity distance is defined as the distance between two quad trees. It represents the number of nodes having the same identifier and different values divided by the cardinal of the union of node identifiers.

## **6. EDGE DETECTION**

ED represents a terminology in image processing, which refer to algorithms that have the scope to identify points in a digital image at which the image brightness changes sharply or has discontinuities. The next section will present some of the methods used for edge detection.

### **Robert's Cross Operator**

The Robert's cross operator performs a simple, quick to compute, 2 – D spatial gradient measurement on an image. It thus highlights regions of high spatial frequency which often corresponds to edges. Pixel values at each point in the output represent the estimated

absolute magnitude of the spatial gradient of the input image at that point. In theory the operator consists of a pair of  $2 \times 2$  convolution kernels as shown below. One kernel is simply the other rotated by  $90^\circ$ .

$$G_x = \begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix} \quad G_y = \begin{bmatrix} 0 & +1 \\ -1 & 0 \end{bmatrix}$$

These kernels are designed to respond maximally to edges running at  $45^\circ$  to the pixel grid, one kernel for each of the two perpendicular orientations. The edge gradient magnitude has the formula:

$$|G| = \sqrt{G_x^2 + G_y^2} \cong |G_x| + |G_y|$$

The angle of orientation of the edge giving rise to the spatial gradient is given by:

$$\theta = \arctan\left(\frac{G_x}{G_y}\right) - \frac{3\pi}{4}$$

### Sobel Operator

The Sobel operator performs a 2-D spatial gradient measurement on an image and typically it is used to find the approximate absolute gradient magnitude at each point in an input image. In theory the operator consists of a pair of  $3 \times 3$  convolution kernel as shown below.

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

These kernels are designed to respond maximally to edges running vertically and horizontally relative to the pixel grid, one kernel for each of the two perpendicular orientations.

The edge gradient magnitude is given by:

$$|G| = \sqrt{G_x^2 + G_y^2} \cong |G_x| + |G_y|$$

The angle of orientation of the edge giving rise to the spatial gradient is given by:

$$\theta = \arctan\left(\frac{G_x}{G_y}\right)$$

### Prewitt Compass Edge Detector

Compass edge detection is an alternative approach to the differential gradient edge detection. When using compass edge detection the image is convolved with a set of (in general 8) convolution kernels each of which is sensitive to edges in a different orientation.

For each pixel the local edge gradient magnitude is estimated with the maximum response of all 8 kernels at this pixel location.

$$|G| = \text{Max}(|G_i|; i = 1 \text{ to } n)$$

Where  $G_i$  is the response of the kernel 'i' at the particular pixel location and 'n' is the number of convolution kernels. The local edge orientation is estimated with the orientation of the kernel that yields the maximum response. Two templates out of the set of 8 for the Prewitt kernel is given below.

$$\begin{bmatrix} -1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & 1 & 1 \end{bmatrix} \quad \theta = 0^\circ \qquad \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & -1 & 1 \end{bmatrix} \quad \theta = 45^\circ$$

The whole set of 8 kernels is produced by taking one of the kernels and rotating its coefficients circularly. Each of the resulting kernels is sensitive to an edge orientation ranging from 0 to 315 insteps of 45, where

0 corresponds to a vertical edge.

### Zero Crossing Detector

The zero crossing detector looks for places in the Laplacian of an image where the value of the Laplacian passes through zero, i.e. points where the Laplacian changes sign. Such points often occur at edges in images, i.e. the points where the intensity of the changes rapidly, but they also occur at places that are not as easy to associate with edges. It is best to think of the zero crossing detector as some sort of feature detector rather than as a edge detector.

The core of the zero crossing detector is the Laplacian of Gaussian filter. Once the image has been Laplacian of Gaussian filtered, it only remains to detect the zero crossings. This can be done in several ways. The simplest is to threshold the Laplacian of Gaussian output to zero and a more accurate approach is to perform some kind of interpolation to estimate the position of the zero crossing to sub-pixel precision.

### Canny Edge Detector

The Canny operator was designed to be an optimal edge detector. The Canny operator works in a multistage process. First of all the image is smoothed by Gaussian convolution, then a derivative operator is applied. The resultant edges pass through a first threshold and in order to connect broken edges, pixels that are able to fill the spaces are passed through a second, lower, threshold. Using the edge directions, the edges are thinned to a pixel wide. The effect of the Canny operator is determined by three parameters: the width of the Gaussian kernel, the lower and the upper thresholds used by the tracker.

## BINARIZATION

A binary image is an image that only uses two values for its pixel values. Binary images are often used in digital image processing as masks and are the result of certain operations such as edge detection, segmentation, thresholding. The following section will present different binarization methods [13].

- **Histogram-based methods:** These methods determine the binarization threshold by analyzing the shape properties of the histogram, such as the peaks and valleys. Pavlidis [11] constructs a histogram by using gray-image pixels with significant curvature, or second derivative, and then selects a threshold based on the histogram.
- **Clustering-based methods:** The threshold is selected by partitioning the image's pixels into two clusters at the level that maximizes the between-class variance, or minimizes the misclassification errors of the corresponding Gaussian density functions [9].
- **Entropy-based methods:** These methods employ entropy information for binarization [8].
- **Object attribute-based methods:** These methods select the threshold based on some attribute quality (e.g., edge matching of Hertz and Schafer [10]) or the similarity measure between the original image and the binarized image.
- **Spatial binarization methods:** These methods binarize an image according to the higher-order probability or the correlation between pixels.
- **Hybrid methods:** Local, adaptive methods [12]; algorithms for improving mentioned methods [13]

## 7. IMPLEMENTATION

The project implementation is in Java using Hadoop Map Reduce for image processing.

The first step of the algorithm was to apply the Canny edge detection algorithm on a image. Then we used an entropy-based binarization algorithm as presented in Binarization section to remove the noise. We spread the photons on the edges using an iterative algorithm. The photons were spread based on the fact that the edges will absorb the light, therefore the photons will model the edges. The image processing was done using map reduce to improve the computational speed. We also maintained a quad tree containing the image which provides access with a higher level of detail once the tree is traversed.

## 8. ACKNOWLEDGEMENT

The authors would like to thank Costin-Anton Boiangiu for his original idea (*paper under review*) onto which the described system is based, for support and assistance with this paper.

## 9. REFERENCES

- [1] J.B.A. Mainz, M.A. Viergever: "A survey of medical image registration", *Conf Proc IEEE Eng Med Biol Soc.* 2004;2:1298-301.
- [2] Yu SN, Chiang CT, "Similarity searching for chest CT images based on object features and spatial relation maps".

- [3] Yihong Gong Intelligent Image Databases: Towards Advanced Image Retrieval, Robotics Institute, Carnegie Mellon University.
- [4] Euripides G.M. Petrakis, "Design and evaluation of spatial similarity approaches for image retrieval", *Image and Vision Computing*, Volume 20, Issue 1, 1 January 2002, Pages 59–76.
- [5] Louis M. Castanier, "An Introduction To Computerized X-Ray Tomography For Petroleum Research", May 1989, Stanford University Petroleum Research Institute.
- [6] Alexandrina Orzan, Adrien Bousseau, Holger Winnemoller, Pascal Barla, Joelle Thollot, David Salesin, "Diffusion Curves: A Vector Representation for Smooth-Shaded Images".
- [7] Tian Xia, Binbin Liao, Yizhou Yu, "Patch-Based Image Vectorization with Automatic Curvilinear Feature Alignment".
- [8] J.N. Kapur, P.K. Sahoo, A.K.C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram", *Computer Vision, Graphics, and Image Processing*, Volume 29, Issue 3, March 1985, Pages 273–285.
- [9] J. Kittler, J. Illingworth, "Minimum error thresholding", *Pattern Recognition*, Volume 19, Issue 1, 1986, Pages 41–47.
- [10] Lois Hertz, Ronald W. Schafer, "Multilevel thresholding using edge matching", *Journal Computer Vision, Graphics, and Image Processing archive*, Volume 44 Issue 3, Dec. 1988, Pages 279-295.
- [11] Pavlidis, T. "Threshold selection using second derivatives of the gray scale image", *Document Analysis and Recognition*, 1993.
- [12] Costin-Anton Boiangiu, Alexandra Olteanu, Alexandru Victor Stefanescu, Daniel Rosner, Alexandru Ionut Egner (2010). „Local Thresholding Image Binarization using Variable-Window Standard Deviation Response" (2010), *Annals of DAAAM for 2010, Proceedings of the 21st International DAAAM Symposium*, 20-23 October 2010, Zadar, Croatia, pp. 133-134.
- [13] Costin-Anton Boiangiu, Andrei Iulian Dvornic, Dan Cristian Cananau, „Binarization for Digitization Projects Using Hybrid Foreground-Reconstruction", *Proceedings of the 2009 IEEE 5th International Conference on Intelligent Computer Communication and Processing*, Cluj-Napoca, August 27-29, pp.141-144, 2009.
- [14] Costin-Anton Boiangiu, "The Beta-Shape Algorithm for Polygonal Contour Reconstruction", *CSCS14 – The 14th International Conference on Control System and Computer Science*, Bucharest, July 2003.
- [15] Costin-Anton Boiangiu, Bogdan Raducanu. "3D Mesh Simplification Techniques for Image-Page Clusters Detection". *WSEAS Transactions on Information Science, Applications*, Issue 7, Volume 5, pp. 1200 – 1209, July 2008.