Chapter 1

Data Acquisition and Preprocessing on Three Dimensional Medical Images

Yuhua Jiao¹, Liang Chen², Jin Chen³*

¹MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, MI 48824
yuhjiao@msu.edu

²Department of Neurosurgery, Huashan Hospital, Fudan University, Shanghai, P.R.China 200433
clclcl95@shmu.edu.cn

³MSU-DOE Plant Research Laboratory and Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824
jinchen@msu.edu

Three dimensional (3D) medical imaging technique is a new field that is rapidly evolved over the last years, leading to a major improvement in patient care. Starting from 3D models of the patient, anatomical structures can be identified and extracted, and diagnosis and surgical simulation shall be supported. Given that the computational reconstruction of 3D models is critical for medical diagnosis and treatment, and the complex imaging processing requires considerable resources and advanced training, we introduce 3D image acquisition, segmentation and registration in this chapter, for all of which, the main purpose is to effectively transform the unstructured image data to structured numerical data for further data mining tasks.

1. Introduction

Data acquisition and preprocessing are the essential steps in data mining process, especially when applying data mining to 3D images. The quality of 3D imaging in healthcare fields is inferior to that in other computer vision fields in following aspects: (1) Automatic data preprocessing is highly required to obtain high quality data for biomedical data mining, since

*Corresponding author.
biomedical 3D images collection systems are primarily designed for doctor’s manual diagnosis; (2) Much of healthcare data are often inconsistent or non-standardized, such as pieces of medical 3D images in different formats from different data sources. These data cannot be directly used for data mining without significant efforts on data preparation, including the utilization of two main techniques: 3D image segmentation and 3D image registration.

Image segmentation is the task of partitioning the data into contiguous regions representing individual anatomical objects.\(^1\) It is a prerequisite for further investigations in many computer-assisted medical applications in that it can classify different tissues to different logical classes in medical images. A central problem in 3D medical image segmentation is how to effectively distinguish interested objects from noisy background. To integrate different forms of 3D image-derived information to serve further data mining tasks, 3D datasets need to be automatically registered. 3D image registration is the process of determining the point-by-point correspondence between two 3D images of a scene. Two medical images may differ by any amount of rotation or translation in any direction, and they may also differ in scale. A central problem in 3D medical image registration is to define the search space of the registration problem, ranging from nonlinear transformation with a virtually infinite degree of freedom, to rigid registration with six degrees of freedom.\(^2\) By registering two medical images, the fusion of multi-modality information becomes possible, and regions of abnormal function can be recognized.

2. Three Dimensional Image Acquisition Techniques

Data processing programs based from radiologic examinations have been widely adopted in clinical application to better visualize specific structures or region of interests in the human body.\(^3\) One of the best examples of such applications is to demonstrate the intracranial arteries and veins in 3D or even four dimensional patterns.\(^4\) That is extremely valuable for doctors to find out the vascular abnormalities and give treatment.

Catheterized cerebral angiography\(^5\) is the most often used imaging technique under plane film x-rays to demonstrate the intracranial vessels inside patients. A catheter is placed in the bilateral external carotid artery, internal carotid artery and the vertebral artery one by one via femoral artery or radial artery puncture. Then a jet of contrast medium from the catheter will make its vascular pathway opaque to x-ray, thus we can see the ven-
Data Acquisition and Preprocessing on Three Dimensional Medical Images

sel shadow in the x-rays machine. Vascular abnormalities in the neck and brain will be realized after all the 6 vessels have been catheterized and angiographed. But the irregular skull base bony structures and the bone fissures will also be demonstrated as linear or massive shadows, making the angiograph obscure. The resulted ambiguous diagnosis would put the treatment in dilemma. To properly settle this problem, digital subtraction angiography (DSA) has been developed.

X-ray features both before and after the jet of contrast medium are acquired and stored in the digital manner. Then the shadows before contrast admission are taken as the noise background and subtracted from the angiography. Theoretically, only the shadow of vessels is left. So DSA is the gold-standard for the diagnosis of vascular abnormalities. But pitfalls for this technique are still prominent: both the patient and the doctor are exposed under x-ray during the process, which may continue for a few minutes or over an hour. The dose of radiation is cumulatively limited and it is harmful especially for a pregnant woman or a child. The technique to safely put the catheter in the right place needs years of training, but unexpected complications might still occur even under those experienced hands. Second of all, blood clot from the catheter might result in cerebral infarction and bleeding from puncture site is still not rare to see. Putting together, less invasive techniques are needed to visualize brain vessels.

Computerized tomography (CT) developed by Hounsfield and Cormack, Nobel Prize winners in 1979, divides a human head into a series of slides and each slide into properly arranged small cubes, namely matrix. The x-ray absorption value of each cube is calculable and each slide is shown as a grey-level image. The obvious enhancement of vessels makes them distinguishable from surrounding structures. But those are only short segments of vessels. With the availability of multi-detector-row CT scanner and development of data processing programs, 3D vascular reconstructions can be acquired, which is called computerized tomographic angiography (CTA). Its non-invasive, less radiation hazard, fast and convenient, and is now replacing traditional DSA. Similar slides of vascular images can be obtained from magnetic resonance (MR) scans and 3D reconstructed at the image post-processing workstation. This technique is called MR angiography (MRA).

In comparison of CTA and MRA, CTA is usually clearer because of higher resolution of raw scan images. Another priority of the CTA is that the vessels and surrounding other structures can be simultaneously demon-
Fig. 1. Examples of (A) the CT image and (B) the MR image.

Fig. 2. Three dimensional image of a localized, blood-filled balloon-like sac (aneurysm).

strated after proper threshold adjustment, which is valuable for doctors to choose a proper way to access the vascular disease. But when the vessel is close to bony structure, such as the internal carotid artery near the anterior clinoid process, or the vertebral artery close to the posterior fossa, partial volume effect would make the artery indistinguishable from the bone as they are both of high density. To adjust the window center or window width can’t settle this problem. Shorter scan time is also an important priority of the CTA. MRA has the priority of non-radiation hazard, being unaffected by bony structures, but the image is influenced by the direction of blood flow. Turbulent flow in the stenosed vessel often makes the vessel seem to be more severely stenosed or even occluded.

In Figure 2, there is a localized, blood-filled balloon-like sac, which is
Data Acquisition and Preprocessing on Three Dimensional Medical Images

An aneurysm is an abnormal bulge in the wall of an artery due to its weakness and hemodynamic force. Rupture of the aneurysm can result in high rates of morbidity and mortality. As the 2D CT or MRI slice images are acquired in a regular pattern (e.g., one slice every millimeter) and all image pixels are regularly numbered, they actually compose a complete 3D data set. Actually, we always need better image processing programs to maximally utilize the raw data acquired from CT or MR scans. Maximum intensity projection (MIP) and volume rendering (VR) are the most often used techniques to render a 2D projection of the 3D data set. The capability of four-dimensional (4D) time-resolved CTA and MRA will be helpful in hemodynamic analysis.

Besides the 3D techniques of inside objects, 3D surface imaging techniques, for example, 3D digital stereo-photogrammetry, have been popularly applied to reconstruct the 3D coordinates of points on a body from a pair of stereo images, especially in craniofacial clinics, such as orthodontics, oral and maxillofacial surgery. Structured light is another technique to detect the depth and surface information of an object. Different to stereophotogrammetry, only one camera is needed in structured light 3D scanner. Note that calibration is required in both techniques, and the reconstructed data should be evaluated to guarantee the optimal data collection.

The formats of the 3D image data can be generally categorized into two categories: 1) the measured surface is usually represented as a point cloud or polygonal mesh and is displayed as a smooth picture with surface rendering algorithms; and 2) volumetric data is a common representation of 3d-CT and 3d-MRI images. Medical images are usually coded in the DICOM (digital imaging and communications in medicine) format.

Graph-based mining algorithms could be applied directly on the point cloud or polygonal mesh, but the acquired 3D images are often noisy and unstructured for most data mining algorithms. To make the 3D data ready for knowledge discovery, segmentation and registration algorithms, the two fundamental 3D image preprocessing steps, are commonly applied.

3. Three Dimensional Image Segmentation

To extract the “ready-to-data-mine” regions from 3D images, computer-assisted automatic segmentation need to be accurate, repeatable and quantitative. In particular, medical images are often corrupted by noise, which can cause considerable difficulties when applying traditional low-level seg-
mentation techniques such as edge detection and thresholding. Consequently, classical image segmentation tends to generate infeasible object boundaries.

3.1. Deformable Model

To overcome these difficulties, deformable models have been extensively studied and widely used in medical image segmentation, with promising results. 3D deformable models are object outlines that move under the influence of forces and constraints minimizing internal energy, which is defined within the outline itself, and external energy, which is computed from the image data. The internal energy is designed to keep the model close to the shape during deformation. The external energy is defined to move the model toward an object boundary. By constraining boundaries to be smooth within the range of an explicit domain learned from a training set, deformable models are robust to both image noise and boundary gaps, and boundary elements in models are in coherent and consistent mathematical description which can then be readily used by subsequent applications, such as object tracking. For example, Xu et al. reconstructed human cerebral cortex from magnetic resonance images using a deformable model.\textsuperscript{22}

Deformable models have been highly recognized with the seminal paper “Snakes: Active Contours” by Kass, Witkin, and Terzopoulos\textsuperscript{23} and have grown to be one of the most active and successful research areas in image segmentation. Different names, such as snakes, active contours or surfaces, balloons, and deformable contours or surfaces, have been used in literatures to refer to deformable models. The models are generally grouped into two types: parametric deformable models\textsuperscript{23–26} and geometric deformable models.\textsuperscript{27–30} Parametric deformable models represent object outlines explicitly in their parametric forms during deformation, so that direct interaction with the model is feasible and a compact representation for fast real-time implementation can be developed. However, model topology operation, such as splitting or merging parts during the deformation, can be difficult when using parametric models. The segmentation result shown in Figure 3 is a parametric surface. Geometric deformable models, on the other hand, can handle topological changes naturally. These models, based on the theory of curve evolution\textsuperscript{31–34} and the level set method,\textsuperscript{35,36} represent curves and surfaces implicitly as a level set of a higher-dimensional scalar function. Their parameterizations are computed only after complete deformation, thereby allowing topological adaptivity to be easily accom-
modated. Despite this fundamental difference, the underlying principles of both methods are very similar.

3.2. Three Dimensional Image Segmentation with Deformable Model

Given the successful 2D deformable models in 2D image segmentation, segmenting 3D image volumes can be done by applying 2D deformable model slice by slice. However, it is labor intensive and requires an accurate post-processing step to connect the sequence of 2D contours into a continuous surface, because the resulting surface reconstruction would contain inconsistencies if the post-processing is not well done. True 3D deformable surface model is faster and more robust resulting in a globally smooth and coherent surface between image slices.

Deformable surface models in 3D were first proposed by Terzopoulos\textsuperscript{37} for computer vision. For the demanding needs in medical image analysis, the use of deformable surface models have explored for segmenting structures in medical image volumes. Miller\textsuperscript{38} constructs a polygonal approximation to a sphere and geometrically deforms this balloon model until the balloon surface conforms to the object surface in 3D CT data. The segmentation process is the minimization of cost functions each of which measures costs with local deformation. The cost function that associates with every vertex of the polygonal is a weighted sum of three terms: a deformation potential that expands the model vertices towards the object boundary, an image term that identifies features such as edges and opposes the balloon expansion, and a term that maintains the topology of the model by constraining each vertex to remain close to the centroid of its neighbors.

Cohen and Cohen\textsuperscript{39,40} and McInerney and Terzopoulos\textsuperscript{41} use finite element and physics-based techniques to implement an elastically deformable cylinder and sphere, respectively. The models are used to segment the inner wall of the left ventricular (LV) of the heart from MR or CT image volumes\textsuperscript{41} (as shown in Figure 3).

These deformable surfaces are based on a thin-plate under tension surface spline, which controls and constrains the deformation of the surface. Lagrangian equations of motion through time is used to fit the models to data dynamically and to adjust the deformational degrees of freedom. In addition, the models are represented with the finite element method as a continuous surface in the form of weighted sums of local polynomial basis functions. Unlike Miller’s\textsuperscript{38} polygonal model, the finite element method
provides an analytic surface representation which uses high-order polynomials so that fewer elements are required to accurately represent an object. Pentland and Sclaroff and Nastar and Ayache also develop physics-based models but use a reduced modal basis for the finite elements. Staib and Duncan describe a 3D surface model used for geometric surface matching to 3D medical image data. The model uses a Fourier parameterization that decomposes the surface into a weighted sum of sinusoidal basis functions. Several different surface types are developed such as tori, open surfaces, closed surfaces and tubes. Surface finding is an optimization problem in gradient ascent that attracts the surface to strong image gradients in the vicinity of the model. A wide variety of smooth surfaces can be described with a small number of parameters by using the Fourier parameterization. Because the basis functions in a Fourier representation are orthonormal and the higher indexed basis functions represent higher spatial variation. Therefore, the series can be truncated and still represent relatively smooth objects accurately.

Besides aforementioned approaches, Szeliski et al. use a dynamic, self-organizing oriented particle system to model surfaces of objects. Small flat disks defined as the oriented particles evolve according to Newtonian mechanics and interact through external and internal forces. The external forces keep the particles to the data, while internal forces attempt to group the particles into a coherent surface (model). Objects reconstructed with the particles are with complex shapes and topologies by flowing over the data, extracting and conforming to meaningful surfaces. A triangulation is then performed to connect the particles into a continuous global model that is consistent with the inferred object surface. Other notable work involving

Fig. 3. (a) Deformable balloon model embedded in volume image deforming towards LV edges. (b) Reconstruction of LV.
3D deformable surface models and medical image applications can be found in.30,46–48

3.3. A Case Study in Segmentation with Active Appearance Models

Active Appearance Models (AAM) is a group of highly flexible deformable models firstly introduced in 1998.49 AAMs model an 3D object as an combination of a model of shape and a model of texture;50 where “shape” is the vertex locations of the 3D mesh, and “texture” also being termed as “appearance” can be pattern of intensities or colors across an image patch. Both shape and texture are represented in AAMs as means with their eigen variations, by applying Principal Component Analysis (PCA) on training data that have been hand annotated by experts.

Mathematically, a labelled training example $i$ is defined by its shape $s_i = (x_0, y_0, z_0, ..., x_{S-1}, y_{S-1}, z_{S-1})$ containing $S$ corresponding 3D surface points $(x, y, z)$ and its texture $g_i = (g_0, ..., g_{T-1})$ containing $T$ corresponding appearance patterns, such as colors. PCA decomposes shape and texture into their means and eigenvectors, given in Eq.1 and Eq.2, where, $\bar{s}$ is the mean shape, $\bar{g}$ is the mean texture in a mean shaped patch, and $Q_s, Q_g$ are eigen matrices that describe the modes of variation derived from the training data. $c_s$ and $c_g$ are parameter vectors for parameterizing shape and texture.

$$s = \bar{s} + Q_sc_s$$  \hspace{1cm} (1)  

$$g = \bar{g} + Q_gc_g$$  \hspace{1cm} (2)

AAM based segmentation is to fit an AAM to an 3D object by iteratively updating parameter vectors, minimizing the error between the input 3D object and the closest model instance; i.e. solving a nonlinear optimization problem. Texture residual (defined in Eq.3) is a commonly used error measure, where $p$ are the parameters of the model, $g_m$ is the modeled texture as a function of the current parameters $p$ (Eq.2) and $g_s$ is the texture, sampled from the current object.

$$r(p) = g_s - g_m$$  \hspace{1cm} (3)

AAM based segmentation is fully automated and successful in practice. The detected contours follow the ground truth quite well. In Leung et al.51 the authors developed a 3D echocardiography segmentation approach, which has been evaluated on 99 patients.
4. Three Dimensional Image Registration

3D images are often acquired with different image modalities (as introduced in Section 2). To integrate different forms of 3D image-derived information to serve the further data mining tasks, the point-by-point correspondence between two or multiple 3D medical images of a scene need to be automatically determined. Such process is called 3D image registration.

4.1. Image Registration Methods

For the two 3D images to register, the image that is kept unchanged is called “reference image”, whereas the image that is resampled to register the reference image is called “sensed image”. Mathematically, let \( T \) denote the spatial transformation that maps coordinates (spatial locations) from a sensed image \( A \) to a reference image \( B \), and \( P_A \) and \( P_B \) denote coordinate points (pixel locations) in images \( A \) and \( B \), respectively, the image registration problem is defined as to determine \( T \) such that the mapping \( T : p_A \rightarrow p_B \Leftrightarrow T(p_A) = p_B \) results in the best alignment of \( A \) and \( B \).

Two medical images may differ by any amount of rotation or translation in any direction, and they may also differ in scale. The nature of the \( T \) transformation characterizes the search space of the registration problem, ranging from nonlinear transformation with a virtually infinite degree of freedom, to rigid registration with six degrees of freedom.

By registering two medical images, the fusion of multi-modality information becomes possible, and regions of abnormal function can be recognized. A registration algorithm usually has three components: a matching criterion of how well two images match; the transformation model, which specifies the way in which the sensed image can be modified to match the reference image; the optimization process that varies the parameters of the transformation model to maximize the matching criterion.

For DICOM images, a preliminary registration can be performed using the geometrical features such as the position and orientation of the image with respect to the acquisition device and the patient, as well as to the voxel size. To align the corresponding features of two or more images, the preliminary feature based registration approach, consisting of two components which are a number of features selected from the images and the correspondence established between them, can be applied. Knowing the correspondences, a transformation function is then found to resample the sensed image to the geometry of the reference image. The live example of
image registration of the CT and MR sample images in Figure 1 is shown in Figure 4.

When domain knowledge or known corresponding geometric function is not applicable, the voxel intensity-based registration approach, developed mainly based on the concept of mutual information (MI), can be applied, implying the comparison of image gray levels to be registered. If one image provides some information about the second one, then $MI(A, B) > 0$, otherwise, we have $MI(A, B) = 0$, meaning that the two images are independent. The MI is related to the image entropy by

$$MI(A, B) = H(A) + H(B) - H(A, B),$$

where $H(x)$ is the entropy of $x$.

The voxel intensity-based registration aligns the images by optimizing the MI values of the corresponding voxel pairs. Because no assumption is made about the nature of the relation between the image intensities, MI is robust for both multimodal and unimodal registration, and it does not depend on the specific dynamic range or intensity scaling of the images.

Inspired from the optical flow equations, Thirion proposed the demons registration that considers a diffusion process to best match the boundary of a reference image to that of a sensed image. Demons (entities) are defined in the reference image, typically on its contour points, to attract the deformation of a subject image for best alignment. The demon registration can be thought of as an approximation to fluid registration.

For large-scale applications such as brain image registrations, images may have different orientations, brightness, sizes, evenness of staining, morphological damage and other types of image noise, which requires a robust algorithm. However, such image registration approaches usually require a lot of computational power and computational time. This means that these algorithms are difficult to use where real time constraints are introduced, as in virtual reality applications. In addition, in one comparison of several widely used methods for registration, all the tested methods yielded unsatisfactory alignments at a rate that make them unsuitable for use in a pipeline that involves thousands of high-resolution 3D laser scanning microscope images. In the end, a major problem of the existing registration approaches is the lack of an effective quality index to be used when assessing the quality of the registration process. To tackle these problems, a non-linear non-rigid mapping schemes called BrainAligner was proposed in Peng et al.
4.2. BrainAligner: A Case Study in 3D Image Registration

Much of the current work on 3D medical image registration involves non-rigid registrations, which is needed to take into account the differences of two patients or one patient and an atlas, especially in acquisition protocols, resolution, etc. BrainAligner is a 3D image non-rigid registration program that automatically finds the corresponding landmarks in a sensed brain image and maps it to the coordinate system of the reference brain via a deformable warp. With BrainAligner, the reference channel for each sensed brain image is mapped to a reference image using a nonlinear geometrical warp, so that the patterns in multiple brain images can be compared in the same coordinate space for the identification of intersecting patterns in various anatomical structures. Note that, in BrainAligner, the percentage of the reference landmarks that are automatically reliably matched can be used to score how many image features are preserved in the automatic registration.

BrainAligner has two steps: global 3D affine transformation and nonlinear local 3D alignment. The global alignment process sequentially optimizes
the displacement, scaling and rotation parameters of an affine transform from sensed to reference to maximize the correlation of voxel intensities between two images. First, the center of mass of a sensed image is aligned to that of the reference image. Then a sensed image is rescaled proportionally, so that its principal axis had the same length with that of the reference image. Finally, a sensed image is rotated around its center of mass and thus detected the angle for which the reference image and the rotated subject image had the greatest overlap.

In the local alignment step, a reliable landmark matching (RLM) algorithm was designed to detect corresponding 3D feature points (i.e. a landmark) in every reference-sensed image pair, as shown in Figure 5. For each reference landmark, at least two independent matching criteria are used in RLM to locally search for matching landmarks in the sensed image. These criteria include mutual information, inverse intensity difference, correlation and similarity of invariant image moments. A match confirmed by a consensus of these criteria is defined as a preliminary landmark match (pre-LM). Next RLM uses a random sample consensus algorithm to remove the outliers that violate the smoothness constraint or the relative location relationship.

In summary, BrainAligner is one of the state-of-the-art automated 3D image registration methods on large-scale datasets. In BrainAligner, the RLM method, which compares the results produced using different criteria and only uses results that agree with each other, can be viewed as an optimized combination of several existing methods.
The main purpose of 3D medical image processing is to effectively transform the unstructured image data to structured numerical data for further data mining tasks. Algorithms for image acquisition, segmentation and registration are the essential steps to provide high quality data for biomedical data mining, which is ultimately important in that the power of data mining models are largely determined by the quality of the data.

References

14. B. Shetty, A. Raju, and V. Reddy, Three dimensional imaging an impact


