VISUALIZATION TECHNIQUES FOR 3D MULTIMODAL MEDICAL DATASETS: A SURVEY

BY

PAUL CORNELIU HERGHELEGIU*

“Gheorghe Asachi” Technical University of Iași,
Faculty of Automatic Control and Computer Engineering

Received: August 26, 2012
Accepted for publication: September 15, 2012

Abstract. Accurate medical diagnosis based on images acquired with various medical imaging techniques commonly requires multiple images inspection. These images can be obtained using the same scanning technique but with different scanning parameters. Alternatively, the physicians can use images obtained with different scanning equipments. The visualization of multimodal data in the same rendering scene currently represents a growing avenue of research, with multiple results. This paper aims to give the reader an overview of recent innovations in multimodal visualization techniques used nowadays by physicians in their clinical practices. Several types of data combining methods are presented, such as CT-MRI, MRI-fMRI, CT-SPECT or 3D ultrasound-CT or MRI, each one being suited for a typical type of pathology. Since the structure of the combined data usually presents significant differences, there is no general method that can be applied. The theoretical concepts of combining the mentioned types of data are presented along with their use in the medical imaging field. A short introduction to commonly used medical scanning techniques is also presented.

Key words: multimodal visualization, medical multimodal data, review.

2010 Mathematics Subject Classification: 65D18, 68U05.

*Corresponding author; e-mail: pherghelegiu@cs.tuiasi.ro
1. Introduction

In order to provide an accurate medical diagnosis based on images acquired from various non-invasive scanning techniques, physicians have to inspect a large number of images. Multiple images depicting one specific region of the human body can present a large number of differences depending on the parameters used in the scanning process. The simplest parameter to modify is the orientation of the scanning planes. More advanced parameters refer to modifying the scanning sequence while using the same equipment for all the scans. The most challenging task is to inspect the images acquired from different scanning techniques. This approach is usually required when investigating evolving pathologies such as tumors. The term multimodality refers to the inspection of more than one image in the same scene with others captured using different scanning parameters. Most used scanning modalities include Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Functional Magnetic Resonance Imaging (fMRI), Positron Emission Tomography (PET), Single Photon Emission Computed Tomography (SPECT) or 3D ultrasound (Bushberg et al., 2001).

Capturing images using different modalities requires that the patient is moved between scans. As a perfect alignment to the same coordinate system cannot be achieved, a method of registering the obtained images has to be applied. This step is required prior to using any multimodal visualization technique. In order to solve the registration problem, a range of techniques have been developed, mostly divided into two main areas, rigid and non-rigid (Wyawahare et al., 2009; Mattes et al., 2001; Fischer & Modersitzki, 2008; Ferré et al., 2004).

As there is no scanning technique that provides all the anatomical and spatial features of any scanned region of the human body, the used scanning techniques are driven by the investigated medical condition. Most of the conditions will not require more than one scan, but there is a large number of cases where multimodal data is required. The most used scanning techniques are the CT and MRI that can be combined in order to acquire data regarding both the soft tissue and the bones.

Multimodal data is also needed when trying to visualize PET scan results. PET data provides no visualization cues about the spatial context of the scanned area, so the problem of visualizing this data in an anatomical context had to be solved. Multimodal data can also be used for better segmentation results (Passat et al., 2007). For visualization purposes, most multimodal rendering techniques employ a Direct Volume Rendering approach (Levoy, 1988; Engel et al., 2005).

In this paper we present the most recent work developed in the multimodality rendering area. Section II provides a short overview of the most common used medical scanning techniques. Section III presents the techniques
used in rendering multimodal data as well as their applications. Section IV concludes this paper.

2. Image Acquisition Techniques

Recently, many medical practices seek to use non-invasive scanning methods for establishing an accurate diagnosis. The best way to achieve this is by using scanning techniques that depict the inner structures of a patient. Medical scanners and image acquisition techniques have developed in many directions, trying to provide specific images for each investigated disease. Next we present the most common image acquisition techniques and their common usage in medicine.

Computed Tomography (CT)

CT works by taking a series of X-Rays from different scanning planes from an object and back-projecting them into a 3D volume. CT scans are commonly used when investigating bone structures. The drawback of this technique is represented by the high level of radiations the patient is exposed to. Also, the images are obtained only along the axial plane, allowing only a slight modification of the scanning angle. Because of the very low scanning time, CT is indicated to use in emergency situations. Current developments involve the use of multiple X-ray detectors for reducing the acquisition time. Also, detectors that work with different energy levels are being developed, for depicting both bone and soft tissue, leading to the introduction of Dual Energy CT.

Magnetic Resonance Imaging (MRI)

MRI is based on the property of nuclear magnetic resonance to produce images of human tissues due to their different response to an applied magnetic field. MRI scans are used mainly for depicting soft tissues (brain, muscles, etc.). Because it relies on the water content of the scanned body parts, the MRI cannot provide adequate data regarding the bone structures. As there is no radiation involved, MRI is also used when certain medical condition has to be scanned repeatedly and tracked over time (Kalender, 2011).

Functional Magnetic Resonance Imaging (fMRI)

Functional MRI detects changes in the blood flow related to neural activity in the brain. An increased flow of blood in an area of the brain is determined by an activation of that specific area, which corresponds to some cognitive task effectuated by the patient. During the scan, the patient is asked to perform various cognitive tasks, such as looking at images, speaking about a subject, etc (Brown & Semelka, 2010).
**Positron Emission Tomography (PET)**

PET is a nuclear medical imaging technique that depicts the functional processes in the body. A radioactive substance has to be injected into the patient body. It is used as a scanning method in oncology or for the diagnosis of certain brain diseases. The drawback of this technique is represented by the fact that the positrons emitted by the injected substance are annihilated by electrons that can be positioned up to a few millimeters away, leading to a fuzzy result (Saha, 2010).

**Single Photon Emission Computed Tomography (SPECT)**

SPECT is an imaging technique that uses gamma rays. A gamma-emitting radioisotope has to be injected into the patient. A gamma camera that is rotating around the patient will capture the gamma radiations directly. SPECT is commonly used to complement other imaging technique, where a true 3D representation is needed (Bailey & Meikle, 2008).

**3D ultrasound**

This medical imaging method uses high frequency sound waves in order to produce images. The main advantages of this method are represented by the small and cost effective hardware used and the real time image acquisition. The quality of the output result is not as good as in CT or MRI scans. The ultrasound technology has evolved toward 3D image acquisition, providing as an output result the spatial position of the outputted image (Suri et al., 2008).

### 3. Techniques for Multimodal Visualization

Beside the registering problem mentioned before, other problems that address the multimodal visualization have been formulated (Ferré et al., 2004):

- data modeling that addresses the problem of finding the structures that can depict different properties at multiple resolutions, while skipping the irrelevant structures;
- multimodal rendering that deals with finding the appropriate strategies needed for rendering in a single visualization window of the images acquired from different scanning modalities.

For every pixel in the final image, two or more sampled voxels from the used volumes have to be combined. This process represents another aspect that has to be taken into consideration when dealing with multimodal data. Different rendering per pixel strategies can be applied (Ferré et al., 2004):

- single property per pixel with two possible combinations, complementary data (rendered properties are the same for all pixels) and supplementary data (for each pixel a decision is made what property should be rendered);
• multiple properties per pixel: it is assumed that each property corresponds to one material, and a combination of these materials is made in the rendering process.

We present next several different types of fusions between the properties of each corresponding voxel belonging to the volumes that will be rendered:

• property fusion: different values of properties are combined at the beginning of the rendering process;
• property and gradient fusion: for a sample point, an unique property of the material is computed depending on the property and gradient values;
• material fusion: for a sample point, the property and gradient of each volume involved in the multimodality rendering are computed. The results are then combined into one single material property;
• illumination fusion that combines the optical properties of pixels;
• color fusion that combines the colors obtained after an illumination model is applied.

The simplest approach for visualizing multimodal CT and MRI data is based on using a threshold. Based on this threshold, a decision upon which data should be used in the rendering process is made. The main drawback of this method is that data from both used volumes is lost. Inclusive opacity can be used to enhance this approach. This method is based on computing the final color using samples from all used volumes (Ghosh et al., 2003). A color coding scheme that assigns different values to red, green and blue channels can be used in order to identify overlapping voxels and to prevent losing data.

New multimodal rendering techniques commonly make use of the computational power of modern GPUs that present a high level of parallelism very well suited for the visualization algorithms (Nguyen, 2008). One such algorithm provides a spatial context for fMRI data using a pre-registered MRI volume (Rößler et al., 2006). Because the fMRI data is very hard to interpret, a statistical tool is usually used for obtaining a 3D volume from each fMRI sequence (SPM, 2005). The algorithm is based on the GPU and uses the DVR technique. The anatomical reference volume is reconstructed from the MRI data. This volume is registered with the reconstructed volumes from fMRI data.

The rendering of fMRI volumes uses the GPU features to store volumes as 3D textures (Cabral et al., 1994). Volume data is saved in a number of slices belonging to the texture. Multi-volume rendering technique is based on slicing the volumes followed by a merging operation according to a pre-defined geometrical order. By applying this technique, multiple volumes can be saved onto a single 3D texture. The final rendered volume is obtained by crossing the volume in a back-to-front order. If the current slice does not belong to the volume previously rendered, the algorithm state is changed, allowing swapping through different rendering techniques. Also a swap between multiple shaders programs can be performed.
Pre-integrated volume rendering via DVR and rendering illuminated volumes are used as rendering techniques (Engel et al., 2001). The use of light sources provides an increased 3D perception of the rendered volume (Westermann & Ertl, 1998). In order to improve the visualization of inner structures, clipping planes can also be implemented. Also, a transfer function can be used for applying different colors and opacity settings to different activated areas of the brain.

Another approach for anatomical data visualization (MRI) along with the functional data (fMRI) has been proposed as an extension of the previously presented algorithm (Schafhitzel et al., 2007). The idea is to replace the undesired effects caused by occlusion by representing the surfaces as a series of sparse lines. The preservation of geometrical information has also been a key requirement for this approach. Similar to (Rößler et al., 2006), the MRI data is needed for providing a spatial context. DVR is used for the visualization of activated regions of the brain and Surface Shaded Display (SSD) is used for rendering data that concern the structure of the brain. The Line Integral Convolution (LIC) algorithm (Cabral & Leedom, 1993) is used for generating replacing lines. This algorithm employs the integration along curves defined by a vector field for generating the lines that will replace the surface. These lines are mainly used for modulating the intensities of activated brain regions that will represent the structure of an isosurface (Weiskopf & Ertl, 2004). The method is implemented on GPU, with resulting frame rates ranging from 25.57 fps to 42.17 fps.

The multimodal rendering algorithm applies a special blending function that takes into consideration the lines obtained from the LIC computation, along with other parameters such as the illuminated isosurface and the pre-rendered brain activation areas. The blending function used is defined as (Schafhitzel et al., 2007):

\[ C_{out} = \alpha C_{iso} + (1 - \alpha)(1 - I_{LIC})C_{act} \]  

where \( C_{out} \) is the final computed color, the functional data is represented by \( C_{act} \), \( C_{iso} \) – the anatomical data, and \( \alpha \) is the opacity of the isosurface. The final image is also altered by \( I_{LIC} \) that represents the intensities of the curvature lines.

For implementing the curve illumination, the algorithm changes slightly. A real time gradient computation for the curvature lines has been applied. The illumination component changes eq. (1) to (Schafhitzel et al., 2007):

\[ C_{outL} = \beta (N \cdot L) + (1 - \beta)C_{out} \]  

where \( N \) represents the normal vector, \( L \) – the light source position and \( \beta \) – the light contribution for emphasizing the line structures.
Another volume rendering technique that uses 3D textures for rendering multimodal datasets has been introduced in (Abellán & Tost, 2008). A 3D texture is created for each pre-registered volume that will be used. GLSL shaders and 3D mapping techniques are used to achieve shading and fusion. The proposed approach has been tested and used with SPECT-MR, PET-MR and PET-CT data.

Two voxel models are used, along with a distance map that references the known transformations needed for registering the two volumes. The algorithm can use as input two types of texture data: value indexed and gradient plus value indexed textures. For each of these types of textures, three different shading techniques can be applied: emission plus absorption, surface shading or a combination of both. Fusion is applied by using a transfer function obtained from merging the two volumes according to a pre-defined fusion-weight value. The fragment shader checks if a rendered point falls inside of any of the two textures. If the point is inside of only one texture, then no fusion is made. If it is inside both volumes, the algorithm performs three tasks:

• samples the two corresponding points as a function of the weight value;
• computes the shading according to the gradients;
• combine the two colors.

For enhancing the rendering results, depth cueing techniques and clipping planes are implemented.

Most surgical interventions require pre-operative planning. It is usually performed to allow the physicians to better understand and to visualize how the actual medical intervention will take place. Many multimodal visualization techniques have been introduced that allow the combination of various data in order to provide the needed pre-operative support (Beyer et al., 2007).

The basis for the framework described in (Beyer et al., 2007) is a represented by a GPU ray caster (Scharsach et al., 2006). The volumes are combined on a per sample basis during the ray-casting algorithm. A decision process has been implemented to decide how the samples from each volume will contribute to the final color of the ray. For each volume a specific transfer function is used, in order to provide a better flexibility.

Each rendered voxel will correspond to one of the combined volumes. No function for combining colors is used. A decision has to be made just for establishing which volume will provide the color of the point. The novelty of the proposed approach is the use a 3D texture for recording the corresponding object for each rendered voxel. A 1D texture is then used for storing the volume each object corresponds to (Hadwiger et al., 2003). Another novelty introduced is represented by the data and memory management. A bricked volume rendering approach is used. This approach subdivides each volume into sub-volumes of size 16x16x16 or 32x32x32. Only these sub-volumes will be held into GPU memory. The object id per voxel is saved on a single texture. Each data volume (CT, PET) has its own lookup texture that will be used for address
translation between the cache volume and data volume and one 3D texture with its currently allocated cache.

If brain visualization is required, rendering it without any prior segmentation is a difficult task due to the presence of occluding structures (skin, skull, etc). Different methods of peeling outer layers that are hard to use due to their high dependency on parameter changes have been developed (Mullick et al., 2000; Rezk-Salama & Kolb, 2006). The approach presented in (Beyer et al., 2007) aims to develop a more easy to use skull peeling algorithm. The result is a clear image of the unoccluded brain, with the skin and bone removed. This is achieved simply by taking advantage of the multi-volume rendering, with no prior segmentation necessary. The algorithm uses preregistered data from CT and MRI, and exploits the advantages of each technique, i.e. CT is used for depicting bone structures while MRI depicts the brain. Using a straight-forward approach, the ray-casting algorithm decides by sampling both volumes if a sample point is part of the brain or part of the skull. When the ray hits a bone for the first time, the accumulated values are reset. After the bone area is exited, accumulation starts again and the brain is revealed. Certain extensions for managing the case when a brain portion is not occluded by the brain were implemented.

A fusion method based on voxel properties is used to combine multimodal volumes. The contributions of each individual volume are accumulated by combining the values for each corresponding sample location after applying the transfer function.

The visualization of inner structures is usually required when trying to understand specific parts of the human body or certain pathologies. A large amount of research has been conducted in the area of generating suitable clipping planes. These planes separate the hidden voxels from the visible ones, allowing the depiction of the region of interest. The main problem of using such planes is that no inner structure of the human body follows an exact plane. In order to solve this problem, various cutting surfaces have been introduced (Manssour et al., 2002).

The main advantage of this method is the possibility to eliminate regions irrelevant for diagnosis determination, without requiring a pre-processing segmentation step.

Another area of research in dealing with multimodal medical data implies the integration of 3d ultrasound images in CT or MRI data. The task of registering 3D freehand ultrasound images with other medical data presents some typical issues and many algorithms for registering these types of data have been developed in the past (Wein et al., 2005; Wein et al., 2008). Because an ultrasound dataset contains slices captured along different spatial orientations, the algorithms presented so far cannot be applied for the visualization of this type of data. A 3D ultrasound dataset usually has a lot of redundancy data, as the physician cannot take images along parallel planes like in the case of CT or
MRI. The problem of visualizing this kind of multimodal data is reduced to integrating one pre-registered ultrasound image into a reconstructed 3D volume. Simple clipping planes can be used for a spatial positioning of the ultrasound image, but the interest have expanded to accurate positioning of the image into a volume with the possibility of letting some structures of interest overlap it.

Another method of visualizing multimodal data using cutting surfaces introduced the concept of contextual cutaway views (Burns et al., 2007). This method assigns an importance factor to each surface and the decision regarding which surfaces to be removed is based on this factor. Although the rendering of MRI data tends to be problematic because intensities are not proportional with tissue densities, the method introduced in (Burns et al., 2007) provides satisfactory results and can be used in pre-interventional visualization. A previously introduced technique of combining multiple modalities using smooth brushing and linked scatter plots have been used (Doleisch et al., 2003).

Different parameters were implemented for controlling the importance of pre-segmented structures. Shading emphasis can also be applied for guiding the viewer attention to the most important part of an image. This technique is applied by modifying the color, opacity and shading styles of materials. As color and opacity can be defined by transfer functions, only the modulation of shading is necessary. This is done by limiting the contribution of a given shading model to the final color of the ray, so important structures will be shaded more than others.

Another approach of integrating ultrasound images into a previously reconstructed MRI volume is the implementation of post-processing pipeline for the guided visualization of saved ultrasound images (Angelelli et al., 2010). This allows the physicians to have access to an easier analyzing tool for the ultrasound data. The concept relies on computing a degree of interest (DOI) for each ultrasound image. DOI is computed with respect to structures the physician wants to see. This step is followed by an integration of an ultrasound image into a DOI volume selected by the user. This technique can be used especially for the cases when a re-inspection of previous ultrasound data is necessary.

Some other methods of integrating ultrasound images with MRI data have also been introduced (Rick et al., 2003; Nikas et al., 2003). There are also commercial applications that provide fused visualization of different types of data including ultrasound-CT or MRI. Most of techniques used in such commercial applications operate mainly on data level (www.sonowand.no).

4. Conclusions

When non-invasive images depicting structures of the human body are required, various scanning techniques are used. Multiple images acquired from different scanning techniques are commonly inspected for determining an accurate diagnosis. Because there is no perfect scanning method that provides
all the data required about a specific body structure, multiple techniques are used. When the evolution of a medical condition is evaluated, multiple scans using the same techniques but at different timeframes are performed. This is also the case when scanning with the fMRI method, which provides multiple datasets, one for each cognitive task the patient is required to perform. Based on this needs, an area of research dealing with such multimodal data has expanded recently. The computer graphics community tries to provide means of visualizing medical multimodal data in ways that will help physicians in their practices.

Because of the various misalignments problems, a registration algorithm has to be applied prior to any multimodal visualization. After the registration process is finished, approaches that will display rendered data in a comprehensible manner can be applied.

In this paper we presented some of the most recent techniques for rendering multimodal data in the same scene. Although the size of medical data is increasing exponentially, the results are obtained at an interactive rate, due to the use of recent GPU technology. Multiple rendering techniques have been presented, as well as their advantages and their use. Theoretical elements of combining data were presented in the beginning of this paper. Most of them can be applied on multiple kinds of multimodal data. Specific combinations of data like CT-MRI, PET-MRI or fMRI-MRI required specific algorithms for combining and rendering. Also, algorithms for rendering 3D ultrasound with CT or MRI data were also presented in this paper.

Methods can rely on a simple threshold algorithm or on a complex texture caching for displaying large data sets, but all of them aim to help the understanding of human anatomy. As a commonly applied method for visualizing the inner structures of the human body, clipping planes are often used, providing good visualization support.

Recently, a new imaging technique has been developed for use in the medical field: Dual Energy CT (DECT). This method uses two energy levels in the same CT scan, for acquiring data both from bones and soft tissue. Methods for combining the data are also already developed and included into the medical framework.

REFERENCES


STUDIU PRIVIND TEHNICILE DE VIZUALIZARE A SETURILOR DE DATE 3D MULTIMODALE OBȚINE ÎN URMA UNOR SCANĂRI MEDICALE

(Rezumat)

Stabilirea unui diagnostic precis, bazat pe imagini ale unei structuri ale corpului uman obținute non-invaziv, necesită în mod uzual obținerea și inspecția mai multor serii de imagini. Aceste serii pot fi obținute utilizând fie aceeasi metodă de scanare medicală dar cu parametri diferiți, fie metode diferite. Termenul multimodal se referă la diferite serii de imagini medicale. Vizualizarea datelor multimodale implică afișarea în aceeași scenă a mai multor volume reconstruite din serii diferite. Deoarece poziția de scanare a pacientului nu poate fi aceeași de-a lungul mai multor scanări efectuate eventual prin diferite metode, un algoritm de aliniere a datelor folosite la același sistem de coordonate trebuie aplicat înaintea utilizării oricăror algoritmi de vizualizare.

Deoarece nu există o metodă de scanare ce poate furniza toate aspectele anatomic și spațiale cu privire la un anumit organ al corpului uman, în mod frecvent sunt necesare mai multe scanări ale structurii de interes. Cele mai frecvente date multimodale sunt obținute de la Computer Tomograf (CT) și Rezonanță Nucleară (RMN) și pot fi folosite concomitent pentru a furniza informații atât despre structurile osoase cât și despre țesuturile moi. Lucrarea de față prezintă ultimele realizări din domeniul vizualizării datelor multimodale. Sunt descrise atât considerentele teoretice, cât și aplicațiile din lumea medicală. Metodele prezentate cuprind tehnici de combinare și vizualizare a datelor CT-RMN, RMN-RMN funcțional, CT-SPECT, RMN-PET sau date ecografice 3D combinate cu date CT sau RMN. De asemenea, lucrarea prezintă succint și principalele metode de achiziție a imaginilor utilizate în prezent în medicină.