A good display is one that emphasizes the diagnostic features and minimizes the camouflaging effect of the nondiagnostic (anatomic) features.

H.L. Kundel, 1990

Chapter 3

Integration of Functional and Anatomical Brain Images: A Survey of Approaches

Abstract

This chapter reviews the literature on integration of functional and anatomical volumetric brain images. Integration consists of two steps: matching or registration, where the images are brought into spatial agreement, and fusion or simultaneous display where the registered multimodal image information is presented in an integrated fashion. Approaches to register multiple brain images are divided into extrinsic methods based on artificial markers, and intrinsic matching methods based solely on the patient related image data. Only rigid registration (*i.e.*, translation and rotation) is considered. The various methods are compared by a number of characteristics, which leads to a clear preference for one class of intrinsic methods, *viz.* voxel based matching. As for image display, an overview of existing methods to simultaneously visualize registered multiple images is presented. Both 2D and 3D approaches are discussed. Several techniques seem quite appropriate for multimodal brain image fusion. Yet, a general task-dependency of the visualization and—in addition—the variability of observation procedures and observer characteristics precludes such a generic conclusion.

3.1 Introduction: Purpose and scope

Integration of images from multiple modalities has rapidly evolved into a substantial area of research in medical imaging. There are two major causes for this development. First, performing the calculations involved in registering and visualizing two 3D datasets has become feasible on present-day computers. This has paved the way, not only for novel matching approaches that are based on the full contents of the images rather than on just a few points from artificial markers or anatomical landmarks, but also for novel visualization techniques aimed at efficiently presenting the multimodal information. Second, there is a steadily growing demand from the clinic for multimodality integration, in particular in neurosurgery and radiation treatment planning and evaluation.

The scope of the present chapter is now outlined. As the title indicates, this chapter covers integration of functional and anatomical brain images, a subclass of multimodal image-to-image integration. This excludes image-to-image integration of single modality image data and image-to-atlas integration. Image-to-image integration of brain data generally deals with information from the same patient, whence it is natural to consider only rigid transformations (translations and rotations). The survey will furthermore be limited to integration of two volumetric images; matching of a 3D image with a 2D image and of time series of images are not dealt with.

Brain imaging modalities that produce anatomic volumetric data are CT, CTA, MRI, and MRA. Functional volumetric brain imaging modalities are SPECT, PET, fMRI, Perfusion Weighted MRI, Diffusion Weighted MRI, and MRSI. In addition, functional volumetric information may be inferred from EEG or MEG by means of mathematical modelling (source localization).

The purpose of the chapter is to give an overview of methods for integration of volumetric brain images from functional and anatomical modalities. The different extrinsic and intrinsic approaches to image registration are described and compared by a number of characteristics. It will be shown that these criteria are appropriate to clearly select a class of approaches as being superior. The second issue, integrated visualization of the registered data, will be discussed more extensively, and will feature our own experiences with several of these approaches as applied to simultaneous display of SPECT/MRI, PET/MRI, and fMRI/MRI. The question of which technique is the most suitable is strongly task—and operator—dependent which prevents a general evaluation of visualization methods.

3.2 Multimodality image registration

In Van den Elsen et al. (1993) a classification of image registration methods is given according to a number of discerning criteria. The main criteria are: Dimensionality (2D/3D/4D), nature of matched properties (extrinsic methods using artificial objects as stereotactic frames, head or dental molds, skin markers; and intrinsic methods using image data only), elasticity of the transformations (rigid/affine/projective/curved), and interaction (interactive/semi-automatic/automatic).

This section discusses rigid registration of multimodal volumetric brain images. In keeping with the categorization, the approaches are divided into extrinsic and intrinsic matching. Extrinsic registration methods are subdivided according to the type of artificial marker employed, intrinsic methods are subdivided according to the property of the image data used for matching.

3.2.1 Extrinsic matching

The three approaches classified under extrinsic matching have in common that they do not admit of retrospective matching, which entails that the clinical protocols must take account of the requirements of the matching procedures. Consequently, an image that was acquired before the necessity of multimodality integration is recognized, cannot be included in the matching procedure if extrinsic approaches are used.

3.2.1.1 Stereotactic frame / skull screws

In stereotactic neurosurgery, a rigid frame is attached to the head of the patient to guide the surgical instruments. In the image acquisition stage, localizer frames containing point markers or line markers (rods) are attached to the sterotactic frame in order to provide a reference system for all imaging modalities. Consequently, an accurate registration of all multimodal images for surgery planning is ensured. We use the terminology stereotactic frame exclusively for a frame which is fixated to the skull by screws (Lunsford 1988, Vandermeulen 1991). Non-invasive molds and adapters are treated under the next heading. Stereotactic frame based registration is the least patient friendly of all image integration approaches. It has a high degree of reproducibility, and applying it is rather labor intensive. For a long time, stereotactic frame based matching has been the gold standard for image integration. This holds no longer true. Even though the accuracy of stereotactic frame based methods can be increased by knowledge based methods (Wang et al. 1994), there is increasing evidence that intrinsic methods can attain a higher accuracy, while they are also more attractive by any other of the criteria used. This issue will be further discussed below.

3.2.1.2 Non-invasively fixated mold or frame

Non-invasively fixated alternatives to the stereotactic frame are a thermoplastic mask (Schad et al. 1987), a dental mold (Hawkes et al. 1992), a combination of these (Greitz et al. 1980), and an adapter with a nasion support and ear plugs (Laitinen et al. 1985). These devices are all slightly less accurate than the stereotactic frame, generally more labor intensive because individual molds have to be made for each patient, and provide less reproducible matching results. On the other hand, the methods are more patient friendly and more generally applicable; both limitations to general application of the stereotactic frame hold to a lesser extent for these devices. Non-invasive frames are suited primarily for radiotherapy purposes (registration of CT with treatment beams); they have little use in functional imaging, and thus in functional-anatomical image integration.

3.2.1.3 Skin markers

Image registration using skin markers is patient friendly and applicable to all clinical imaging modalities. The reproducibility is good for brief time intervals, in which case the reference points of the markers can be marked with (if necessary, invisible) ink; for long time intervals, the reproducibility is at best fair. The accuracy of point marker based matching may be quite good under ideal circumstances, *i.e.*, when the tomographic image slices are thin and interslice gaps are narrow or absent, and when the reference points are inside the scanned volume. The method is not very labor intensive if the number of markers is limited to four or five (Lehmann et al. 1991). For image protocols with thick slices and/or large interslice gaps, the accuracy of point marker based registration is poor. Arrow-shaped skin markers were introduced in Van den Elsen et al. (1991) to combine EEG or MEG derived 3D dipole data with tomographic (CT, MRI) image data of the same patient (see also Figure 3.7). Subsequently, this type of marker was used for various image-to-image matching procedures (Knufman et al. 1992, Van den Elsen and Viergever 1994). The main advantage of arrow-shaped markers over point-shaped markers is that they can be located in tomographic images with subslice accuracy, which makes them superior especially in matching datasets with an inferior sampling in the axial direction. Furthermore, the markers can indicate points slightly outside the scanned volume.

3.2.2 Intrinsic matching

The three methods classified under intrinsic matching have two properties in common. The first is the retrospective nature of the match; the imaging protocols need not make provisions for the matching procedure. The second common property is, accordingly, the extreme patient friendliness of the approaches. The key problem of intrinsic matching methods is the selection of the image properties on which the match is based. These have to be derived from quite dissimilar images, which poses a challenging task.

3.2.2.1 Anatomical landmarks

Image registration using anatomical landmarks is generally a rather labor intensive process, since the landmarks have to be pointed out interactively (see *e.g.*, (Evans et al. 1989, Schiers et al. 1989, Hill et al. 1991)). While a first guess may be provided by automatic means, no fully automated landmark extraction algorithms have been reported. Consequently, the approach has a low degree of reproducibility. The accuracy is fair and increases with the number of landmarks used until a certain limit is reached (typically at around 20 - 25 landmarks (Timmens 1991)). Landmark matching is applicable to all tomographic imaging modalities and can readily be extended to nonlinear (curved) matching. This latter property is shared only by voxel based methods.

3.2.2.2 Structures / objects

Object based image matching has become popular by the work of Pelizzari and coworkers (Chen et al. 1987, Pelizzari et al. 1989, Levin et al. 1989). Their method defines objects by contour detection in the 2D slices of the tomographic set; most commonly the external surface of the skin is used for registration. In one image modality, these contours are stacked to generate a surface, the 'head', onto which a 'hat', consisting of a set of points derived from the contours in the other modality, is fitted by means of an optimization procedure. The method is quite accurate and applicable to all tomographic imaging modalities (although for application to PET and SPECT the availability of a transmission scan is desirable); the robustness is questionable, though, owing to the dependency on a high-level object definition. The main disadvantage of the method is that user interaction is required to steer the optimization process, both by identifying the parts of the 'head' and 'hat' to be used and by selecting and adapting the transformation parameters. Several attempts have been made to improve upon the original concept of Pelizzari, *e.g.*, increasing the accuracy of the match by removing outliers (Jiang et al. 1992) and combination with anatomical landmark matching into one algorithm (Collignon et al. 1994), or rendering object matching free of user interaction by basing it on automatic segmentation (Van Herk and Kooy 1994).

Object definition is a high-level task, which is difficult to automate without endangering the accuracy. Instead, multimodality matching may be based on low-level binary feature images, *e.g.*, using a second order invariant (Van den Elsen et al. 1992, Maintz et al. 1996b), or using higher order geometrical features (Liu et al. 1994). The first method compares favorably with registration using arrow-shaped markers, but later proved to be inferior to registering the full feature images. The second method (*also known as* core matching) is a theoretically promising intrinsic registration method (Fritsch et al. 1994), but the practical results of the method are not yet convincing, both because of the size of the registration errors and because of the highly interactive nature of the method (Liu et al. 1994).

3.2.2.3 Voxel properties

Multimodality registration methods based on voxel properties have only recently appeared on the scene, but they have nonetheless taken the lead in brain image

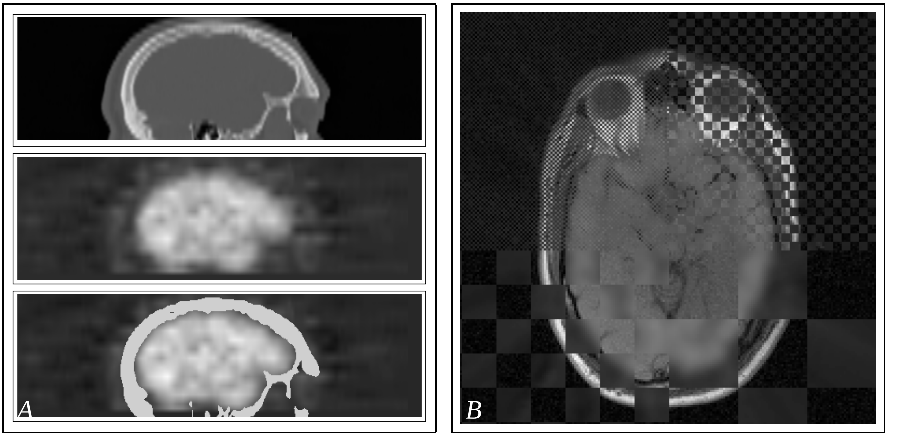

Figure 3.1 *Examples of integrated 2D displays of registrations based on Mutual Information. Frame (A): Adjacent display of a CT image (top), PET image (middle), and PET image with superimposed skull from the CT image (bottom). Frame (B): 'Checkerboard' display of a PET and MR image where the four quadrants have different checkers size (top left: 1 pixel (= alternate pixel display), top right: 4 pixels, bottom left: 16 pixels, and bottom right: 32 pixels).*

matching. This type of methods shares the advantages of being retrospective, patient friendly, and generally applicable with the other classes of intrinsic matching approaches. In addition, most voxel based methods do not require user interaction and thus are both labor extensive and reproducible; they can also be extended to curved (usually referred to as elastic) matching, which is a desired property in image-to-atlas matching and intersubject matching. On top of all this, early algorithms of voxel based type already produced results of surprisingly high accuracy (Van den Elsen 1993, Woods et al. 1993). In consequence, research efforts in multimodality image registration have focused on this class of approaches. Of the recent methods (Hill et al. 1994, Studholme et al. 1995, Collignon et al. 1995, Collignon et al. 1995b, Viola and Wells III 1995, Wells III et al. 1995, Van den Elsen et al. 1995, Maintz et al. 1995, Studholme et al. 1996, Maintz 1996, Maes et al. 1997), the Mutual Information approach described in detail by Maes et al. (1997) appears the most promising (see Figure 3.1).

Table 1 gives a condensed overview of the classes of multimodality image registration techniques vs. seven quality criteria. The latter class of methods, voxel based intrinsic matching, clearly outperforms the other approaches.

3.2.3 Evaluation aspects

An issue which so far has remained undiscussed is the evaluation of registration algorithms as regards accuracy. This poses a serious problem in clinical practice, since

	Accuracy	Patient friendly	Repro-ducible	Labor exten-sive	Retro-spective	Extensible to curved match-ing	Extensible to intra-operative matching
Extrinsic matching							
Frame / screws	+	–	+	±	–	–	+
Mold / adapter	±	±	±	–	–	–	+
Skin markers	±	+	±	+	–	–	+
Intrinsic matching							
Anatomic landmarks	±	+	–	–	+	+	–
Surfaces / objects	±	+	±	±	+	±	±
Voxel properties	+	+	+	+	+	+	±

Table 3.1 *Comparative overview of multimodality image matching approaches. Accuracy: + high, ± satisfactory. Patient-friendliness: + patient friendly, ± bothersome, but non-invasive, – invasive. Reproducibility: + good, ± satisfactory, – questionable. Labor extensiveness: + labor extensive, ± intermediate, – labor intensive. Retrospectivity: + retrospective, – prospective. Extensibility: + readily extensible, ± limited extensibility, – not extensible.*

the best transformation is an unknown. Therefore, the accuracy can only be assessed qualitatively by visual inspection, and quantitatively by composing the found transformation to ones obtained upon employing other registration techniques. Even in the latter case the accuracy measure is not an absolute one, since it is relative to a transformation that inevitably has its own intrinsic registration error.

Different presentation techniques have been used to evaluate registration accuracy, *e.g.*, Van den Elsen (1993) reformatted the usually transaxially acquired tomographic data to coronal and sagittal slices in which directions registration can be much more error prone. By zooming in on specific structures, a detailed account of all misregistrations can be obtained. The display can be improved considerably by making use of selective integration (see section 3.3.1.3) of, *e.g.*, bone or the contour of the brain (see Figures 3.1*A* and 3.5)). Basically all techniques listed in Section 3.3.1 have been applied to assess registration accuracy (see also (Chen et al. 1987, Hill et al. 1993, Pietrzyk et al. 1996)). Use of visualization techniques allows the conclusion that a registration is accurate, but it leaves the problem of how to interpret minor differences.

A novel setup to provide an independent gold standard is the use of cadaver studies, where rigid fiducial tubes are inserted prior to imaging. This has shown to be a viable approach to evaluate CT-MRI registrations (Hemler et al. 1995b, Hemler et al. 1995c). For instance, it clearly shows that the accuracies reported for surface based matching do not hold true for internal brain structures. An extensive evaluation technique was proposed and demonstrated by West et al. (1997). Here, the registration results of many different retrospective methods are compared to a gold standard based on screw-mounted markers. The differences between registration results are evaluated in clinically relevant regions only, and the images used (PET, CT, and MRI) are drawn from a large database of clinically obtained images. Although the actual

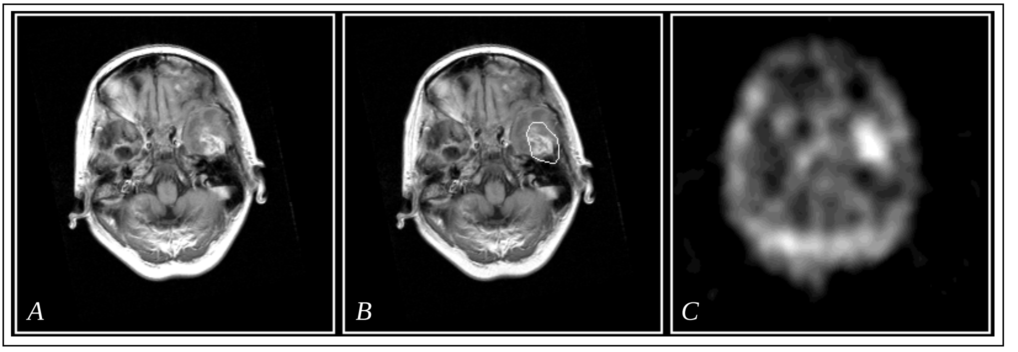

Figure 3.2 *Registration of Thallium–SPECT and MRI using Mutual Information. The image data are taken from a patient with a left temporal lobe abnormality which is indicated by the contour from the abnormal Thallium–SPECT region superimposed onto the MR image. Frame (A) MR image, (B) integrated image, (C) Thallium–SPECT image.*

accuracy of the gold standard used cannot be ascertained, methods like these have the potential to provide a more quantitative measure of accuracy than visual inspection can.

3.3 Integrated image display

When the multimodality images have been matched, the question of how to optimally convey the integrated information remains. This problem is task-driven. For instance, in order to provide an anatomical frame of reference for SPECT using MRI slices with approximately the same slice thickness, a combined MRI/SPECT image may be presented best by a 2D grey value or color display of the original SPECT slices, with the contours of relevant structures (cortex, ventricles, lesions) as derived from the corresponding—resampled—MR slice outlined by a white, black, or color overlay (see Figure 3.5). If, however, for some clinical indication the MR image is the primary source of information with high resolution in the three dimensions and the SPECT image serves to provide additional diagnostic value, the same MRI/SPECT combination may be presented best by the reverse order, *i.e.*, a 2D grey value display of original MR slices with an outline of the corresponding—resampled—SPECT distribution, *e.g.*, indicating an area with abnormal activity (see Figure 3.2). In consequence, in evaluating a specific integrated display technique or in comparing two or more display techniques, the detection task must be well specified. Furthermore, experience with different visualization techniques indicates a significant operator dependency, *e.g.*, superimposed contours in functional data lead to reactions ranging from praise to sheer dislike (see Chapter 6).

We now discuss integrated display approaches that have been proposed in the recent image processing literature and present experiences with several of these techniques for PET/CT, fMRI/MRI, PET/MRI and SPECT/MRI visualization. We distinguish between 2D presentation methods in which one or more tomographic slices through the 3D image dataset, or one or more (usually orthogonal) projections of this dataset are shown, and 3D presentation methods in which a volumetric display of one or more structures is offered. Illustrations of several options for integrated image display are presented. We have used the images from a patient with a right frontal lobe astrocytoma for all following illustrations in this chapter. Note that in some of the methods the presentation is restricted to two image modalities, whereas in other methods simultaneous display of three or even more modalities is supported. Furthermore, several of the methods discussed can be advantageously combined with each other.

3.3.1 2D integrated visualization

A major advantage of the computer is the ability to interact with the data. Visualization is therefore not limited to adjacent display as with the light box, but with the aid of computers data can be processed with the objective to make the display of the information more effective (Kundel 1990, Hill et al. 1991, Hill et al. 1992).

The first logical step beyond adjacent display of the images is integration of multiple 2D images into one image. We distinguish two categories which we denote as 'non-selective' and 'selective' integration.

3.3.1.1 Adjacent display

Adjacent display presents corresponding slices of two (or more) modalities on multiple screens—or in multiple windows on one monitor —, with separate controls for contrast and brightness in each image. It is the simplest, but also one of the most effective types of integrated 2D display, especially when the display is extended with a linked cursor (Hawkes et al. 1990) indicating corresponding locations in the image slices of different modalities (see Figure 3.3).

3.3.1.2 Non-selective integration

Non-selective integration comprises various methods that do not require any user interaction or a priori knowledge. These operations are carried out on the entire images and use *all* available information without making a decision based on the content of a pixel.

- Arithmetic integration:
  This class of methods involves pixelwise addition, subtraction, multiplication, etc. of images (see Figure 3.4).

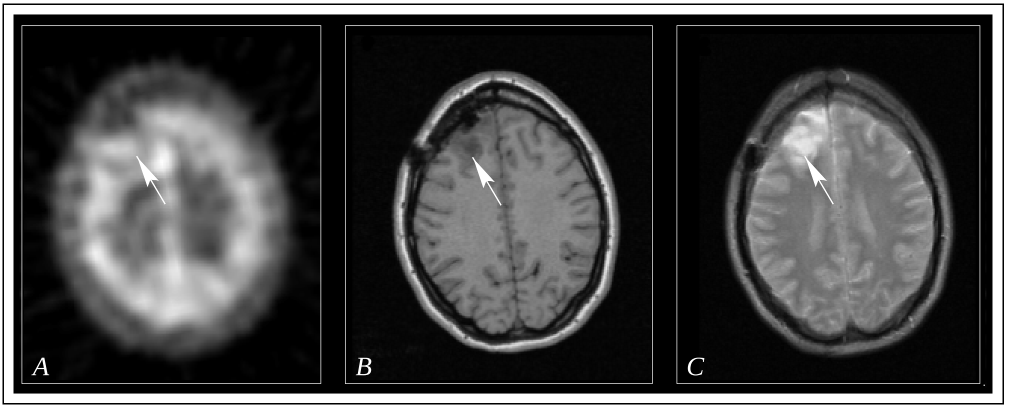

Figure 3.3 *Adjacent display with a linked cursor: Corresponding slices from the original images of three modalities. (A) SPECT image, (B) MR–T1 image, (C) MR–T2 image.*

- Transparency or opacity weighted display:
  This approach stems from the classic work of Porter and Duff (1984). A transparency value is assigned to each pixel (the so-called α-channel or α-value), which determines its contribution to the final image. In theory, the transparency value can be dependent on the content of the voxel or attributed to specific parts by an observer, but to our knowledge only a global value has been applied (for PET/MRI see (Evans et al. 1996)).
- 'Checkerboard' or 'chessboard', alternate pixel and split-screen display:
  Alternate pixels in the display are assigned grey values and/or color to represent the two different modalities (Hawkes et al. 1990, Rehm et al. 1994, Van Herk and Kooy 1994, Hemler et al. 1995b, Robb and Hanson 1996) (see Figure 3.1*B*). Caution is called for when using color, because problems occur when the size of the checker is small and neighboring pixels are perceptually mixed (see (Hawkes et al. 1990) on 'color smearing' and Section 5.2.4.)
- Color modeling:
  Color models can be applied to more effectively convey the information of the multiple datasets to our visual system (Alfano et al. 1992, Kundel 1990). The 'color wash' technique (Pelizzari et al. 1989, Holman et al. 1991) adds a color corresponding with functional information to a grey value corresponding with anatomical data. Furthermore, color information can be encoded as a three-parameter space (*e.g.*, the RGB color model or the HSV color model), with which it is possible to make a mapping of two or three parameters to unique colors. This technique has been amply employed, *e.g.*, for integrated visualization of MRI acquisitions (Weiss et al. 1987, Kamman et al. 1989, Alfano et al. 1995, Robb and Hanson 1996), PET/MRI (Levin et al. 1988), SPECT/CT (Van Herk

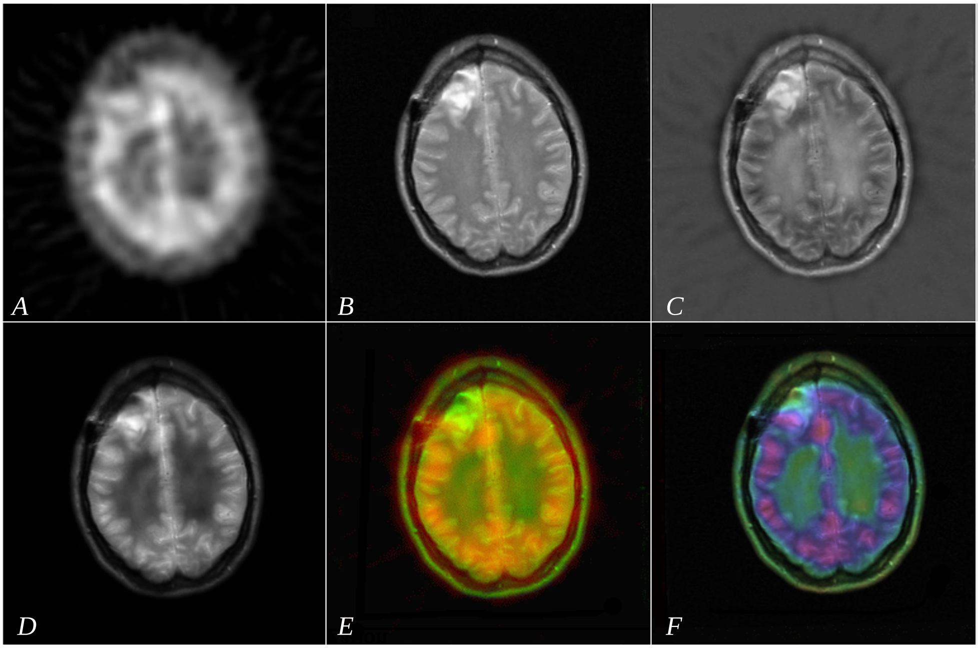

Figure 3.4 *Non-selective integration of (A) original SPECT image and (B) original MR–T2 image. Frames (C), (D), (E), and (F) are the (rescaled) integrated images of (A) and (B) with (C) MRI–T2 minus SPECT grey values, (D) SPECT multiplied with MRI–T2 grey values, (E) RGB integration; SPECT represented by a green color scale fused with MRI–T2 information represented with a red color scale, and (F) HSV integration; the hue component represents the SPECT data and the value component represents the MRI–T2 data.*

and Kooy 1994), and SPECT/MRI (Hawkes et al. 1990, Holman et al. 1991).

With the methods we evaluated, we found that non-selective integration poses the danger of obscuring relevant data by irrelevant information, thereby degrading the overall diagnostic quality of the fused image (see Figure 3.4). This is possibly not problematic for the assessment of registration accuracy, but clinical observation tasks might well be impaired.

In our work with SPECT/MRI we tested both RGB and HSV integration (see Figures 3.4*E* and 3.4*F*, where each of the two modalities is represented by a different component of the RGB or HSV color space). We found RGB integration difficult to interpret for several reasons; *i)* relevant data are camouflaged by irrelevant data (as noted earlier in this section), and *ii)* use of color to present MRI information is not intuitive, traditionally grey values are used.

We found the HSV scheme for integration of SPECT with MRI much more intuitive, with the hue component describing the SPECT data (color encoding is a frequently applied technique in the investigation of SPECT data), and the value component (= achromatic information or shades of grey) describing the MRI data. Manipulation of the remaining saturation component using an appropriate image display tool is a means to change the contribution of color to the image. This enables a gradual increase of the contribution of the functional information to the image from zero to overwhelming.

3.3.1.3 Selective integration

Selective integration requires that certain characteristic features are extracted from one or more modalities and subsequently integrated with one or more other modalities. Contrary to non-selective integration a decision is made by an observer or based on the contents of the images.

- Windowed display:
  The displayed slice is divided into a number of parts, each showing the grey scale contents of one of the involved modalities (Kooy et al. 1994, Soltys et al. 1995). The distinction with the checkerboard display is prominent in the required user interaction for 'steering' the selection.
- Feature display:
  Geometrical features, *e.g.*, points or contours, or objects, *e.g.*, bone, from one or more modalities replace the original information from another modality (Stokking et al. 1994, Farrell et al. 1995, Hemler et al. 1995b).

Extracting well defined structures from SPECT or PET poses problems because the resolution is poor (Kundel 1990). In routine clinical SPECT and PET diagnosis, as much as possible information is used (*e.g.*, comparison with contralateral information is vital), which implies that no functional information should be removed for the integrated visualization. MRI or CT, on the other hand, do allow definition of features of various kind that can be transferred to the functional image to supply an anatomical framework.

In our work with SPECT and MRI we found segmentation of the MR images a good approach to convey anatomical information to the functional SPECT images (Stokking et al. 1994). The outline of the brain, the boundary of the white matter, and the ventricles can be segmented from MRI–T1 data and additional line segments or contours can be drawn to denote landmarks like characteristic sulci, brain areas, and the plane through the longitudinal fissure. Furthermore, the location and the extent of an abnormal region can be outlined and transferred to corresponding images.

Figure 3.5 shows two indicators, a contour for the brain and a filled area for the tumor, which were extracted from MRI and transferred to the SPECT image. The results indicate that selective integration is useful to provide the SPECT data with

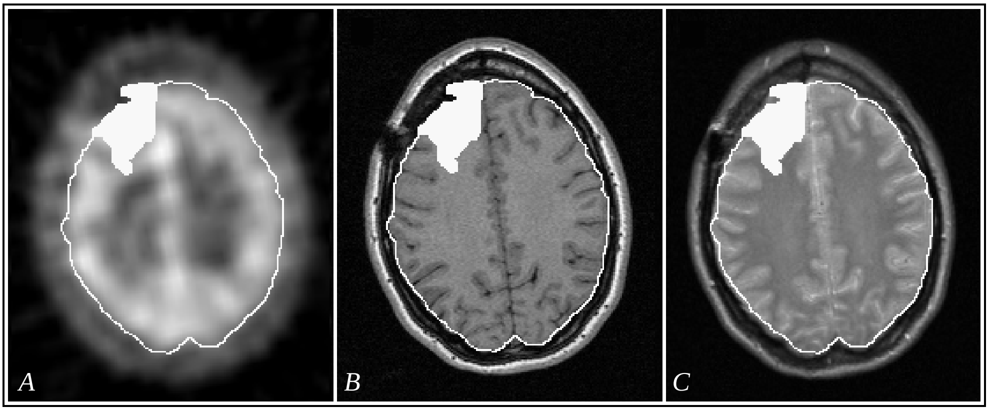

Figure 3.5 *Selective integration using manual segmentation. The brain contour drawn in the MR–T1 image (B) and the damaged region rendered white in the MR–T2 image (C) are transferred to the corresponding SPECT image (A).*

an anatomical framework (see Chapter 6), without loss of relevant information. To illustrate the waste of relevant SPECT data we deliberately applied a filled region for the tumor area. The use of filled regions obviously calls for caution, which is why we feel that it should only be used when non-relevant SPECT information is replaced, *e.g.*, outside the brain or with good anatomical landmarks like the ventricles. Accordingly, for the case of Figure 3.5 we would replace the area by a contour in the clinical application.

3.3.1.4 Discussion of 2D integrated visualization

Adjacent display of registered functional and anatomical images presents all acquired information in an easy and intuitive way, where the use of a linked cursor is a very simple, but effective tool to assist in the integration of the information. Technically, however, the information is still presented separately and the observer has to perform the mental integration.

A common denominator in our experiments with non-selective integration for functional-anatomical visualization was that no or little additional diagnostic information could be conveyed compared to adjacent display, while valuable features were often camouflaged by non-diagnostic information. Consequently, while these techniques can be attractive thanks to their ease-of-use and high speed, they are not very effective for the integrated presentation of information from functional and anatomical data. The integration using color models is an exception, since our visual system employs color more effectively than grey levels (Weiss et al. 1987, Kundel 1990, Gouras 1991). Of the employed color models, RGB was not satisfactory for SPECT/MRI and PET/MRI integration, but HSV gave promising results (see also

Chapter 5 for use in 3D visualization). The HSV integration scheme will be subject to further research, but as yet we favor both adjacent display with a linked cursor and selective integration to integrate only the diagnostic relevant information.

With functional-anatomical visualization, the use of selective integration offers the possibility to display the relevant data from the different modalities more effectively. The approach also has an inherent drawback, *viz.* that a selection of features has to be made to perform a specific integration task. For some applications, delineation of the brain contour may be sufficient, but more complex cases may require the definition of additional features. In our work with SPECT and MRI, we used manual and semi-automated (thresholding and region growing) segmentation routines to extract several features (brain contour, tumor area, ventricles, etc.). The speed, accuracy and reproducibility with which the resulting visualizations were obtained, led us to favor semi-automated segmentation. For future work we are inclined to use more sophisticated, (near-) automatic segmentation methods (*e.g.*, those based on morphology (Robb and Hanson 1996) or multiresolution (Vincken et al. 1995)), where suitable choices and proper tools should reduce the workload significantly. Furthermore, it is apparent from our own work presented in Chapter 2 that various segmentation tasks allow complete automation. Consequently, an observer can be offered reproducible integrated visualizations that require no user interaction.

The most obvious handicap of the described 2D multimodality visualization methods is the inherent lack of 3D information (see also (Keyes Jr. 1990, Maisey et al. 1992)). The observer must study consecutive slices to mentally reconstruct the 3D picture needed for proper diagnosis, treatment planning, and communication with the referring physician or surgeon. The following section will discuss several techniques for 3D multimodality visualization of functional and anatomical brain data that alleviate this task.

3.3.2 3D integrated display methods

The application of volume visualization techniques to display 3D image information is steadily increasing in routine clinical work. For a rapid assessment of functional and anatomical relations nuclear medicine physicians investigate image data using, *e.g.*, a MIP or a Bull's eye display (Wallis 1992) and radiologists use a MIP or a 3D surface visualization. Furthermore, a 3D representation can improve communication with the referring specialist (see also (Wallis 1992)). The efficient presentation of information from multiple sources is even more demanding as mental integration of the wealth of information is nigh impossible. To assist clinicians in extracting all relevant information from the data, several integrated 3D visualization techniques have been reported. In this section, we review and present techniques for 3D integrated visualization of functional and anatomical brain images.

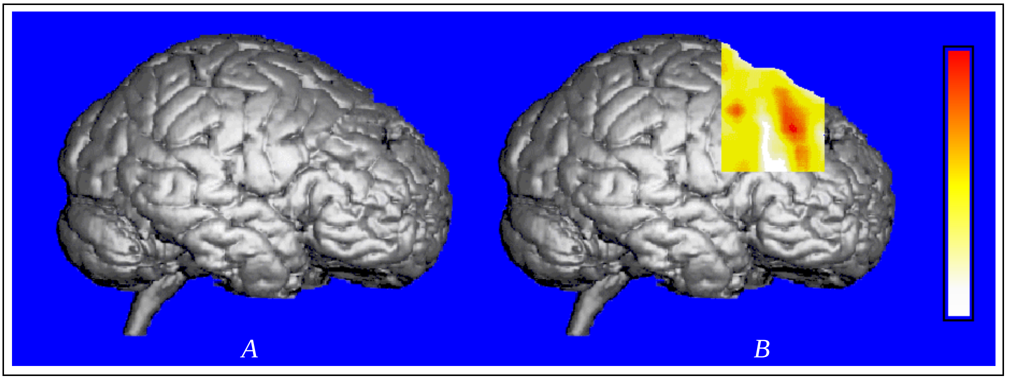

Figure 3.6 *Multimodal window display. Frame (A) shows a volumetric rendering of the brain from the MR–T1 images (voxel gradient shading). In Frame (B) part of the rendering is replaced by a MIP of the corresponding SPECT data (colored insert, lookup table on the right).*

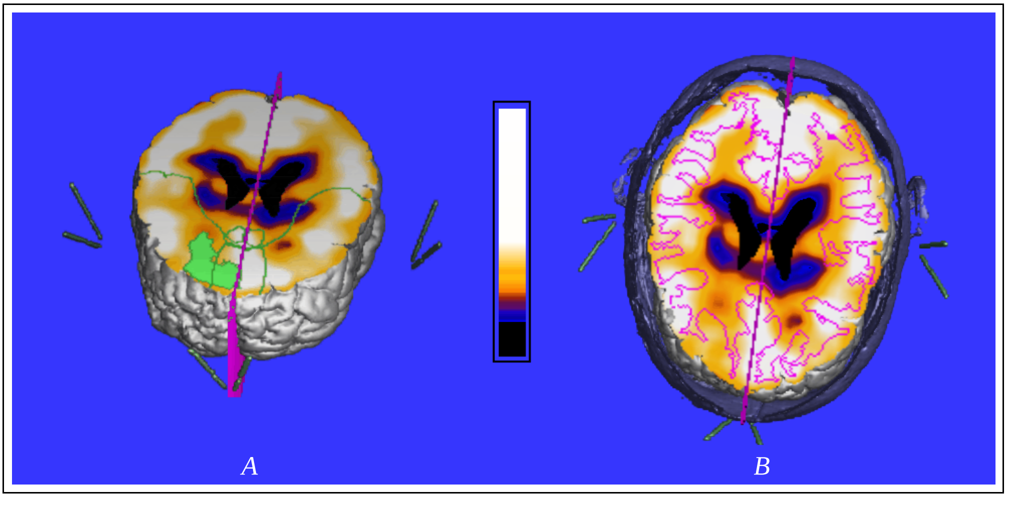

Figure 3.7 *Multimodal cutplane visualization of a plane through the frontal lobe tumor area, with a voxel gradient shaded rendering of the cortex from MRI–T1 as a reference frame. The cutplane shows SPECT activity with a number of diagnostic features transferred from the corresponding MRI–T1 and T2 slices. Frame (A) shows the tumor area and ventricles as filled regions and different parts of the brain outlined by contours, Frame (B) shows the boundary between grey and white matter plotted in purple. The SPECT lookup table is shown in the middle. In both images the three arrow-shaped markers used for the registration of these images are clearly visible (see Section 3.2.1.3).*

3.3.2.1 Linked feature display

The linked cursor discussed with 2D adjacent display can be extended to 3D by correlating a location in a 3D rendering to the corresponding position in a 2D image of another modality. Features like points, lines (Levin et al. 1989, Hu et al. 1990), or planes can be used. The possibilities to interact with the display may be absent or fairly limited, *e.g.*, when a rendering from one of the six directions of a cube of data contains a line indicating one of the image slices of another modality. However, when a task and/ or an observer demands full interaction with, *e.g.*, a plane and a rendering to indicate an oblique image slice in another modality, more intricate techniques are required.

3.3.2.2 Integrated data display

Volumetric structures as derived from various modalities are integrated into one dataset and subsequently displayed by standard rendering techniques, *e.g.*, as opaque or transparent surfaces (Viergever et al. 1992, Evans et al. 1996) or points and contours as single, or multicolor overlays (Van den Elsen et al. 1991).

3.3.2.3 Multimodal window display

A 3D volume visualization, although it creates a 3D illusion, is basically a 2D representation of the image data and all techniques for integrated 2D display can be applied. We will discuss the multimodal window approach (see section 3.3.1.3) as this seems the most promising of these 'semi-3D' methods for 3D integrated visualization, but also to illustrate the consequences of volume visualization for 3D integrated visualization.

A 'window' in a volume visualization of anatomical data reveals the corresponding part of the functional data. Once the 3D renderings of the functional and anatomical volumes are available, actual integration is very fast, as it only involves the use of 2D cut-and-paste routines.

Figure 3.6 shows a voxel gradient shading of the cortex from MRI–T1 with an insert from a maximum intensity projection of the SPECT data of the right hemisphere. A tumor area is indicated by the deteriorated cortical surface structure (in front of the window), and the SPECT insert indicates a strip of increased blood perfusion in the vicinity of the damaged region. Several researchers have presented similar images, *e.g.*, for MRI/PET (Levin et al. 1989) and MRI/CT (Stimac et al. 1988).

A wide range of techniques can be applied for multimodal window visualization of the functional and anatomical images. 'Hot' or 'cold-spot' visualization (Wallis 1992) and projection techniques can be employed for 3D SPECT and PET visualization. However, hot or cold-spot visualization requires a subjective threshold, and the projection techniques, like the MIP we applied here, generally fall short in conveying real 3D information (Ehricke and Laub 1990) and represent a viewpoint-dependent area of image information, rather than local data. New techniques for improved 3D functional visualization, like the method for 3D SPECT visualization presented by Hashikawa et al. (1995), will consequently also improve the multimodal window display.

3.3.2.4 Multimodal cutplane display

The use of cutplanes has become a standard technique in (single modality) volume visualization. Its established use in volume visualization indicates that it is a powerful technique for the investigation of clinical datasets. Landmarks visible in a volumetric rendering can be used to position a cutplane for close investigation of corresponding data. In multimodality visualization, cutplanes have been used, *e.g.*, in a volumetric visualization of skin from MRI and skull from CT, with two cutplanes representing the original CT and MRI grey values (Schiers et al. 1989) or in a volume visualization of the brain from MRI with a cutplane representing functional information (Payne and Toga 1990). In Stokking et al. (1994) functional data from SPECT images and anatomical features from MR images are presented on a multimodal cutplane in an anatomical framework supplied by a 3D rendering of the brain from the MR images. A cutplane is basically a 2D image. We used the results of 2D multimodality visualization, where we favored selective integration of features, to guide the integration of SPECT and MRI data on the cutplane. In Figure 3.7 the SPECT data on the cutplane are color encoded through a lookup table and the MRI–T1 and T2 characteristics are represented by conspicuous colors not used in the SPECT lookup table.

3.3.2.5 Surface texturing and mapping

In functional brain research, one of the primary regions of interest is the grey matter of the folded surface layer (about 2-10 mm thick). Integration of information can be performed by texture mapping functional information onto a surface (Payne and Toga 1990), but this will only present a small part of the interesting functional information in one comprehensive image. A technique proposed by Valentino et al. (1991) first maps a neighborhood of functional information onto an anatomical volume followed by rendering of the combined volume. Other methods aim more specifically at mapping underlying grey matter activity onto the brain surface rendered from anatomical data. Levin et al. (1989) uses a technique that samples functional activity below the surface along the viewing direction, which introduces (localization) artefacts. The Normal Fusion technique (see also Chapter 4) overcomes these artefacts by spawning a secondary ray along the reverse gradient (inward normal) at a surface (in volume visualization the gradient is generally calculated to determine shading (Höhne et al. 1990)) to sample local functional activity and, subsequently, color encode the sampled information onto the surface rendered from anatomical data. The

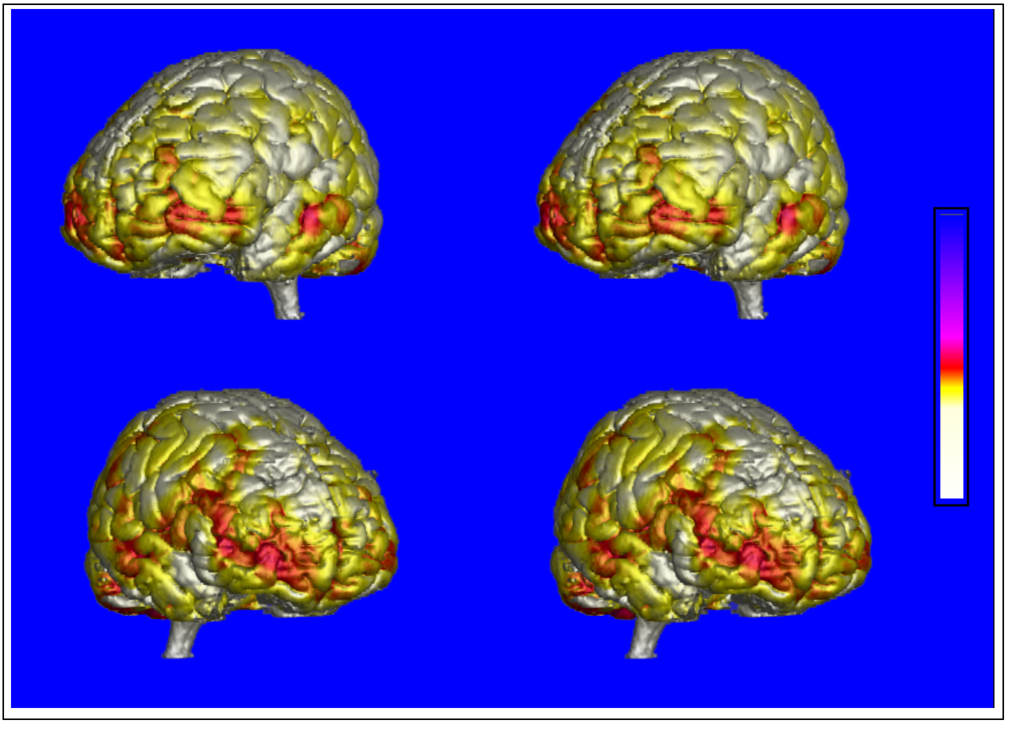

Figure 3.8 *Multimodal SPECT/MRI visualization using normal fusion. The maximum intensity of the SPECT samples along the direction of the normal was used to color encode the brain surface via a lookup table (shown on the right). Top: in stereo (cross-eyed viewing) a left frontal view of the healthy side of the brain. Bottom: in stereo a right frontal view of the inflicted hemisphere. The tumor is located in the right frontal lobe, indicated by the deteriorated structure of gyri and sulci. A comparison of the left and right frontal lobes clearly shows an area of increased cerebral blood perfusion surrounding the tumor.*

approach has been quite successful in simultaneous visualization of SPECT, PET, or fMRI with structural MRI (see Figure 3.8), and recent results with SPECT and MRI data have confirmed that this method adds diagnostic value to straightforward interpretation of the individual slices (see Chapters 4 and 6).

The renderings presented in Figure 3.8 can be used for left-right comparison of the hemispheres. This comparison may be alleviated by using different projection approaches thereby supplying roundabout views of the brain in one image. We have applied two comparable approaches using cylinder and spherical projection (for an example of the cylinder map see Figure 3.9). Others have applied similar techniques in investigation of skull from CT images (Vannier et al. 1994), integration of EEG data with an unfolded brain from MRI (Holländer 1995), supplying a sperical view for radiotherapy planning based on CT images (Bendl et al. 1995), or unfolding of

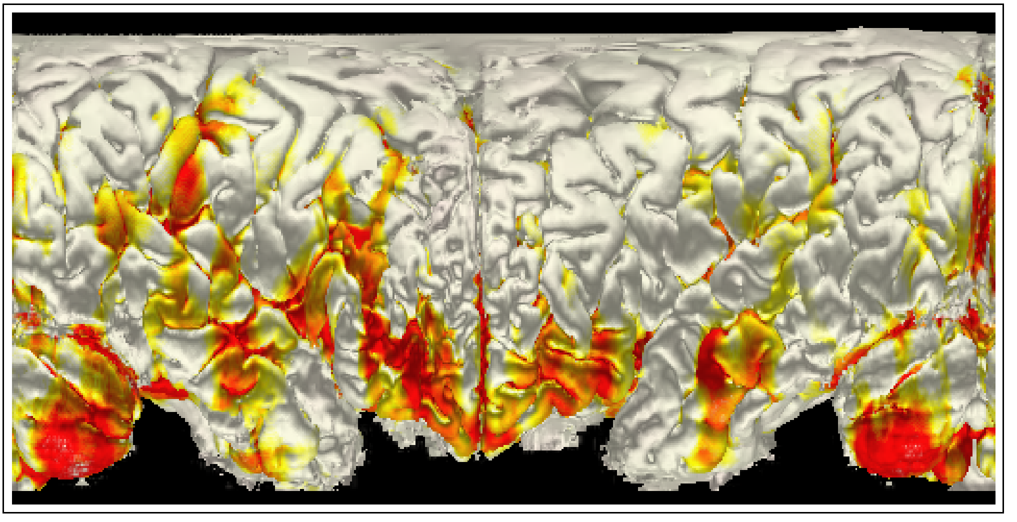

Figure 3.9 *Cylinder projection (frontal view) for integrated visualization of SPECT and MRI data.*

SPECT data for the heart (Lamoureux et al. 1990, Wallis 1992).

3.3.2.6 Discussion of 3D integrated visualization

Linked feature display is an extension of the linked cursor for 2D display and appears very effective to indicate corresponding positions in different modalities. The intricacy of the applied interaction technique is both task and observer dependent.

Integrated data display can be a simple and effective technique to combine information from different sources, as long as the interesting information from one of the modalities has been preprocessed and binarized using a point, line, or other geometrical feature.

Multimodal window display is hampered by problems generally encountered in volume visualization of functional data (see also (Links and Devous Sr. 1995)). However, it offers a quick assessment (since it is essentially an integration of 2D images) of functional information in an anatomical framework.

Multimodal cutplane display is capable of providing selective information on integration of functional and anatomical data in one comprehensive image that visualizes the plane of interest with respect to a 3D anatomical framework supplied by volumetric rendering.

Surface texturing appears to present little information and clinical use is highly questionable. The surface mapping techniques where the viewing direction is used for integration may result in some appealing images, but localization problems seriously hampers general use (as is indicated by Hu et al. (1990)). Accurate localization is obtained using the Normal Fusion technique (see Chapter 4), where the local curvature of the surface is used for integration. This method presents information about the blood perfusion in the surface layer of cortical grey matter within an anatomical frame of reference.

The techniques described in this chapter can be further enhanced by making use of stereo or animated images. The improved depth perception from stereo images is demonstrated in Figure 3.8. Animation strongly contributes to 3D perception and we refer to our web-site (http://www.cv.ruu.nl).

3.4 Conclusions

Research of multimodality image registration has made huge progress in the last few years. Especially, intrinsic methods based on voxel similarity criteria have improved significantly with maximization of Mutual Information as the most promising approach.

There is a wide variety of methods for integrated display. Which method is best in a specific situation depends on the image understanding task to be performed and individual preferences of the observer. In many applications, a 2D display in grey or color of one or more slices of the primary diagnostic modality, overlaid by relevant structures (points, contours, objects) from another modality appears to be quite appropriate. Volumetric (3D) integrated display is called for in some cases, but is certainly not optimal for all interpretation tasks.

Acknowledgements

We would like to thank our colleagues P.C. Anema, F.J. Beekman, P.A. van den Elsen, J.H. de Groot, J.W. van Isselt, R.T.M. Jonk, J.B.A. Maintz, L.C. Meiners, M. Metselaar, L.J.A. Peeters, J. Pluim, P.P. van Rijk, and G.R. Timmens for their contributions.