Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Introduction to Clinical and Laboratory (Small-Animal) Image Registration and Fusion

Pat B. Zanzonico and Sadek A. Nehmeh

Abstract—**Imaging has long been a vital component of clinical medicine and, increasingly, of biomedical research in small-animals. Clinical and laboratory imaging modalities can be divided into two general categories, structural (or anatomical) and functional (or physiological). The latter, in particular, has spawned what has come to be known as "molecular imaging." Image registration and fusion have rapidly emerged as invaluable components of both clinical and small-animal imaging and has lead to the development and marketing of a variety of multi-modality, e.g. PET-CT, devices which provide registered and fused three-dimensional image sets. This paper briefly reviews the basics of image registration and fusion and available clinical and small-animal multi-modality instrumentation.**

I. Introduction

Since the discovery of x-rays by Wilhelm Roentgen, imaging has been a vital component of clinical medicine. Increasingly, in vivo imaging of small laboratory animals, i.e. mice and rats, has emerged as an important aspect of basic biomedical research. Clinical and laboratory imaging modalities can be divided into two general categories, structural (or anatomical) and functional (or physiological). Anatomical modalities, i.e. depicting primarily morphology, include x-rays (plain radiography), CT (computed tomography), MRI (magnetic resonance imaging), and US (ultrasound). Functional modalities, i.e. depicting primarily information related to underlying metabolism, include (planar) scintigraphy, SPECT (single-photon emission computed tomography), PET (positron emission tomography), MRS (magnetic resonance spectroscopy), and fMRI (functional magnetic resonance imaging); planar scintigraphy, SPECT, and PET are components of nuclear medicine. The functional modalities form the basis of the rapidly advancing field of "molecular imaging," defined as the direct or indirect non-invasive monitoring and recording of the spatial and temporal distribution of in vivo molecular, genetic, and/or cellular processes for biochemical, biological, diagnostic or therapeutic applications[1].

Since information derived from structural and functional images is often complementary, e.g. localizing the site of an apparently abnormal metabolic process to a pathologic structure such as a tumor, integration of this information may be helpful and even critical. In addition to anatomic localization of "signal" foci, image registration and fusion provides: intra- as well as inter-modality corroboration of diverse images; more accurate and more certain diagnostic and treatment-monitoring information; image guidance of external-beam radiation therapy; and potentially, more reliable internal radionuclide dosimetry, e.g. in the form of radionuclide image-derived "isodose" contours superimposed on images of the pertinent anatomy.

The initial step in this integration process is to bring the respective images into spatial alignment in a common coordinate system, a procedure referred to as *registration*. After registration, *fusion* is required for the integrated display of these aligned images. This paper briefly reviews the basics of image registration and fusion and available clinical and small-animal multi-modality instrumentation and provides illustrative examples of registered and fused multi-modality images in both clinical and laboratory settings.

II. Technical Aspects of Image registration and Fusion

A. Registration

The first step in registration of multiple image sets is re-formatting of one image set (the "floating," or secondary, image) to match that of the other image set (the reference, or primary, image). (Alternatively, both image sets may be transformed to a new, common image format.) Three-dimensional (3D), or tomographic, image sets are characterized by: the dimensions (e.g. in mm), i.e. the length (ΔX), width (ΔY), and depth (ΔZ), of each voxel; the image matrix, XxYxZ = number of rows, X x number of columns, Y x number of tomographic images, Z; and the image depth (e.g. in bytes), which defines the dynamic range of signal display-able in each voxel (e.g. a word-mode, i.e. one word- or two byte-"deep," PET image can display up to 2^{16} = 65,536 counts per voxel for 16-bit words). The foregoing image parameters are provided in the image "header," a block of data which may either be in a stand-alone text file associated with the image file or incorporated into the image file itself. Among the image sets to be registered, either the finer matrix is re-formatted to the coarser matrix by combining of voxels or the coarser matrix is re-formatted to the finder matrix by interpolation of voxels. One of the resulting 3D image set is then magnified or minified to yield

Manuscript submitted May 26, 2006. This work was supported in part by National Institutes of Health Public Health Service grants 5 P50 CA86438 and R24 CA83084.

Pat Zanzonico and Sadek Nehmeh are with the Department of Medical Physics, Memorial Sloan-Kettering Cancer Center, New York, NY 10021 (corresponding author is Pat Zanzonico - phone: 212-639-2487; fax: 212-717-3442; e-mail: ZanzoniP@MSKCC.org). Dr Nehmeh's e-mail is NehmehS@MSKCC.ORG.

primary and secondary images with equal voxel dimensions. Finally, the "deeper" image is re-scaled to match the depth of the "shallower" matrix. Usually, the higher-spatial resolution and finer-matrix structural (e.g. CT) image is the primary image and the functional (e.g. PET) image the secondary image.

The second step in image registration is the transformation of the re-formatted secondary image set to spatially align it, in three dimensions, with the primary image set. Such transformations are commonly characterized as either "rigid" or "non-rigid" [2,3].In a rigid transformation, the secondary image is translated and/or rotated with respect to the primary image. The Euclidean distance between any two points (i.e. voxels) within an individual image set remains constant, however. In non-rigid, or deformable, transformations (commonly known as "warping"), selected sub-volumes within the image set may be expanded or contracted and/or their shapes altered. Translations and/or rotations may be performed as well. Such warping is therefore distinct from the magnification or minification described in the re-formatting step, where distances between points *all* change by the same *relative* amount. Unlike rigid transformations, which may be either manual or automated, non-rigid cannot be implemented manually and thus are always automated.

Registration transformations are often based on alignment of specific landmarks visible in the image sets; this is sometimes characterized as the "feature-based" approach. Such landmarks may be either intrinsic, i.e. one or more well-defined anatomic structure(s), or extrinsic, i.e. one or more fiducial markers placed in or around the patient or animal. Other registration algorithms, sometimes characterized as "intensity-based" approaches, include alignment of the respective "centers of mass" or principal axes calculated for each image set. A widely used automated registration algorithm is based on the statistical concept of mutual information[4]. The mutual information, or transinformation, of two random variables X and Y is a quantity that measures the mutual dependence of the two variables. Mutual information measures the information about X that is shared by Y. If X and Y are independent, then X contains no information about Y and vice versa and their mutual information is therefore zero. Conversely, if X and Y are identical, then all information conveyed by X is shared with Y and their mutual information is maximized. If a patient or animal is imaged by two different modalities, there is presumably considerable mutual information between the spatial distribution of the respective signals in the two images sets no matter how diverse. For example, the distribution of fluorine-18-labeled fluoro-deoxyglucose (FDG) visualized in a PET scan is, at some level, dictated by (i.e. dependent on) the distribution of different tissue types imaged by CT. Accurate spatial registration of two such image sets thus results in maximization of their mutual information and vice versa.

The third and fourth steps are, respectively, evaluation of the accuracy of the registration of the primary and transformed secondary images and adjustment, iteratively, of the secondary-image transformation until the registration is optimized. The evaluation and adjustment may be as simple as visual (i.e. qualitative) inspection of the aligned images and a judgment by the user that the registration is or is not "acceptable." A more objective, and preferably quantitative, evaluation of the accuracy of the registration is, of course, preferable. One "similarity metric," for example, is the sum of the Euclidean distances between corresponding fiduciary markers (or anatomic landmarks) in the two image sets; the optimum alignment then corresponds to the transformation yielding the minimum sum of distances. Another similarity metric, as noted, is the mutual information: when the mutual information between the two image sets is maximized, they are optimally aligned.

B. Fusion

As noted, fusion is the integrated display of registered images. This may be as simple as simultaneous display of images in a juxtaposed format. A more common, and more useful, format is an overlay of the registered images, where one image is displayed in one color table and the second image in a different color table. Typically, the intensities of the respective color tables as well as the "mixture" of the two overlayed images can be adjusted. Adjustment (e.g. with a slider) of the mixture allows the operator to interactively vary the overlay so that the designated screen area displays only the first image, only the second image, or some weighted combination of the two images, each in its respective color table.

III. Practical Approaches to Image registration and Fusion

A. General Considerations

In both clinical and laboratory settings, there are two practical approaches to image registration and fusion, which may be characterized as the "software" and "hardware" approaches, respectively; both approaches are, of course, dependent on both the appropriate software and hardware. In the software approach, images are acquired on separate devices, imported into a common image-processing workstation, and registered and fused using the appropriate software. In the hardware approach, images are acquired on a single, multi-modality device and transparently registered and fused with the manufacturer's integrated software. Because the respective images may be acquired on different makes of imaging devices, each employing some proprietary format, the hardware approach is dependent on software sufficiently robust to recognize and import diverse image formats. The availability of industry-wide standard formats, such as the ACR-NEMA DICOM standard (i.e. the American College of Radiology (ACR) and National Electrical Manufacturers Association (NEMA) for Digital

Imaging and Communications in Medicine (DICOM) standard[5-7], is therefore critical, not only for image registration and fusion but also for picture archiving and communication systems (PACS).

Robust image registration and fusion software, running on diverse computer platforms, is now widely available.

B. *Multi-modality Devices*

Software to register images acquired on separate devices has been notably successful in the brain because of the ability to reliably immobilize and position the head, the pronounced contrast between the bony skull (an intrinsic landmark) and the brain, and the lack of motion or deformation of internal structures. Outside the brain, however, software registration is difficult and often unsuccessful because of the many degrees of freedom of the torso and its internal structures when imaged by two different devices in two different positions at two different times. For example, depending on the variable degree of filling of the bladder with urine or the intestines with gas, pelvic and abdominal structures may be significantly displaced from one imaging study to the next. At best, the registration process is labor intensive and not routine.

Multi-modality, e.g. PET-CT, devices overcome many of these difficulties. The major manufacturers of PET and CT scanners now also market multi-modality scanners[8,9], combining high-performance state-of-the-art PET and CT scanners in a single device. These instruments provide near-perfect registration of images of in vivo function (PET) and anatomy (CT) and are already having a major impact on clinical practice, particularly in oncology. PET-CT devices are currently outselling "PET-only" systems by a two-to-one ratio[10]. Although generally encased in a single seamless housing, the PET and CT gantries in such multi-modality devices are separate; the respective fields of view are separated by a distance of the order of 1 m and the PET and CT scans are performed sequentially[11]. In one such device (Gemini, Philips Medical), the PET and CT gantries are actually in separate housings with a sizable space between them; this not only provides access to patients but also may minimize anxiety among claustrophobic subjects.

In addition to PET-CT scanners, SPECT-CT scanners are now commercially available. The design of SPECT-CT scanners is similar to that of PET-CT scanners in that the SPECT and CT gantries are separate and the SPECT and CT scans are acquired sequentially, not simultaneously. In such devices, the separation of the SPECT and CT scanners is more apparent because the rotational and other motions of the SPECT detectors effectively precludes encasing them in a housing with the CT scanner.

Multi-modality imaging devices for small animal (i.e. mice and rats) - PET-CT, SPECT-CT, and even SPECT-PET-CT devices - are now commercially available as well.

Multi-modality devices simplify image registration and fusion - conceptually as well as logistically – by taking advantage of the fixed geometric arrangement between the PET and CT scanners or the SPECT and CT scanners in such devices. Further, because the time interval between the sequential scans is short (i.e. a matter of minutes), it unlikely that a patient's geometry will change significantly between the PET or SPECT scan and the CT scan. Accordingly, a rigid transformation matrix (i.e. translations and rotations in three dimensions) can be used to align the PET or SPECT and the CT image sets. This matrix can be measured using a "phantom," i.e. an inanimate object with PET- or SPECT- and CT-visible landmarks arranged in a well-defined geometry. The transformation matrix required to align these landmarks can then be stored and used to automatically register *all* subsequent multi-modality studies, since the devices mechanics and therefore this matrix should be fixed.

IV. Examples of Registered and Fused Multi-modality Imaging Studies

To illustrate the utility of registered and fused multi-modality imaging studies in both clinical and laboratory settings, examples are presented in Figures 1 and 2.

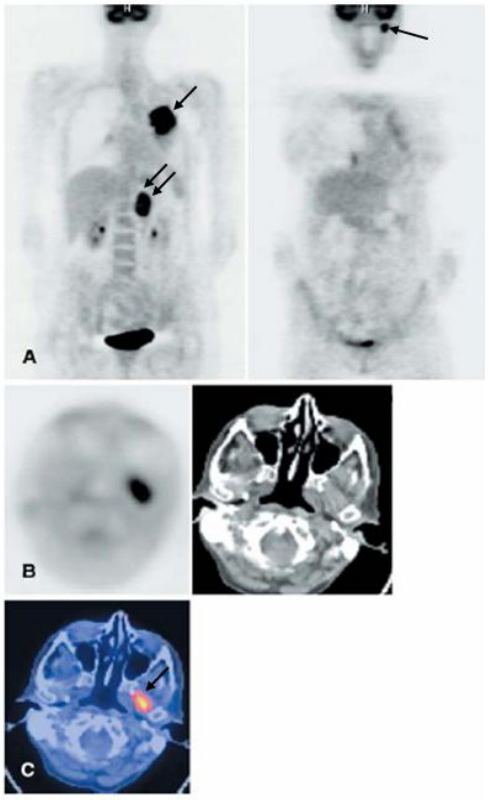

Figure 1. Registered and fused FDG PET and CT scans of a patient with lung cancer and an adrenal gland metastasis. (a) Coronal PET images show typically increased FDG uptake in the primary lung tumor (*single arrow in left panel*) and in the metastasis in the left adrenal gland (*double arrow in left panel*) but also in the left side of the neck (*arrow in right panel*). (b) Transaxial PET and CT images through this neck

lesion. Reading these images separately or in the juxtaposed format shown, it is difficult to definitively identify the anatomic site (i.e. tumor versus normal structure) of the focus of activity in the neck. (c) The registered and fused PET-CT images, using the overlay display discussed in Section II. B., unambiguously demonstrate that the FDG activity is located within muscle, a physiological normal variant. Because it is best visualized using the original color display, an arrow is used to identify the location of this unusual, but non-pathologic, focus of FDG activity on the fused images. Therefore, the FDG activity in the neck was *not* previously undetected disease, a finding which would significantly impact the subsequent clinical management of the patient. Adapted from reference [12] with permission of the authors.

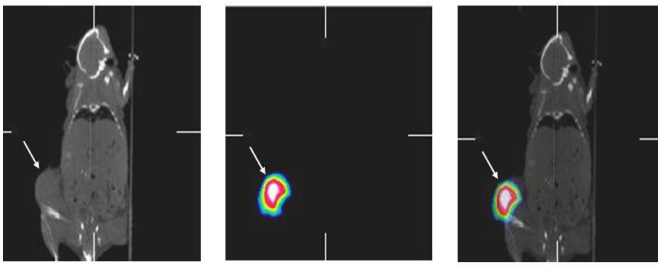

Figure 2. Registered and fused SPECT-CT images (coronal views) of a mouse with a LAN1 neuroblastoma tumor xenograft in its hindlimb (arrow). The radiotracer was iodine-125-labeled 3F8, an antibody directed against the ganglioside 2 (GD2) antigen, which is overexpressed on neuroblastomas, including LAN1. The images were acquired at 2 days post-injection. The CT image shows the tumor as a space-occupying structure along the contour of the animal (*left panel*). The specific targeting of the radiolabeled 3F8 to the GD2-expressing tumor xenograft is demonstrated by the high-contrast SPECT image (*middle panel*). The registered and fused PET-CT images, again using the overlay display discussed in Section II. B., unambiguously demonstrates that the 3F8 activity is located in the tumor, confirming that the focus of activity represents specific tumor-targeting by this antibody and not, for example, excreted activity in the urinary bladder or radioactive contamination.

V. Concluding Remarks

Image registration and fusion has rapidly emerged as an invaluable component of both clinical and small-animal imaging and has lead to the development and marketing of a variety of multi-modality devices which provide registered and fused three-dimensional image sets. To date, such multi-modality devices have been restricted almost exclusively to PET-CT and SPECT-CT scanners. While MRI-CT scanners might have little practical advantage, since both MRI and CT are both anatomic imaging modalities, PET-MRI and SPECT-MRI devices would be highly attractive. Combining PET or SPECT and MRI remains problematic, however, because the magnetic fields proximal to an MRI scanner interfere with the scintillation detection process in all current-generation PET and SPECT scanners.

References

[1] RSNA News, *RSNA, SNM urge interdisciplinary cooperation to advance molecular imaging,* 2005.

[2] J.B.A. Maintz and M.A. Viergever, "A survey of medical image registration," *Med Image Anal* 1998, vol. 2, pp. 1-36, 1998.

[3] J.V. Hajnal, D.L.G. Hill DLG, and D.J. Hawkes DJ, Eds., *Medical image registration*, Boca Raton, FL: CRC Press, 2001.

[4] P. Viola and W.M. Wells III. "Alignment by maximization of mutual information," *Inter J Computer Vision*, vol. 22, pp. 137-154, 1997.

[5] American College of Radiology, National Electrical Manufacturers Association, "ACR-NEMA Digital Imaging and Communications Standard," *NEMA Standards Publication No. 300-1985*, Washington, DC, 1985.

[6] American College of Radiology, National Electrical Manufacturers Association, "ACR-NEMA Digital Imaging and Communications Standard: Version 2.0," *NEMA Standards Publication No. 300-1988*, Washington, DC, 1988.

[7] American College of Radiology, National Electrical Manufacturers Association, "Digital Imaging and Communications in Medicine (DICOM): Version 3.0," *Draft Standard,* ACR-NEMA Committee, Working Group VI, Washington, DC, 1993.

[8] T. Beyer T, D.W. Townsend, T. Brun, et al. "A combined PET/CT scanner for clinical oncology," *J Nucl Med*, vol. 41, pp.1369-79, 2000.

[9] D.W. Townsend, J.P.J. Carney, J.T. Yap. et al. "PET/CT today and tomorrow," *J Nucl Med*, vol. 445, pp.4S-14S, 2004.

[10] J Nucl Med (Newsline), *PET on display: Notes from the 59th SNM Annual Meeting*, pp. 24N-26N, 2003.

[11] J.T. Yap, J.P.J. Carney, N.C. Hall. et al. "Image-guided cancer therapy using PET/CT," *Cancer J*, vol. 10, pp.221-223, 2004.

[12] H. Schoder, Y. Erdi, S. Larson, et al. PET/CT: a new imaging technology in nuclear medicine, *Eur J Nucl Med Mol,* vol. 30, pp. 1419-1437, 2003.