Medical Dosimetry, Vol. 33, No. 2, pp. 149-155, 2008

0958-3947/08/$–see front matter
ELSEVIER
doi:10.1016/j.meddos.2008.03.001

MULTIMODALITY IMAGE FUSION AND PLANNING AND DOSE DELIVERY FOR RADIATION THERAPY

CHENG B. SAW, PH.D., HUNGCHENG CHEN, M.S., RON E. BEATTY, M.ENGR., and HENRY WAGNER, JR., M.D.
Division of Radiation Oncology, Penn State Hershey Cancer Institute, Hershey, PA; and Department of Radiation Oncology, UPMC Cancer Centers – McKeesport, McKeesport, PA

(*Received* 1 *November* 2007; *accepted* 29 *February* 2008)

Abstract—Image-guided radiation therapy (IGRT) relies on the quality of fused images to yield accurate and reproducible patient setup prior to dose delivery. The registration of 2 image datasets can be characterized as hardware-based or software-based image fusion. Hardware-based image fusion is performed by hybrid scanners that combine 2 distinct medical imaging modalities such as positron emission tomography (PET) and computed tomography (CT) into a single device. In hybrid scanners, the patient maintains the same position during both studies making the fusion of image data sets simple. However, it cannot perform temporal image registration where image datasets are acquired at different times. On the other hand, software-based image fusion technique can merge image datasets taken at different times or with different medical imaging modalities. Software-based image fusion can be performed either manually, using landmarks, or automatically. In the automatic image fusion method, the best fit is evaluated using mutual information coefficient. Manual image fusion is typically performed at dose planning and for patient setup prior to dose delivery for IGRT. The fusion of orthogonal live radiographic images taken prior to dose delivery to digitally reconstructed radiographs will be presented. Although manual image fusion has been routinely used, the use of fiducial markers has shortened the fusion time. Automated image fusion should be possible for IGRT because the image datasets are derived basically from the same imaging modality, resulting in further shortening the fusion time. The advantages and limitations of both hardware-based and software-based image fusion methodologies are discussed.

Key Words: **IGRT, Image registration, Image fusion, Mutual information.**

INTRODUCTION

As the practice of medicine evolves, medical imaging modalities have become a vital component of modern medicine for the diagnosis and anatomical localization of diseases. These medical imaging modalities offer a noninvasive mechanism of mapping the internal anatomical structures and providing biological functioning information regarding the patients. Because of the differences in the principles of detection in these imaging modalities; for example, computed tomography (CT) uses linear attenuation and magnetic resonance imaging (MRI) uses magnetic moments of protons, the mapping reveals different aspects of the internal anatomical structures of the patients.[1] On the other hand, positron emission tomography (PET) and single-photon emission tomography (SPECT), which rely on the administration of radiotracers that are metabolically and physiologically distributed, reveal biological functional information within the patients. Clinicians may use 2 or more imaging modalities to extract relevant information to diagnose the symptoms the patient experiences. To arrive at the proper diagnoses, clinicians have developed techniques of synthesizing these images from different imaging modalities in 3 dimensions mentally. With recent advances in computer technology, these mental processes have been supplemented fully or partially with automated hardware and/or software techniques referred to as image fusion or image registration.

Reprint requests to: Cheng B. Saw, Ph.D., Division of Radiation Oncology, Penn State Hershey Cancer Institute, 500 University Drive - H063, Hershey, PA 17033–0850. E-mail: cbsaw2003@yahoo.com

IMAGE FUSION CONCEPT

Image fusion is the process that matches 2 or more image datasets, resulting in a single merged image dataset. The concept of "matches" or "merged" is very specific in reference to patient anatomical alignment. The obvious outcome of image fusion is the creation of a merged image dataset that will enhance the clinical interpretation or diagnosis of symptoms of the patients. For example, if one image dataset from CT (Fig. 1) and another image dataset from PET (Fig. 2) are merged, the fused image dataset will show the geographical flow or biological uptake of radiotracer with respect to the patient anatomy (Fig. 3), giving the clinician a better appreciation of the involvement of the disease in the tissue or organ in the patient. If the image datasets are from the same medical imaging modality but acquired at different times, the fused image dataset give an assessment on the progression of the diseases. For example, the comparison of MRI image dataset taken before and after treatment through image fusion technique allows clinicians to eval-

Fig. 1. A CT transaxial image showing anatomical structure of a patient.

uate the effectiveness of the treatment. If it is desired to irradiate a pre-chemotherapy volume of a tumor that has responded to induction chemotherapy, it will be necessary to fuse images that contained the “old” pathology to the patient’s current anatomy. The fusion of image datasets acquired using different imaging modalities is referred to as multimodal fusion.[2] If the image datasets are

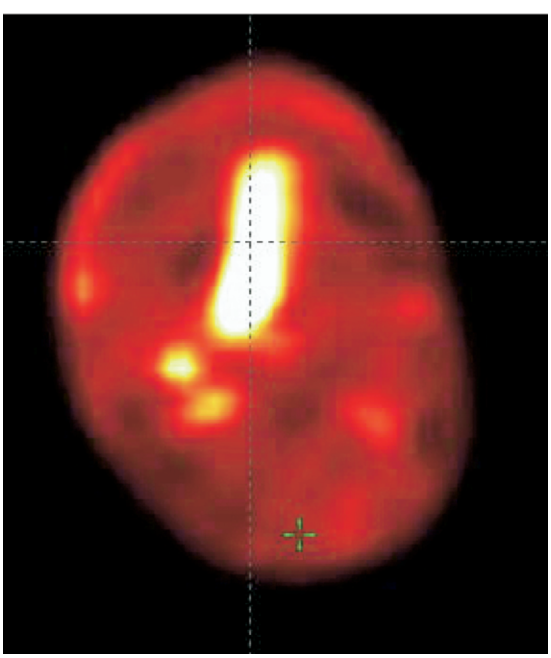

Fig. 2. A PET transaxial image showing radiotracer uptake but lacks anatomical structure of a patient.

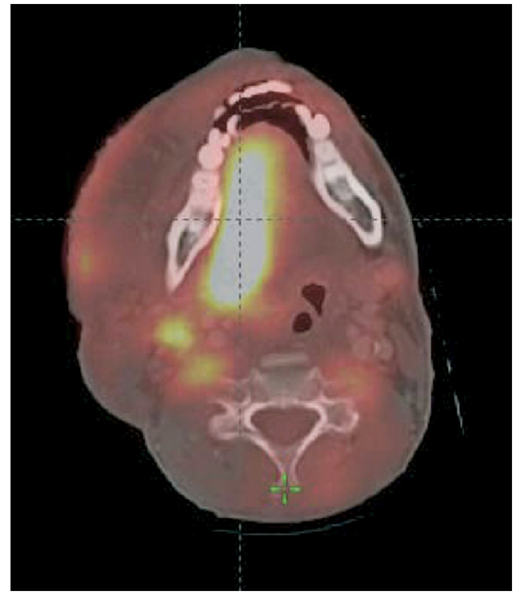

Fig. 3. A fused PET and CT image showing the uptake in relation to the anatomy of the patient.

taken from the same imaging modality, the fusion process is referred to as monomodal fusion. In summary, image fusion is the coregistration of 2 or more image datasets obtained from different imaging modalities or at different times for the purpose of enhancing clinical interpretation.

IMAGE FUSION METHODS

Image registration methodology can be characterized as either hardware-based or software-based image fusion.[3] The hardware-based approach relies on having the patient in the same position for 2 image datasets acquisition using 2 different imaging modalities. Software-based image fusion relies on manipulating digital image datasets using mathematical algorithms to produce the desired results. It is not restricted to any particular imaging modality and can be expanded to other or newer imaging modalities. Unlike hardware-based image fusion, it also allows for temporal registration where image dataset are acquired at different times such as before and after patient treatment. Although the 2 approaches of image fusion are different, the goal is basically the same—to provide an image dataset that yields critical information for diagnosis and anatomical localization of the diseases. These 2 methods can coexist, and using both technologies can result in a greater degree of fusion accuracy and improved utility.

HARDWARE-BASED IMAGE FUSION

In the hardware-based image fusion method, 2 imaging modalities are coupled into a single device. This device is called a combined, or multimodality, or hybrid scanner. At present, the most common hybrid modality is the combination of CT and PET, although a prototype PET-MRI hybrid scanner is being proposed.[4] With the current scanners, studies are not performed simultaneously but sequentially, one study immediately following the next, without repositioning the patient on the scanning couch. For example, in the PET-CT or SPECT-CT scanners, the CT study precedes the PET or SPECT study. The CT study maps the patient anatomy, and is also used to perform attenuation correction for the radiotracer studies. Because the patient is in the same position for both studies, the fusion of the image datasets is relatively simple, based on the external coordinates system (couch coordinates) to the patient. By using an external reference system, the accuracy of the fusion is improved significantly. The precise location of the uptake in relation to the patient anatomy is crucial to superior clinical diagnosis.

The PET-CT scanner is currently the most popular type of the hybrid scanner. The merged image dataset offers an overlay of the metabolic information onto the anatomical structures and hence provides superior clinical interpretation compared to the side-by-side comparison.[5] It also offers excellent co-registration, faster scanning time, and better attenuation correction. Its popularity reflects the emerging role of 18FDG-PET as a favored modality for molecular imaging. Abnormally high metabolic activity areas seen in PET images are often associated with cancerous cells, although this correlation cannot be assumed because infectious and inflammatory conditions can also produce intense 18FDG uptake. When this information is used in combination with CT images that provide structural anatomy that is lacking in PET images, it offers clinicians the chance to interpret the uptake in the context of the patient's biological system.

In addition to the simplicity of image fusion, hybrid scanners such as the PET-CT scanner offer other advantages. Both studies can be performed in a single appointment, and hence minimize patient inconvenience. Hybrid scanners also allow for better patient throughput because they usually require shorter study times. In a PET scanner alone, transmission scan is usually taken for photon attenuation correction. In the hybrid scanner, the CT study is used for the same purpose and hence reduces the total study session time.[6] The third advantage of hybrid scanners is the reduction of floor space because the 2 imaging modalities are combined into 1 device.

SOFTWARE-BASED IMAGE FUSION

Software-based image fusion refers to the various available image-processing methods. Fusion software can vary in the level of sophistication and functionality; therefore, understanding the capabilities of the fusion software package is fundamental to assessing its usefulness. They have developed independently from the hardware-based image fusion technology and are general sold separately from the imaging modalities. The software used to operate the hybrid scanner and to display their output is generally not considered image fusion software. However, the equipment operational software may include rudimentary fusion tools. Image fusion software can co-register image datasets from the same modality acquired at different times or different imaging modalities to produce a fused image dataset. The image fusion software can work despite the fact that the patient's position may be different or the patient's body shape may change, although drastic anatomical differences such as prone vs. supine patient positions, or arms up vs. arms down will result in poor image fusion. The image fusion software can be considered more versatile compared to hardware-based image fusion with flexibility in the choice and combination of imaging modality. In addition, the advanced image fusion software product can compensate for misalignments in the images caused by differences in patient posture or positioning.

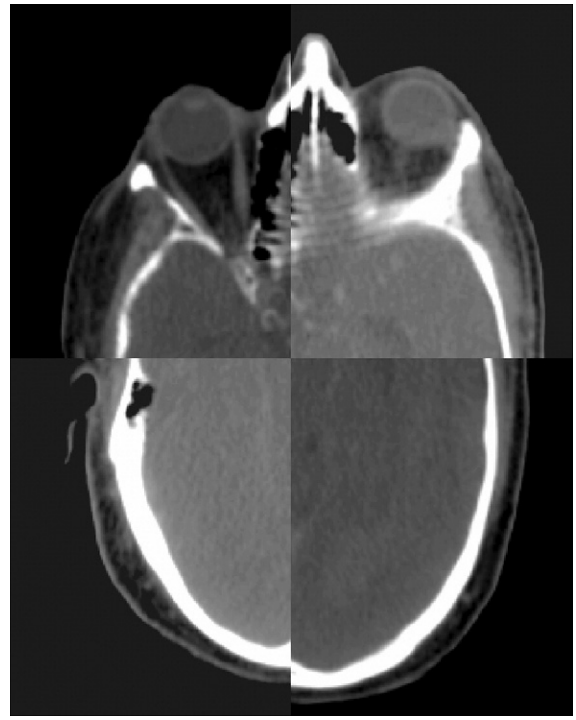

Fig. 4. A merged CT - CT image showing the overlay of one image onto another.

The simplest form of software-based image fusion is the manual image registration method where the operator directly overlays one image onto another. By adjusting the transparency of the image on the upper layer, both images can be viewed simultaneously in the same space. The transparent image can be moved left and right, up and down, and front and back, giving motion in 3 dimensions. Alternatively, the 2 images can be split in quadrants for display and manipulation, as shown in Fig. 4. In addition, the image can be rotated in some image fusion software.

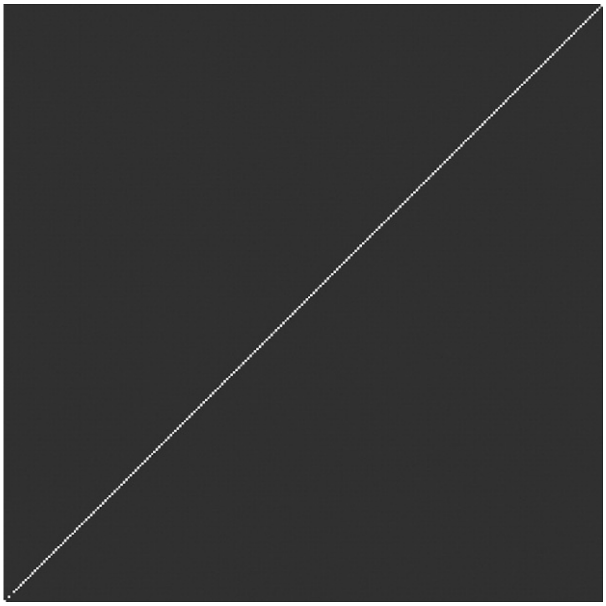

Fig. 5. Mutual information correlation can be visually assessed by plotting pixel intensities of one image vs. another.

The quality of this image fusion is very dependent on the experience of the operator.

Another method of software-based image fusion performs landmark image registration by requiring the operator to identify a few landmarks. Once the landmarks are identified, the software automatically aligns the 2 image datasets based on these landmarks. The quality of the image fusion is therefore dependent on the number and the location of the landmarks chosen. The landmarks must be clearly seen on each image set and can be a part of the patient anatomy such as the spinous processes, tip of the nose, tip of the chin, or dental filling. Landmarks can also be defined outside or embedded inside the patient, using distinct materials called fiducial markers. Gold markers with a diameter of less than 1 mm and 1.0- to 1.5-cm long have been implanted to serve as fiducial markers for the localization of the prostate in image-guided radiation therapy (IGRT). Implanted fiducial markers have been successful in extracranial irradiation, as reported in this special volume.[7,8] However, there is a tendency to move away from the use of fiducial markers because the implantation is an invasive technique, which can introduce complications, and the fiducial markers can migrate after implantation. The quality of image fusion using landmarks is also dependent on the skill of the operator in choosing the appropriate landmarks for image registration.

Fully automated image registration is the most sophisticated software-based image fusion method. The image fusion software uses complex mathematical algorithms and statistical techniques that operate independent of the imaging modalities to align the image datasets. The fusion time varies depending on the efficiency of the mathematical algorithms and the parameters chosen for optimization. Those parameters chosen are generally described in terms of numerical values that are similar in both image datasets referenced as "mutual information."[3,9,10] Although the mathematical algorithms and principles of mutual information are complex, the following gives some basic understandings of mutual information. The brightness or intensity of each pixel in an image is described using numerical values. One of the image datasets can be transformed through translation, rotation, and/or deformation to give a maximum overlap of common regions. The quality of image fusion is measured using the mutual information correlation function. A visual inspection of mutual information correlation function is illustrated in Figs. 5 and 6. It is a plot of intensity of all pixels from one image vs. another. Figure 5 shows a linear graph for an ideal image registration, where every pixel on one image has the corresponding pixel in the other image with the same intensity. Data points that deviate from this line indicate the lack of correlation, as shown in Fig. 6, where the image is shifted diagonally by 1 mm to illustrate misalignment. The correlation or best fit is mathematically defined using a cost function that will be minimized for optimization. It should be emphasized that mutual information does not recognize the anatomy of the patient or the underlying physiological characteristics of the tissues and/or organs being imaged. The image fusion software may also use internal parameters to handle potential local mismatches to increase the reliability of the method. The effectiveness of image fusion is dependent on the type of anatomical structures under study. In addition, the qual-

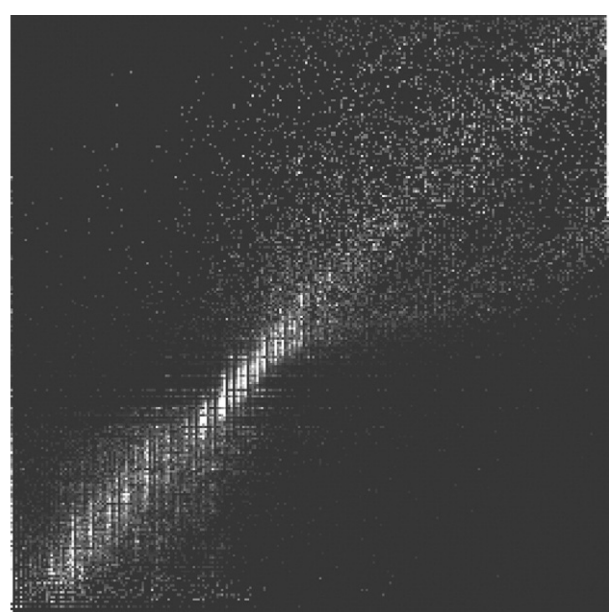

Fig. 6. Misaligned image fusion of 1 mm in the diagonal direction showing data dispersion from the ideal line (see Fig. 5) in the mutual information correlation graph.

ity of images and the type of image modalities also influence the effectiveness of the image fusion.

Automated fusion also allows for both rigid and deformed image registration. Rigid image registration refers to the transformation that is applied to each pixel of the image uniformly. On the other hand, the deformed image registration treats every pixel individually to allow for localized image registration. The recommended process is to initially perform rigid image registration, which is satisfactory for most clinical case. Deformed image registration can then be applied, if needed, to refine the fused image data set. However, care should be taken to ensure that the fused image makes sense because automatic deformable fusion can produce confusing results. In selective cases, landmarks or reference points may be identified to ensure consistency through the deformation processes.

Manual image registration is the most appropriate when (a) the misalignment between the images are extreme, (b) the quality of the images is poor, (c) one of the image dataset lacks easily recognizable features such as in the PET images, or (d) the fusion of a small-image dataset to another with a large-volume image dataset. Automated image registration is best used when the image datasets are not dramatically out of alignment and the anatomical range is similar. For example, an automatic fusion of a CT and MR image dataset of the patient's head would give a high-quality fused image data set because the mapping of the head is basically the same for the same anatomical region. Because of the similarity between the image datasets, the automatic image registration process is fast. On the other hand, automatic image registration will give poor results if the quality of one image dataset is poor compared to the other.

The advantages of software-based image fusion are its flexibility and lower cost. Software-based image fusion can be applied to any image datasets from any imaging modality. Currently, imaging modalities include CT, MR, PET, and SPECT. The image fusion software can also be extended to future applications such as 3D ultrasound and/or triple imaging modalities fusion, while hybrid scanners are limited to fusing the designated image datasets. Software-based image fusion will evolve at a faster rate compared to hardware-based image fusion. Software-based image fusion has been used to compare studies of the same imaging modality over time, thereby enabling correction for slight differences in the patient positioning.

LIMITATIONS OF IMAGE FUSION TECHNIQUES

Although image fusion techniques have been successful and are routinely used in clinics, there are potential limitations to their applications.[3] Hardware-based image fusion techniques assume that the patient anatomy remains relatively static during each study and between studies. This condition is difficult to maintain due to physiological movements such as swallowing, coughing, and breathing. However, breathing can be limited using various techniques such as coaching and/or forced-breathing. Evidence of breathing artifact can be seen by observing the position of the diaphragm on CT images. There are gating techniques referenced as 4D scanning techniques being introduced to address the issue of breathing motion. In addition, there must be standardized scanning protocols with respect to patient positioning. Inconsistent patient scanning protocols, such as hands over the head in one study, and on the side in another study, would result in poor-quality image fusion.

Because of the length of study time (about 1 minute for a CT study and 45 minutes for a PET study), there will be invariable patient movement during the study and between the studies in the hardware-based image fusion technique. In addition, deformation registration is not available to correct for the effect of patient motion. Hardware-based image fusion is limited to those imaging modalities incorporated into the device and cannot be expanded to other modalities. As discussed above, hardware-based image fusion does not allow for temporal registration.

Software-based image fusion also has its limitations in that the quality of the fusion is as good as the experience of the operator and/or the fusion algorithm. In manual image registration, the skill of the operator is critical and automated image fusion generally leads to a better fusion. There is a tendency to use skin marks, which routinely move as landmarks. This would lead to inaccurate image fusion. Automatic image fusion will result in misaligned fused images if (a) the mutual information is inconsistent, (b) poor quality acquired images are involved, and (c) extreme misalignment in the image datasets occurs. The fusion of image datasets from different imaging modalities are generally very challenging when the mutual information is inconsistent (or lacking) due to the use of different detection principles.

Aside from the concerns on the degree of alignment of image datasets, consideration should also be given to the optimal display of relevant clinical information. Typically, the de-facto standard of display of PET has been a colorwash applied to the grayscale CT images, as depicted in Fig. 3. Such display may not be effective when compared to side-by-side image presentations. This is because the overlay can obscure subtle detail, even though it produces dramatic-looking highly contrasted fused images. The degree of brightness displayed in the PET images represents the state of the metabolic activity of tissues at a particular location. This brightness has been misinterpreted as the size of the lesion.

IMAGE FUSION IN TREATMENT PLANNING

The paradigm of radiation therapy has drastically changed since the introduction of three-dimensional (3D) treatment planning systems in the 1980s. Patient data for

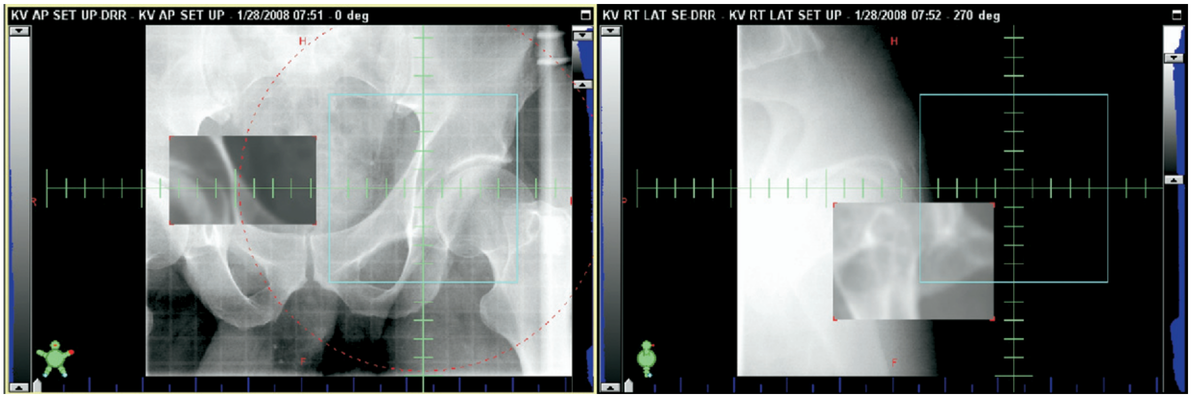

Fig. 7. Fusion of orthogonal live radiographs and DRRs for performing patient setup.

3D treatment planning are now acquired using primarily CT imaging modality. Hence, CT-simulator scanners are an integral part of the modern radiation oncology department. After a patient has been scanned, the CT images are pushed and downloaded into a 3D treatment planning system for individualized treatment planning. During the treatment planning process, the external contour of the patient, the critical structures, and the volume of known suspected disease must be delineated. While CT images provide anatomical structure in details, at times the diseased regions may not be evident. Other imaging modalities have been explored with interests in MRI and PET. Recently, PET and PET-CT have shown increased use, in particular, in the detection of miniature tumors and tumor extension, and hence have been advocated for the management of cancer patients.[5] Because of these sensitivities, PET and PET-CT image datasets have been used to delineate diseased regions and fused to CT image datasets to realize the uptake in relation to the patient's anatomy. The relationship of the diseased region with respect to patient anatomy in the treatment position is especially important to deliver the prescribed dose to the target while minimizing dose to normal structures as the patient undergo radiation therapy. With the availability of 3D treatment planning systems, it is possible to perform conformal radiation therapy (CRT) and intensity modulated radiation therapy (IMRT) dose delivery techniques.[11–14] These dose delivery techniques require a higher degree of accuracy and precision and hence the deployment of IGRT for treatment verification.[14–19] After the individualized treatment planning is completed, the image datasets are reconstructed into digital reconstructed radiographs (DRR) and exported to the dose delivery system, which is typically the medical linear accelerator to perform IGRT.

IMAGE-GUIDED RADIATION THERAPY

Besides delineating the diseased region in individualized treatment planning, imaging modalities are being used to aid in patient setup prior to dose delivery with the interest of reducing setup uncertainties. This approach, referred to as IGRT, requires the integration of one or more imaging modalities into the radiation dose delivery system. A number of positioning and tracking devices have been introduced among which are the radiographic imaging of fiducial markers, ultrasound imaging of the patient anatomy, detection of radiofrequency from beacons, video-based surface tracking, in-room CT imaging, megavoltage (MV) cone-beam CT imaging, and kilovoltage (kV) cone-beam CT imaging.

At Penn State Cancer Institute, an on-board kilovoltage imager is used to perform IGRT. The on-board imager is incorporated in a Varian linear accelerator model Trilogy having (a) 2 photon beams with potentials of 6 MV and 10 MV, (b) 6 electron energies from 4 to 20 MeV, (c) an additional 6-MV photon beam with specially designed flattening filter and high dose rate at 1000 MU per min to perform stereotactic radiosurgery, and (d) Millennium MLC-120 multileaf collimation system. The on-board imager consists of a kilovoltage (kV) x-ray source and a large-area flat panel amorphous silicon (aSi) detector that are mounted onto the gantry of the linear accelerator in orthogonal direction to the megavoltage beam and electronic portal imaging device (EPID). The physical attributes of this equipment have been described elsewhere.[17–19] The IGRT capabilities of this linear accelerator include live radiograph, kilovoltage cone-beam computed tomography (kV-CBCT), and fluoroscopic acquisitions. Prior to dose delivery, live radiographs are typically acquired in orthogonal directions and are fused to the respective DRRs to determine the difference in the patient setup. The DRRs provide information on the initial patient setup at the time of CT acquisition for treatment planning. The comparison can be made manually by fusing the images or performing automatic image fusion as shown in Fig. 7. The difference is converted into numerical values to adjust or shift the treatment couch and hence the patient position. If soft-tissue alignment is required, the kV-CBCT technique can be used.[18] In the kV-CBCT technique, the data are acquired volumetrically and reconstructed to obtain images in the 3 axial planes for comparison to CT images acquired for treatment planning. The kV-CBCT images has the possibility to be exported to the treatment planning system for dose calculations to obtain the overall effective dose distribution over a number of fractions allowing for adjustment of total prescribed dose before completing the course of treatment.

REFERENCES

1. Bushberg, J.T.; Seibert, J.A.; Leidholdt, E.M., Jr.; *et al.* The Essential Physics of Medical Imaging. Baltimore, MD: Williams & Wilkins; 1994.
2. Maisely, M.; Leong, J. Introduction to image fusion. In: Maisely, M.; Leong, J., editors. Medical Image Fusion.Albuquerque, NM: American Society of Registered Technologist; 2003:1–12.
3. Maisely, M.; Leong, J. Comparing hardware-based and software-based medical image fusion. In: Maisely, M.; Leong, J., editors. Medical Image Fusion. Albuquerque, NM: American Society of Registered Technologist; 2003:17–23.
4. Zaidi, H.; Mawlawi, O.; Orton, C.G. Simultaneous PET/MR will replace PET/CT as the molecular multimodality imaging platform of choice. *Med. Phys.* **34:**1525–8; 2007.
5. Heron, D.E.; Smith, R.P.; Andrade, R.S. Advances in image-guided radiation therapy – the role of PET-CT. *Med. Dosim.* **31:**3–11; 2006.
6. Burger, C.; Goerres, G.; Schoenes, S.; *et al.* PET attenuation coefficients from CT images: Experimental evaluation of the transformation of CT into PET 511-keV attenuation coefficients. *Eur. J. Nucl. Med.* **29:**922–7; 2002.
7. Gerszten, P.C.; Burton, S.A. Clinical assessment of stereotactic IGRT: Sinal radiosurgery. *Med. Dosim.* **33:**107–116;2008.
8. Saw, C.B.; Chen, H.; Wagner, Jr., H. Implementation of fiducial-based image registration in the Cyberknife Robotic system. *Med. Dosim.* **33:**155–159; 2008.
9. Maurer, C.R., Jr.; West, J.B. Medical image registration using mutual information. In: Cyberknife Radiosurgery Practical Guide 2.Heilbrun, M.P., editor. Sunnyvale, CA: The Cyberknife Society; 2006:23–34.
10. Hill, D.L.G.; Batchelor, P.G.; Holden, M.; *et al.* Medical image registration. *Phys. Med. Biol.* **46:**R1–45; 2001.
11. Purdy, J.A. Intensity-modulated radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **35:**845–6; 1996.
12. Saw, C.B.; Ayyangar, K.M.; Enke, C.A. MLC-Based IMRT – Part II. *Med. Dosim.* **26:**111–2; 2001.
13. Ezzell, G.A.; Galvin, J.M.; Low, D.; *et al.* Guidance document on delivery, treatment planning, and clinical implementation of IMRT: Report of the IMRT subcommittee of the AAPM Radiation Therapy Committee. *Med. Phys.* **30:**2089–115; 2003.
14. Saw, C.B.; Heron, D.E.; Huq, M.S.; *et al.* Target delineation and localization (IGRT) – Part I. *Med. Dosim.* **31:**1–2; 2006.
15. Saw, C.B.; Heron, D.E.; Yue, N.J.; *et al.* Cone-beam imaging and respiration motion (IGRT) – Part II. *Med. Dosim.* **31:**89–90; 2006.
16. Saw, C.B.; Heron, D.E.; Huq, M.S. Stereotactic body radiation therapy (IGRT) – Part III. *Med. Dosim.* **32:**69–70; 2006.
17. Pawlicki, T.; Kim, G.Y.; Hsu, A.; *et al.* Investigation of linac-based image-guided hypofractionated prostate radiotherapy. *Med. Dosim.* **32:**71–9; 2006.
18. Saw, C.B.; Yang, Y.; Li, F.; *et al.* Performance characteristics and quality assurance aspects of kilovoltage cone-beam CT on medical linear accelerator. *Med. Dosim.* **32:**80–5; 2007.
19. Huntzinger, C.; Friedman, W.; Bova, F.; *et al.* Trilogy image-guided stereotactic radiosurgery. *Med. Dosim.* **32:**121–33; 2007.