

2D/3D Registration Based on Volume Gradients

Wolfgang Wein¹, Barbara Röper² and Nassir Navab¹

¹Computer Aided Medical Procedures (CAMP) Group, TU Munich
Boltzmannstr. 3, 85748 Garching, Germany

²Clinic and Policlinic of Radiation Oncology,
Klinikum rechts der Isar, TU Munich,

Ismaninger Str. 22, 81675 Munich, Germany

Email: {wein,navab}@cs.tum.edu, barbara.roeper@lrz.tum.de

January 9, 2005

ABSTRACT

We present a set of new methods for efficient and precise registration of any X-Ray modality (fluoroscopy, portal imaging or regular X-Ray imaging) to a CT data set. These methods require neither feature extraction nor 2D or 3D segmentation. Our main contribution is to directly perform the computations on the gradient vector volume of the CT data, which has several advantages. It can increase the precision of the registration as it assesses mainly the alignment of intensity edges in both CT and X-Ray images. By using only significant areas of the gradient vector volume, the amount of information needed in each registration step can be reduced up to a factor of 10. This both speeds up the registration process and allows for using the CT data with full precision, e.g. 512^3 voxels. We introduce a *Volume Gradient Rendering (VGR)* as well as a *Volume Gradient Correlation (VGC)* method, where the latter one can be used directly for computing the image similarity without DRR generation.

1. INTRODUCTION

2D-3D Registration has numerous applications in computer-aided diagnosis and therapy, including intraoperative navigation with fluoroscopy, patient positioning for radiation therapy and multimodal data fusion for diagnosis and therapy planning. These applications usually involve a two-dimensional X-Ray projection image and a preoperative CT data set. The task is to define a common coordinate frame, in a way that corresponding structures of the patient's anatomy are properly aligned in both data sets. The general workflow of a 2D-3D registration algorithm is to simulate X-Ray images by computing two-dimensional projections of the CT volume at an estimate of the pose of the real X-Ray image. This digitally reconstructed radiograph (DRR) is iteratively compared to the real X-Ray image and the DRR pose is altered. Finding a rigid pose which maximizes the similarity of two images is a non-linear optimization problem, where adequate algorithms can be used in order to find the maximum as fast as possible.

2. RELATED WORK

For performing an intensity-based 2D/3D registration, numerous DRRs have to be created, each of them requiring processing of the full volumetric data set. Plenty of research has been done on speeding up DRR computation. Contrary to the classical back-projection approach, *Ray Casting*, algorithms like *Shear-Warping*,^{1,2} precomputation of DRR rays,³ light field rendering⁴ and hardware-accelerated 2D Texture rendering³ are more efficient in some orders of magnitude, however they all use simplifications which reduce the DRR quality and produce artifacts which may affect the registration accuracy. Using the 3D texturing feature of modern graphics hardware, combined with rendering to floating-point color buffers, it is possible to accumulate slices trilinearly interpolated from a 3D volume. Based on this feature, the 3D texture volume rendering algorithms are today the methods of choice, because they are efficient and produce high-quality DRRs without artifacts. However, their implementation imposes a lot of technical difficulties, and designated software usually works only for a specific graphics board manufacturer.

The DRR and X-Ray images to be compared can be considered as two different modalities, mainly because of different energy spectrums used for the acquisition of CT slices and X-Ray images. This difference can be enormous if Portal Images with MegaVolt energies are acquired.⁵ By applying an adequate transfer function

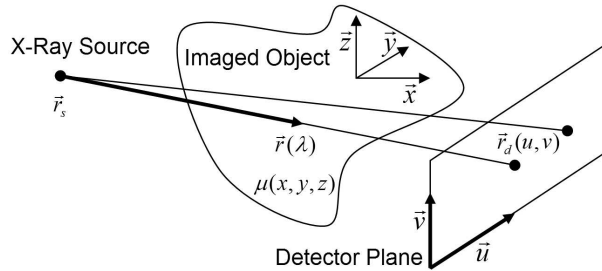


Figure 1. Scheme of X-Ray projection imaging

on the volume this can be taken care of partly. Nevertheless the images will have different structures, and it is tough to assess the quality of alignment automatically. In this context many *Similarity Measures* have been developed. Some are using just sum-of-squares or correlation of the intensity differences, some are working on two-dimensional gradient images. Information theory approaches assess the alignment based on the properties of histograms and joint histograms, as the very popular *Mutual Information* measure,⁶ or *Correlation Ratio*.⁷ Correlation of small image regions weighted with local variances is another very stable measure.³ Various researchers compared the performance of 2D/3D registration using different similarity measures^{8,9} and.¹⁰

Our approach yields efficient use of volume gradients, in order to obtain either the two-dimensional DRR image gradients or directly the gradient based similarity measures. The underlying idea of using volume gradient vectors has been described in two other recent papers. Tomazevic et al.¹¹ backproject the gradients from the X-Ray image into the volume and compute a similarity based on their correspondence with 3D gradient vectors, which are defined on the surface of (previously segmented) bony structures. Livyatan et al.¹² use, as final stage in their registration algorithm, the correspondence of the 2D gradients on the X-Ray image with ray-casted sums of 3D gradient vectors in the volume. These correspondences are only evaluated for border areas in the 2D image, which they define by applying a Canny edge detector algorithm on the X-Ray image. In contrast to those algorithms, we are using the CT gradient information in order to perform a purely intensity-based registration, without prior feature extraction.

3. METHODS

3.1. Foundation for DRR Computation

The pixel values in a X-Ray image originate from the number of particles reaching the detector plane after passing the imaged object. The basic equation is

$$I(u, v) = \int_0^{E_{max}} I_0(E) \exp\left(-\int_{r(u,v)} \mu(x, y, z, E) dr\right) dE \quad (1)$$

where u, v defines a pixel of the X-Ray detector, $r(u, v)$ is the ray from the source to this pixel, $\mu(x, y, z, E)$ is the attenuation coefficient at a specific position in space and a X-Ray energy E (figure 1). As X-Ray sources are always polychromatic, the attenuation has to be integrated not only along the ray, but also over the incident energy spectrum $I_0(E)$. Those spectrums unfortunately vary greatly depending on the imaging devices. Therefore attenuation coefficients are being treated in a simplified manner, denoting the attenuation of a monochromatic X-Ray source at an effective energy E_0 :

$$I(u, v) = I_0 \exp\left(-\int_{r(u,v)} \mu(x, y, z, E_{eff}) dr\right) \quad (2)$$

Most X-Ray imaging devices measure the logarithm of equation 2, thus the pixel intensities are reduced to a simple integral term, which we will use for our computations.

A heuristic means to compensate for the different effective energies in CT and X-Ray images is to apply a scaling

and truncation of the CT intensities. The CT values are cropped with a user-defined window/level setting, the line integrals are evaluated (e.g. with a ray-casting algorithm) and the resulting pixels are scaled in order to fully cover the gray-scale range of the X-Ray image. This process can be referred to as *radiometric calibration*,⁵ it has to be done only once for a specific combination of CT and X-Ray imagers. This even allows to compute highly realistic DRRs of MegaVolt Portal Images from CT.

3.2. Determining Quality of Alignment

In order to search for the best DRR pose, the DRR and X-Ray images have to be compared iteratively and the quality of alignment has to be assessed by some measure of similarity. *Gradient Correlation*⁸ is a very powerful one, its robustness originating from the use of gradient images. It uses *Normalized Cross Correlation*, which expresses the linear dependency between the intensities in the images I_1 and I_2 :

$$\begin{aligned} NCC(I_1, I_2) &= \frac{1}{\sigma_1 \sigma_2} \frac{1}{n} \sum_{u,v} (I_1(u, v) - \bar{I}_1) \cdot (I_2(u, v) - \bar{I}_2) = \\ &= \frac{\sum (I_1(u, v) - \bar{I}_1) \cdot (I_2(u, v) - \bar{I}_2)}{\sqrt{\sum (I_1(u, v) - \bar{I}_1)^2} \cdot \sqrt{\sum (I_2(u, v) - \bar{I}_2)^2}} \end{aligned} \quad (3)$$

Gradient Correlation is then the mean of the NCC values for both pairs of the horizontal and vertical gradient images, respectively:

$$GC = \frac{1}{2} \left(NCC \left(\frac{\partial I_1}{\partial u}, \frac{\partial I_2}{\partial u} \right) + NCC \left(\frac{\partial I_1}{\partial v}, \frac{\partial I_2}{\partial v} \right) \right) \quad (4)$$

It reaches its maximum value 1 if both the horizontal and vertical components of the DRR and X-Ray gradients are fully linearly dependent. This assumes a looser dependency on the image content itself,¹⁰ which is especially beneficial if the data to be registered has been taken at very different energies, as the is case with Portal Images / CT.

3.3. Direct Computation of DRR Gradient

A DRR pixel is the attenuation along the ray originating in the X-Ray source \vec{r}_s and passing through the location of the pixel in the image plane $\vec{r}_d(u, v)$. Each position on this ray can be parameterized:

$$\vec{r}(u, v, \lambda) = \vec{r}_s + \lambda (\vec{r}_d(u, v) - \vec{r}_s), \quad \lambda \in [0 \dots 1] \quad (5)$$

The refined equation for the resulting intensity value at the pixel u, v (corresponding to the logarithm of equation 2) is

$$I(u, v) = \int \mu(\vec{r}(u, v, \lambda)^T) d\lambda \quad (6)$$

For computing the similarity measure, equation 4, we need the partial derivatives of $I(u, v)$ with respect to the pixel locations u and v , i.e. $\frac{\partial I}{\partial u}(u, v)$ and $\frac{\partial I}{\partial v}(u, v)$. Given the three-dimensional gradient $\nabla \mu(x, y, z)$ of the attenuation in space, we can directly compute this information:

$$\frac{\partial I}{\partial u}(u, v) = \left(\int \lambda \nabla \mu(\vec{r}(u, v, \lambda)^T) d\lambda \right)^T \cdot \vec{u} \quad (7)$$

$$\frac{\partial I}{\partial v}(u, v) = \left(\int \lambda \nabla \mu(\vec{r}(u, v, \lambda)^T) d\lambda \right)^T \cdot \vec{v} \quad (8)$$

The volumetric gradient $\nabla \mu(x, y, z)$ has to be incorporated into the line integral by scaling it with λ . This is due to the fact that the rays belonging to neighbored pixels are further apart from each other proportional to the distance from the X-Ray source (see figure 1).

The length of the gradient vectors is to a big extent close to zero. It is therefore applicable to consider only gradient vectors of significant size, e.g. 1% of the maximum size. Neglecting all smaller vectors has the side effect of improving the registration quality further, as the important edges from rigid, bony anatomy are emphasized, while too small ones, likely to belong to deformable tissue, are discarded.

3.3.1. Gradient Ray Casting

Equations 7 and 8 can be directly evaluated using a ray-casting technique. For each pixel u, v in the DRR gradient image, the line equation 5 is sampled at constant intervals, while $\nabla\mu(x, y, z)$ is computed at the respective positions by trilinear interpolation from the gradient volume. For comparison purposes, we also implemented equation 6 in the same way, which produces regular DRRs. It is important to mention that both approaches take the same computation time, if we make use of Single Instruction Multiple Data (SIMD) machine commands, which are available on today's ix86 CPUs. Hence summation and interpolation of a 3-vector takes the same number of operations than using just scalar data. We used the GNU C++ Compiler under Linux, where Intel's SSE2 instructions are supported.

For skipping the huge number of unimportant small gradient vectors, we implemented a binary Octree data structure. For every sampling step along a ray, the octree is queried. The largest bounding cube that is empty (side width $32, 16, \dots, 1$ voxels) is skipped, and the first sampling position on the ray is set, which is outside of this cube. In order to efficiently implement this skipping, we use an algorithm similar to the one originally proposed by Amanatides and Woo¹³ for ray tracing purposes. We refer to our rendering method as *Volume Gradient Rendering (VGR)*.

3.3.2. Gradient Splatting

Splatting is a forward-projection method, where the volume is traversed voxel-wise, and a footprint of each projected voxel is created in the 2D image. Thus a great reduction in computation time can be achieved by just considering the gradient voxels larger than a specific magnitude. The footprint is usually a gaussian kernel, its size varying with the distance of the voxels from the image plane. For each voxel, the pixel intensities of all pixels affected by the kernel have to be updated. Unless hardware-accelerated methods are used (e.g. drawing the kernels as OpenGL textures), this method is too slow in order to be superior to ray-casting, even if only a small percentage of the gradient voxels is considered.

Therefore, if the objective is DRR generation, splatting would not be the preferred approach. However, a smart algorithm could take advantage of the splatting formulation to directly use the 3D gradient data for the similarity computation, without DRR generation.

We can leave the use of kernels out, given that the size of a projected voxel is negligible with respect to the pixel size in the 2D image. Each gradient voxel then adds to the intensity of just one pixel:

$$\frac{\partial I}{\partial u}(u, v) = \left(\sum_i \lambda_i \nabla \mu_i \right)^T \cdot \vec{u} \quad (9)$$

$$\frac{\partial I}{\partial v}(u, v) = \left(\sum_i \lambda_i \nabla \mu_i \right)^T \cdot \vec{v} \quad (10)$$

Let $I_1(u, v) = \frac{\partial I}{\partial u}(u, v)$ be the resulting vertical gradient values, and $I_2(u, v)$ the respective values in the X-Ray image. The Normalized Cross Correlation equation (4) can then be rewritten:

$$\frac{1}{\sigma_1 \sigma_2} \frac{1}{n} \left(\sum I_1(u, v) I_2(u, v) - \overline{I_1 I_2} \right) = \quad (11)$$

$$\frac{1}{\sigma_1 \sigma_2} \frac{1}{n} \left(\sum \sum [((\lambda_i \nabla \mu_i)^T \cdot \vec{u}) I_2(u, v)] - \overline{I_1 I_2} \right) \quad (12)$$

The expression in squared brackets can be directly evaluated for each voxel, independently of the sum of projected voxels resulting in pixel values. We therefore compute this correlation by multiplying each projected gradient voxel with the intensity $I_2(u, v)$ of the X-Ray image in full resolution, i.e. 1024^2 pixel. This approach directly creates the correlation values, without producing a high-quality DRR. As we do not assemble ray integrals, we do not have access to the mean and standard deviation of the simulated image, though.

This results in a very efficient computation of gradient correlation values without normalization, which uses both all significant gradient voxels directly, without interpolation and kernel functions, and the full content of the X-Ray image. We denote this technique *Volume Gradient Correlation (VGC)*.

3.4. Implementation

A prototypical application was developed, which demonstrates the advantages of the described methods. A CT volume is loaded with full precision (12 bit) into memory. The volume intensities are then cropped with a defined window/level setting. Successively the gradient volume $\nabla\mu(x, y, z)$ is computed using a three-dimensional Sobel filter cube, which takes 27 values into account for each voxel, resulting in very smooth gradient information. For the VGR method, a binary octree is constructed from the gradient volume. For the VGC method, the gradient voxels above the threshold are saved in a run-length encoded structure for fast traversal.

The actual registration is performed by iteratively computing the DRR/X-Ray similarity with one of the described methods. The pose is refined according to a simple best neighbor search strategy, which has proven to be stable and sufficiently fast.¹⁰ At the same time it allows to distribute the computations on multiple CPUs, as (for a 6 DOF optimization problem) 12 cost function evaluations can be done in parallel, before the pose for the next iteration is defined. We implemented a distributed version of the described registration methods using the *Message Passing Interface* (MPI), allowing for transparent distribution on any number of nodes and/or CPUs. Several client processes receive job messages with a pose description, and return a single scalar similarity measure value to the coordinating process. Therefore the communication overhead is very small and the computation speed scales directly with the number of CPUs used.

4. EXPERIMENTAL RESULTS

The subsequent registrations were computed on a cluster comprising 3 nodes with each 4 Intel Itanium-2 CPUs running at 1.3 GHz*. Accordingly, 12 processors were used simultaneously, resulting in a speedup of 9.7 compared to the original sequential implementation, executed on a Pentium M 1.6 GHz Notebook.

4.1. Phantom Data

For assessing the registration robustness and accuracy, we used CT and portal image data acquired from a Rando Body Phantom, including Ground Truth information.^{5,10} The registration was launched iteratively from a pose randomly displaced (up to 20 mm / 20° in each degree of freedom) from the ground truth pose. The following table denotes the mean displacement with the respective standard deviation of the translational and rotational components (in mm and degrees, respectively) of the resulting pose. As we primarily assessed the registration robustness with respect to various pose parameters, we preferred this error description over a target registration error (TRE) value, especially as we were dealing with phantom data. We compared our novel methods to a standard registration using the 2D gradients of a DRR rendered with a regular volume rendering technique (VRT), i.e. software based Ray-Casting, which yields very high accuracy whilst performing relatively slow.

method	Δ_{trans}	Δ_{rot}	σ_{trans}	σ_{rot}	time
VRT	1.31	0.41	0.08	0.09	32 s
VGR	0.79	0.64	0.12	0.28	3.2 s
VGC	1.24	0.48	0.10	0.14	2.5 s

4.2. Patient Data

A second set of experiments was performed on real pelvis data of two prostate cancer patients, comprising a pretherapeutic CT scan and scanned simulator X-Ray images for radiotherapeutic purposes. Before the gradient computation the CT data was thresholded in order to restrain the registration to the pelvis bone, successively a gradient threshold was applied. This resulted in just 1% of the voxels being used for the similarity computation, using our new methods. Ground Truth information was not available, thus we depict the robustness in terms of the standard deviation of the resulting pose parameters, registered from randomly displaced starting poses, as above. The alignment was visually assessed by a physician and declared as optimal for all three methods (figure 2).

*InfiniBand Cluster at TU Munich, <http://infiniband.in.tum.de>

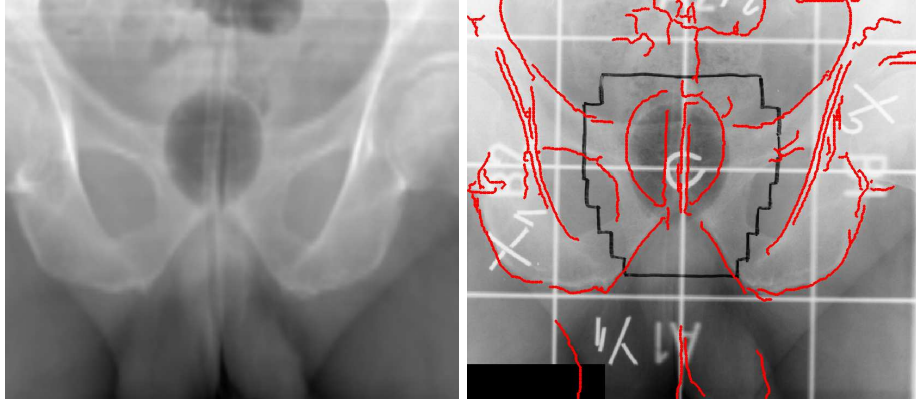


Figure 2. DRR image registered with VGR, X-Ray image with edges from DRR

method	σ_x	σ_y	σ_z	σ_{alpha}	σ_{beta}	σ_{gamma}	time
VRT	0.07	0.44	0.05	0.06	0.05	0.08	22.2 s
VGR	0.18	2.73	0.18	0.23	0.16	0.15	1.8 s
VGC	0.15	4.12	0.41	0.47	0.10	0.21	3.1 s

While the standard deviations (in mm and $^\circ$, resp.) are exceedingly small with regular volume rendering (VRT), the values for our new methods are still well within the range to be considered robust, apart from σ_y . The y values reflect the out-of-plane translation, and therefore are higher due to the small field of view (9°) of the images. This can be compensated by concurrently registering a second, perpendicular X-Ray image. The inferior results of the VGC method probably originates from the missing normalization of the correlation, in conjunction with the relatively small region of interest of these data sets (due to the positioning grid in the X-Ray images). Even though the pelvis was registered correctly, the highly deformable balloon inserted manually into the patients rectum was displaced as expected (figure 2). Therefore both the thresholding of intensity / gradient voxels and the use of a gradient-based similarity measure was essential for the successful registration, as the balloon and tissue had to be excluded for registration.

5. CONCLUSION

We presented a set of new techniques for performing 2D/3D registration of X-Ray modalities, which are based on the immediate use of gradient information. The volumetric data can be used with both full precision (16 bit intensities) and full size for highly precise registration. By thresholding the gradient magnitudes, the amount of information used for the registration can be reduced dramatically, delineating only significant (bony or contrasted) structures. These are in fact well adapted to the use of a rigid registration algorithm, as other parts of the patient’s anatomy often undergo deformable transformations between the image acquisitions (which is the case in experiment 4.2). Note moreover that no segmentation is necessary to achieve this. The VGC method computes the gradient correlation values directly from the thresholded gradient voxels and the full-resolution X-Ray image, therefore using all of the available, pertinent image information.

We tested the proposed methods on both phantom (CT / Portal) and patient (CT / X-Ray) data. In terms of robustness, they are slightly inferior to an algorithm comprising DRR generation with highest possible quality (which can be considered the “Gold Standard” of intensity-based registration lacking Ground Truth information), while they are an order of magnitude faster in execution. Using a purely CPU-based registration solution, we were also able to exploit the speed advantage of distributed computations, which resulted in overall registration times of a few seconds on a system with 12 processors. This leads toward new intra-operative and real-time applications of automatic image registration. In addition, it motivates further research on parallel software-based registration solutions, as opposed to GPU-based techniques, which are difficult to distribute.

6. ACKNOWLEDGEMENTS

This research was funded by an academic grant from Siemens Corporate Research, Princeton, USA. The authors would like to thank in particular Frank Sauer and Ali Khamene for their continuous support.

REFERENCES

1. J. Weese, R. Goecke, G. P. Penney, P. Desmedt, T. M. Buzug, and H. Schumann, "Fast voxel-based 2D/3D registration algorithm using a volume rendering method based on the shear-warp factorization," 1999.
2. P. Lacroute and M. Levoy, "Fast volume rendering using a shear-warp factorization of the viewing transformation," *Computer Graphics* **28**(Annual Conference Series), pp. 451–458, 1994.
3. D. A. LaRose, *Iterative X-Ray/CT registration using accelerated volume rendering*. PhD thesis, Carnegie Mellon University, May 2001.
4. D. B. Russakoff, T. Rohlfing, D. Rueckert, R. Shahidi, D. Kim, and C. R. M. Jr., "Fast calculation of digitally reconstructed radiographs using light fields," *Proceedings of SPIE* **5032**, 2003.
5. A. Khamene, P. Bloch, W. Wein, F. Sauer, and M. Svatos, "Automatic portal image based patient positioning for radiation therapy," *Journal of Medical Image Analysis*, 2004.
6. P. A. Viola, *Alignment by Maximization of Mutual Information*. PhD thesis, Massachusetts Institute of Technology, June 1995.
7. A. Roche, G. Malandain, X. Pennec, and N. Ayache, "Multimodal image registration by maximization of the correlation ratio," *INRIA Technical Report*, 1998.
8. G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. Hill, and D. J. Hawkes, "A comparison of similarity measures for use in 2D-3D medical image registration," *Lecture Notes in Computer Science* **1496**, 2001.
9. J. H. Hipwell, G. P. Penney, T. C. Cox, J. V. Byrne, and D. J. Hawkes, "2D-3D intensity based registration of DSA and MRA - a comparison of similarity measures," *Lecture Notes in Computer Science* **2489**, pp. 501–508, 2002.
10. W. Wein, "Intensity based rigid 2D-3D registration algorithms for radiation therapy," Master's thesis, Technische Universität München, Dec. 2003.
11. D. Tomazevic, B. Likar, T. Slivnik, and F. Pernus, "3-D/2-D registration of ct and mr to x-ray images," *IEEE Transactions on Medical Imaging*, Nov 2003.
12. H. Livyatan, Z. Yaniv, and L. Joskowicz, "Gradient-based 2-D/3-D rigid registration of fluoroscopic x-ray to ct," *IEEE Transactions on Medical Imaging*, Nov 2003.
13. J. Amanatides and A. Woo, "A fast voxel traversal algorithm for ray tracing," in *Eurographics '87*, pp. 3–10, Elsevier Science Publishers, Amsterdam, North-Holland, 1987.