

Progressive Attenuation Fields: Fast 2D-3D Image Registration Without Precomputation

Torsten Rohlfing¹, Daniel B. Russakoff^{1,2}, Joachim Denzler³, and Calvin R. Maurer, Jr.¹

¹ Image Guidance Laboratories, Stanford University, Stanford CA, USA, rohlfiing@stanford.edu, calvin.maurer@igl.stanford.edu

² Computer Science Department, Stanford University, Stanford CA, USA dbrussak@stanford.edu

³ Fakultät für Mathematik und Informatik, Universität Passau, Passau, Germany denzler@fmi.uni-passau.de

Abstract. This paper introduces the progressive attenuation field (PAF), a method to speed up computation of digitally reconstructed radiograph (DRR) images during intensity-based 2D-3D registration. Unlike traditional attenuation fields, a PAF is built on the fly as the registration proceeds. It does not require any precomputation time, nor does it make any prior assumptions of the patient pose that would limit the permissible range of patient motion. We use a cylindrical attenuation field parameterization, which is better suited for medical 2D-3D registration than the usual two-plane parameterization. The computed attenuation values are stored in a hash table for time-efficient storage and access. Using a clinical gold-standard spine image dataset, we demonstrate a speedup of 2D-3D image registration by a factor of four over ray-casting DRR with no decrease of registration accuracy or robustness.

1 Introduction

The rate-limiting step in intensity-based 2D-3D image registration is the computation of digitally reconstructed radiograph (DRR) images, which are synthetic x-ray images generated from CT images. A common method to speed up the process is to precompute projection values, which can then be looked up from a table during the registration phase. For storing the projection values, LaRose [1] used a data structure called “Transgraph” based on the “Lumigraph” introduced by Gortler *et al.* [2] for rendering. Similarly, Russakoff *et al.* [3,4] used an “attenuation field” (AF), which is based on the principle of “light field” rendering introduced by Levoy & Hanrahan [5] simultaneously with Gortler *et al.* [2].

Two major issues arise in the use of precomputation methods: they require time for precomputation, and, for memory efficiency, assumptions have to be made regarding expected viewing angles. Substantially more projections are precomputed than are actually used. It is our experience that only about 10 percent of all rays in an AF are typically used, and the rays which are used are, on average, accessed about 1,000 times (unpublished results). In the present paper we introduce the concept of a progressive AF (PAF), which is not precomputed but instead constructed on the fly as the registration proceeds.

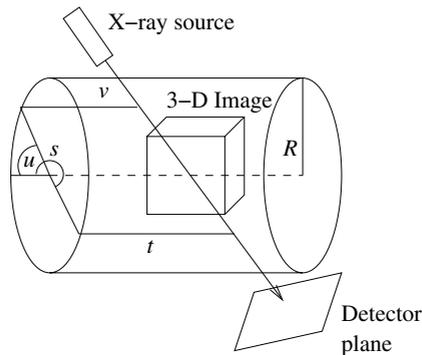


Fig. 1. Cylindrical attenuation field parameterization. Any valid ray through the image volume is parameterized by four parameters. Parameters u and s are rotation angles around the axis of the coordinate system (*dashed line*). Parameters v and t are distances along the axis with respect to an arbitrary reference level.

A PAF effectively acts as a cache for projection values once they are computed, rather than as a lookup table for precomputed projections like standard AFs.

Our novel method eliminates the drawbacks of traditional precomputation methods mentioned above. These benefits come at the cost of higher computational cost during the actual registration compared to precomputed attenuation fields. However, we demonstrate in this paper that there is still a substantial performance increase over un-accelerated computation of DRRs. Also, the speedup provided by our method increases in subsequent serial registrations, which is important for patient tracking applications.

2 Methods

In an AF, each ray through the 3D image is parameterized by a tuple of four real numbers, (u, v, s, t) . In traditional light fields in computer graphics, these numbers are most commonly defined as coordinates in two parallel planes where the ray intersects the respective plane [5]. The PAF, like a light field, samples the space of parameter 4-tuples at discrete intervals. Let $\delta_u, \delta_v, \delta_s,$ and δ_t be the step sizes of the discretization of parameters u through t . Then a ray with continuous parameters (u, v, s, t) is stored in the PAF at the discrete index

$$(U, V, S, T) = (\lfloor u/\delta_u \rfloor, \lfloor v/\delta_v \rfloor, \lfloor s/\delta_s \rfloor, \lfloor t/\delta_t \rfloor). \tag{1}$$

Projection values are only computed and stored for rays with parameters that are multiples of the discretization factors and that coincide with grid points in the discretized parameter space. Projection values for all rays in between are obtained by quadrilinear interpolation.

2.1 Cylindrical Parameterization of Attenuation Field

For 2D-3D registration, we define the AF parameterization using a cylinder of radius R around the image volume, the cylinder axis aligned with the body axis. Each ray of the

projection geometry is parameterized by four coordinates, which correspond to the two intersection points where the ray enters and exits this cylinder (Fig. 1). Two parameters, u and s , are rotation angles around the cylinder axis, the other two, v and t , are distances along the cylinder axis. Let L be a ray from \mathbf{a} to \mathbf{b} , parameterized as $L(c) = \mathbf{a} + \lambda \mathbf{d}$ with $\mathbf{a} = (a_x, a_y, a_z)$, $\mathbf{b} = (b_x, b_y, b_z)$, and $\mathbf{d} = \mathbf{b} - \mathbf{a}$. The AF parameters of R are computed as follows. First the parameters λ_{uv} and λ_{st} of the intersecting points of R with the coordinate cylinder are computed by solving the system

$$R^2 = (a_x + \lambda d_x)^2 + (a_y + \lambda d_y)^2 \quad (2)$$

for λ . The resulting two solutions λ_{uv} and λ_{st} describe the two intersecting points as

$$\begin{aligned} \mathbf{p}_{uv} &= \mathbf{a} + \lambda_{uv} \mathbf{d}, \\ \mathbf{p}_{st} &= \mathbf{a} + \lambda_{st} \mathbf{d}. \end{aligned}$$

From these, the actual AF parameters of the ray can be computed by inverting the following expressions for computing points on the cylinder from a parameter 4-tuple:

$$\begin{aligned} \mathbf{p}_{uv} &= (R \cos u, R \sin u, v), \\ \mathbf{p}_{st} &= (R \cos s, R \sin s, t). \end{aligned}$$

The coordinate system is fixed with respect to the 3D image. The cylinder axis is parallel to the z coordinate and runs through the image center. Note that since the cylinder is infinite, the reference level along the axis can be defined arbitrarily, and therefore the absolute values of v and t are irrelevant. The radius R is chosen as the difference between the center of the 3D image and its projection onto the initial location of the x-ray projection plane. Note that we assume here a stationary 3D image and moving 2D projection devices. In reality, one typically encounters the opposite situation, that is, a moving 3D image that is tracking the patient movement in space, while the x-ray projections are stationary (or at least in a known position). For rigid motion, both situations are easily translated into each other by applying the inverse 3D image transformation to the projection geometry, and vice versa.

Compared to the common two-plane parameterization, the cylindrical coordinate space is capable of describing full 360 degree rotations around the cylinder axis. A spherical parameterization would also be possible and provide full rotational freedom. However, two faces of a medical image volume are virtually always truncation surfaces. Rays entering or exiting through these surfaces are not physically possible and therefore irrelevant to the computation (see Fig. 2). The cylindrical parameterization is naturally suited for rotating imaging systems such as gantry-based systems and C-arm fluoroscopes.

2.2 Storage Data Structure

An efficient data structure is the key to a successful implementation of a PAF. Unlike the case of a precomputed AF, it is important that not only lookup but also storage operations are time-efficient. Memory efficiency is also important, in particular since we cannot

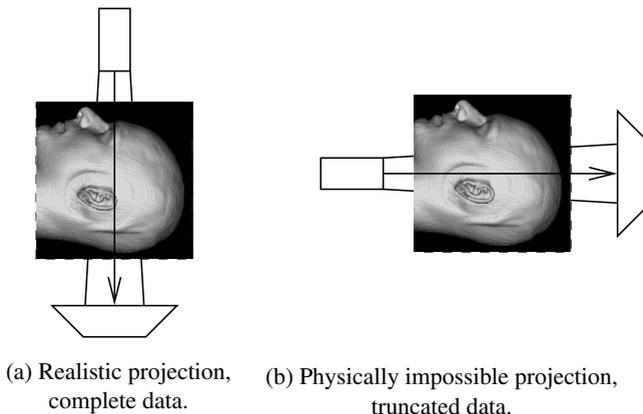


Fig. 2. Illustration of physical limitations to projection geometry and data truncation.

use data compression, which prevents a preallocated 4D array. A hash table allows both lookup and insertion operations in constant time and can grow incrementally as necessary. To store the ray attenuation for a given ray, we use the binary-OR composition of the respective lowest 8 bits of the integer values of U , V , S , and T as the hash function. In order to minimize collisions, the binary representation of V is shifted by 8 bits, that of S by 16 bits, and that of T by 24 bits.

For applications with two or more projection geometries, one can either use a single combined PAF for all of them, or one PAF per projection. Both options have their merits, depending on the actual setup. As an example, when two projections are (nearly) orthogonal, it is virtually impossible that the AFs for both will ever intersect. One should therefore use two separate AFs, one for each projection, and thereby reduce the likelihood of hash collisions. On the other hand, when the projection geometries are arranged closer to each other, or when excessive patient motion is possible, it may be advantageous to use only a single AF to share computed attenuation values among all available projections.

2.3 Multi-resolution Progressive Attenuation Fields

Generating a DRR from a PAF takes more time as the sampling of the AF parameters becomes finer. At the same time, the coarser these parameters, the more blurred are the resulting DRR images. It is conceivable to exploit this in an additional layer of multi-resolution computation by starting the registration process with a relatively coarse AF sampling and successively refining it. On the other hand, however, it is desirable to continue using the information stored in the coarse AF when switching to a finer one. Finally, converting the AF to finer coordinate samplings takes time and should be avoided.

Let (U, V, S, T) be the discrete hash table indices for the continuous parameter tuple (u, v, s, t) , discretized with samplings δ_u , δ_v , δ_s , and δ_t as before. Also, let $c \in \mathbb{N}^+$ be any positive integer. Then

$$(U', V', S', T') = (c \lfloor u/c\delta_u \rfloor, c \lfloor v/c\delta_v \rfloor, c \lfloor s/c\delta_s \rfloor, c \lfloor t/c\delta_t \rfloor) \quad (5)$$

is a super-discretization of the previous AF with upsampling factor c . This means that we can initially start building a PAF with an upsampling factor $c > 1$ and later refine the discretization to $c = 1$ without having to convert any entries in the hash table. Note also that c can be different for each of the four parameters.

3 Evaluation Study

3.1 Methods

We evaluate the usefulness of the PAF algorithm by using it to compute DRRs in the intensity-based 2D-3D registration of spine images. We compare the execution speed and the registration accuracy and robustness to that obtained by computing DRRs with traditional ray casting. The registration algorithm has been described previously [3]. In summary, the algorithm searches for the six parameters of the rigid transformation that produces the DRR that is most similar to the real x-ray projection image. We crop the projection images to a region of interest that includes the vertebra of interest plus the two adjacent vertebrae. We optimize the mutual information (MI) image similarity measure, which we have previously found to be both accurate and robust [6]. There are two x-ray projection images and two corresponding DRRs; the similarity measure is the sum of the MI for each of the real-synthetic image pairs. The transformation parameters are optimized using a simple best neighbor search strategy [8]. The registration process is performed in two passes, the first with smoothed versions of the projection images, and the second with the original images.

We use spine image data from two patients who underwent treatment of spinal cord lesions with the CyberKnife Stereotactic Radiosurgery System (Accuray, Sunnyvale, CA). These patients had four to five metal fiducial markers implanted percutaneously before pre-treatment CT scanning in the posterior bony elements of the vertebral levels adjacent to the lesions. For each patient, we obtained: 1) a pre-treatment CT image, 2) a pair of orthogonal projection x-ray images obtained with flat-panel silicon x-ray cameras (Flashscan 20, dpiX, Palo Alto, CA), 3) the camera calibration model and parameters for the two x-ray cameras, 4) positions (3D) of four metal fiducial markers in the CT image, and 5) positions (2D) of the four markers in the projection x-ray images. Additional details can be found in [3,6].

The metal fiducial markers were used to compute a gold-standard transformation. The target registration error (TRE) of an intensity-based registration transformation being evaluated is computed as the difference between the positions of a target mapped by the evaluated transformations and the gold-standard transformation. The TRE values are computed for each voxel inside a rectangular box bounding the vertebra and then averaged. Initial transformations were generated by perturbing the gold-standard reference transformation by adding randomly generated rotations and translations. The initial transformations were characterized by computing the TRE for the transformation and grouped into eight 2 mm intervals: 0–2, 2–4, . . . , 14–16 mm. We run the registration algorithm twice for each of twenty random initial transformations from each interval group, once using DRRs computed using PAF and once using DRRs computed using traditional ray casting.

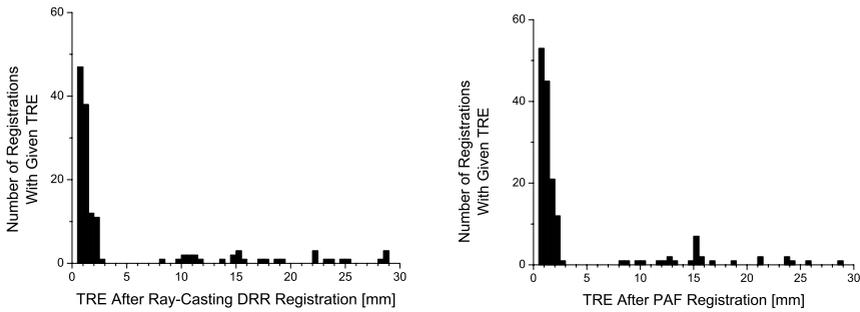


Fig. 3. Histograms of TRE after registration using ray-casting DRR (*left*) and PAF (*right*). Bin width is 0.5 mm in both histograms.

Table 1. Target registration errors (mean \pm standard deviation) for patient #1 in millimeters after registration with ray-casting DRR and PAF. Statistics include 20 registrations with random initial transformation per class (failed registrations excluded).

Initial TRE	0–2 mm	2–4 mm	4–6 mm	6–8 mm
DRR	0.95 \pm 0.16 mm	1.33 \pm 0.55 mm	1.37 \pm 0.55 mm	1.36 \pm 0.45 mm
PAF	0.99 \pm 0.21 mm	1.19 \pm 0.50 mm	1.28 \pm 0.44 mm	1.43 \pm 0.34 mm

Initial TRE	8–10 mm	10–12 mm	12–14 mm	14–16 mm
DRR	1.17 \pm 0.38 mm	1.51 \pm 0.70 mm	1.20 \pm 0.77 mm	1.29 \pm 0.39 mm
PAF	1.31 \pm 0.46 mm	1.66 \pm 0.73 mm	1.70 \pm 0.61 mm	1.12 \pm 0.29 mm

3.2 Results

Computational Efficiency. A comparison of computation times between ray-casting DRRs and PAF on a per-case basis is not very meaningful, because the number of steps in the search for the optimum transformation varies between the two methods, and so does the number of DRR computations. Over all registrations, the average registration time using ray-casting DRR was 2,357 seconds (standard deviation 1,168 seconds). The average time to complete a PAF registration was 604 seconds (std. dev. 312 seconds). The maximum time using ray-casting DRR was 13,772 seconds, and 2,584 seconds using PAF.

Accuracy and Robustness. Registration can sometimes fail, and the further away from the correct transformation the algorithm is started, the more likely failure becomes. Since the “accuracy” of failed registrations is basically random, these registrations need to be excluded from the accuracy analysis and instead treated in a separate robustness analysis. As Fig. 3 shows, final TRE values using both PAF and ray-casting DRR are either below 3 mm or above 8 mm. We therefore select a final TRE of 3 mm as the threshold for successful vs. failed registrations.

Fig. 4 shows the relative number of failed registrations vs. the initial TRE. All registrations were successful, based on the above threshold of 3 mm, for both PAF and

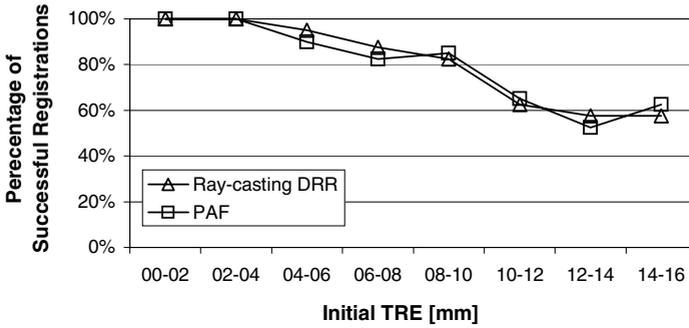


Fig. 4. Robustness vs. initial TRE for registration using ray-casting DRR computation and PAF. The graph shows the relative number of successful registrations (TRE less than 3.0 mm) for initial TREs from 0 through 16 mm (20 registrations with random initial transformation per class and per patient).

Table 2. Target registration errors (mean ± standard deviation) for patient #2 in millimeters after registration with ray-casting DRR and PAF. Statistics include 20 registrations with random initial transformation per class.

Initial TRE	0–2 mm	2–4 mm	4–6 mm	6–8 mm
DRR	1.61±0.07 mm	1.61±0.06 mm	1.60±0.06 mm	1.61±0.08 mm
PAF	1.62±0.06 mm	1.63±0.05 mm	1.60±0.05 mm	1.63±0.06 mm

Initial TRE	8–10 mm	10–12 mm	12–14 mm	14–16 mm
DRR	1.61±0.05 mm	1.61±0.05 mm	1.61±0.05 mm	1.65±0.21 mm
PAF	1.61±0.05 mm	1.64±0.06 mm	1.62±0.09 mm	1.62±0.06 mm

ray-casting DRR as long as initial TRE was less than 4 mm. For initial TREs between 4 mm and 8 mm PAF produced slightly more failed registrations than ray-casting DRR, while between 8 mm and 12 mm it produced slightly less. Overall, the robustness of registration using ray-casting DRR and PAF are approximately the same. The accuracy statistics of all successful registrations are shown in Table 1 for patient #1 and in Table 2 for patient #2. Again, the resulting values are virtually identical for both methods.

Note that for patient #2, the vast majority of TRE values are very close to 1.6 mm. Since the registration results appear visually convincing, this is probably due to inaccuracy of the gold standard transformation rather than errors of the intensity-based registration method. For medical reasons, fiducial markers in actual patients cannot be implanted in mathematically optimal locations that would maximize the accuracy of the computed transformation. Unfortunately, there is currently no more accurate method to compute “correct” coordinate transformations from clinical data than fiducial markers either. The spine as a whole moves nonrigidly, which can only be approximately detected using a small number of markers that are placed on different vertebrae.

4 Discussion

We have described a novel approach to increase the computational efficiency of intensity-based 2D-3D registration with no precomputation and virtually no restriction of the permissible patient movement. Both accuracy and robustness of the registration using PAF are essentially the same as those of registration using ray-casting DRR.

Our method is easily implemented and performs on average about four times faster than registration using ray-casting DRR. Our absolute computation times are still fairly large, which is due to a sub-optimal ray-casting DRR implementation. However, since in our method all attenuation values are computed as they are needed, rather than in a precomputation step, it benefits directly from faster ray casting methods during the registration phase. For example, Mori *et al.* [9] recently presented a fast CPU-based volume rendering method that computes DRRs several times faster than our relatively sub-optimal ray-casting DRR implementation.

Acknowledgment. This research was performed as part of a collaboration established with support from the Bavaria California Technology Center (BaCaTec).

References

1. D. LaRose, *Iterative X-ray/CT Registration Using Accelerated Volume Rendering*, Ph.D. thesis, Robotics Institute, Carnegie Mellon University, 2001.
2. S. J. Gortler, R. Grzeszczuk, R. Szeliski, M. F. Cohen, "The lumigraph," in *Proc. 23rd Annu. Conf. Computer Graphics and Interactive Techniques*, New York, NY, 1996, ACM SIGGRAPH, pp. 43–54, ACM Press.
3. D. B. Russakoff, T. Rohlfing, C. R. Maurer, Jr., "Fast intensity-based 2D-3D fluoroscopy-to-CT registration of clinical data using light fields," in *Ninth IEEE Int. Conf. Computer Vision*, 2003, pp. 416–422.
4. D. B. Russakoff, T. Rohlfing, D. Rueckert, C. R. Maurer, Jr., "Fast calculation of digitally reconstructed radiographs using light fields," in *Medical Imaging: Image Processing*, 2003, vol. 5032 of *Proc. SPIE*, pp. 684–695.
5. M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. 23rd Annu. Conf. Computer Graphics and Interactive Techniques*, New York, NY, 1996, ACM SIGGRAPH, pp. 31–42, ACM Press.
6. D. B. Russakoff, T. Rohlfing, A. Ho, *et al.* "Evaluation of intensity-based 2D-3D spine image registration using clinical gold-standard data," in *Biomedical Image Registration – WBIR 2003*, 2003, vol. 2717 of *LNCS*, pp. 151–160, Springer-Verlag.
7. G. P. Penney, J. Weese, J. A. Little, *et al.* "A comparison of similarity measures for use in 2D-3D medical image registration," *IEEE Trans. Med. Imag.*, vol. 17, no. 4, pp. 586–595, 1998.
8. C. Studholme, D. L. G. Hill, D. J. Hawkes, "Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures," *Med. Phys.*, vol. 24, no. 1, pp. 25–35, 1997.
9. K. Mori, Y. Suenaga, J. Toriwaki, "Fast volume rendering based on software optimization using multimedia instructions on PC platforms," in *Computer Assisted Radiology and Surgery*, Berlin, 2002, pp. 467–472, Springer-Verlag.