Radiographic 2D/3D image registration with pre-integrated random rays

Edmund Prager

2nd December 2024

1 DRR-based 2D/3D image registration

The standard method for 2D/3D image registration is an optimisation who's objective function is to formed of the following stages:

1. A 3D transformation is parametrised by the input values.

2. A DRR is generated through the 3D CT data at this transformation.

3. A similarity metric is used to compare the DRR with the fixed X-ray image for a scalar output value.

The computational expense of the similarity metric calculation can vary a lot depending on the choice of the similarity metric, but the expense of the DRR generation is never insignificant, so reducing time spent on DRR generation has the potential to be useful.

Another difficulty of this method is the difficulty of the optimisation problem, which features a highly multi-modal landscape with a narrow basin of attraction around the global optimum. To get around this, a common approach is mult-scale registration where the images are downsampled to various levels and registration is performed sequentially on each level, from more to less downsampled. The idea being that, with downsampled images, the optimisation landscape is much smoother with a larger basin of attraction around the, admittedly likely less accurate, global optimum.

2 Derivation of the 'random ray' approach

The first observation is that there are many conceivable pairs of transformations which will result in DRRs that have pixels who's rays will share identical radiological paths through the CT data, and therefore result in the same intensity value. This could be exploited to reduce the total number of radiological paths that need to be calculated for a given optimisation problem.

The feasibility of this is not clear; trivially, no number of distinct DRRs could be generated that would between them contain radiological paths that exactly correspond to a further distinct DRR. Taking the perspective of the potential for re-use of a single radiological path, the locus of transformations who's resulting DRRs contain a pixel that aligns exactly with that radiological path is a cloud of infinitesimal points in the transformation space.

The second observation is that for intensity-based registration methods (i.e. those that involve similarity metrics which are agnostic to the image structure of the data they are comparing; they simply map a set of pairs of scalar values to a single scalar value $S : \{\mathbf{x}_i \in \mathbb{R}^2 : i = 1 \ldots N\} \to \mathbb{R}$) a full DRR need not be constructed for comparison with the X-ray:

- only a subset of the X-ray pixels need be paired with corresponding values from the DRR for comparison, and

- with interpolation of the X-ray image, any radiological path that intersects the simulated X-ray detector array may have its corresponding X-ray image intensity sampled for use in the comparison.

With this, our ray's aforementioned locus is grown into a (still infinitesimally thin) surface through the transformation space, and evaluation of image similarity is feasible without a full structured set of radiological paths.

We make the assumption that a ray may be usefully re-used if a new transformation is near its surface in the transformation space, and so the 'random ray' method reveals itself:

- Radiological paths are initially calculated for a large number of rays at random in the CT space.

- The objective function of the optimisation is then:

    1. A 3D transformation is parametrised by the input values.

    2. The rays are transformed into X-ray space with this transformation.

    3. Rays that are near enough to being used in DRR generation at that transformation have their intersection positions with the simulated detector array evaluated, and their corresponding intensity values in the X-ray image sampled.

    4. Each ray then provides a pair of values for the similarity metric: pre-calculated radiological path and X-ray image sample. The similarity metric is evaluated with these for the scalar output value.

The third observation is that an intensity based similarity metric may be extended to incorporate its input values according to weights. For example, the zero-normalised cross-correlation (ZNCC) is defined as follows:

$$\text{ZNCC}(X, Y) = \frac{1}{N\sigma_X\sigma_Y} \sum_i (x_i - \mu_X)(y_i - \mu_Y) = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N\sum x_i^2 - (\sum x_i)^2}\sqrt{N\sum y_i^2 - (\sum y_i)^2}}$$

Which we can modify into a 'weighted' zero-normalised cross-correlation (WZNCC):

$$\text{WZNCC}(\{x_i, y_i, w_i\} : i = 1 \dots N) = \frac{1}{\sigma_X\sigma_Y \sum_i w_i} \sum_i w_i(x_i - \mu_X)(y_i - \mu_Y),$$

$$\mu_X = \frac{\sum_i w_i x_i}{\sum_i w_i}, \quad \sigma_X = \frac{1}{\sum_i w_i} \sum_i w_i (x_i - \mu_X)^2$$

$$\Rightarrow \text{WZNCC} = \frac{\sum w_i \sum w_i x_i y_i - \sum w_i x_i \sum w_i y_i w}{\sqrt{\sum w_i \sum w_i x_i^2 - (\sum w_i x_i)^2}\sqrt{\sum w_i \sum w_i y_i^2 - (\sum w_i y_i)^2}}$$

This allows us to modify our objective function to be continuous (and therefore differentiable) by weighting rays according to their proximity to the exact rays for a given transformation's DRR. We define this proximity as the distance d between the ray and the simulated X-ray source in X-ray space.

The corresponding weights w are then calculated using some sort of drop-off function, parametrised by a drop-off rate α:

$$w = \exp\left(-\frac{d^2}{\alpha}\right)$$

For a given value of α and number of rays used N, there will be a different expected value of $\sum_i w_i$. In the analogous 1D/2D case:

$$\mathrm{E}\left(\sum_i w_i\right) = \frac{N\alpha}{2r\sqrt{\pi}},$$

and in the 2D/3D case:

$$\mathrm{E}\left(\sum_i w_i\right) = \frac{N\alpha^2}{4r^2},$$

where r is the distance from the patient to the X-ray source. Increasing the value of α has a similar effect on the similarity value to downsampling the images: the optimisation landscape is smoothed out, but the optimum is less accurately located at the alignment of the images. Too small a value of α will result in a larger number of rays required for a reasonable value of $\sum_i w_i$.

2.1 Blurring the fixed image

As an increased value of α will effectively blur the DRR of which the rays are providing an approximating subset, a Gaussian blur of some corresponding standard deviation can be applied to X-ray before sampling, or equivalently the fixed image can be downsampled a corresponding amount for reduced computational and memory expense.

2.2 Clipping rays to the detector array

So far, the discussion of ray weights has considered only the proximity of the rays to the X-ray source position, but a ray that doesn't pass through the X-ray detector array, and therefore has no corresponding intensity value in the X-ray image, ought to be excluded from the similarity calculation as well. This can be achieved by determining a weight multiplier from the X-ray image sample position, e.g. a linear drop off from 1 to 0 between the outer rows/columns of X-ray image and the outside of the image. Incorporating this results in an optimisation landscape that:

- is slightly less smooth,

- is slightly higher, except at the global minimum, where it is more similar, and

- has stronger peaks outside of the basin of the global minimum.

This makes sense as most rays that pass outside of the X-ray detector array will have small radiological path integrals, so will generally match closely with the 0s that would be sampled outside of the X-ray.

3   Maths Needs significant polishing

An intensity-based similarity metric s is a function that maps an arbitrary number N of pairs of values $\{\mathbf{x}_i \in \mathbb{R}^2 = \{x_i, y_i\} : i = 1 \ldots N\}$ to a scalar value $S \in \mathbb{R}$, where the images are $X = \{x_i : i = 1 \ldots N\}$ and $Y = \{y_i : i = 1 \ldots N\}$.

A metric of the similarity between two images can still be calculated without all the values from both images, each value used just needs to have its corresponding value in the other image.

- Let the fixed X-ray image $X : \mathbb{Z}^2 \to \mathbb{R}$ be approximately interpolated by the function $f : \mathbb{R}^2 \to \mathbb{R}, \ f(i, j) \approx x_{ij}$.

- Consider the position of the X-ray source to be fixed relative to the X-ray detector array.

- A ray R through the CT space can be defined as the locus of points

$$\mathbf{p}(\lambda) = \mathbf{p}_0 + \lambda \hat{\mathbf{d}}, \ \ \forall \lambda \in \mathbb{R},$$

  parametrised by $\{\mathbf{p}_0, \hat{\mathbf{d}}\}$.

- The integral of μ through the CT space can be calculated for such a ray, giving $I \in \mathbb{R}$. This value can be considered an attribute of the ray: $R = \{\mathbf{p}_0, \hat{\mathbf{d}}, I\}$.

- Parametrising the transformation T of the CT in X-ray space, as an anti-clockwise rotation through r followed by a translation by (t_x, t_y):

$$T = \begin{bmatrix} R & \mathbf{t} \\ 0 & 1 \end{bmatrix}$$

- The corresponding ray in X-ray space is T applied to the ray in CT space:

$$\mathbf{p}'(\lambda) = \mathbf{p}'_0 + \lambda \hat{\mathbf{d}}' = T \begin{bmatrix} \mathbf{p}_0 \\ 1 \end{bmatrix} + \lambda T \begin{bmatrix} \hat{\mathbf{d}} \\ 0 \end{bmatrix}$$

- The ray will correspond to a valid ray for DRR generation if it passes through the X-ray source position $\mathbf{p}_s$ in X-ray space:

$$\left\| \left(\mathbf{p}_s - \mathbf{p}'_0 \right) \times \hat{\mathbf{d}}' \right\| < \epsilon$$

- Note that:

$$\mathbf{p}'_0 = R\mathbf{p}_0 + \mathbf{t}, \quad \hat{\mathbf{d}}' = R\hat{\mathbf{d}}$$

$$\Rightarrow (\mathbf{p}_s - \mathbf{p}'_0) \times \hat{\mathbf{d}}' = \mathbf{p}_s \times R\hat{\mathbf{d}} - \mathbf{t} \times R\hat{\mathbf{d}} - R\mathbf{p}_0 \times R\hat{\mathbf{d}} = R\Big(\big(R^{-1}(\mathbf{p}_s - \mathbf{t}) - \mathbf{p}_0 \big) \times \hat{\mathbf{d}} \Big)$$

$$\Rightarrow \left\| (\mathbf{p}_s - \mathbf{p}'_0) \times \hat{\mathbf{d}}' \right\| = \left\| \left(T^{-1} \begin{bmatrix} \mathbf{p}_s \\ 1 \end{bmatrix} - \mathbf{p}_0 \right) \times \hat{\mathbf{d}} \right\| = \left\| (\mathbf{p}_{s,\mathrm{CT}} - \mathbf{p}_0) \times \hat{\mathbf{d}} \right\|$$

- If it is valid, the location (u, v) on the DRR it corresponds to will be given by:

$$(u, v) = \left(p'_{0x} - \frac{p'_{0z}}{\hat{d}'_z} \hat{d}'_x, \ p'_{0y} - \frac{p'_{0z}}{\hat{d}'_z} \hat{d}'_y \right)$$

- Where the DRR is located on the x, y-plane, with $u = 0, v = 0$ corresponding to its centre.

- Given a transformation T, any ray $R = \{\mathbf{p}_0, \hat{\mathbf{d}}, I\}$ can provide an intensity pair for the similarity function as follows:

$$\{x_i', y_i'\} = \left\{ f\left(p_{0x}' - \frac{p_{0y}'}{d_y'}d_x'\right), I \right\} \text{ if } \left\|(\mathbf{p}_\mathrm{s} - \mathbf{p}_0') \times \hat{\mathbf{d}}'\right\| < \epsilon, \ \mathbf{p}_0' = T\begin{bmatrix}\mathbf{p}_0 \\ 1\end{bmatrix}, \ \hat{\mathbf{d}}' = T\begin{bmatrix}\hat{\mathbf{d}} \\ 0\end{bmatrix}$$

- This can be made continuous by considering the number of intensity pairs as a continuous value $N = \sum_i n_i$, to which each can contribute $n \in [0, 1]$:

$$n = \exp\left(-\alpha \left\|(\mathbf{p}_\mathrm{s} - \mathbf{p}_0') \times \hat{\mathbf{d}}'\right\|^2\right)$$

for some scaling value α.

- So, a number $M \sim 10^6$ of rays can be calculated, $\mathbf{p}_0$ and $\hat{\mathbf{d}}$ for each generated randomly, and I calculated for each through the CT data. Each ray will need about 28 bytes, so a total memory use of the order of $100\,\mathrm{MB}$.

- For an iteration of the optimisation at a transformation T, the source position is calculated in CT space, and used to calculated the weight of each ray in parallel:

$$n_i = \exp\left(-\alpha \left\|(\mathbf{p}_\mathrm{s,CT} - \mathbf{p}_{0i}) \times \hat{\mathbf{d}}_i\right\|^2\right)$$

- For $\alpha \sim 3000$, rays further than $3\,\mathrm{cm}$ from the source position will have small n_i, so for a patient - detector distance of $\sim 1\,\mathrm{m}$, and assuming a uniform distribution of ray intersections through the $1\,\mathrm{m}$ radius spherical surface centred on the patient, about $\frac{\pi \cdot 0.03^2}{4\pi \cdot 1^2} \approx \frac{1}{4000}$ of the rays $\sim \frac{M}{10^3} \sim 10^3$ will be significant.

- The value pair contributions $\{x_i', y_i', n_i\}$ are then calculated for these rays only:

$$\{x_i', y_i', n_i\} = \left\{ f\left(p_{0x}' - \frac{p_{0z}'}{\hat{d}_z'}\hat{d}_x', p_{0y}' - \frac{p_{0z}'}{\hat{d}_z'}\hat{d}_y'\right), I, \exp\left(-\alpha \left\|(\mathbf{p}_\mathrm{s,CT} - \mathbf{p}_0) \times \hat{\mathbf{d}}\right\|^2\right) \right\}$$

$$\mathbf{p}_0' = T\begin{bmatrix}\mathbf{p}_0 \\ 1\end{bmatrix}, \ \hat{\mathbf{d}}' = T\begin{bmatrix}\hat{\mathbf{d}} \\ 0\end{bmatrix}$$

- And the final similarity S is then calculated as

$$S = S(\{x_0, y_0, n_0\}, \{x_1, y_1, n_1\}, \dots), \ x_i = n_i x_i', \ y_i = n_i y_i'$$

- Consider the transformation to be parametrised by some values $\boldsymbol{\theta}$.

- The derivatives of each value w.r.t. to the transformation parameters are:

$$\frac{\partial x_i}{\partial \theta_j} = x_i' \frac{\partial n_i}{\partial \theta_j} + n_i \frac{\partial x_i'}{\partial \theta_j}$$

$$\frac{\partial x_i'}{\partial \theta_j} = \frac{\partial f}{\partial u}(u, v)\frac{\partial u}{\partial \theta_j} + \frac{\partial f}{\partial v}(u, v)\frac{\partial v}{\partial \theta_j}$$

$$u = p_{0x}' - \frac{p_{0z}'}{\hat{d}_z'}\hat{d}_x', \quad v = p_{0y}' - \frac{p_{0z}'}{\hat{d}_z'}\hat{d}_y'$$

$$\frac{\partial u}{\partial \theta_j} = \frac{\partial p_{0x}'}{\partial \theta_j} - \beta\hat{d}_x' - \frac{p_{0z}'}{\hat{d}_z'}\frac{\partial \hat{d}_x'}{\partial \theta_j}, \quad \frac{\partial v}{\partial \theta_j} = \frac{\partial p_{0y}'}{\partial \theta_j} - \beta\hat{d}_y' - \frac{p_{0z}'}{\hat{d}_z'}\frac{\partial \hat{d}_y'}{\partial \theta_j}$$

$$\beta = \frac{\partial}{\partial \theta_j}\left(\frac{p'_{0z}}{\hat{d}'_z}\right) = \frac{\hat{d}'_z \frac{\partial p'_{0z}}{\partial \theta_j} - p'_{0z}\frac{\partial \hat{d}'_z}{\partial \theta_j}}{\hat{d}'^2_z}$$

$$\frac{\partial n_i}{\partial \theta_j} = -2\alpha n_i\Big((\mathbf{p}_{s,CT} - \mathbf{p}_0) \times \hat{\mathbf{d}}\Big) \cdot \left(\frac{\partial \mathbf{p}_{s,CT}}{\partial \theta_j} \times \hat{\mathbf{d}}\right)$$

$$= -2\alpha n_i\left((\mathbf{p}_{s,CT} - \mathbf{p}_0) \cdot \frac{\partial \mathbf{p}_{s,CT}}{\partial \theta_j} - \Big((\mathbf{p}_{s,CT} - \mathbf{p}_0) \cdot \hat{\mathbf{d}}\Big)\left(\frac{\partial \mathbf{p}_{s,CT}}{\partial \theta_j} \cdot \hat{\mathbf{d}}\right)\right)$$

- The derivative of the similarity metric w.r.t. a transformation parameter θ_j is

$$\frac{\partial S}{\partial \theta_j} = \sum_i\left(\frac{\partial S}{\partial x_i}\left(x'_i\frac{\partial n_i}{\partial \theta_j} + n_i\frac{\partial x'_i}{\partial \theta_j}\right) + \frac{\partial S}{\partial y_i}\left(y'_i\frac{\partial n_i}{\partial \theta_j} + n_i\frac{\partial y'_i}{\partial \theta_j}\right) + \frac{\partial S}{\partial n_i}\frac{\partial n_i}{\partial \theta_j}\right)$$

- A common similarity metric is the zero-normalised cross-correlation:

$$\text{ZNCC}(X,Y) = \frac{1}{N\sigma_X\sigma_Y}\sum_i(x_i - \mu_X)(y_i - \mu_Y) = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N\sum x_i^2 - (\sum x_i)^2}\sqrt{N\sum y_i^2 - (\sum y_i)^2}}$$

Which we can modify into a 'weighted' zero-normalised cross-correlation:

$$\text{WZNCC}(\{x'_i, y'_i, n_i\}\dots) = \frac{1}{\sigma_X\sigma_Y\sum_i n_i}\sum_i n_i(x'_i - \mu_X)(y'_i - \mu_Y)$$

$$= \frac{\sum n_i \sum n_i x'_i y'_i - \sum n_i x'_i \sum n_i y'_i}{\sqrt{\sum n_i \sum n_i x_i^2 - (\sum n_i x_i)^2}\sqrt{\sum n_i \sum n_i y_i^2 - (\sum n_i y_i)^2}}$$

- Add stuff about blurring fixed image according to alpha

- Zero weight when outside X-ray?