

Henry Westmacott

henryjw@mrc-lmb.cam.ac.uk Department of Computer Science, Durham University Durham, UK MRC Laboratory of Molecular Biology Cambridge, UK Ioannis Ivrissimtzis

ioannis.ivrissimtzis@durham.ac.uk Department of Computer Science, Durham University Durham, UK

Tobias Weinzierl

tobias.weinzierl@durham.ac.uk Department of Computer Science, Durham University Durham, UK

ABSTRACT

We produce thickness and bone to soft tissue ratio estimations from a single, 2D medical X-ray image. For this, we simulate the scattering of the rays through a model of the object and embed this simulation into an optimiser which iteratively adjusts the model to match the X-ray simulation to the observed X-ray image. Utilising a combination of different techniques, first, a CNN-based image segmentation serves as a regularisation term to the underlying cost function to guide the descent, while domain knowledge about physical parameter correlations is injected by additional penalty terms. Next, the optimiser is embedded into a multilevel framework which, similar to multi-grid's philosophy, successively improves the model on varying resolutions while individual resolutions focus on particular terms of the cost function. Initial results suggest that we can obtain meaningful thickness and material estimations.

CCS CONCEPTS

• Mathematics of computing \rightarrow Optimization with randomized search heuristics; • Computing methodologies \rightarrow Regularization; Reconstruction.

KEYWORDS

X-ray simulation, model reconstruction, material estimation, gradient descent, regularization

ACM Reference Format:

Henry Westmacott, Ioannis Ivrissimtzis, and Tobias Weinzierl. 2024. A multiscale optimisation algorithm for shape and material reconstruction from a single X-ray image. In 2024 the 7th International Conference on Image and Graphics Processing (ICIGP 2024), January 19–21, 2024, Beijing, China. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3647649.3647690

1 INTRODUCTION

X-ray imaging and the subsequent analysis and interpretation of these images is one of the fundamental diagnostic tools of modern medicine. Inferring spatial quantitative information from a single X-ray image challenges data-based artificial intelligence algorithms.

This work is licensed under a Creative Commons Attribution International 4.0 License.

ICIGP 2024, January 19–21, 2024, Beijing, China © 2024 Copyright held by the owner/author(s). ACM ISBN 979-8-4007-1672-0/24/01 https://doi.org/10.1145/3647649.3647690 First, bone and tissue thicknesses are not primary output quantities of X-ray imaging, and it is thus difficult to create labelled training data. Second, the radiation dosage should be minimised, meaning that ideally we would only have a single X-ray intensity measurement with considerable noise. Finally, scatter yields artefacts in the images. Scatter is a naturally occurring phenomenon in which photons in flight between source and detector are deflected from their initial direct path by the object that is being imaged, henceforth referred to as "the object". This deflection creates a fog-like haze across the image where these scattered photons landed, degrading image quality. Crucially, different to noise, scatter is not random. It encodes spatial information.

We propose a physics-based approach to reconstruct the model of the object. Let the input $I \in \mathbb{R}^{n_x \times n_y}$ be the $n_x \times n_y$ -dimensional X-ray image. We want to estimate the thickness t of the scanned object, and the alloy (bone/soft tissue) ratio μ within. Let $L(\mathbf{t}, \mu)$ be a simulated X-ray image corresponding to the object's model, parameterised through the vector variables t and μ , represented in our implementation by 2-dimensional arrays. The challenge then reads as:

$$\min_{(\mathbf{t},\boldsymbol{\mu})} J(\mathbf{t},\boldsymbol{\mu}) \quad \text{with} \quad J(\mathbf{t},\boldsymbol{\mu}) = \frac{1}{2} \|I - L(\mathbf{t},\boldsymbol{\mu})\|_{X}^{2} \quad (1)$$

with a suitable norm $\|.\|_x$ over $\mathbb{R}^{n_x \times n_y}$, and approach it as a classic optimisation problem subject to a gradient descent algorithm which minimises the difference between the simulated and the true image.

The X-ray simulator, here the Geant4 [13], produces a projection $L(\mathbf{t}, \boldsymbol{\mu})$ of the object's model onto the imaging plane. Even though we spatially parameterise our model through a 2D array of thickness values, and thus work in 2.5D rather than a real 3D representation, the projection $L(\mathbf{t}, \boldsymbol{\mu})$ is much more complex, and contains more information about the object's shape, than the standard orthographic projection that we get if we ignore scatter and consider direct beam only. Thus, where other reconstruction methods ignore scatter and require two ([24]), or more ([9]) X-ray images, we are able to reconstruct the object from a single X-ray image using a realistic physics simulator to compute projections of a 2.5 model with extra material information, onto a 2D imaging plane.

As (1) is ill-posed, the minimisation requires further regularisation. First, we introduce a weak admissibility criterion guided by segmentation: A neural network identifies regions within the image I which are free of bone or contain a combination of bone and tissue. This material knowledge helps to yield an initial guess for (t, μ) , and is then added as a penalty term to (1). Second, we introduce a penalty term which enforces smoothness and continuity over (t, μ) . Finally, we introduce a regulariser to penalize values that fall outside our valid ranges, thus enforcing physical constraints. To the best of our knowledge, this triad of optimisation ingredients, plus a rigorous multiscale formalism, have not been proposed before.

Various limitations however remain. Our model employs a gradient descent method which is robust, but is not yet tuned towards fast convergence. A systematic optimisation and balancing of our hyper-parameters (primarily, learning rate and regulariser parameter weights) was deemed out of scope. Moreover, we work solely with geometrically constructed multi-scale operators. As many fine grid effects can not be represented on coarse resolution levels if we only average and interpolate geometrically, the integration of geometric inter-grid transfer operators, or more advanced homogenisation techniques, is a natural follow-up step. The quality of our results was further limited by the low resolution of the 2.5D thickness/alloy models we were able to use. At each level of spatial resolution of the thickness/allow model, the use of the Geant4 simulator requires a calibration process which takes into account the energy level at which the X-ray machine was operated at. In practice, that meant that we could only work with a small number of coarse already calibrated resolutions.

The remainder is organised as follows. After a brief literature review in Section 2, we discuss our physical and numerical model which leads to a formulation of the optimisation algorithm in Section 3. This model is then augmented by a segmentation-guided initial guess and the proposed regularisation techniques in Section 4, before introducing the multi-scale variant in Section 5. Numerical results in Section 6 highlight the algorithm's properties on an artificial benchmark and real-world problems, before a quick outlook and summary in Section 7 close the discussion.

2 LITERATURE REVIEW

2.1 Ill-posed inverse problems

The inverse problem we seek to solve is ill-posed according to [22]. The ill-posed nature of tomography, that is, the separation of a 2D image space into 3D components, is well documented [27], and iterative reconstruction algorithms for X-ray tomography imaging such as in CT are thought by some to be reaching the limits of their clinical value [31]. In the general field of solving ill-posed problems, [5] proposes a iterative regularization gradient technique for such problems, [10] presents a deep, CNN-based method for producing super-resolution images for a single, low-resolution image, while [41] presents an analogous technique for image denoising.

The incorporation of domain knowledge has been proposed in [30]. This work reformulates the matrix-operations inherent in CT imaging and factors them into a neural network, thereby injecting prior knowledge about the physics of the problem into the solution. Similar ideas are applied to CBCT imaging from incomplete data in [39]. Finally, as we attempt to optimise a non-convex function, we note that non-convex optimisation has taken on renewed research interest in recent years due to its applicability to common machine learning problems [21].

2.2 Regularisation

[33] discuss the deep neural networks' potential for regularizing inverse problems, proposing that the entire regularization process be carried out by a neural network. [25] use neural networks for regularization in ill-posed problems relating to structure determination in biological macromolecule imaging. [3] describes a variational approach to solving inverse problems in the presence of noise, minimising a convex function that is regularised using knowledge about the forward-problem and the type of solutions we might expect. Comparable approaches are proposed in [7] specifically for astronomical interferometry images. Physics-informed techniques also fall under the principle of regularisation when knowledge about the physics of the underlying system is used to constrain an optimization [23, 40]. [26] also seeks to resolve the 3D structure of objects through regularization from prior knowledge, however, they use several images, at different focus points, to infer depth information.

[1] consider a technique for functional regularized reconstruction of CT data via learned gradients, aiming at developing techniques which will use both classical regularization and deep learning. [6] seeks to replace altogether traditional functional regularization in image reconstruction with CNN based methods, while [29] investigates how functional regularization encoding knowledge about the model and the forward operator can be learned by GANs.

[14] questions the robustness of techniques that seek to learn the forward operator in inverse problem solving. They propose a technique for model adaptation to make the learned models more accurate under *model drift*. [28] proposes a neural network alternative to classical Tikhonov regularization, and applies it to the problem of low-dose, incomplete CT imaging. [4] uses a combination of deep, learned regularization techniques and classical regularization for ill-posed problems in medical imaging, utilising Deep Image Priors, introduced in [38] for ill-posed image restoration problems.

2.3 Multi-scale

Here, we borrow ideas from multiscale analysis, making use of both the error-smoothing and coarse-grid correction that multigrid can offer us, as outlined in chapter 2 of [37]. The concept of algebraic multigrid described in [36] is also of interest to us, as they allow a better use of the extra information provided by a segmentation of our image, which will be discussed in more detail in 5.

In [34], grey scale native-resolution images are analysed alongside their down-sampled counterparts, as it is recognised that different resolutions will better represent information encoded at different frequency features. This is a feature that we try to harness the value of as well. Recent work has been carried out to apply multi-grid techniques for efficiency in medical image processing, specifically patch-based scatter correction in X-ray imaging [15]. In [15] multi-grid is used to process medical images at various resolutions, making computationally feasible their technique of correcting X-ray scatter based on relatively large patch sizes and search areas. [18] uses multiscale methods in inverse problem solving, specifically, CT reconstruction with the use of neural networks incorporating handcrafted features of domain knowledge, as was in the case in [30].

2.4 X-ray simulation

The X-ray simulation of proposed the thickness/alloy model uses an IBEX Ltd technology, developed over Geant4 simulation kit [11– 13]. Geant4 is a software toolkit for physically realistic, stochastic simulation of particle transport, developed at CERN. It has been more intensely developed, and it is more widely used than its direct competitors, such as EGS5 [19], and PENELOPE [32]. Geant4 has been utilised in numerous projects and applications, ranging from simulations of targeted particle therapy at biological cell scale [20], to atmospheric simulations at planetary scale [16].

3 PHYSICAL AND NUMERICAL MODEL

3.1 The thickness / alloy model

Our model of the X-rayed object consists of a single layer of $n_x \times n_y$ voxels. Each voxel x, y has a particular thickness $\mathbf{t}_{xy} \ge 0$ and accommodates a homogeneous mixture of two materials, soft tissue and bone, at a specific for that voxel ratio $\mu_{xy} \in [0, 1]$, referred to as the voxel's alloy. A 0 value of μ means pure soft tissue, while 1 means pure bone.



Figure 1: An X-ray beam hits a voxel and is scattered or pass through while its energy decreases. The voxels have different thicknesses and alloy compositions.

Our spatial 2.5D model is similar to models used in shallow water equations, or some ice sheet models erecting a volumetric model over a 2D domain. It has been designed to facilitate computational efficiency and be compatible with the simulator. In the simulation, X-rays hit the object orthogonal to the voxel layer and then hit the imaging plate orthogonally, too, (Fig. 1). The function

$$L: \left(\mathbb{R}_0^+ \times [0,1]\right)^{n_x \times n_y} \mapsto \left(\mathbb{R}_0^+\right)^{n_x \times n_y}$$

accepts a thickness field $\mathbf{t} \in (\mathbb{R}_0^+)^{n_x \times n_y}$, and a corresponding alloy field $\boldsymbol{\mu} \in [0, 1]^{n_x \times n_y}$, returning a $n_x \times n_y$ simulated X-ray image.

3.2 Gradient computation

As $L(t, \mu)$ is a black-box function, no analytic second derivative is available. We thus make (1) subject to a simple gradient descent based on numerical evaluations, computing a sequence of approximations of the model $(\mathbf{t}, \boldsymbol{\mu})^{(1)}, (\mathbf{t}, \boldsymbol{\mu})^{(2)}, (\mathbf{t}, \boldsymbol{\mu})^{(3)}, \dots$ through

$$(\mathbf{t}, \boldsymbol{\mu})^{(n+1)} = (\mathbf{t}, \boldsymbol{\mu})^{(n)} - \gamma (\nabla_{\delta_{xy}}) J((\mathbf{t}, \boldsymbol{\mu})^{(n)})$$

where

$$(\nabla_{\boldsymbol{\delta}_{xy}})J((\mathbf{t},\boldsymbol{\mu})^{(n)}) = \frac{J((\mathbf{t},\boldsymbol{\mu})^{(n)} + \boldsymbol{\delta}_{xy}) - J((\mathbf{t},\boldsymbol{\mu})^{(n)})}{|\boldsymbol{\delta}_{xy}|} \cdot \frac{\boldsymbol{\delta}_{xy}}{|\boldsymbol{\delta}_{xy}|}$$

We randomly pick a direction δ_{xy} to walk along. The real gradient $(\nabla_{\delta_{xy}})J$ along this direction is approximated by simple finite differences of the values of J at $(\mathbf{t}, \boldsymbol{\mu})^{(n)}$ and at $(\mathbf{t}, \boldsymbol{\mu})^{(n)} + \delta_{xy}$. The modification term δ_{xy} has the same dimensions as the model $(\mathbf{t}, \boldsymbol{\mu})$ and value zero everywhere, except for one randomly chosen voxel (x, y), where one or both of the thickness \mathbf{t}_{xy} and alloy $\boldsymbol{\mu}_{xy}$ values is equal to a small, random, non-zero quantity. The code randomly chooses between changes to thickness, alloy, or a combination of the two. Once we have randomly picked a search direction, J is evaluated both for $\pm \delta_{xy}$, i.e., we look both "left" and "right". Finally, the learning rate γ is a small hard-coded scalar quantity.

The data in *I* and $L(t, \mu)$ reflect raw data pulled from X-ray devices and results of particle simulations, respectively. Intensities are measured in approximate photon counts, typically on the scale of 10⁴ counts per pixel. The thickness **t** is given in centimeters, and in our experiments, we pick it from 0cm–30cm. An a priori calibration of the simulator ensures that the values of $L(t, \mu)$ fit to those of *I*.

4 REGULARIZED GRADIENT DESCENT

As (1) is of potentially high dimension, likely non-convex (in $L(\mathbf{t}, \boldsymbol{\mu})$) and constrained, it is hard to solve. It also is a coarse approximation of reality: the Geant4 simulates the scatter behaviour of the rays approximately; the reference image *I* contains measurement noise, as well as error from the calibration; most importantly, we use of a low resolution 2.5D model of only two material types.

For these reasons, we introduce various regularisers plus a biased randomised choice of gradient directions. These ingredients have two purposes: they ensure that the explored (t, μ) choices represent physically meaningful objects, and speed up the computation. We note, however, that even with the use of regularisers, still we cannot guarantee global convexity or the existence of a global minimum. We only can experimentally verify that the problem is tractable.

4.1 Physical admissibility constraints

Besides $t \ge 0$ and $0 \le \mu \le 1$, we also introduce $t \le t_{max}$, where t_{max} is the maximum object thickness the scanner is able to handle. Instead of enforcing physical admissibility by hard-coding the constraints, we add instead penalty terms to the cost (1). Specifically, we have

$$J_{adm}(\mathbf{t}, \boldsymbol{\mu}) = \frac{1}{2} \|I - L(\mathbf{t}, \boldsymbol{\mu})\|_{x}^{2} + \lambda_{adm} \|\min\{\mathbf{t}, 0\} + \max\{0, \mathbf{t} - \mathbf{t}_{max}\}\|_{fro} + \lambda_{adm} \|\min\{\boldsymbol{\mu}, 0\} + \max\{0, \boldsymbol{\mu} - 1\}\|_{fro}$$
(2)

for a fixed value of $\lambda_{adm} \ge 0$.

Any $\lambda_{adm} < \infty$ permits slight violations of the admissibility constraints, as we enforce them only weakly. Since the simulator

ICIGP 2024, January 19-21, 2024, Beijing, China

cannot handle values that do not satisfy these constraints, as for example negative thicknesses, we approximate (t, μ) to the nearest admissible values before passing them to the simulator.

4.2 Segmentation-guided admissibility and initialisation

Deep learning algorithms have been firmly established as the current state-of-the-art in image segmentation, outperforming earlier approaches, [35], based on shallow networks. Here, we use the deep learning X-ray segmentation algorithm *XNet* [8], which segments an X-ray image into three classes: open beam (void), soft tissue only, contains bone. The computed segmentation is encoded by two booleans per pixel

$$seg: I \mapsto (\{0,1\} \times \{0,1\})^{n_x \times n_y} \tag{3}$$

where values of 1 in seg_1 and seg_2 denote the presence of soft tissue and bone, respectively. As $seg_2 = 1 \Rightarrow seg_1 = 1$, i.e., there is no bare bone without tissue around it, seg classifies the image voxels into three classes.

Where the segmentation has estimated an alloy of just soft tissue, i.e., we should have $\mu_{xy} = 1$, the cost function is increased by a value proportional to $1 - \mu$. Where the segmentation has estimated that bone is present, i.e., we should have $\mu_{xy} < 1$, the cost function is increased by a constant value for any pixel with $\mu \ge 1$. This gives the further regularized cost function:

$$J_{seg}(\mathbf{t}, \boldsymbol{\mu}) = J_{adm}(\mathbf{t}, \boldsymbol{\mu}) + \lambda_{seg1} ||seg_1 \cdot (1 - \boldsymbol{\mu})||_{fro} + \lambda_{seg2} ||seg_2 \cdot [\boldsymbol{\mu} > 1]||_{fro}$$
(4)

where \cdot denotes element-wise matrix multiplication, 1 is the matrix with all values 1, and $[\mu > 1]$ is the matrix with values 1 at the places where $\mu_{x,y} > 1$, and 0 elsewhere.

The segmentation (3) is also used for the initialisation of the algorithm. Thickness is initialised to $t \approx t_{max}/2$ where we see some material, and 0 elsewhere. The alloy is initialised to $\mu = 0.5$ in the presence of some bone, and 0 elsewhere. Void voxel entries are omitted from the stochastic gradient trials δ .

4.3 Smoothness constraints

Anatomies are organic shapes which rarely exhibit non-smooth changes in thickness or alloy in any dimension, an issue that makes modelling them a challenging task [17]. The exception to this rule can be found along the edges of the object and the bone inside, where we expect a sharp change from non-zero thickness to zero thickness, and changes of 100% soft-tissue to < 100%. We apply a zero-cross edge detection filter on the seg_1 and seg_2 binary masks of the segmented image (3), identifying the edges of the just the bone region, and the edges of the imaged object. This yields a further tuple

$$edge: I \mapsto (\{0,1\} \times \{0,1\})^{n_x \times n_y} \tag{5}$$

per pixel, which flags the transition from void into tissue and from tissue-only into a tissue-and-bone region. We use it to define a further regularization term which penalizes large second derivatives, deactivated along edges where we expect to sharp changes



Figure 2: Why a smoothness regularisation is important. To model the scattering in (a), an intensity increase in the lower right voxel requires either a significant reduction of the voxel's height (b) or a moderate of several voxels around (c). The desirable model of a homogeneous object would be (c), however, stochastic gradient tends to modify single voxels as in (b).

$$J_{\Delta}(\mathbf{t}, \boldsymbol{\mu}) = J_{seg}(\mathbf{t}, \boldsymbol{\mu}) + \lambda_{\Delta} \frac{1 - edge(I_{\mathbf{t}})}{2} ||\Delta \mathbf{t}||_{fro}^{2} + \lambda_{\Delta} \frac{1 - edge(I_{\boldsymbol{\mu}})}{2} ||\Delta \boldsymbol{\mu}||_{fro}^{2}$$
(6)

where Δt is the sum of the absolute values of the convolutions of \boldsymbol{t} by the matrices

-1	2	-1]		-1	-4	-1	
-4	8	-4	and	2	8	2	
-1	2	-1		-1	-4	-1	

and $\Delta \mu$ is defined similarly. The motivation for this regulariser is illustrated in Fig. 2. Finally, we note that this regulariser, using a second derivative, is resolution dependent. The finer the resolution, the higher the value of λ_{Δ} should be.

4.4 Continuity constraints and the complete regulariser

As we switch off smoothness considerations along edges, it is possible for sharp peaks to arise along them. The smoothness regulariser (6) then causes this error to propagate to adjacent pixels, polluting the reconstructed model. To help prevent this, we introduce a further regulariser based on the C^0 norm, which applies to all voxels. Its aim is to penalize values that are significantly higher or lower than their adjacent values.

We convolve the thickness and alloy images with the left and right first order differences filters, [0, -1, 1] and [-1, 1, 0] to take a measure of continuity along the *x* dimension, and measure continuity in the *y* dimension similarly. Let Δ_h denote the maximum of the absolute values of the responses of these filters. We define our final regulariser:

$$J_{reg}(\mathbf{t},\boldsymbol{\mu}) = J_{\Delta}(\mathbf{t},\boldsymbol{\mu}) + \lambda_{C^0} ||\Delta_h \mathbf{t}||_{fro} + \lambda_{C^0} ||\Delta_h \boldsymbol{\mu}||_{fro}$$
(7)

where λ_{C^0} is a constant.

5 MULTISCALE

Following a fundamental paradigm in multigrid community [36], our strategy is to first obtain a reasonable guess over the large, smooth areas, and then concentrate on the regions of material transitions, where errors are effectively eliminated by fine resolution processing. At each resolution, we work with a pre-computed database of radially symmetric scatter kernels. Specifically, from coarse to fine, we work at resolutions DS89, DS45, DS23, and DS11, where DS89 means that one voxel of the thickness/alloy model corresponds to 89×89 pixels of the input X-ray.

Different to classic multigrid, we introduce resolution switch criteria which change the resolution depending on the *improvement factor* of the algorithm. The cost function J is used to compute the improvement factor γ between iterations n and n + p

$$\gamma = \frac{J_{reg}(\mathbf{t}, \boldsymbol{\mu})^{(n+p)}}{J_{reg}(\mathbf{t}, \boldsymbol{\mu})^{(n)}}$$
(8)

where p depends on the number of scattering voxels of the model, and it is a pre-defined constant on each resolution. Acknowledging that non-linear optimisation problems are vulnerable to overfitting and thus require resolution switches in both directions of the coarse to fine hierarchy, we distinguish between four cases:

- (1) $\gamma < 0.8$: significant improvement is being made on the current resolution. We expect that further updates would reduce *J* further, so we continue on the present resolution.
- (2) 0.8 ≤ γ < 0.95: some improvement on the current resolution, but it is slow. We assume that there is a significant amount of high frequency error which cannot be easily eliminated at this resolution. We trigger a V-cycle, that is, move to the the next coarser level and then back again.
- (3) 0.95 ≤ γ < 1.05: no significant improvement has been made. We switch to the next finer level.
- (4) 1.05 ≤ γ: a significant deterioration has occurred. We assume that is is due to inconsistencies between coarse estimations of the object and the object's true shape. We blacklist the coarsest allowed resolution, meaning that it will no-longer be reached in future V-cycles, and carry out a V-cycle to reset the recent deterioration.

6 **RESULTS**

In section 6.1, we present results from the proposed optimisation algorithm. The input images were X-rays of limb phantoms made of PMMA plastic and aluminium rods, and X-rays of a human cadaver. In Section 6.2, we present the results of an ablation study with the regularisers turned off.

6.1 Results

Based on the X-rays of four limb phantoms, Fig. 3 shows estimates of the overall thickness, and estimates of the bone thickness computed as the product of the overall thickness by the alloy values.

Fig. 4 shows the results of the algorithm on cadaver images of a wrist and a shoulder. Compared to the limb phantoms, these are more complex anatomies, with many smaller, overlapping bones. We note that in the wrist model, the greater volume of bone at the joint between hand and arm was successfully identified, as well as the increase in bone thickness where the ulna and radius overlap near the elbow. The overall accuracy of the thickness estimates is unverifiable, but in both cases the broad structure of the estimated models is visually appropriate.

6.2 Ablation study on the use of regularisers

The ablation study was performed on an X-ray image of the simple limp phantom (SLP), which has the shape of a partially-rounded cuboid, with a single aluminium rod running down the centre. It has the same thickness t = 5cm everywhere, except the far left and right edges, and in the middle it has a value of $\mu = 0.4$, that is, 40% of its depth is aluminium. As we explicitly know the t and μ ground truth values, we can compare the ground truth cost function *J* and compare it with the *J* values during optimisation.

We use removed all regularisers and used three setups: a systematically bad initial guess, where all thicknesses are set to 1% of the maximum supported thickness, while the alloy is 50% everywhere; a randomised initial guess; a good initial estimate. The latter uses initial thickness value of t = 5*cm* wherever the segmentation identifies non-open-beam voxels, and μ = 0.5 wherever the segmentation identifies some bone.

Fig. 5) (left and middle) shows that bad initial guesses were not compensated by the gradient descent over the first 80 updates. Moreover, as expected, the impact of a bad initialisations correlates with the image resolution. That is, the gradient descent's capability vanishes for higher image resolutions.

We note that a good initial guess could be directly provided by the segmentation, or be implicitly realised via penalty terms added to the objective function J. The main observation here is that the design of the regularisers should aim at pushing the solution into a state that agrees with the outcome of the image segmentation.

We also note that while, in line with our previous discussion, a $\theta = 0.5$ outperforms $\theta = 1$, nevertheless, all descents stagnate in the higher resolutions, or even start to increase *J* again after a certain iteration count. The unregularised gradient descent is effective only for a few updates and if and only if we work in a coarse resolution. Furthermore, the gradient descent's outcome is off from the optimal solution (ground truth) by almost a factor of eight; in the best case.

7 CONCLUSIONS

We developed a cost function for the ill-posed inverse problem of thickness and material estimation in single-scan X-ray images, where the forward function is a physically realistic particle simulator. We use a neural network for image segmentation and regularization to feed-in domain knowledge. Moreover, we developed a multiscale version of the algorithm to obtain better solutions and speed-up the process.

In this paper, we demonstrated the effectiveness of our method with relatively coarse resolution voxel models. Naturally, a next step would be to expand to higher resolutions. A major constraint for this is the required calibration between the Geant4 particle

Henry Westmacott, Ioannis Ivrissimtzis, and Tobias Weinzierl



Figure 3: Surface plots of the final thickness and material estimations for four limb-analogue phantoms.

simulator and the X-ray hardware system, which should be done for each resolution separately; one cannot just construct resolutions on the fly.

A second promising direction of future work is the use of better, machine learning informed, priors to start the optimisation process. At the moment, we only use the segmentation, and in the absence of other information, we begin from a state that is perfectly flat, with the bone and soft-tissue regions assigned the same thickness, and just different alloy values. Recent research, [2], has shown how advances in the field of depth estimation from single images can be used to enhance algorithmic performance in downstream tasks. Thus, we can imagine the use of a second neural network trained to predict better thickness priors, enabling our algorithm to converge faster.

Finally, we would also like to test the system on a wider range of X-ray hardware systems, even though we do not expect significant issues, since the algorithm is constructed around a system-agnostic simulation algorithm, courtesy of IBEX Innovations.

ACKNOWLEDGMENTS

We thank IBEX Innovations Ltd for their assistance in providing the development source code for their TrueView[®] scatter correction algorithm.

ICIGP 2024, January 19-21, 2024, Beijing, China



Figure 4: Overall thickness (left) and bone thickness (middle), estimated from a single X-ray image (right).



Figure 5: No regularisation used. Three different initialisations: uniformly wrong initial estimate (left), random initial choice (middle), good initial estimate (right). Different colours indicate different resolutions, and continuous or dashed lines correspond to different θ .

REFERENCES

- Jonas Adler and Ozan Öktem. 2017. Solving ill-posed inverse problems using iterative deep neural networks. *Inverse Problems* 33, 12 (2017), 124007.
- [2] Yatharth Ahuja, Tushar Nitharwal, Utkarsh Sundriyal, Sreedevi Indu, and Anup K Mandpura. 2022. Depth Reconstruction Based Visual SLAM Using ORB Feature Extraction. Journal of Image and Graphics 10, 4 (2022), 172–177.
- [3] Simon Arridge, Peter Maass, Ozan Öktem, and Carola-Bibiane Schönlieb. 2019. Solving inverse problems using data-driven models. Acta Numerica 28 (2019), 1–174.
- [4] Daniel Otero Baguer, Johannes Leuschner, and Maximilian Schmidt. 2020. Computed tomography reconstruction using deep image prior and learned reconstruction methods. *Inverse Problems* 36, 9 (2020), 094004.
- [5] Anatoly B Bakushinsky, Mihail Yu Kokurin, and Alexandra Smirnova. 2010. Iterative methods for ill-posed problems. de Gruyter.
- [6] Mohsin Bilal and Muhammad Arif. 2019. Multivariate Regression CNN for illposed Inverse Reconstruction of Satellite Images. *Proceedia Computer Science* 163 (2019), 312–318. 16th Learning and Technology Conference 2019Artificial Intelligence and Machine Learning: Embedding the Intelligence.
- [7] Jasleen Birdi, Audrey Repetti, and Yves Wiaux. 2017. A regularized tri-linear approach for optical interferometric imaging. Monthly Notices of the Royal

ICIGP 2024, January 19-21, 2024, Beijing, China

Henry Westmacott, Ioannis Ivrissimtzis, and Tobias Weinzierl

Astronomical Society 468, 1 (Feb 2017), 1142-1155.

- [8] Joseph Bullock, Carolina Cuesta-Lázaro, and Arnau Quera-Bofarull. 2019. XNet: a convolutional neural network (CNN) implementation for medical x-ray image segmentation suitable for small datasets. In Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging, Vol. 10953. SPIE, 453–463.
- [9] S. Delorme, Y. Petit, J. A. de Guise, H. Labelle, C. . Aubin, and J. Dansereau. 2003. Assessment of the 3-D reconstruction and high-resolution geometrical modeling of the human skeletal trunk from 2-D radiographic images. *IEEE Transactions on Biomedical Engineering* 50, 8 (2003), 989–998.
- [10] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. 2015. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern* analysis and machine intelligence 38, 2 (2015), 295–307.
- [11] J. Allison et al. 2006. Geant4 developments and applications. IEEE Transactions on Nuclear Science 53, 1 (2006), 270–278.
- [12] J. Allison et al. 2016. Recent developments in Geant4. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 835 (2016), 186–225.
- [13] S. Agostinelli et al. 2003. Geant4-a simulation toolkit. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 506, 3 (2003), 250-303.
- [14] Davis Gilton, Gregory Ongie, and Rebecca Willett. 2021. Model adaptation for inverse problems in imaging. IEEE Transactions on Computational Imaging 7 (2021), 661–674.
- [15] Yingying Gu, Jun Zhang, and Ping Xue. 2018. Multi-grid nonlocal techniques for x-ray scatter correction. In *Medical Imaging 2018: Image Processing*, Elsa D. Angelini and Bennett A. Landman (Eds.), Vol. 10574. International Society for Optics and Photonics, SPIE, 427 – 433.
- [16] Jingnan Guo, Saša Banjac, Lennart Röstel, Jan C Terasa, Konstantin Herbst, Bernd Heber, and Robert F Wimmer-Schweingruber. 2019. Implementation and validation of the GEANT4/AtRIS code to model the radiation environment at Mars. Journal of Space Weather and Space Climate 9 (2019), A2.
- [17] Normi Abdul Hadi, Arsmah Ibrahim, Fatimah Yahya, and Jamaludin Md Ali. 2013. Centroid based on branching contour matching for 3D reconstruction using Beta-spline. *Journal of Image and Graphics* 1, 3 (2013), 138–142.
- [18] Andreas Hauptmann, Jonas Adler, Simon Arridge, and Ozan Öktem. 2020. Multiscale learned iterative reconstruction. *IEEE transactions on computational imaging* 6 (2020), 843–856.
- [19] Hideo Hirayama, Yoshihito Namito, Walter R Nelson, Alex F Bielajew, Scott J Wilderman, and U Michigan. 2005. *The EGS5 code system*. Technical Report. United States. Department of Energy.
- [20] Sebastien Incerti, Michael Douglass, Scott Penfold, Susanna Guatelli, and Eva Bezak. 2016. Review of Geant4-DNA applications for micro and nanoscale simulations. *Physica Medica* 32, 10 (2016), 1187–1200.
- [21] Prateek Jain, Purushottam Kar, et al. 2017. Non-convex optimization for machine learning. Foundations and Trends® in Machine Learning 10, 3-4 (2017), 142–363.
- [22] Sergey I Kabanikhin. 2011. Inverse and ill-posed problems. de Gruyter.
- [23] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. 2021. Physics-informed machine learning. *Nature Reviews Physics* 3, 6 (2021), 422–440.
- [24] Yoni Kasten, Daniel Doktofsky, and Ilya Kovler. 2020. End-to-end convolutional neural network for 3D reconstruction of knee bones from bi-planar X-ray images. In International Workshop on Machine Learning for Medical Image Reconstruction. Springer, 123–133.
- [25] Dari Kimanius, Gustav Zickert, Takanori Nakane, Jonas Adler, Sebastian Lunz, C-B Schönlieb, Ozan Öktem, and Sjors HW Scheres. 2021. Exploiting prior knowledge about biological macromolecules in cryo-EM structure determination. *IUCrJ* 8, 1 (2021), 60–75.
- [26] HN Latha and Rajiv R Sahay. 2020. Joint estimation of depth map and focus image in SFF: an optimization framework by sparsity approach. In 2020 International Conference on Communication and Signal Processing (ICCSP). IEEE, 0656–0660.
- [27] Mikhail M Lavrent'ev, Sergei M Zerkal, and Oleg E Trofimov. 2014. Computer modelling in tomography and ill-posed problems. De Gruyter.
- [28] Housen Li, Johannes Schwab, Stephan Antholzer, and Markus Haltmeier. 2020. NETT: Solving inverse problems with deep neural networks. *Inverse Problems* 36, 6 (2020), 065005.
- [29] Sebastian Lunz, Ozan Öktem, and Carola-Bibiane Schönlieb. 2019. Adversarial Regularizers in Inverse Problems. arXiv:1805.11572 [cs.CV]
- [30] Andreas K Maier, Christopher Syben, Bernhard Stimpel, Tobias Würfl, Mathis Hoffmann, Frank Schebesch, Weilin Fu, Leonid Mill, Lasse Kling, and Silke Christiansen. 2019. Learning with known operators reduces maximum error bounds. *Nature machine intelligence* 1, 8 (2019), 373–380.
- [31] Achille Mileto, Luis S Guimaraes, Cynthia H McCollough, Joel G Fletcher, and Lifeng Yu. 2019. State of the art in abdominal CT: the limits of iterative reconstruction algorithms. *Radiology* 293, 3 (2019), 491–503.
- [32] Francesc Salvat, José M Fernández-Varea, and Josep Sempau. 2006. PENELOPE-2006: A code system for Monte Carlo simulation of electron and photon transport. In PENELOPE workshop proceedings. OECD Publishing, 7.

- [33] Johannes Schwab, Stephan Antholzer, and Markus Haltmeier. 2019. Big in Japan: Regularizing networks for solving inverse problems. *Journal of Mathematical Imaging and Vision* 62, 3 (2019), 445–455.
- [34] Pedro M Silva and Joao B Florindo. 2019. Using down-sampling for multiscale analysis of texture images. *Pattern Recognition Letters* 125 (2019), 411–417.
- [35] Manoj Kumar Singh. 2013. Image reconstruction and edge detection based upon neural approximation characteristics. *Journal of Image and Graphics* 1, 1 (2013), 12–16.
- [36] Klaus Stüben. 2000. Multigrid. Academic Press, Chapter Appendix A, 413–532.
- [37] U Trottenberg. 2001. Multigrid. Academic Press, San Diego, Calif. ; London.
 [38] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. 2018. Deep image prior. In Proc. CVPR. IEEE, 9446-9454.
- [39] Tobias Würfl, Mathis Hoffmann, Vincent Christlein, Katharina Breininger, Yixin Huang, Mathias Unberath, and Andreas K Maier. 2018. Deep learning computed tomography: Learning projection-domain weights from image domain in limited angle problems. *IEEE transactions on medical imaging* 37, 6 (2018), 1454–1463.
- [40] Yibo Yang and Paris Perdikaris. 2018. Physics-informed deep generative models. arXiv:1812.03511 [stat.ML]
- [41] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. 2017. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Transactions on Image Processing* 26, 7 (Jul 2017), 3142–3155.