

The *Mokume* Dataset and Inverse Modeling of Solid Wood Textures

MARIA LARSSON, The University of Tokyo, Japan

HODAKA YAMAGUCHI, Gifu Pref. Research Institute for Human Life Technology, Japan and Nihon University, Japan

EH SAN PAJOUHESHGAR, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

I-CHAO SHEN, The University of Tokyo, Japan

KENJI TOJO, The University of Tokyo, Japan

CHIA-MING CHANG, The University of Tokyo, Japan

LARS HANSSON, Luleå University of Technology, Sweden and Norwegian University of Science and Technology, Norway

OLOF BROMAN, Luleå University of Technology, Sweden

TAKASHI IJIRI, Shibaura Institute of Technology, Japan

ARIEL SHAMIR, Reichman University, Israel

WENZEL JAKOB, The University of Tokyo, Japan and École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

TAKEO IGARASHI, The University of Tokyo, Japan

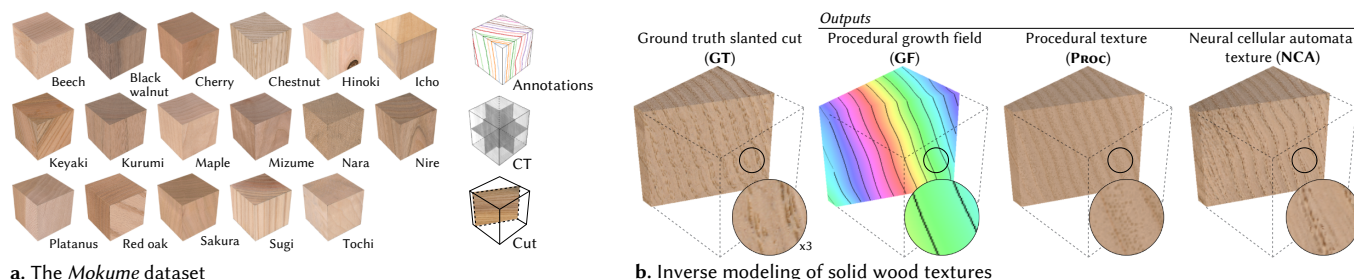


Fig. 1. **a.** Overview of the diverse wood species covered by the *Mokume* dataset. **b.** This data is used for training and evaluation of an inverse modeling pipeline for solid wood textures. Our method first evaluates a neural model that converts the exterior photographs into 2D annotations of the annual ring pattern. We then compute a compatible procedural *growth field* (GF) that assigns a time to every 3D position, denoting when the associated material was added during the tree's growth. The annual-rings are iso-curves/surfaces of this growth field. We showcase two ways to transform this representation into a detailed 3D texture: an efficient inverse procedural model (Proc) with support for point-wise evaluation, and a versatile but resource-intensive neural cellular automaton (NCA).

We present the *Mokume* dataset for solid wood texturing consisting of 190 cube-shaped samples of various hard and softwood species documented by

Authors' addresses: Maria Larsson, ma.k.larsson@gmail.com, The University of Tokyo, Japan; Hodaka Yamaguchi, yamaguchi-hodaka@rd.pref.gifu.jp, Gifu Pref. Research Institute for Human Life Technology, Japan and Nihon University, Japan; Ehsan Pajouheshgar, ehsan.pajouheshgar@epfl.ch, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland; I-Chao Shen, jdilyshen@gmail.com, The University of Tokyo, Japan; Kenji Tojo, knjtojo@g.ecc.u-tokyo.ac.jp, The University of Tokyo, Japan; Chia-Ming Chang, info@chiamingchang.com, The University of Tokyo, Japan; Lars Hansson, lars.hansson@ltu.se, Luleå University of Technology, Sweden and Norwegian University of Science and Technology, Norway; Olof Broman, olof.broman@ltu.se, Luleå University of Technology, Sweden; Takashi Ijiri, takashi.ijiri@riken.jp, Shibaura Institute of Technology, Japan; Ariel Shamir, arik@idc.ac.il, Reichman University, Israel; Wenzel Jakob, wenzel.jakob@epfl.ch, The University of Tokyo, Japan and École Polytechnique Fédérale de Lausanne (EPFL), Switzerland; Takeo Igarashi, takeo@acm.org, The University of Tokyo, Japan.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM 0730-0301/2025/8-ART

<https://doi.org/10.1145/3730874>

high-resolution exterior photographs, annual ring annotations, and volumetric computed tomography (CT) scans. A subset of samples further includes photographs along slanted cuts through the cube for validation purposes.

Using this dataset, we propose a three-stage inverse modeling pipeline to infer solid wood textures using only exterior photographs. Our method begins by evaluating a neural model to localize year rings on the cube face photographs. We then extend these exterior 2D observations into a globally consistent 3D representation by optimizing a procedural growth field using a novel iso-contour loss. Finally, we synthesize a detailed volumetric color texture from the growth field. For this last step, we propose two methods with different efficiency and quality characteristics: a fast inverse procedural texture method, and a neural cellular automaton (NCA). We demonstrate the synergy between the *Mokume* dataset and the proposed algorithms through comprehensive comparisons with unseen captured data. We also present experiments demonstrating the efficiency of our pipeline's components against ablations and baselines. Our code, the dataset, and reconstructions are available via <https://mokumeproject.github.io/>.

CCS Concepts: • **Computing methodologies** → **Volumetric models**.

Additional Key Words and Phrases: procedural texturing, neural cellular automaton

ACM Reference Format:

Maria Larsson, Hodaka Yamaguchi, Ehsan Pajouheshgar, I-Chao Shen, Kenji Tojo, Chia-Ming Chang, Lars Hansson, Olof Broman, Takashi Ijiri, Ariel

Shamir, Wenzel Jakob, and Takeo Igarashi. 2025. The *Mokume* Dataset and Inverse Modeling of Solid Wood Textures. *ACM Trans. Graph.* 44, 4 (August 2025), 18 pages. <https://doi.org/10.1145/3730874>

1 INTRODUCTION

Wood is an ubiquitous material in real and virtual environments. Natural wood surfaces exhibit mesmerizing textural detail and color variation that impart a sense of warmth and comfort to humans. In this work, we seek to characterize the space of textures originating from the abundance of tree species and the complex relationship between appearance and the angle and position of cuts in wood.

Previous wood datasets imaged localized 2D regions for species identification and tree-ring dating [De Blaere et al. 2023; Fabijańska et al. 2017]. In contrast, the ability to map wood texture onto objects without causing distortions depends on 3D *solid* wood representations that assign an albedo value to positions in space. Comprehensive datasets of this nature are not available, hence prior wood texturing methods have relied on few examples, often unique to each study [Larsson et al. 2022; Liu et al. 2016; Marschner et al. 2005]. We present the *Mokume* dataset to provide this critical ingredient.

Mokume, named after the Japanese word for wood grain, comprises 190 physical cube samples with a side length of 4.0 cm. The samples were taken from 17 different hard- and softwood species, and exhibit diverse anatomical patterns, imperfections, and cut-out placements. Each sample is documented by external surface photographs with annotations of the annual rings, and volumetric computed tomography (CT) scans. A subset of 38 samples were cut to photograph an internal surface for validation purposes. The *Mokume* dataset can support data-driven approaches and enables comprehensive evaluations and robust method comparisons.

Building on this new dataset, we further propose and evaluate an inverse modeling pipeline that infers solid wood textures based on exterior photographs. We address this challenge in three stages: the first localizes 2D annual rings in each input photograph. The resulting information is used in the second step to construct a continuous procedural *growth field* (GF) that maps 3D positions to a scalar value representing the time when the associated material was added during the tree’s growth process. This field roughly encodes the distance from the tree’s centerline (“pith”) and further accounts for distortions due to uneven growth. The annual rings are iso-surfaces of this field. The final step uses the growth field to synthesize solid wood textures.

Unlike previous methods for inverse wood texturing that make assumptions about the location of the pith axis [Lefebvre and Poulin 2000; Nindel et al. 2023], or require manually traced numbered annual rings as inputs [Larsson et al. 2024], our method can directly and robustly infer the growth field from external photographs and annotations of knot centers (if present). We trained a U-Net on a portion of the *Mokume* dataset to extract annual rings from the external photographs, to which we fit a volumetric procedural growth field. In addition, unlike forward procedural wood textures that use noise to create distortions [Liu et al. 2016], our model is parameterized by the tree’s varying radial growth speed, giving sufficient control to enable fitting the model to reference data. Moreover, we propose a differentiable *iso-contour loss* that measures the deviation of growth

field values along extracted target annual rings, which is effective for inferring the pith axis and field distortions.

With the growth field at hand, we showcase two approaches to synthesize a detailed solid wood texture: an extended procedural model that further accounts for detailed features such as pores and rays, and a neural cellular automaton (NCA) that learns an update rule to iteratively create the desired structure [Mordvintsev et al. 2020]. We optimize both models using style-based losses to produce detailed and realistic textures. The two models occupy opposite quadrants of the quality/efficiency design space and serve to demonstrate the varied applications of the *Mokume* dataset: the procedural model is memory-efficient, editable, and requires a single pass, while the NCA is highly parameter-efficient and can reproduce fine details if significant computation and memory usage are acceptable.

We present qualitative results by reconstructing solid wood textures for each of the 38 samples in our test portion of the dataset, comparing the inferred texture of both models to the slanted cuts through the physical cubes. We further evaluate specific components of our framework in a series of quantitative experiments and ablations.

2 RELATED WORK

This section discusses prior work on datasets (Section 2.1), texturing (Section 2.2), and wood modeling and analysis (Section 2.3).

2.1 Datasets

2.1.1 Material appearance. Most prior appearance datasets were curated to support (SV-)BRDF reconstruction of diverse material types such as leather and cloth [Aittala et al. 2015; Henzler et al. 2021; Nielsen et al. 2015]. Marschner et al. [2005] fit a specialized wood SVBRDF to a dataset comprising five samples. In contrast, our study focuses on volumetric structure and color, leaving 5D spatio-directional reflectance modeling for future work.

2.1.2 Wood analysis. Existing wood datasets cater to applications such as species identification, abnormality detection, and dendrochronological measurement (tree-ring dating). Datasets for species identification typically include classified optical scans of small patches of end-grain surfaces [De Blaere et al. 2023], tree bark [Warner et al. 2024], or CT scans [Kobayashi et al. 2019]. Another dataset was developed to analyze the effect of surface preparation (the grit of sand paper and type of saw) [Ravindran et al. 2023]. Datasets for abnormality detection are typically made up of labeled images of different surface defects of solid wood or veneer, such as knots and cracks [Kodytek et al. 2022; Shi et al. 2020b]. Finally, dendrochronological datasets provide strip images in the outward direction from the pith on cross-section of a stem, often including annual ring annotations [Fabijańska et al. 2017; Wu et al. 2023].

Unlike existing datasets, we provide comprehensive cross-sectional coverage, capturing all sides of samples taken in various orientations relative to the pith.

2.2 Texture Modeling

2.2.1 Procedural texturing. Procedural texturing, first introduced by Peachey [1985] and Perlin [1985], involves generating a material texture through an algorithmic process rather than referring to

explicitly stored data. This approach yields a compact and editable representation capable of producing variations at a low computational cost, and remains a widely used technique. Given these advantages, procedural texturing is particularly suitable for solid textures that exhibit high natural variation, such as wood. The functions for generating various wood features are well understood, owing to Liu et al. [2016], who proposed a comprehensive set of forward functions for generating pores, rays, and light reflections from fiber directions, while Larsson et al. [2022] proposed a function for knots. In this paper, we target the inverse problem.

2.2.2 Inverse modeling. Inverse modeling aims to automatically discover optimal parameters so that a generated output matches a given target. In computer graphics, this technique has been widely explored; for example, in inverse differentiable rendering [Kato et al. 2020; Li et al. 2018], material or scene parameters are optimized through gradients of the rendering function, enabling tasks such as volumetric material capturing [Gkioulekas et al. 2013], caustic design [Nimier-David et al. 2019], color reproduction [Nindel et al. 2021], and geometric optimization [Liu et al. 2018; Nicolet et al. 2021].

Node-graph representations of procedural shaders, have advanced significantly, evolving from classification or regression for predicting node graph parameters [Hu et al. 2019], later extended to differentiable or semi-differentiable pipelines for material learning [Hu et al. 2022a,b; Li et al. 2023; Shi et al. 2020a], and reinforcement learning for parameter prediction [Li et al. 2024]. Beyond parameter fitting, other systems enable the generation of node graphs [Guerrero et al. 2022; Li et al. 2025] and prompt-guided graphs searches [Hu et al. 2023], achieving near-photorealistic reproduction of diverse materials.

2.2.3 Solid texture synthesis. Creating a solid texture from a 2D reference image is a classical problem [Kopf et al. 2007; Kwatra et al. 2005]. Recent work has extended solid texturing approaches to support inverse modeling through gradient-based learning. Henzler et al. [2020] train a coordinate-to-color mapping MLP by feeding multi-frequency noise as input, and Portenier et al. [2020] inject noise into hidden layers for richer outputs. Likewise, Gutierrez et al. [2020] use a 3D CNN that transforms random noise into volumetric textures. However, while these methods focus on synthesizing visually plausible 3D textures from a 2D reference, they do not incorporate wood-specific growth patterns or aim for a faithful volumetric reconstruction. In contrast, our approach leverages wood anatomical priors, such as the pith axis and annual ring distortions, to recover the interior of a real wood sample from its external faces.

2.2.4 Neural cellular automata. Neural cellular automata (NCA), proposed by Mordvintsev et al. [2020], are inspired by reaction-diffusion (RD) systems and classical cellular automata (CA), adapting their local update principle while substituting the traditional hand-crafted rule with a small neural network. The model is unrolled for multiple iterations, and trained end-to-end using backpropagation through time and a style-based loss to synthesize a reference texture. This paradigm is effective for producing self-organizing textures [Niklasson et al. 2021] and has been extended for generating dynamic textures [Pajouheshgar et al. 2023] or texturing 3D meshes

directly [Pajouheshgar et al. 2024b], showcasing the flexibility of NCAs in producing complex patterns from simple local rules. This study demonstrates that a natural extension of NCAs to the 3D volumetric setting, combined with our proposed wood-specific priors, is a highly effective model for synthesizing realistic and faithful solid wood textures.

2.3 Wood Modeling and Analysis

2.3.1 Inverse modeling. Some previous studies tackle the backward procedural wood texturing problem specifically, and show a few successful outputs [Guo et al. 2020; Lefebvre and Poulin 2000; Nindel et al. 2023]. However, none of these handles knots or texture details (pores, rays). Moreover, Nindel et al. [2023] and Lefebvre and Poulin [2000] make strict assumptions about the pith axis location, determining its location analytically by identifying the largest orthogonal distance between annual rings, which holds only for a subset of cutout poses. Finally, Nindel et al. [2023] extracts contours as a first step, using curved Gabor filters, which, while not requiring training data, is dependent on the image gradients and therefore not robust against features superimposed over the annual rings, such as rays.

Larsson et al. [2024] proposed a learning-based method for inferring a (non-procedural) volumetric texture based on external surface information. Although the problem setup is analogous to ours, there are key technical differences. First, while they require manually tracing numbered contours on the input photographs, we circumvent this requirement by automatic extraction of contours. Second, for global structure inference, they train a model on procedurally generated data (without knots), and optimize the parameters of the trained model, creating a neural growth field. In contrast, we optimize the parameters of a procedural growth field directly, owing to our proposed iso-contour loss. Third, for colored volume synthesis, they independently apply 2D style-transfer on each volume slice, which does not guarantee 3D consistency. Our texture synthesis approaches (procedural texturing and NCA) are fully volumetric/3D.

2.3.2 Growth field. A growth field representation of wood material analogous to ours was used in previous works, modeled procedurally [Larsson et al. 2022], using a neural model [Larsson et al. 2024], or by evolution over numerous time-steps [Kratt et al. 2015; Mann et al. 2006; Sellier et al. 2011].

2.3.3 Pith detection. Pith detection is a fundamental task in wood engineering, essential for modeling wood growth patterns to evaluate mechanical properties and optimize material usage. This typically involves estimating the pith point in a 2D image captured in a plane roughly perpendicular to the pith axis, whether a photograph of the end of a log [Kurdthongmee and Suwannarat 2019; Norell and Borgefors 2008; Schraml and Uhl 2013] or a CT image of a radial cross-section [Boukadida et al. 2012; Gazo et al. 2020; Longuetaud et al. 2004]. In contrast, we estimate the 3D axis of the pith based, and do so based on images of external surfaces facing in various directions.

3 WOOD ANATOMY AND TERMINOLOGY

This section presents key terminology to explain and classify key features of wood textures (Figure 2). The most basic wood material geometries and features are as follows:

- **Pith** is the innermost center skeleton of a tree.
- **Annual rings** are formed by alternating layers of early- and latewood. Earlywood, which grows in spring, appears lighter, while latewood, which forms later in the year, appears darker. This alteration creates a pattern of concentric circles in the radial views and parallel bands in the tangential plane.
- **Pores**, also known as vessels, are tube-like structures extending vertically. Pores appear as tiny circular openings in radial view, and elongated openings in the tangential plane.
- **Rays** are thin, ribbon-like structures extending radially from the wood's pith. Rays are visible as fine lines in a radial view and patch-like dots in the tangential plane.

We identify five categories of wood based on the visibility and arrangement of small-scale features (see examples in Figure 4). The abbreviations used below are as follows: SW for softwood, HW for hardwood, DP for diffuse-porous, RP for ring-porous, and R for rays:

- **Softwoods (SW)**. The annual rings are typically distinct from the early- and late-wood color contrast, while both pores and rays are indistinct.
- **Diffuse-porous hardwoods (HW-DP)**. Pores are distributed relatively evenly while the annual rings are distinct from the early-to-late-wood color contrast like in softwood.
- **Diffuse-porous hardwoods with rays (HW-DP-R)**. Same as previous but with distinct rays.
- **Ring-porous hardwoods (HW-RP)**. The annual rings are distinct from the pores that follow them, rather than by early-to-latewood color contrast.
- **Ring-porous hardwoods with rays (HW-RP-R)**. Same as previous but with distinct rays.

Finally, there are non-repetitive features that we term *occasional features*, and which are visible in some cut-outs (refer to Figure 5).

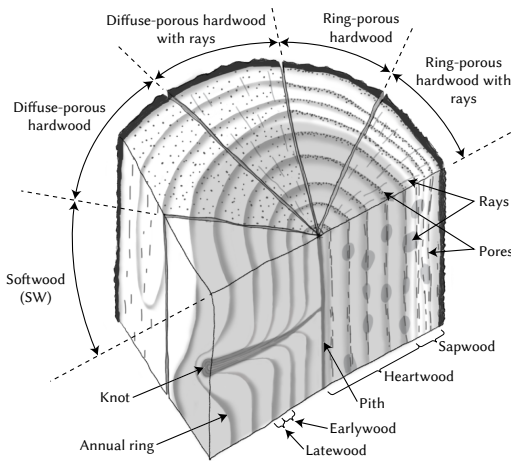


Fig. 2. Terminology of wood features.

- **Knots**. Darker spots formed where branches once extended.
- **Cracks**. Splits or fissures that typically occur during drying.
- **Insect holes**. Small openings caused by boring insects.
- **Gum**. Resinous deposits that seep from the wood.
- **Heart-to-sapwood color transition**. Heartwood is the older, inner core of the tree that is often darker. Sapwood is the outer, often lighter-colored portion of the tree. We list the heart-to-sapwood color transition under occasional features because it is visible only in some cut-outs (those placed at the border).

4 THE MOKUME DATASET

This section introduces the content of our dataset: the characteristics of the physical samples (Section 4.1) and how they were collected (Section 4.2) and documented (Section 4.3). For implementation details, refer to Appendix A.1.

4.1 Physical Samples

The *Mokume* dataset comprises 190 physical wood cubes with a side length of 4.0 cm, which are diverse in terms of species and trees, anatomical features, and cut-out poses.

4.1.1 Species and trees. The samples come from 17 different hard- and softwood species (Table 1). There are 10 samples per species except for the softwoods *Sugi* and *Hinoki*, both types of cypress, for which we included 20 samples because of the frequent presence of knots, which adds significant variety. To capture diversity within each species, we collected the samples from different trees, obtaining 1–10 samples from a total of 91 unique trees (Table 1).

4.1.2 Anatomical features. Species from five categories of small-scale features combinations (see Section 3) are represented in our dataset (Figure 3-left, Figure 4). Over a third of the samples (74 of 190) have at least one occasional feature: insect holes, visible

Table 1. **Species in the Mokume dataset.** N_S is the number of cube samples. N_T is the number unique trees from which the samples were cut.

Code	Name	Scientific Name	N_S	N_T
B	Beech	<i>Fagus spp.</i>	10	10
BW	Black walnut	<i>Juglans nigra</i>	10	10
CH	Cherry	<i>Prunus serotina</i>	10	4
CN	Kuri	<i>Castanea crenata</i>	10	5
H	Hinoki	<i>Chamaecyparis obtusa</i>	20	2
IC	Icho	<i>Ginkgo biloba</i>	10	2
K	Keyaki	<i>Zelkova serrata</i>	10	10
KR	Kurumi	<i>Juglans mandshurica</i>	10	4
MP	Maple	<i>Acer spp.</i>	10	5
MZ	Mizume	<i>Betula grossa</i>	10	5
N	Nara	<i>Quercus crispula</i>	10	10
NR	Nire	<i>Ulmus davidiana</i>	10	7
P	Platanus	<i>Platanus occidentalis</i>	10	1
RO	Red oak	<i>Quercus rubra</i>	10	8
S	Sakura	<i>Prunus spp.</i>	10	4
SG	Sugi	<i>Cryptomeria japonica</i>	20	2
TC	Tochinoki	<i>Aesculus turbinata</i>	10	2
			190	91

heart-to-sapwood transitions, cracks, knots, or gum (Figure 3-right, Figure 5).

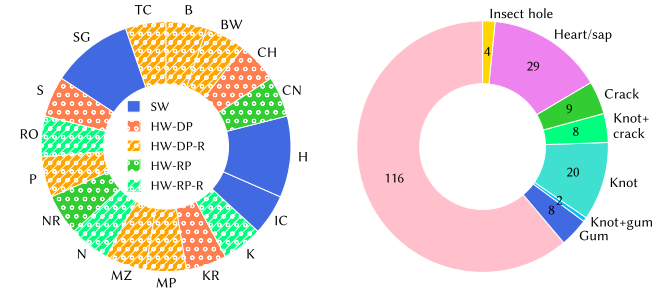


Fig. 3. **Left:** regular pattern categories of each species in the dataset. Pores are classified according to Murayama and Murayama [2020]. Ray classifications are based on our observations. For category codes, refer to Section 3. **Right:** occasional feature counts based on our observations.

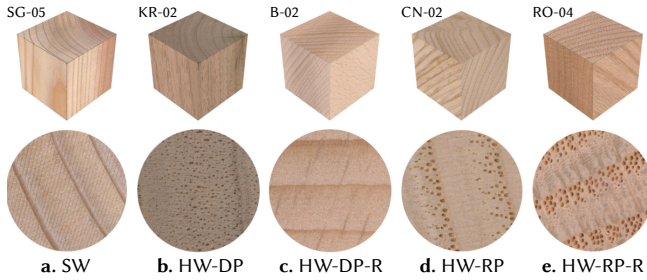


Fig. 4. Examples of regular pattern categories. Note that the pores of B-02 are small and just barely visible.

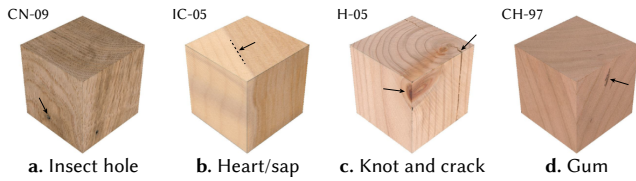


Fig. 5. Examples of occasional features.

4.1.3 Cut-out placements. The anisotropic nature and 3D structure of wood cause surface patterns to vary based on the cut-out location and orientation in relation to the tree. For instance, nearer to the pith (where the growth starts), the annual rings have stronger curvatures.

To capture these types of diversities, we varied the cut-out *locations*, that is, closer or further from the pith of the tree, and *orientations*, that is, straight or angled in relation to the pith (Figure 6). We estimate the actual cutout locations and orientations by running our pith inference optimization (refer to Section 5.2) on the full dataset based on the annotated annual rings. This analysis reveals many samples are relatively straight and positioned at a 5–10 cm distance from the pith. However, cut-out poses vary significantly

with orientations ranging from 0 to 53 degrees and pith distances ranging from 1.5 to 33 cm. (Figure 7). Pith distances less than 2.0 cm indicate that the pith itself is included in the 4.0 cm cube sample, as the distance is measured from the sample’s center.

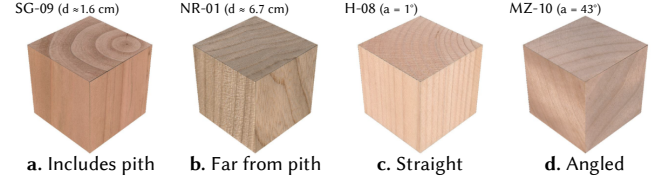


Fig. 6. Examples of cut-out *locations* (a-b) and *orientations* (c-d).

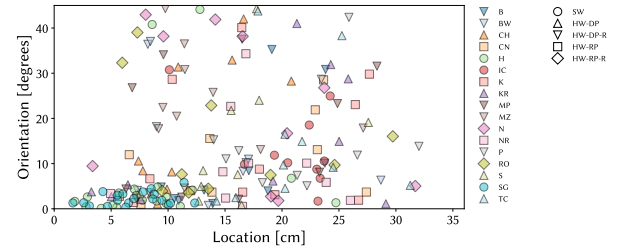


Fig. 7. Estimated cut-out locations and orientations. The x-axis indicates the *location*: the distance between the sample center and its nearest point on the pith axis. The y-axis indicates the *orientation*: the angle between the pith axis and its most closely aligned orthogonal axis.

4.2 Collection and Preparation

Hardware and online stores typically sell unfinished wood cubes from a few common species (e.g., cypress and oak) with small variation between specimens. Therefore, we sourced dry wood directly from wood mills and cut them into cubes ourselves, enabling a diverse sample set. The wood mills visited were those cutting wood for furniture production and building, thus using relatively high-grade materials of species likely to be found in interior settings.

We used a table saw to cut the blocks, producing smoother surfaces than a band saw. However, some samples, particularly softwoods, show visible cut marks from the table saw. We did not sand or finish the surfaces due to various possible techniques and available products, which would require a separate extensive study of its own to cover a sufficient range [Ravindran et al. 2023]. We chose a cube side length of 4.0 cm, as it is the largest commonly available thickness, with thicker lumber being rare due to its proneness to cracking when drying.

4.3 Documentation

Here, we introduce the modes of documentation of external and internal information of the wood cubes (Figure 8).

4.3.1 External photographs. We obtained photographs of the six sides by scanning them in batches together with a color scale using a high-resolution (1,200 dpi) office desktop 2D scanner (EPSON DS-G20000) (Figure 9). Individual photographs were cropped out

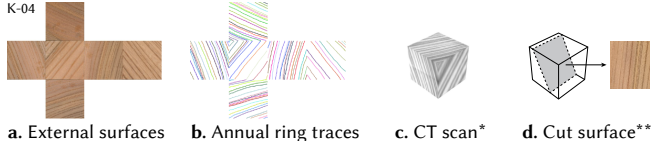


Fig. 8. Dataset documentation overview (available for *185/190 and **38/190 samples, respectively).

from the batch scans, resulting in an image size of approximately 1,880 px × 1,880 px. There are six external surfaces per cube and 190 samples, giving a total of 1,140 photographs.

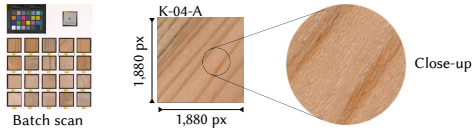


Fig. 9. External surfaces capturing.

4.3.2 Annual ring annotations. For each photograph, we provide manually created annotations in the form of numbered traces of the annual rings. The traces are stored as monochromatic 256^2 image with integer pixel values representing the year up to an arbitrary global offset. Inside knots, the annual ring annotations are largely left blank because they have highly dense annual rings, which are difficult to annotate in the set resolution. The annotations were created using a drawing interface with a stylus pen and a touch screen, and adds up to a total of 420 m.

4.3.3 CT scans. We scanned wood samples using an industrial CT scanner for timber (Microtec Mito) to obtain 3D volumes (Figure 8c).¹ Although we do not use them in our own experiments, we include the CT scans to enable future applications requiring additional 3D ground truth data. Each scan approximately covers 128^3 voxels of size $[.3 \text{ mm}]^3$.

4.3.4 Photographs of cut surfaces. After dividing the samples of each species into a training- and testing portion (refer to Section 5.1), we physically cut the 38 test samples and photographed the exposed cut surfaces to create ground truth references for comparison with inferred textures. There are two types of slanted cuts (refer to Supplement A.1.4). We did not perform this documentation on all samples because cutting the samples destroys them, keeping as many specimens as possible intact for future experiments.

5 INVERSE MODELING METHOD

Building on the *Mokume* dataset, we propose an inverse modeling pipeline that infers a volumetric color texture from the external photographs of a wood sample (Figure 10). It comprises three phases: converting photographs into a 2D annual ring localizations (Section 5.1), fitting a volumetric procedural growth field to the extracted localized annual rings (Section 5.2), and synthesizing detailed volumetric color textures (Section 5.3). Only the first phase uses the

¹The *Mokume* dataset includes CT scans of all samples except 5 (NR-6 to 10) that were added after the CT scans had been performed.

Mokume dataset as training data. For implementation details, refer to Appendix A.2.

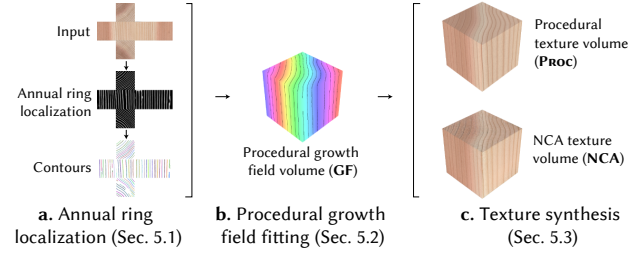


Fig. 10. High-level overview of our reconstruction pipeline.

5.1 2D Annual Ring Localization from Photographs

Annual rings are often subtle and obscured by overlapping features like rays and pores, making their detection in photographs challenging. Standard edge detection alone is insufficient (see Section 6.3.3). Therefore, we approach this as an image-to-image translation problem, converting wood color images into annual ring localization images, enabling reliable contour extraction (Figure 11).

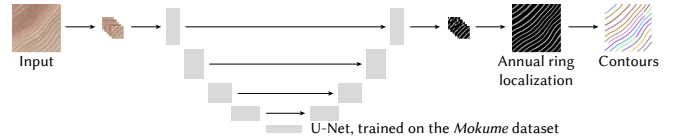


Fig. 11. Overview of the 2D annual ring localization from photographs.

Similar to prior work on annual ring detection [Fabijańska and Danek 2018], we train a U-Net for its ability to recognize image features across different scales. Its input consists of 1cm^2 patches with resolution 64×64 . The details of the network architecture and patch-based training follow Texler et al. [2020]. To further augment the training data and improve generalization to unseen wood samples of potentially different sizes, we augment the data by mirroring, scaling, and transforming the color space of the patches. During inference, we evaluate the network on a sliding window and reconstruct an image by blending the overlapping areas. Additional details on these steps can be found in Appendix A.2.1.

5.1.1 Training data. We used 80% of the *Mokume* dataset for training. Each training pair consists of an input RGB photo patch and a monochromatic target patch describing the associated annual ring structure. To create this target, we interpolate the hand-annotated numbered traces (Section 4.3.2) to every pixel to create a fractional 2D growth field g value, to which we assign an intensity value using the following formula:

$$I(g) := (\cos(2\pi g)/2 + 1/2)^{10}. \quad (1)$$

Raising the result to the 10^{th} power sharpens the peaks (annual rings locations). This transformation is invariant to integer shifts², localizes annual rings, and creates smooth oscillatory output (Figure 12).

²This is important because the year labels of annual ring traces are only known up to an integer offset.

Section 6.3.3 showcases the benefits of this mapping compared to other alternatives. Training minimizes the L_1 loss between the target patches and the U-Net's output.

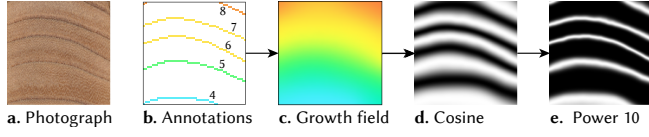


Fig. 12. Training data construction. Each training pair consists of a photo patch and an oscillatory image of the annular ring structure. We create the latter by interpolating integer-valued traces into a fractional 2D growth field that we subsequently transform using a cosine and power function.

5.1.2 Contour detection. The final step applies a threshold followed by a Canny [1986] edge detector. We trim detections at the image border and connect the remainder into sets of contour curves. This process creates two contours surrounding each annual ring, which is not an issue since the next pipeline stage (Section 5.2.2) only depends on them being iso-curves of the growth field.

5.2 Fitting a 3D Procedural Growth Field

In the next pipeline stage, we extend the 2D annual ring localization images and their extracted contours into a 3D procedural growth field. This section introduces the underlying procedural model (Section 5.2.1), loss function (Section 5.2.2), and optimization scheme (Section 5.2.3).

5.2.1 Procedural model. The function $G(\omega, p)$ evaluates the growth field at a 3D position $p \in \mathbb{R}^3$. It further depends on parameters $\omega = [O, V, R]$ (Figure 13), where

- the position and direction $O, V \in \mathbb{R}^3$ specify the *pith axis* representing the start of the tree's radial growth process. This axis establishes a cylindrical coordinate system, in which positions p can alternatively be expressed in terms of their height h , azimuth ϕ , and radius r .
- $R \in \mathbb{R}^{n_h \times n_a \times n_r}$ characterizes *rate of growth* resulting from varying environmental conditions. In particular, the values R_{ijk} discretize the *radial derivative* $\partial/\partial_r G$ of the growth field at n_h heights, n_a azimuths, and n_r radii.

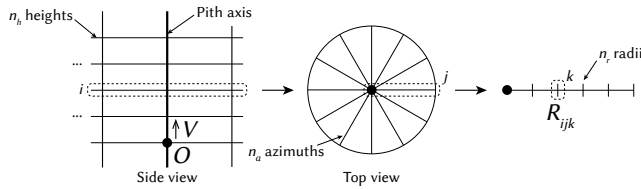


Fig. 13. Associated geometry of the growth field parameters $\omega = [O, V, R]$.

Integrating R_{ijk} along the r axis starting from the pith yields the actual value of the growth field at each discretization point. We employ trilinear interpolation to evaluate the growth field at other positions and clamp $R_{ijk} < 0$ when optimizing to prevent inversions.

5.2.2 Iso-contour loss. The 2D annual ring contours from Section 5.1.2 correspond to iso-contours of the 3D growth field, enabling their use as effective optimization constraints (Figure 14). Given a contour C_i (a set of pixels included in a connected component), we define an *iso-contour loss* that minimizes the variation of values of G along the contour C_i , by minimizing the difference between each point's value to the average value along the contour:

$$\mathcal{L}_{ic}(\omega, C_i) := \frac{1}{|C_i|} \sum_{p \in C_i} \left| G(\omega, p') - \frac{1}{|C_i|} \sum_{p' \in C_i} G(\omega, p') \right|. \quad (2)$$

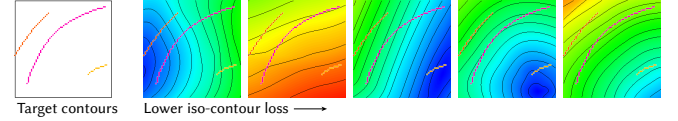


Fig. 14. Iso-contour loss promotes convergence to a 3D growth field, whose level sets on faces align with a prescribed set of target contours.

5.2.3 Optimization. The recovery of the pith axis and annual ring distortions occur in sequence (Figure 15).

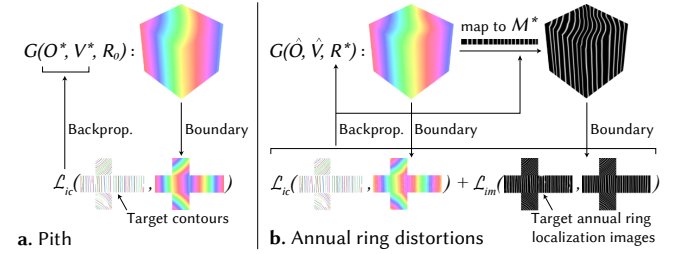


Fig. 15. Overview of the growth field fitting. **a.** Optimization of the pith axis (\hat{O}, \hat{V}), and **b.** the radial growth speeds (\hat{R}) and a gray-map (\hat{M}).

Step 1: pith. We first optimize the pith axis (O and V) of an *undeformed* (i.e., $R_{ijk} \equiv 1$) growth field to find parameters that best match the prescribed set of contours $C = \{C_1, C_2, \dots, C_N\}$. Specifically, we solve the following optimization using a discontinuous coarse grid search followed by continuous gradient-based descent:

$$\hat{O}, \hat{V} = \arg \min_{O, V} \sum_{i=1}^N \mathcal{L}_{ic}([O, V, 1], C_i) \quad (3)$$

Step 2: distortions. Annual rings can be heavily distorted, which the previous step fails to consider. Intuitively, we should be able to reproduce this distortions by optimizing the tree's spatially varying growth rate R via the loss

$$\mathcal{L}_1(R) := \sum_{i=1}^N \mathcal{L}_{ic}([O, V, R], C_i) \quad (4)$$

While this indeed corrects the iso-contour alignment, it does not position annual rings in an absolute sense. One way to recover their position would be to apply the oscillatory mapping I (Equation 1,

Figure 12) to generate images that can be compared to the U-net’s output (denoted I_{ref}):

$$\mathcal{L}_2(R) \equiv \|I(G([O, V, R], \boxplus)) - I_{\text{ref}}\|_1, \quad (5)$$

where “ \boxplus ” indicates simultaneous evaluation on all cube faces. However, this constrains annual rings to integer values of G , which we found to be exceedingly challenging to optimize. For example, inserting a new annual ring requires a concerted update of all entries of R to preserve the other ring positions at their integer positions. In practice, we often observe convergence to suboptimal local minima.

Rather, we replace I with a piecewise linear function $M: \mathbb{R} \rightarrow \mathbb{R}$ that assigns a gray value to each growth field value. That parameterized function becomes a part of the optimization:

$$\mathcal{L}'_2(R, M) \equiv \|M(G([O, V, R], \boxplus)) - I_{\text{ref}}\|_1. \quad (6)$$

This indirection makes it possible to place annual rings at arbitrary locations, not limited to integer values of G . We finally optimize a linear combination of both losses

$$\hat{R}, \hat{M} = \arg \min_{R, M} \lambda_1 \mathcal{L}_1(R) + \lambda_2 \mathcal{L}'_2(R, M)$$

The annual ring positions are easy to extract as a post-process by locating global maxima of \hat{M} . Specifically, we apply a threshold and extract the location of the centers of each detected area. The final output comprises the volumetric growth field ($G(\hat{\omega})$) and the iso-curve values ($A \in \mathbb{R}$) of annual rings in the field.

5.2.4 Knots. In the rare case where a sample contains a knot, we introduce an additional parameter set ω_k , denoting the parameters of the growth field of the knot (as opposed to the pith of the stem), and a smoothness parameter $s \in \mathbb{R}$, which controls the shape of the stem-to-knot union. We use Larsson et al. [2022]’s method that evaluates a smooth minimum of growth fields around individual skeleton strands:

$$G_k(\omega, \omega_k, s, p) = \text{smooth_min}(G(\omega, p), G(\omega_k, p), s). \quad (7)$$

To find the optimal parameters ($\hat{\omega}$, $\hat{\omega}_k$, \hat{s}) of a sample with a knot, we apply the same optimization framework as before (refer to Section 5.2.3) after replacing $G(\omega)$ with $G_k(\omega, \omega_k, s)$. Importantly, $[O_k, V_k] \in \omega_k$ requires a close-to-optimal initialization, which is constructed by manually marking the center points of a knot on the external images. There are typically two center points (the enter and exit points) because a knot typically penetrates through a sample, from which we construct the initial knot axis.

5.3 Synthesizing 3D Color Textures

After inferring the 3D procedural growth field, our next goal is to produce an RGB color volume of the wood texture that matches the appearance of the real wood as closely as possible, including details of the texture such as rays and pores. We apply two different techniques—inference of the parameters of a procedural texture (Section 5.3.1) and NCA (Section 5.3.2). Both techniques use the VGG-based style loss proposed by Kolkin et al. [2019].

5.3.1 Approach 1: inverse 3D procedural texturing. We introduce a function $F(G(\hat{\omega}), A, \phi)$ that takes the inferred global structure ($G(\hat{\omega})$ and A) and a set of additional parameters (ϕ) as input, and produces a color volume. We model the annual rings, fibers, pores,

and rays based on Liu et al. [2016]’s formulation of the wood procedural model, with some simplifications and extensions. In particular, we simplify by not modeling surface depth nor complex light interactions (Figure 16). Furthermore, we model the fiber direction as parallel to the pith axis. The extensions include two color-map parameters (size 64×3) that control the radial-outward color variations of the early- and late-wood areas, respectively, facilitating the inverse modeling of sap-to-hardwood color transitions. For knots, we employ a similar color map. Overall, ϕ includes 436 values (or 639 values with one knot).

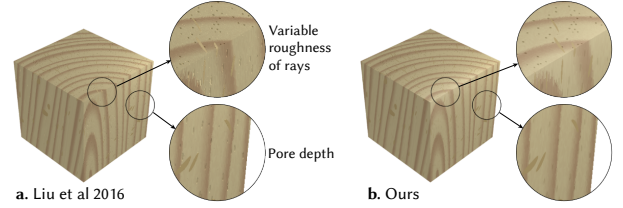


Fig. 16. **a.** Procedural wood texture based on Liu et al. [2016]. **b.** Our version which does not model complex light–material interactions based on variable reflectivity or depth (normal direction).

To determine the optimal parameter set $\hat{\phi}$, we initialize annual ring colors using a continuous optimization and L1 image loss, then switch to style loss, perform a coarse grid search of fiber, pore, and ray sizes, followed by a continuous optimization of all parameters in ϕ .

5.3.2 Approach 2: 3D neural cellular automata. To generate finer details of the wood texture, we first evaluate the growth field on a $H \times W \times D$ grid of coordinates, giving a growth field tensor G . We then fit an average wood color to each growth field value, using a loss similar to Equation 6, yielding a base color tensor B . Although the base color approximates the reference, it lacks fine details like pores and subtle grains. To recover this missing complexity, we adapt the 2D NCA framework from Pajouheshgar et al. [2024a; 2023] to a 3D volumetric setting and train the NCA to learn a residual to enhance the base color B .

Each cell in our 3D grid of size $H \times W \times D$ maintains a high-dimensional state $S \in \mathbb{R}^{H \times W \times D \times C}$, where the number of channels $C = 12$ in our experiments and the first three channels store the residual RGB color. As wood’s fine details depend strongly on anatomical context such as ring structure, we incorporate wood-specific priors into NCA by concatenating the base color B and growth field G with the cell state S , thereby forming an augmented state tensor. In Section 6.5, we show that conditioning the NCA on B and G strongly boosts synthesis quality; ablations omitting these priors lead to less realistic growth-ring continuity and color variation in the volumetric texture.

Each cell then gathers information from its local $3 \times 3 \times 3$ neighborhood by convolving the augmented state tensor with five fixed 3D convolution kernels: identity, three Sobel filters (approximating spatial gradients $\nabla_x, \nabla_y, \nabla_z$), and a 27-point stencil approximation of the Laplacian operator ∇^2 . This yields a $(5C + 10)$ -dimensional

vector for each cell, which is then concatenated with the cell's 3D coordinates (x, y, z) to form a $(5C + 13)$ -dimensional input to a multi-layered perceptron (MLP) with two layers, which calculates the residual cell state update (Figure 17). By unrolling this update rule for many iterations and applying a VGG-based style loss on the cube's visible faces, we train the NCA model parameters via backpropagation through time. We adopt the open-source implementation, training strategies, and style-loss functions from Pajouheshgar et al. [2024a].

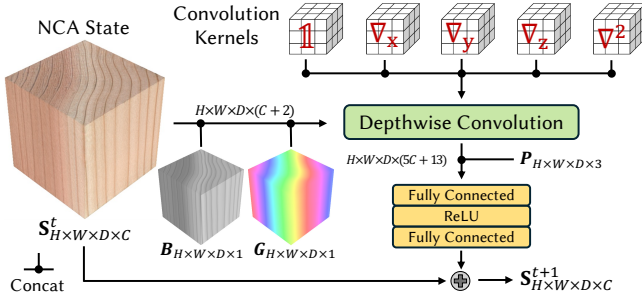


Fig. 17. Overview of 3D NCA architecture for wood texture synthesis.

6 RESULTS

In this section, we show visual outputs of our method compared to ground truth (Section 6.1) and a related work (Section 6.2), and present evaluations of individual components of our pipeline: the structural information extraction (Section 6.3), the global inference (Section 6.4), and the NCA texture synthesis (Section 6.5).

6.1 Qualitative Evaluation of Synthesized Textures

Figures 18-19 present visualizations of the 38 test cubes with slanted cuts constructed from ground truth photographs alongside the outputs of our pipeline: the growth-field and the two texture synthesis methods (procedural texture and NCA). Both methods often successfully reproduce heart-to-sapwood color transitions (e.g., IC-04) and local color differences (H-14), and likewise for pore distributions (e.g., CN-01) and ray orientations (e.g., RO-03). Owing to its flexibility, the NCA reproduces dense and complex ray patterns for which the procedural model finds a close but not fully satisfactory approximation within its expressivity (e.g., RO-09). The NCA model also reproduces a gum inclusion on the cut surface of MZ-03, though not in the ground truth location, showing its ability to reproduce irregular features beyond that which is included in our (or any existing) procedural wood model. However, the NCA model does not recognize saw marks as a 2D surface artifact, and partially reproduces them on the cut surface unrealistically (e.g., H-01). Finally, the NCA shows a relatively better capacity to recover when the growth field reconstruction is poor (e.g., N-01).

6.2 Visual comparison

Figure 20 shows the outputs from Larsson et al. [2024]’s method, showing that it does not reconstruct knots (H-11) or heart-to-sapwood transitions (P-06) as well as our method does. Moreover, note that

their method takes contours *numbered-by-year* (Figure 21a) as input, and thus it cannot we cannot apply it directly on our automatically extracted unordered edges (Figure 21b, refer to Section 5.1.2), as would be required for fair comparison. Instead, we applied it to the *manually annotated* annual ring trances of our dataset (refer to Section 4.3.2) because they are numbered. However, the manually annotated traces are naturally more complete and accurate compared to the automatically extracted edges that our results in Figures 18-19 are based on.

6.3 2D Annual Ring Localization

In this section, we show the species-wise performance of the annual ring localization image generation (Section 6.3.1) and a generalization test (Section 6.3.2). We also evaluate the accuracy of the edges extracted from the annual ring localization image compared to other methods (Section 6.3.3).

6.3.1 Performance. The image losses of the annual ring localization images vary when comparing different species (Figure 22). Moreover, we observe that the model performs best on the most regular cases when annual rings are distinct and complex patterns are not present (e.g., MP-07-B in Figure 23). The model also typically successfully ignores superimposed visual patterns like sap-to-heart wood borders (e.g., IC-04-B), cut marks (e.g., H-14-D), and distinct rays and pores (e.g., N-01-C in Figure 23). Extremely subtle, barely visible annual rings, however, appear to be more challenging (e.g., S-08-C and TC-04-B in Figure 23). Occasionally, there are block-like black solid artifacts due to patch blending (e.g., towards upper left in H-14-D in Figure 23).

6.3.2 Generalization test. We conducted an experiment to evaluate the ability of the model to generalize to unseen species. The 17 species were randomly divided into five groups of three species each and one group of two species. For each group, we trained the network on all samples from the species outside the group and evaluated its performance on the unseen species within the group (Figure 24). The median image loss across the six experiments is 0.20, which is somewhat higher compared to 0.17 of our standard setting, where the network is trained on 80% of the samples from each species and evaluated on the remaining 20%.

6.3.3 Contour extraction. We evaluate the quality of the contours extracted from our generated annual ring localization, and compare it to one ablation and two baselines. The ablation entails omitting the power of 10 in the annual ring localization image construction (refer to Section 5.1.1). The two baselines are extracting contours from 1) a thresholded gray-scale wood photograph directly, 2) a wood photograph after applying a k-means filter ($k = 2$) [Lloyd 1982]. The quality of extracted contours are measured by their iso-contour error against the ground truth continuous growth field (refer to Figure 12c). Moreover, since the iso-contour error can be minimized with no contours detected, we additionally report the ratios of the ground truth contour length and count to confirm that a sufficient length and amount of contours have been detected, where ground truth contours are those extracted from the ground truth annual ring localization image (refer to Figure 12e). Our method outperforms alternatives in terms of median average iso-contour error

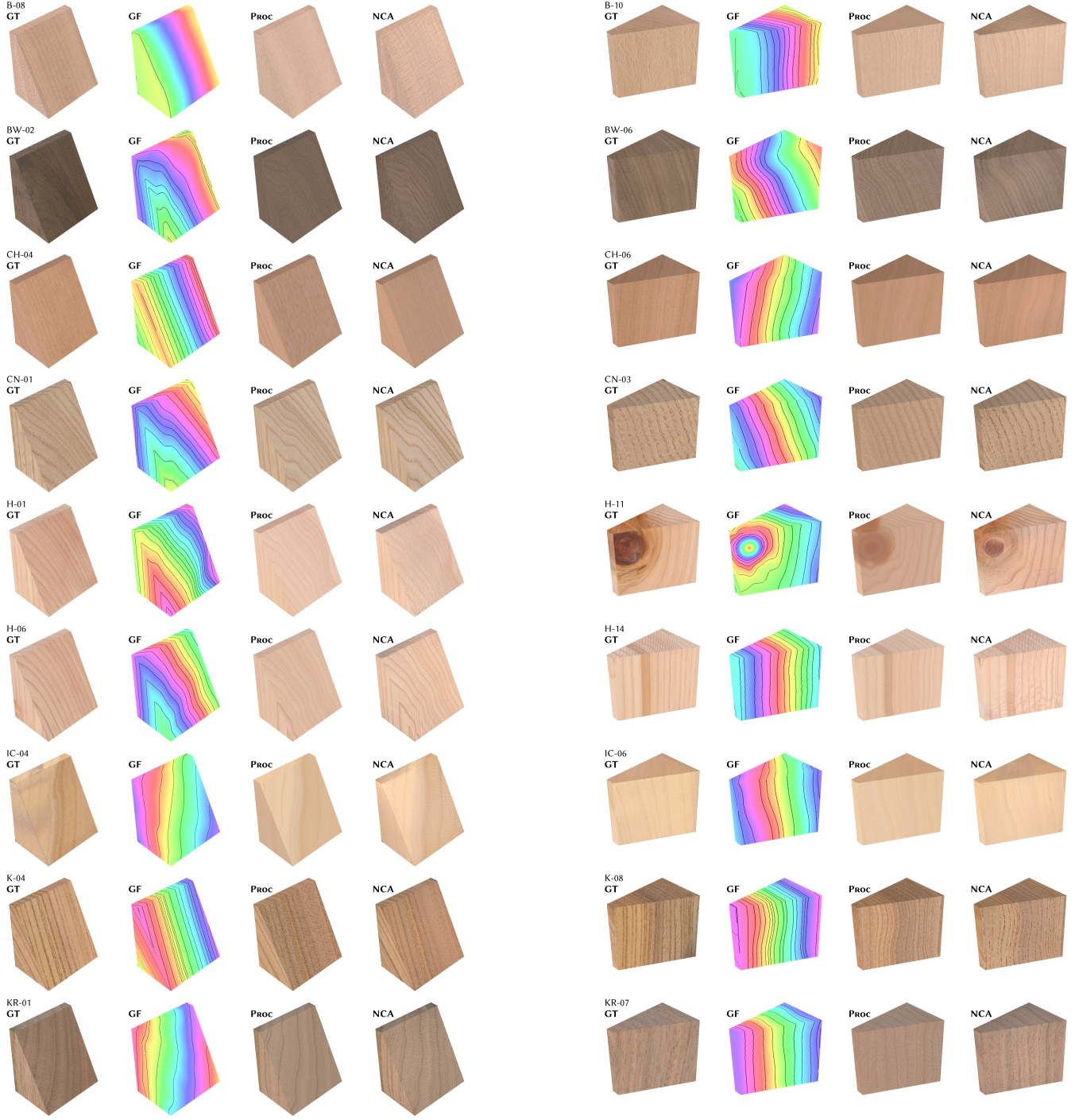


Fig. 18. Qualitative validation (continued on the next page). This matrix covers two samples (columns) from the test subset of each species (rows), showing a held-out ground truth photograph of the physically cut wood cube (sub-column 1), compared to slanted cuts through the reconstructed growth field (sub-column 2) and synthesized solid wood textures (sub-column 3, 4). When comparing the outputs, note that, while it is our objective to closely infer the annual ring locations relative to the reference, we do not expect to reconstruct texture details (pores, rays) with a one-to-one pixel correspondence. Rather, we effectively model their distributions. See Section 6.1 for further discussion of these results. Videos of rotating cubes for all 38 samples can be found in the supplementary materials.

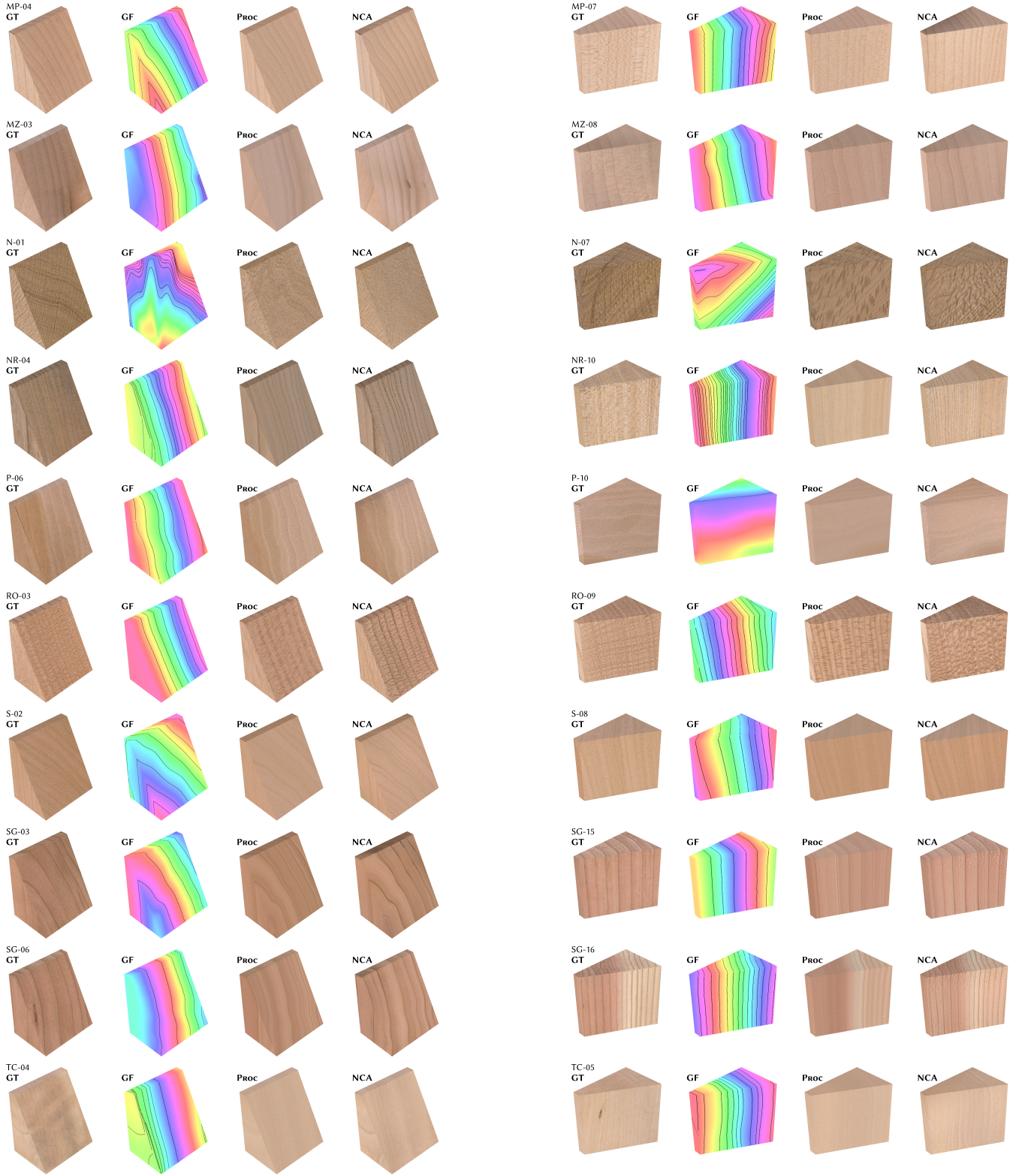


Fig. 19. Continuation of Figure 18.

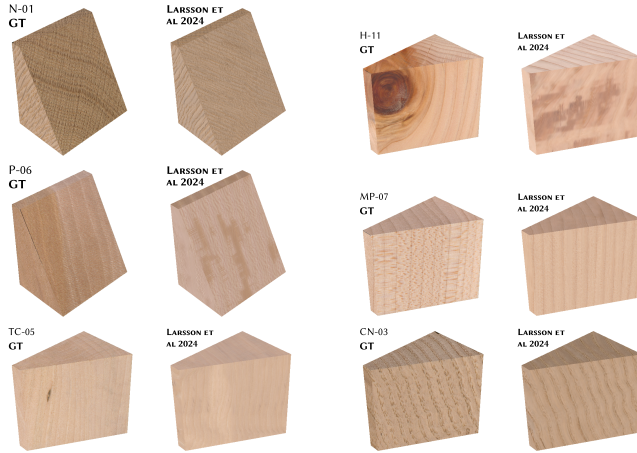


Fig. 20. Qualitative outputs of related work. Outputs of the inference method of Larsson et al. [2024] applied to six samples from the *Mokume* dataset. Note that these were created based on *manually annotated and numbered* annual ring trances, as opposed to automatically extracted from input photographs.

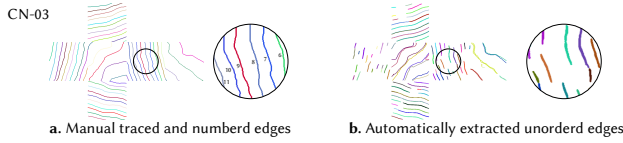


Fig. 21. Edge types. **a.** Manually annotated numbered annual ring traces, based on which the results of Larsson et al. [2024] in Figure 20 are generated. **b.** Automatically extracted, unordered edges, based on which our results in Figures 18–19 are generated.

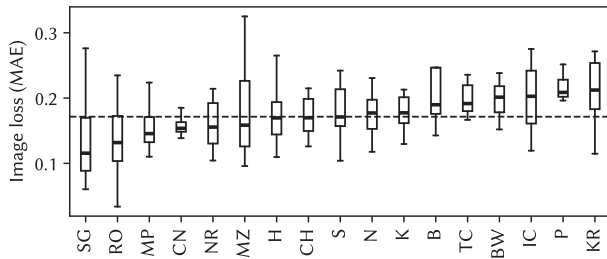


Fig. 22. Performance of the annual ring localization image translation. Image loss of each species of U-Net generated annual ring localization images. The species are organized by ascending order of loss. The dashed line indicates the median image loss across all species.

(Figure 25). Moreover, the length and count ratios of the baselines deviate significantly from 1.0, indicating that extracted contours are either overly few and overly short, or overly long and overly many.

6.4 3D Procedural Growth Field Fitting

This section presents an ablation study focusing on the optimization of annual ring distortions (Section 6.4.1). Analysis of the pith

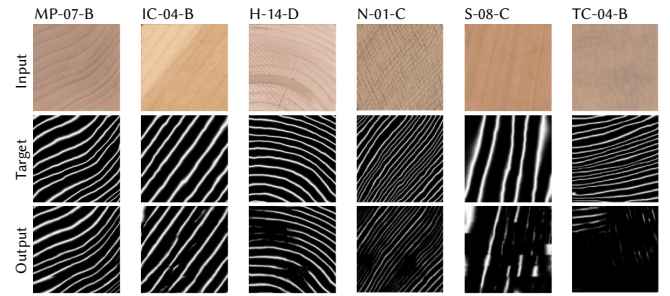


Fig. 23. Example outputs of the annual ring localization image translation. Top: color image of the wood (input). Middle: ground truth annual ring localization images created from manual annotations (target). Bottom: U-Net generated annual ring localization images (output).

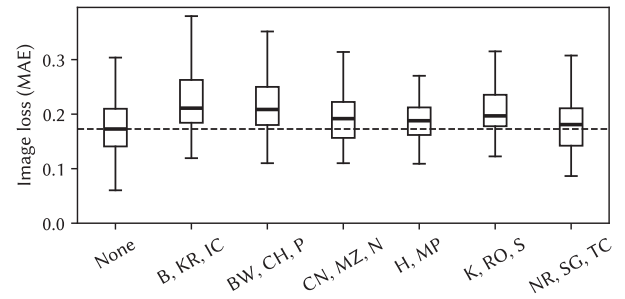


Fig. 24. Generalization test of the annual ring localization image translation. Box 1 ("None") represents the results when the network is trained on 80% of all species and evaluated on the remaining 20%. Boxes 2–7 ("B, KR, IC," and so on) show the results when the network is trained on 100% of the samples of all species except a specific group of 2–3 species and evaluated on the unseen species in that group. The dashed line indicates the median loss of box 1.

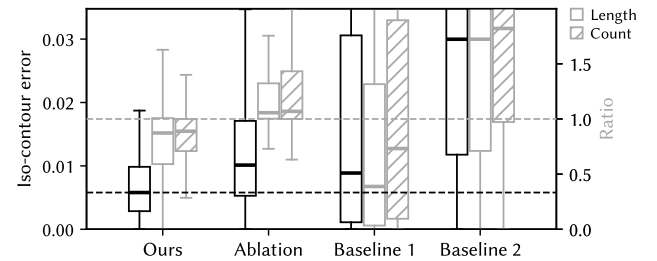


Fig. 25. Quality of extracted contours. Comparison between our method, an ablation, and two baselines. Each box shows the distribution of the average errors and ratios for the 228 images (6×38) in our test set. Lower iso-contour error indicates that the contours better follow the zero-gradient of the ground truth growth field. The dashed black line shows the median average iso-contour error of the proposed method (0.006). The dashed gray line indicates the ideal ratio (1.0) of contour lengths and counts.

estimation step and growth rate (R) parameter resolution can be found in Appendices B.1 and B.2, respectively.

6.4.1 Optimization step 2: distortions. In global structure inference step 2, we optimize the distortions of the growth field, which are adjusted through the control radii ($R \in \mathbb{R}^{n_h \times n_a \times n_r}$), and the gray-map (M) from which the iso-values of the annual rings are extracted (refer to Section 5.2.3). Possible questions includes, is it important that R represents the growth speed at a point rather than the explicit growth value, and is the iso-contour loss necessary? We ran ablation experiments to answer these questions. Specifically, we compare our method to two ablations: 1) with an alternative explicit construction of R , 2) without iso-contour loss (just the image loss between target and output annual ring localization images drives the optimization). We evaluate the results against the ground truth annual ring annotations by two metrics. First, if the annual ring localization is correct, the color values of the annotated pixels should be white, and therefore, we report the average pixel colors at the annotation pixel locations. Second, we measure the iso-contour error of the ground truth annotated annual ring edges in the underlying growth-field from which the annual ring localization is constructed. This measures how well the annual ring pattern in the procedurally generated annual ring patterns follow the ground truth shapes.

Our method creates outputs with a lower median iso-contour errors, while median color values are similar for all methods (Figure 26). This is confirmed by the visual outputs, where we observe that estimated annual rings in the ablations frequently “jump” between different GT annual rings, that is, not following their shape although hitting many peaks (Figure 27). In ablation 1 (explicit construction of R) there are visual artifacts caused by this construction not guaranteeing that an inner radius is smaller than an outer radius. This shows up as annual rings crossing each other, which is anatomically infeasible.

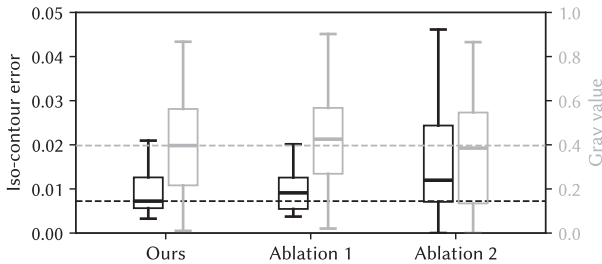


Fig. 26. Ablation study of the annual ring distortions and locations optimization. Each box shows the distribution of losses for the 38 samples in the test dataset. A better estimated annual ring localization image is characterized by a lower iso-contour error (ideally 0.0), and a higher gray value (ideally 1.0, white).

6.5 Ablation of NCA

NCAs generally excel at synthesizing repetitive patterns or textures without strong global structures. Pajouheshgar et al. [2023] showed that positional encodings can help capture some long-range correlations of the target texture. However, by ablating the wood-specific conditioning and omitting B and G from the NCA architecture, we showed that positional encoding alone is insufficient for producing coherent solid wood texture. Without our wood-specific priors, the

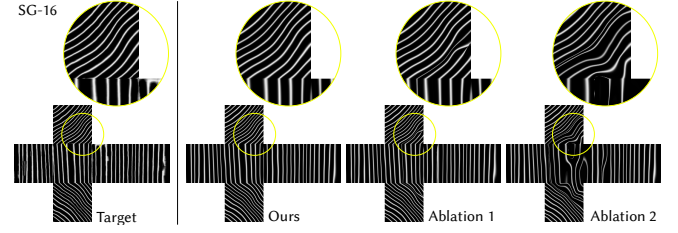


Fig. 27. Visual outputs from the ablation study of the annual ring distortions and locations optimization. The optimization target are the external annual ring localization images generated by the U-Net (along with the contours extracted from it). In the ablations, annual rings tend to jump up or down between different GT annual rings.

annual rings become discontinuous (e.g., CN-01), the rays form in an incorrect direction (e.g., B-08), and the knots fail to form properly in the synthesized textures (e.g., H-11, Figure 28). This highlights that while positional encoding can help NCAs to model some aspects of a global structure, they cannot substitute our global wood structure priors.

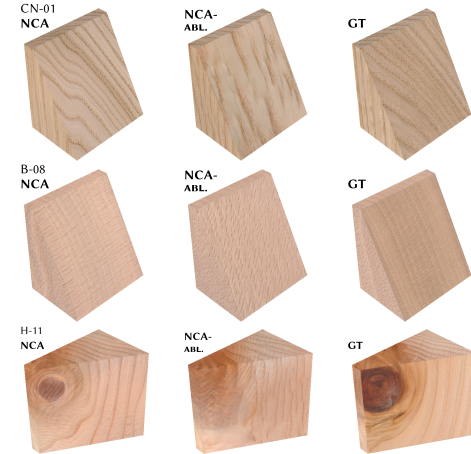


Fig. 28. Visual outputs from the NCA ablation study. Omitting the priors of the global wood structure leads to discontinuous annual rings (CN-01), incorrect orientation of rays (B-08), and indistinct knot structure (H-11).

7 LIMITATIONS

We have not yet captured samples in the *Mokume* dataset under varying lighting conditions, the dataset is limited to one sample size ($[4.0 \text{ cm}]^3$). Moreover, although we included samples from many (17) species, this list is not exhaustive. That being said, the *Mokume* dataset includes species with a wide variety of anatomical features and material imperfections, and the generalization test (Section 6.3.2) indicates that the method could be applied to unseen species with a reasonable expectation of success. For example, birch and acacia are classified as HW-DP-R, a category represented by six species in our dataset (B, BW, MP, MZ, P, and TC), while ash and teak are HW-RP-R, represented by three species (K, N, and RO).

Our annual ring localization step sometimes fails to generate accurate and continuous annual ring contours, leading to downstream errors in reconstitution. In particular, block-like artifacts could potentially be mitigated by a more advanced patching strategy at multiple scales or with confidence weighting.

We approximate the pith as a straight line, and our parametrization of annual ring distortions is based on the assumptions that annual rings are star-shaped relative to the closest point the pith axis (or the knot axis). While a more complex formulation of the pith and distortions is unnecessary for our target samples, it could prove beneficial for handling highly distorted wood instances. While a more complex formulation of the pith and the distortions would be unlikely to yield better results for our target samples, it could prove beneficial for extremely distorted wood instances. Moreover, we model only alive knots (not dead or broken off ones), do not apply growth distortions to the knots, and approximate their central skeleton as a straight line.

As for texture synthesis, our procedural texturing model is a simplification of the state-of-the-art ([Liu et al. 2016] (refer to Section 5.3.1)). The main limitation of the NCA is that a single state tensor ($256 \times 256 \times 256 \times 12$) requires 384 MiB of storage in float16 precision. Backpropagating derivatives through a sequence of NCA updates requires intermediate states to be held in memory, which can saturate the memory of even the largest GPUs on the market.

8 CONCLUSION

The *Mokume* dataset provides a comprehensive new resource for wood structure modeling, offering exceptional diversity by covering various species with high-resolution multimodal data. We demonstrate how the combination of this dataset with computational models has powerful predictive capabilities, evidenced by high-fidelity synthetic wood textures that closely align with unseen ground truth data.

Central to our approach is the use of annual ring patterns as a deterministic feature, while fine-scale details (rays, pores, etc.) are treated as a learnable stochastic distribution. A compelling avenue for future research could involve extrapolating beyond the sample's boundaries. This would necessitate a modified methodology focused on capturing the overall "style" of distortions rather than achieving an exact match.

The potential applications of inverse procedural modeling are numerous and span diverse fields. Examples include predicting the visual outcome of sculptural pieces before physical carving, synthesizing extensive libraries of realistic materials for digital rendering, or the inference of non-visual properties: since the global structure of wood determines its mechanical properties, we envision the use of growth fields to augment simulations of material deformation and strength, using the resulting insights to guide irreversible processes such as milling and cutting. Another potential research area is the directional reflectance behavior caused by the anisotropic nature of wood fibers. While previous research required complex gonio-photometric measurements to infer their properties, we hypothesize that such effects could also be retrofitted to our work, as the fiber direction is inherently tied to the growth field.

Besides serving as a benchmark for procedural and learning-based texture synthesis, we hope that the release of the *Mokume* will catalyze interdisciplinary research to explore the manifold properties of wood.

ACKNOWLEDGMENTS

This research was made possible through a collaboration between The University of Tokyo and Gifu Prefectural Research Institute for Human Life Technology in Takayama, and it was generously supported by JST ACT-X Grant Number JPMJAX210P, JSPS KAKENHI Grant Number JP23K19994, JST AdCORP Grant Number JPMJKB2302, JST ASPIRE JPMJAP2401, and a collaborative research fund between Mercari Inc. R4D and RIISE, The University of Tokyo.

REFERENCES

- Miika Aittala, Tim Weyrich, and Jaakko Lehtinen. 2015. Two-shot SVBRDF capture for stationary materials. *ACM Trans. Graph.* 34, 4, Article 110 (July 2015), 13 pages. <https://doi.org/10.1145/2766967>
- H. Boukadida, F. Longuetaud, F. Colin, C. Freyburger, T. Constant, J. M. Leban, and F. Mothe. 2012. PithExtract: A robust algorithm for pith detection in computer tomography images of wood - Application to 125 logs from 17 tree species. *Comput. Electron. Agric.* 85 (July 2012), 90–98. <https://doi.org/10.1016/j.compag.2012.03.012>
- John Canny. 1986. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8, 6 (1986), 679–698. <https://doi.org/10.1109/TPAMI.1986.4767851>
- Ruben De Blaere, Kevin Lievens, Dieter Hassel, Victor Deklerck, Tom De Mil, Wannes Hubau, Joris Van Acker, Nils Bourland, Jan Verwaeren, Jan Van den Bulcke, and Hans Beeckman. 2023. SmartWoodID—an image collection of large end-grain surfaces to support wood identification systems. *Database* 2023 (05 2023). <https://doi.org/10.1093/database/baad034>
- Anna Fabijańska and Malgorzata Danek. 2018. DeepDendro – A tree rings detector based on a deep convolutional neural network. *Computers and Electronics in Agriculture* 150 (2018), 353–363. <https://doi.org/10.1016/j.compag.2018.05.005>
- Anna Fabijańska, Malgorzata Danek, Joanna Barniak, and Adam Piórkowski. 2017. Towards automatic tree rings detection in images of scanned wood samples. *Computers and Electronics in Agriculture* 140 (2017), 279–289. <https://doi.org/10.1016/j.compag.2017.06.006>
- Rado Gazo, Juraj Vanek, Michel Abdul-Massih, and Bedrich Benes. 2020. A fast pith detection for computed tomography scanned hardwood logs. *Computers and Electronics in Agriculture* 170 (2020), 105107. <https://doi.org/10.1016/j.compag.2019.105107>
- Ioannis Gkioulekas, Shuang Zhao, Kavita Bala, Todd Zickler, and Anat Levin. 2013. Inverse Volume Rendering with Material Dictionaries. *ACM Trans. Graph.* 32, 6, Article 162 (nov 2013), 13 pages. <https://doi.org/10.1145/2508363.2508377>
- Paul Guerrero, Miloš Hasan, Kalyan Sunkavalli, Radomir Měch, Tamy Boubekeur, and Niloy J. Mitra. 2022. MatFormer: a generative model for procedural materials. *ACM Trans. Graph.* 41, 4, Article 46 (jul 2022), 12 pages. <https://doi.org/10.1145/3528223.3530173>
- Y. Guo, M. Hasan, L. Yan, and S. Zhao. 2020. A Bayesian Inference Framework for Procedural Material Parameter Estimation. *Computer Graphics Forum* 39, 7 (2020), 255–266. <https://doi.org/10.1111/cgf.14142> arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14142>
- Jorge Gutierrez, Julien Rabin, Bruno Galerne, and Thomas Hurtut. 2020. On demand solid texture synthesis using deep 3D networks. In *Computer Graphics Forum*, Vol. 39. Wiley Online Library, 511–530.
- Philipp Henzler, Valentin Deschaintre, Niloy J. Mitra, and Tobias Ritschel. 2021. Generative modelling of BRDF textures from flash images. *ACM Trans. Graph.* 40, 6, Article 284 (Dec. 2021), 13 pages. <https://doi.org/10.1145/3478513.3480507>
- Philipp Henzler, Niloy J Mitra, and Tobias Ritschel. 2020. Learning a neural 3D texture space from 2D exemplars. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 8356–8364.
- Yiwei Hu, Julie Dorsey, and Holly Rushmeier. 2019. A novel framework for inverse procedural texture modeling. *ACM Trans. Graph.* 38, 6, Article 186 (nov 2019), 14 pages. <https://doi.org/10.1145/3355089.3356516>
- Yiwei Hu, Paul Guerrero, Milos Hasan, Holly Rushmeier, and Valentin Deschaintre. 2022a. Node Graph Optimization Using Differentiable Proxies. In *ACM SIGGRAPH 2022 Conference Proceedings* (Vancouver, BC, Canada) (SIGGRAPH '22). Association for Computing Machinery, New York, NY, USA, Article 5, 9 pages. <https://doi.org/10.1145/3528233.3530733>
- Yiwei Hu, Paul Guerrero, Milos Hasan, Holly Rushmeier, and Valentin Deschaintre. 2023. Generating Procedural Materials from Text or Image Prompts (SIGGRAPH

- '23). Association for Computing Machinery, New York, NY, USA, Article 4, 11 pages. <https://doi.org/10.1145/3588432.3591520>
- Yiwei Hu, Chengan He, Valentin Deschaintre, Julie Dorsey, and Holly Rushmeier. 2022b. An Inverse Procedural Modeling Pipeline for SVBRDF Maps. *ACM Trans. Graph.* 41, 2, Article 18 (jan 2022), 17 pages. <https://doi.org/10.1145/3502431>
- Hiroharu Kato, Deniz Beker, Mihai Morariu, Takahiro Ando, Toru Matsuoka, Wadim Kehl, and Adrien Gaidon. 2020. Differentiable Rendering: A Survey. *CoRR abs/2006.12057* (2020). arXiv:2006.12057 <https://arxiv.org/abs/2006.12057>
- Kayoko Kobayashi, Sung-Wook Hwang, Takayuki Okochi, Won-Hee Lee, and Junji Sugiyama. 2019. Non-destructive method for wood identification using conventional X-ray computed tomography data. *Journal of Cultural Heritage* 38 (2019), 88–93. <https://doi.org/10.1016/j.culher.2019.02.001>
- P Kodytek, A Bodzas, and P Bilik. 2022. A large-scale image dataset of wood surface defects for automated vision-based quality control processes [version 2; peer review: 2 approved]. *F1000Research* 10, 581 (2022). <https://doi.org/10.12688/f1000research.52903.2>
- Nicholas Kolkin, Jason Salavon, and Gregory Shakhnarovich. 2019. Style transfer by relaxed optimal transport and self-similarity. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10051–10060.
- Johannes Kopf, Chi Wing Fu, Daniel Cohen-Or, Oliver Deussen, Dani Lischinski, and Tien Tsin Wong. 2007. Solid texture synthesis from 2D exemplars. *ACM Transactions on Graphics* (2007). <https://doi.org/10.1145/1276377.1276380>
- J. Kratt, M. Spicker, A. Guayaquil, M. Fiser, S. Pirk, O. Deussen, J. C. Hart, and B. Benes. 2015. Woodification: User-Controlled Cambial Growth Modeling. In *Computer Graphics Forum*, Vol. 34. <https://doi.org/10.1111/cgf.12566>
- Wattananong Kurdhongmee and Korakot Suwannarat. 2019. Locating Wood Pith in a Wood Stem Cross Sectional Image Using YOLO Object Detection. *2019 International Conference on Technologies and Applications of Artificial Intelligence (TAAI)* (2019), 1–6. <https://api.semanticscholar.org/CorpusID:210692891>
- Vivek Kwatra, Irfan Essa, Aaron Bobick, and Nipun Kwatra. 2005. Texture optimization for example-based synthesis. In *ACM SIGGRAPH 2005 Papers* (Los Angeles, California) (SIGGRAPH '05). Association for Computing Machinery, New York, NY, USA, 795–802. <https://doi.org/10.1145/1186822.1073263>
- Maria Larsson, Takashi Ijiri, I-Chao Shen, Hironori Yoshida, Ariel Shamir, and Takeo Igarashi. 2024. Learned Inference of Annual Ring Pattern of Solid Wood. *Computer Graphics Forum* n/a, n/a (2024), e15074. <https://doi.org/10.1111/cgf.15074> arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.15074
- Maria Larsson, Takashi Ijiri, Hironori Yoshida, Johannes A. J. Huber, Magnus Fredriksson, Olof Broman, and Takeo Igarashi. 2022. Procedural texturing of solid wood with knots. *ACM Transactions on Graphics* 41, 4 (7 2022), 1–10. <https://doi.org/10.1145/3528223.3530081>
- Laurent Lefebvre and Pierre Poulin. 2000. Analysis and Synthesis of Structural Textures. In *Graphics Interface 2000*. 77–86.
- Beichen Li, Yiwei Hu, Paul Guerrero, Miloš Hašan, Liang Shi, Valentin Deschaintre, and Wojciech Matusik. 2024. Procedural Material Generation with Reinforcement Learning. *ACM Transactions on Graphics (TOG)* 43, 6 (2024), 1–14.
- Beichen Li, Liang Shi, and Wojciech Matusik. 2023. End-to-end Procedural Material Capture with Proxy-Free Mixed-Integer Optimization. *ACM Trans. Graph.* 42, 4, Article 153 (jul 2023), 15 pages. <https://doi.org/10.1145/3592132>
- Beichen Li, Rundui Wu, Armando Solar-Lezama, Changxi Zheng, Liang Shi, Bernd Bickel, and Wojciech Matusik. 2025. VLMaterial: Procedural Material Generation with Large Vision-Language Models. *International Conference on Learning Representations (ICLR)* (2025).
- Tzu-Mao Li, Miika Aittala, Frédo Durand, and Jaakko Lehtinen. 2018. Differentiable Monte Carlo Ray Tracing through Edge Sampling. *ACM Trans. Graph. (Proc. SIGGRAPH Asia)* 37, 6 (2018), 222:1–222:11.
- Albert Julius Liu, Zhao Dong, Miloš Hašan, and Steve Marschner. 2016. Simulating the structure and texture of solid wood. *ACM Transactions on Graphics* (2016). <https://doi.org/10.1145/2980179.2980255>
- Hsueh-Ti Derek Liu, Michael Tao, and Alec Jacobson. 2018. Paparazzi: Surface Editing by way of Multi-View Image Processing. *ACM Transactions on Graphics* (2018).
- S. Lloyd. 1982. Least squares quantization in PCM. *IEEE Transactions on Information Theory* 28, 2 (1982), 129–137. <https://doi.org/10.1109/TIT.1982.1056489>
- Fleur Longuetaud, Jean-Michel Leban, Frédéric Mothe, Erwan Kerrien, and Marie-Odile Berger. 2004. Automatic detection of pith on CT images of spruce logs. *Computers and Electronics in Agriculture* 44, 2 (2004), 107–119. <https://doi.org/10.1016/j.compag.2004.03.005>
- Joanne Mann, Mike Plank, and Andy Wilkins. 2006. Tree growth and wood formation — applications of anisotropic surface growth. In *Proceedings of the Mathematics in Industry Study Group*. 153–192.
- Stephen R. Marschner, Stephen H. Westin, Adam Arbree, and Jonathan T. Moon. 2005. Measuring and modeling the appearance of finished wood. In *ACM Transactions on Graphics*, Vol. 24. <https://doi.org/10.1145/1073204.1073254>
- Alexander Mordvintsev, Ettore Randazzo, Eyvind Niklasson, and Michael Levin. 2020. Growing Neural Cellular Automata. *Distill* (2020). <https://doi.org/10.23915/distill.00023> <https://distill.pub/2020/growing-ca>
- Tadachika Murayama and Motoharu Murayama. 2020. *Mokuzai-dai-jiten 200-shu*. Seibundo Shinkosha Publishing Co. LTD.
- Baptiste Nicolet, Alec Jacobson, and Wenzel Jakob. 2021. Large Steps in Inverse Rendering of Geometry. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 40, 6 (Dec. 2021). <https://doi.org/10.1145/3478513.3480501>
- Jannik Boll Nielsen, Henrik Wann Jensen, and Ravi Ramamoorthi. 2015. On Optimal, Minimal BRDF Sampling for Reflectance Acquisition. *ACM Transactions on Graphics (TOG)* 34, 6 (November 2015), 186:1–186:11. <https://doi.org/10.1145/2816795.2818085>
- Eyvind Niklasson, Alexander Mordvintsev, Ettore Randazzo, and Michael Levin. 2021. Self-organising textures. *Distill* 6, 2 (2021), e00027–003.
- Merlin Nimier-David, Delio Vicini, Tizian Zeltner, and Wenzel Jakob. 2019. Mitsuba 2: A Retargetable Forward and Inverse Renderer. *Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 38, 6 (dec 2019). <https://doi.org/10.1145/3355089.3356498>
- Thomas K. Nindel, Mohcen Hafidi, Tomáš Iser, and Alexander Wilkie. 2023. Automatic inference of an anatomically meaningful solid wood texture from a single photograph. arXiv:2302.01820 [cs.GR] <https://arxiv.org/abs/2302.01820>
- Thomas Klaus Nindel, Tomáš Iser, Tobias Rittig, Alexander Wilkie, and Jaroslav Krivánek. 2021. A Gradient-Based Framework for 3D Print Appearance Optimization. *ACM Trans. Graph.* 40, 4, Article 178 (July 2021), 15 pages. <https://doi.org/10.1145/3450626.3459844>
- Kristin Norell and Gunilla Borgefors. 2008. Estimation of pith position in untreated log ends in sawmill environments. *Computers and Electronics in Agriculture* 63, 2 (2008), 155–167. <https://doi.org/10.1016/j.compag.2008.02.006>
- Ehsan Pajouheshgar, Yitao Xu, Alexander Mordvintsev, Eyvind Niklasson, Tong Zhang, and Sabine Süsstrunk. 2024b. Mesh Neural Cellular Automata. *ACM Trans. Graph.* (2024). <https://doi.org/10.1145/3658127>
- Ehsan Pajouheshgar, Yitao Xu, and Sabine Süsstrunk. 2024a. NoiseNCA: Noisy Seed Improves Spatio-Temporal Continuity of Neural Cellular Automata (Artificial Life Conference Proceedings, Vol. ALIFE 2024: Proceedings of the 2024 Artificial Life Conference). 57. https://doi.org/10.1162/isal_a_00785 arXiv:https://direct.mit.edu/isal/proceedings-pdf/isal2024/36/57/2461193/isal_a_00785.pdf
- Ehsan Pajouheshgar, Yitao Xu, Tong Zhang, and Sabine Süsstrunk. 2023. DyNCA: Real-time Dynamic Texture Synthesis Using Neural Cellular Automata. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 20742–20751.
- Darwyn R Peachey. 1985. Solid texturing of complex surfaces. In *Proceedings of the 12th annual conference on Computer graphics and interactive techniques*. 279–286.
- Ken Perlin. 1985. An image synthesizer. In *Proceedings of the 12th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '85)*. Association for Computing Machinery, New York, NY, USA, 287–296. <https://doi.org/10.1145/325334.325247>
- Tiziano Portenier, Siavash Arjomand Bigdeli, and Orcun Goksel. 2020. Gramgan: Deep 3D texture synthesis from 2D exemplars. *Advances in Neural Information Processing Systems* 33 (2020), 6994–7004.
- Prabu Ravindran, Frank C. Owens, Adriana Costa, Brunela Pollastrelli Rodrigues, Manuel Chavesta, Rolando Montenegro, Rubin Shmuly, and Alex C. Wiedenhoeft. 2023. Evaluation Of Test Specimen Surface Preparation On Macroscopic Computer Vision Wood Identification. *Wood and Fiber Science* 55, 2 (Nov. 2023), 176–202. <https://doi.org/10.22382/wfs-2023-15>
- Rudolf Schraml and Andreas Uhl. 2013. Pith estimation on rough log end images using local fourier spectrum analysis. In *Proceedings of the IASTED International Conference on Computer Graphics and Imaging, CGIM 2013*. <https://doi.org/10.2316/P.2013.797-012>
- Damien Sellier, Michael J. Plank, and Jonathan J. Harrington. 2011. A mathematical framework for modelling cambial surface evolution using a level set method. *Annals of Botany* 108, 6 (2011). <https://doi.org/10.1093/aob/mcr067>
- Jiahao Shi, Zhenye Li, Tingting Zhu, Dongyi Wang, and Chao Ni. 2020b. Defect Detection of Industry Wood Veneer Based on NAS and Multi-Channel Mask R-CNN. *Sensors* 20, 16 (2020). <https://doi.org/10.3390/s20164398>
- Liang Shi, Beichen Li, Miloš Hašan, Kalyan Sunkavalli, Tamy Boubekur, Radomir Mech, and Wojciech Matusik. 2020a. Match: differentiable material graphs for procedural material capture. *ACM Trans. Graph.* 39, 6, Article 196 (nov 2020), 15 pages. <https://doi.org/10.1145/3414685.3417781>
- Ondřej Texler, David Futschik, Michal kučera, Ondřej jamříška, Šárka Sochorová, Menci-Chai, Sergey Tulyakov, and Daniel Šykora. 2020. Interactive video stylization using few-shot patch-based training. *ACM Trans. Graph.* 39, 4, Article 73 (Aug. 2020), 11 pages. <https://doi.org/10.1145/3386569.3392453>
- Charles Warner, Fanyou Wu, Rado Gazo, Bedrich Benes, Nicole Kong, and Songlin Fei. 2024. CentralBark Image Dataset and Tree Species Classification Using Deep Learning. *Algorithms* 17, 5 (2024). <https://doi.org/10.3390/a17050179>
- Fanyou Wu, Yunmei Huang, Bedrich Benes, Charles C. Warner, and Rado Gazo. 2023. Automated tree ring detection of common Indiana hardwood species through deep learning: Introducing a new dataset of annotated images. *Information Processing in Agriculture* (2023). <https://doi.org/10.1016/j.inpa.2023.10.002>

A IMPLEMENTATION DETAILS

A.1 Dataset

A.1.1 Cube unfolding convention. Figure 29 shows the cube unfolding convention that we employ.

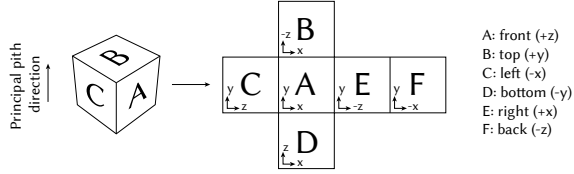


Fig. 29. Cube unfolding convention. Physical samples are oriented such that the principal pith direction is aligned with the B-D direction. The principal pith direction is the orthogonal axis that is *most closely* aligned with the pith direction.

A.1.2 Annual ring annotation data. The annotated data is stored in a 1-channel 256^2 image where the integer value of each pixel corresponds to the year of the ring, except the background value 255 (Figure 30). The year-tags of the traces are numbered relatively rather than absolutely, as determining the exact age of a ring is infeasible without the cutout including the pith.

The total number of annual ring traces in all 1,140 photos of external surfaces are approximately 12,600 (average about 11 traces per external surface). The total length of all traces is approximately 2,300,000 pixels, corresponding to roughly 420 meters (per external surface average of about 2,300 pixels, corresponding to 37 cm).

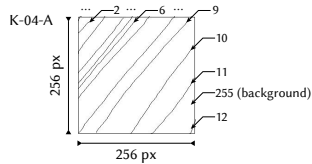


Fig. 30. Data for numbered annual ring traces.

A.1.3 Computed tomography scans. The scanner was set to an acceleration voltage of 180 kV and a current of 5.55 mA. The reconstruction of the images utilized the KVD-1PI algorithm [Katsevich 2002].³ We scanned five samples at a time and cropped out each sample.

A.1.4 Cut surfaces. We used two different types of cuts, shown in Figure 31.

A.2 Inverse Modeling Method

A.2.1 2D Annual ring localization from photographs. As for the image translation (Section 5.1), we train a U-Net on all training data. Since the training data is relatively limited (912 images after setting aside 80% for evaluation), we train on patches. Specifically,

³Alexander Katsevich. 2002. Theoretically Exact Filtered Backprojection-Type Inversion Algorithm for Spiral CT. SIAM J. Appl. Math. 62 (2002), 2012–2026. <https://api.semanticscholar.org/CorpusID:45126757>

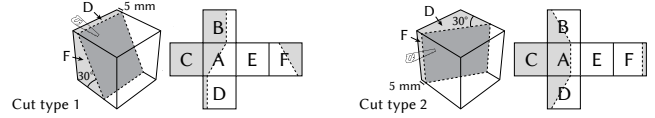


Fig. 31. A sub-set of cubes are split in two parts to photograph the cut surface. There are two types of angular cuts. Type 1 (left) cuts through the principal pith axis. Type 2 (right) is placed along the principal pith axis.

after setting the full image size to 256^2 pixels, we train on patches of size 64^2 pixels (corresponding to 1.0^2 cm) randomly sampled with varying location and orientation (16 patches per image gives $16 \times 912 = 14,592$ patches in total). In addition to increasing the amount and diversity of the data, this approach enables input images with varying aspect ratios, which is crucial for handling input wood blocks with various sizes and proportions. During inference, we evaluate the network on a sliding window and reconstruct an image by blending the overlapping areas.

We set the batch size as 256 and number of epochs to 2,000 and use an adaptive learning rate, starting at 0.005 and decreasing linearly to 0.0005 until the end. The learning rate is updated every 100 epochs. During the last 200 epochs, we lower the batch size to 128, and lock all but the last layers.

A.2.2 Fitting a 3D procedural growth field. The system was implemented using PyTorch and uses Adam [Kingma 2015]⁴ as optimizer. We set the number of height levels n_h to 8, the number of spokes n_a to 16, the number of rings n_r to 16, which were experimentally found to be effective (refer to B.2).

A.2.3 Texture synthesis approach 1: inverse procedural texturing. Below is a list of the parameters of the detailed procedural model (additional to the global structure parameters). The total count is 436 (or 630 with one knot).

- Earlywood color parameters: a base color (size 3) and a 1D color map for outward variation (size 64×3).
- Latewood color parameters: same as above.
- Side color: a color for each external surface (size 6×3).
- Early-to-latewood color transition parameters: start, end, smooth blend, linear blend.
- Fiber parameters: size, color strength, darkness strength, strength on late wood.
- Pore parameters: radius, size of base cell (width/length and height), color (size 3), anisotropy strength, general occurrence ratio, latewood occurrence ratio, occurrence–annual ring correlation factor, scale–annual ring correlation factor.
- Ray parameters: radius, size of base cell (width, length and height), color (size 3), occurrence ratio.
- Knot parameters: 1D color map (size 64×3), and anisotropy strength, knot-to-stem color transition smoothness.

A.2.4 Texture synthesis alternative 2: neural cellular automata. We train our neural cellular automaton (NCA) on a 256^3 cube using a single H100 NVIDIA GPU with mixed-precision to save memory. The NCA maintains 12 channels per cell, and its MLP adaptation

⁴Diederik P. Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In iclr, Yoshua Bengio and Yann LeCun (Eds.). <http://arxiv.org/abs/1412.6980>

network has 128 hidden neurons, for approximately 11, k parameters. We train the model for 1,500 epochs with a batch size of 1, applying the style loss averaged over all 6 cube faces.

To stabilize training, we adopt a pooling technique with a pool size of 32, alongside the overflow loss from Niklasson, et al. [2021].⁵ Gradient checkpointing further helps reduce memory usage, trading off additional compute. We use the Adamp optimizer with an initial learning rate of 0.001, decayed by a factor of 0.3 at epochs 700 and 1,200. We additionally employ the grafting technique introduced by Pajouheshgar, et al. [2024]⁶ to improve training convergence. We first train on sample "B08" then initialize subsequent trainings with the learned weights from that model.

After training completes, we unroll the NCA for 200 update steps and take S^{200} as the final state. We then add these RGB channels to the base color B to obtain the final volumetric texture.

B EXTENDED EVALUATION

B.1 Pith Estimation

Pith estimation refers to the first step in the differentiable texturing framework: estimating the pith origin (O) and its direction (V , refer to Section 5.2.3). Evaluating pith estimation is difficult without ground truth (GT) data, which is available only in samples that include the pith (e.g., SG-09, refer to Figure 6a). In the *Mokume* dataset, most samples—and all test samples—do not include the pith. However, the accuracy of the pith estimation can be systematically evaluated by comparing the shape of the estimated growth field with photographs of the actual sample by observational validation.

Therefore, we evaluated the pith estimation as follows. We ran the pith inference (optimization step 1) against contours extracted from the GT annual ring localization images created from manual annotations (refer to Figure 12), and verify by observation that all output growth fields are reasonable approximations of the real wood patterns. We consider these output pith poses as a "pseudo GT." Subsequently, we ran the pith inference against contours extracted from the annual ring localization images created by image translation (refer to Section 5.1), and measured the difference compared to the pseudo GT (Figure 32). The results show that the distance between pith origins range from 0.0 to 49 cm with a median average of 1.4 cm (Figure 32-top), and the angle between the pith direction vectors range from 0.1 to 85 degrees with a median average of 3.0 degrees (Figure 32-middle).

Relatively poor estimations (B-10, BW-06, CN-03, H-11, KR-07, MZ-03, N-01, NR-04, P-10, RO-03, S-02, SG-06, TC-04, TC-05), can largely be attributed to poorer quality of the U-Net generated annual ring localization images. However, three samples (BW-06, N-01 and NR-04) stands out, having exceptionally high distances between the estimated pith origins (Figure 32-top). In these three samples, the pith origin is far from the cut-out region, resulting in nearly parallel annual rings and ambiguity in pith orientation. Measuring the pith direction relative to the sample center confirms this, with differences of about 180 degrees, indicating opposite orientations

(Figure 32-bottom). This flipped pith origin still offers a reasonable approximation of the volumetric shape of the annual rings in the sample. Possible problems include the modeling of features that are related to the direction of growth, such as knots and pores with direction-dependent distributions.

The existence of the challenging cases raises the question, are there other such cases that are possibly inferred wrongly in the pseudo GT? To analyze this, we go through all samples and manually identify 9 samples (IC-04, IC-06, N-01, NR-04, NR-10, P-06, P-10, S-02, TC-04) with close-to-parallel annual rings. Although it might be possible to analyze the details of the texture in the photographs to determine if the inferred pith is flipped or not, we find it difficult to make a definite conclusion for these samples and mark them as ambiguous (red colored sample names in Figure 32).

Pith estimation: ours compared to pseudo-GT

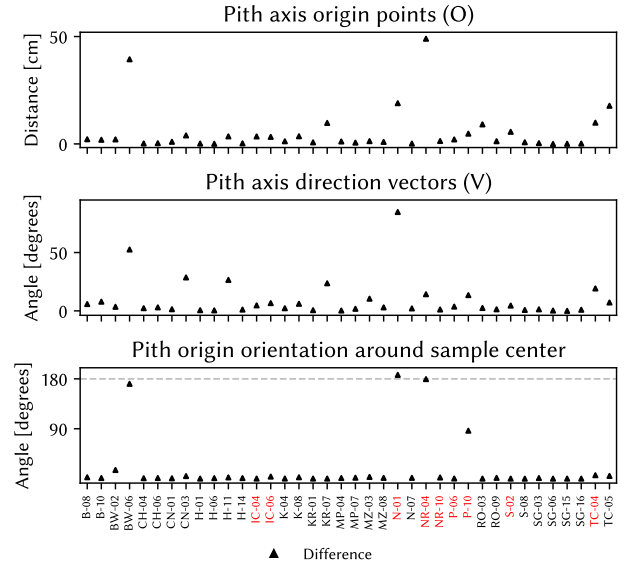


Fig. 32. Pith inference (optimization step 1) analysis comparing pith poses inferred based on edges extracted from annual ring localization images created from annotated data (pseudo GT) and those inferred based on edges extracted from annual ring localization images created automatically by image translation (ours). Samples where the pith locations lies far from the sample center and which thus have nearly parallel annual rings are marked by red, indicating that these are ambiguous cases where we were unable to verify if the origin direction is in one direction or the opposite.

B.2 Resolution of R

The resolution $n_h \times n_a \times n_r$ of R is fixed during the optimization. There is a trade-off between quality (higher resolution allow for more precise approximation) and efficiency (lower resolutions allow for faster calculation and lower memory storage). We performed an experiment evaluating different combinations of n_h , n_a , and n_r . For each of the three resolution axes, we tested five different resolutions—2, 4, 8, 16, and 32—while keeping the other two axis fixed at 8. For each combination and for all 38 test samples, we ran

⁵Eyvind Niklasson, Alexander Mordvintsev, Ettore Randazzo, and Michael Levin. 2021. Self-organising textures. *Distill* 6, 2 (2021), e00027–003

⁶Ehsan Pajouheshgar, Yitao Xu, Alexander Mordvintsev, Eyvind Niklasson, Tong Zhang, and Sabine Süsstrunk. 2024b. Mesh Neural Cellular Automata. *ACM Trans. Graph.* (2024). <https://doi.org/10.1145/3658127> Ehsan, Yitao Xu, and Sa

optimization step 2, and recoded the output losses and computation times. For different height level resolutions, the loss decrease until resolution 8, and then remain rather constant, while computation times are approximately constant throughout, or even decrease slightly (Figure 33-top). For spoke and radius resolutions the losses decrease with higher resolutions, at least until 16, while computation times increase rather linearly (Figure 33-middle and bottom). The relative higher importance of spoke and radius resolution is consistent with our observations that distortions of annual ring patterns are stronger in cross-planes compared to vertical planes. Based on these results, we chose the resolution setting $n_h = 8$, $n_a = 16$, and $n_r = 16$ in our implementation.

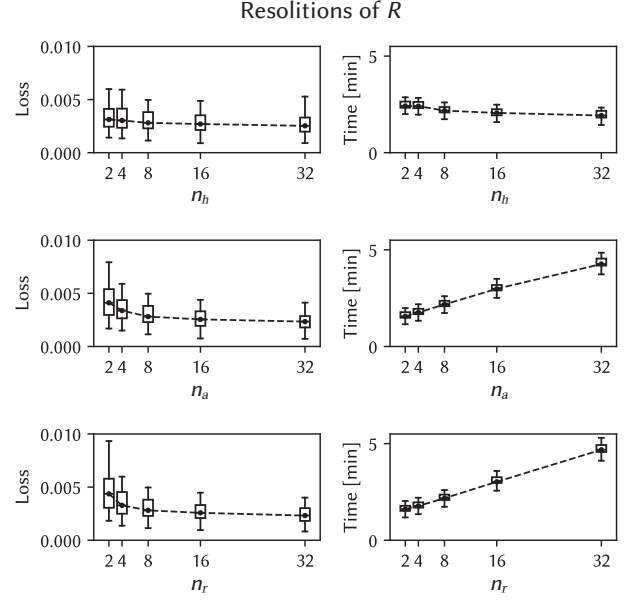


Fig. 33. Plots of losses (left) and computation times (right) considering different values for n_h (top), n_a (middle), and n_r (bottom). For each of the three resolution axes, we test the five values 2, 4, 8, 16, and 32, while keeping the other two axes fixed at 8. Each box shows the distribution of losses or computation times for the 38 samples in the test dataset.