BOTANICAL TREE PERCEIVED REALISM METRIC

by

David Hrusa

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Computer Graphics Technology

West Lafayette, Indiana

May 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Bedrich Benes Ph.D, Chair

Department of Computer Graphics Technology

Christos Mousas Ph.D

Department of Computer Graphics Technology

Tim McGraw Ph.D

Department of Computer Graphics Technology

Approved by:

Nicoletta Adamo-Villani

ACKNOWLEDGMENTS

Many thanks to Bedrich Benes, Tomas Polasek and Martin Cadik for contributing to this research in all aspects of it, dedicating over two years to see it through.

I also deeply appreciate the support of Tim McGraw and Christos Mousas, both in classroom and as members of my graduate committee.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In this thesis a novel metric for evaluating botanical plant perception is proposed. Firstly, utilizing diverse data sets of both synthetic and organic plants a unifying representation is found. This makes the presented approach generalizable to any type of tree model.

Large data set of tree models from various sources both realistic and synthetic was assembled. Through a comparative perceptual study, every plant was sorted, and following an analysis labeled in terms of apparent realism.

Since the specific parameters contributing to realistic perception are not fully understood, this research identifies key features based on observed data. Deep neural network and decision tree classifiers were trained and provided significant results on the data set and can predict perceived realism on new data as illustrated by a secondary study.

1. PROBLEM AND PURPOSE

1.1 Overview

1.1.1 Introduction

The HPCG Laboratory at Purdue has a long history of publishing in the field of botanical trees. Following prompts by Stava et al., 2014 for future work, it occurs to that further insight could be gained due to uneven distribution of existing body of work. Stava et al. have closely studied properties of trees in their work on differences between tree structures. Their findings could serve well even while not comparing two trees against each other.

1.1.2 Statement of the Problem

Simulation of botanical plants has always been a significant topic in computer graphics. However, most applications of this research either focus on large areas of greenery with little focus on individual plant structure (Beneš et al., 2009; Côté et al., 2009) or are intended for the entertainment industry and realism is not a primary focus for them.

Techniques for modeling botanical trees have been known for many decades now and are progressively being improved upon (Aono and Kunii, 1984a). Gradual improvements to representing the influence of environment and underlying biological nature of trees have been made over the years (Livny et al., 2011; Stava et al., 2014). The current trend in simulation is recreation of natural phenomenons which drive growth of plants. Nutrient transport and light exposure play a significant role as opposed to the early fractal and manual approaches. (Oppenheimer, 1986)

An underlying theme of these works is the search for realism. Comparative metrics for measuring distances between plant structures have been proposed by Stava et al., 2014. However, there are few works attempting to link physical features of plants to perceived realism. It could be argued that perception is the most significant part of computer graphics in its pure form.

Meanwhile, in the general field of classification, principal component analysis and projections onto variable subspaces have been a popular approaches to extracting significant

Figure 1.1. Example of a rendered tree skeleton used in the user study.

features and labeling them. Such approaches have been even applied in other fields of graphics such as terrain analysis (Rajasekaran et al., 2019), however, never in the case of plants.

The problem addressed by this thesis is the limited understanding of the causality between botanical tree features and the perceived realism of trees. Understanding of plant perception has not been conclusively defined while intuitive measures are regularly used to compensate for the limited understanding.

1.1.3 Significance of the Problem

In the conclusions of Rajasekaran et al., 2019 it is stated that their experiment design was sound, however terrains have proven difficult to generalize, as they are structures formed over billions of years through varied means (fluvial, thermal, hydraulic, wind erosion, plate tectonics etc.).

Trees on the other hand are well understood down to the cellular level. Utilizing this trial format in a more appropriate setting could produce results that significantly transform the methods of computer perception and provide valid and reliable metric. Furthermore, a metric of this sort could transform the work in entertainment industry and shape it to be situated around quantifiable measures of various emotional responses to digital creations. This would serve as a valuable tool to creators to better predict audience impressions without the expenditures of conducting a user study every time a change is made.

1.1.4 Statement of the Purpose

The purpose of this study is to formulate a metric indicating the perceived level of realism in botanical plants. It will be determined which parameters of a tree structure and observed appearance influence the mean perception of a tree and whether there is a definite relationship to be found. The significance of these parameters will be found through projecting each examined tree onto a multi-dimensional space of parameter values. Human subjects will be prompted to provide their perceived assessment in a comparative study. These results will be utilized to determine an area of perceived realism where trees are predicted to be treated as realistic by observers.

1.2 Research Topic

1.2.1 Research Questions

The question which generates the necessary deliverables to serve the purpose is **Which measurable properties of tree structures influence their perceived realism the most?** Additionally, how can these features be effectively combined into a single classifier? How many features are necessary for reliable prediction?

1.2.2 Significance of the Research Questions

The hypothesis - that such region exists - is significant, because it offers unprecedented insight into an intuitive property that has been successfully measured in other fields of graphics (Rajasekaran et al., 2019), yet remains implied in many present state of the art publications. Plant growth simulation is a topic still currently being actively explored (Thompson et al., 2011) and has a lot of fairly open ended problems. (Qin et al., 2020)

The research question requires the conductor to be familiar with the biological background of plant growth to a degree as to be able to propose and comprehend the relationships being modeled. However, in practice perceptual studies are often conducted with just very rudimentary knowledge of the area modeled. Visual papers often rely on appearances over substance.

The proposition of a novel metric in a still expanding field could have substantial positive impacts on the outcomes of future academic work. Other comparable studies rarely combine synthetic and realistically measured trees at such scale or attempt to draw comparisons in relation to their structure and features. The sheer scale of the dataset gathered, cleaned up and labeled could on its own also serve as a substantial academic contribution for other research teams.

1.2.3 Definitions

The important topics when analysing concepts for the research will be located in several loosely connected categories. The first group will be features significant in botanical trees, since understanding the nature of these will be significant for potential feature detection. Secondly, topics related to general computer perception should be considered. Perception analysis applies to numerous fields of science and many proven procedures are bound to be established. Finally, the area of computer simulation should be researched since that pertains to the methods which will be utilized in generating tree models.

- Phyllotaxis – ”the arrangement of branches (or leaves) on the main axis (stem).” (Aono and Kunii, 1984a)
- Axiality – ”the quantity that represent how clearly the main axis can be detected.” (Aono and Kunii, 1984a)
- Apical dominance – ”the quantity that represents external appearance of the tree.” (Aono and Kunii, 1984a)
- LiDAR scans – environment laser scans gathered in slices around the device. (operational definition)

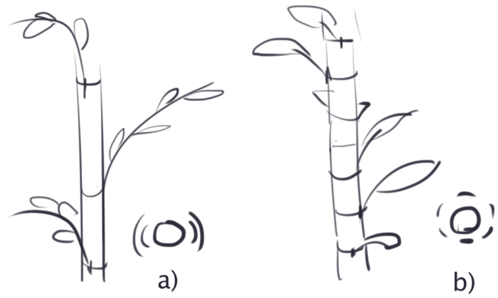

Figure 1.2. Two different phyllotaxis layouts. Branches spiral around the central stem. a) spirals with increments of 180 degrees. b) spirals with a step about 45 degrees.

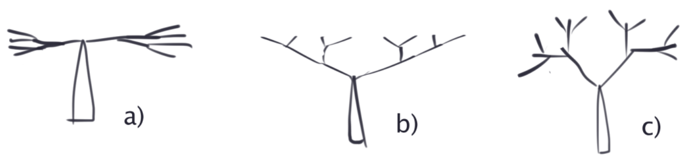

Figure 1.3. Increasing amounts of axiality in the tree structure. (a) minimal to c) maximal). Low axiality graphs are spatially flattened. High axiality graphs have identifiable branching angles.

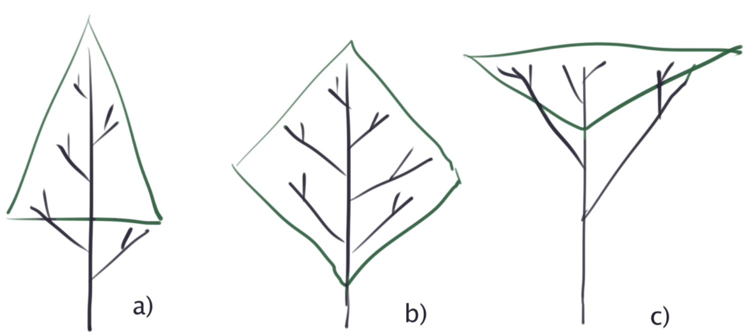

Figure 1.4. Examples of different apical dominance. a) is cone shaped, b) intermediate and c) is inverted cone.

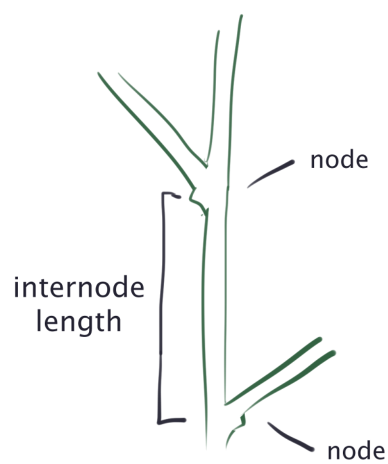

Figure 1.5. Internode length shown as the distance between two structural nodes of the tree graph. Any branching or sprouting point is considered a node.

- Internode length – the distance between two nodes of a tree graph. (operational definition)
- Human subjects – People undertaking a user study. (operational definition)
- Pruning factor – ”the impact of the amount of incoming light on the shedding of branches.” (Stava et al., 2014)
- L-systems – a form of representing graph structures through offset rules. (operational definition)
- Phototropism and Gravitropism – ”The impact of the average direction of incoming light and gravity on the growth direction of a shoot.” (Stava et al., 2014)
- CycleGAN – ”Cycle-consistent Adversarial Networks ... an approach for learning to translate an image from a source domain X to a target domain Y in the absence of paired examples.” (Zhu et al., 2017)
- Feature vector – a vector of various metrics that contribute to perception. Commonly used to represent a particular object in computer vision. (operational definition)
- Human visual system - Is the psychophysical framework describing the ways in which humans receive and parse visual information. It is constantly evolving with new information about the functioning of human body. (Saghri et al., 1989)

1.2.4 Assumptions

The most substantial assumption regarding this project is the hypothesis itself. It is possible that there is no objective metric to indicate based on observable features whether or not a tree appears realistic. It has to be factored in as a possible outcome.

In a more general sense, it would be even possible that the perception of trees is entirely subjective. This however is highly improbable as perceptual research in numerous other fields indicates a certain degree of shared perception emerges in perceptual research.

”In the experiment, we displayed the set of 10 rendered test images ... The subjects are seemingly able to mark strong artifacts quite accurately without seeing the reference, while for perceptually weak artifacts, the reference is needed." (Herzog, Cadik, et al., 2012) Other works have implemented the same approach to comparably successful results. (Rajasekaran et al., 2019)

Additional assumptions are that the respondents will answer truthfully and that the newly introduced skeletal representation reasonably expresses all major features of the studied trees, in the attempt to unify the format of the data while dealing with sources from all around the internet.

1.2.5 Limitations

Herzog, Cadik, et al., 2012 notes several steps they have taken in their work which during this study are bound to be limitations. Namely, they calibrate their monitor and ensure that their respondents have accurate eyesight and are well seated. Due to needing large volumes of data to serve a machine learning approach, this study relies on online survey techniques through third parties. This renders ensuring these conditions impossible by design, though it does not prevent these techniques from being widely used in computer graphics. (Farid and Bravo, 2012)

Another area of limitation is the availability of data. It is possible that only a subset of realistic trees is adequately represented in the overall pool of data on the internet. Furthermore, many models lack leaves, showcase inconsistent topology and overall display characteristics which make comparisons between various sources biased.

Finally, the inherent limitation of the proposed approach is that there is no exhaustive list of tree features. It is possible that there might be a property which simply was not considered and plays a significant role. This limitation is slightly mitigated by the fact that if a property is omitted, it most likely is difficult to compute and thus impractical in application.

1.2.6 Delimitations

In the ideal scenario, in-person controlled study and new data set of scanned trees would be produced from scratch. It could be argued as a consequence that the aforementioned issues are in fact delimitations, however since their solution would involve additional steps being added to methodology, they are not classified as such.

A delimitation of this study's design is the lack of realistic trees across the proposed state space being investigated to serve as ground truth. In the final step, this study is using synthetic botanical trees to sample the proposed state space of features. These trees do not have realistic counterparts due to scope which limits the possible extrapolation of results back onto the real world pool of trees. However, this does not take away from the generalizablity of the findings as the studied sample is any plausible tree structure, not just realistic ones.

1.3 Summary

This chapter introduced the background and context for pursuing a perceptual quality metric. Feature extraction and correlation had been successfully applied in other areas. Key terms have been explained and the primary goal of the study is to identify a metric for evaluating the perceived realism in botanical plants.

It is assumed that there is a relationship between some set of measurable features and the resultant level of perception. No major delimitations stand in the way of conducting any of the required experiments and surveys.

2. REVIEW OF LITERATURE

2.1 Perception

Perception in computer graphics always depends on the conditions stemming from the limitations of Human Visual System as laid out by Ferwerda et al., 1996.

Especially in this particular project, understanding the properties of greyscale images is significant. Prior research has been conducted by Cadik et al., 2007 and K. Smith et al., 2008.

Following that, Grill-Spector, 2003 discuss the decision making process happening during the perception and decoding of objects in the environment. Herzog, Cadik, et al., 2012 introduce a metric for evaluation of image quality without required reference and provide insights on perception limitations in computer graphics.

While exploring computer graphics, studying human perception has proven very beneficial in enhancing various applications such as photo-realistic rendering (Weier et al., 2017), optimizing 3D rendering (Reddy, 1997, 2001), predicting popping artifacts (Schwarz and Stamminger, 2009), tone mapping and high dynamic range imaging (Reinhard et al., 2010), rendering and animating humans (O'Sullivan et al., 2004; Shi et al., 2017), flow of fluids (Bojrab et al., 2013; Um et al., 2017), crowds (H. Wang et al., 2016, 2017). Notably for the purposes of this study, the explicit appeal of subjects has also been measured in photographs (Aydın et al., 2015), fractals (Spehar et al., 2003) and even paintings of fractals (Taylor et al., 2011).

An important concept in relation to perceptual studies are the image quality metrics (IQM) which allow for quantification of perceptual ratings, thus enabling a quantitative study to take place. IQM are subdivided into several types. Full-reference IQMs are generally applied during comparative experiments in image editing operations. They take a before and after image and quantify the difference in perception. (Mantiuk et al., 2011; Z. Wang et al., 2004; Wolski et al., 2018) No-reference metrics on the other hand operate without a reference image for comparison. (Bosse et al., 2018; Herzog, Čadík, et al., 2012; Jung et al., 2002; Moorthy and Bovik, 2010; Suresh et al., 2009; Tang et al., 2011; Ye et al., 2014)

Figure 2.1. a) full-reference IQM comparing a before and after image. b) no-reference IQM scoring a single provided image. c) VQM rating the quality of a sequence of images. If the frames are treated as individual images then VQM could be considered a form of full-reference quality metric despite only receiving a single input.

An extension of these metrics are also Video Quality Metrics which factor in the element of change over time. (Aydın et al., 2010; Winkler and Mohandas, 2008) A common application of these metrics is in studying the effects of compression and rendering artifacts. (Andersson et al., 2020; Aydın et al., 2010; Čadík et al., 2013; Jung et al., 2002; Ramanarayanan et al., 2007; Z. Wang et al., 2004) This means that they are not directly suited for scoring stand alone realism of a studied image. A recent development in the field are deep features by Zhang et al., 2018 which have been shown to provide much stronger results.

The field of 3D model perception is increasingly more complex and has not yet been as deeply and thoroughly explored as the one of 2D images. Several approaches have been demonstrated for raw models (Guo et al., 2015; Lavoué et al., 2016; Nader et al., 2016; Rogowitz and Rushmeier, 2001) as well as textured. (Guo et al., 2016; Yixin Pan et al., 2005)

When it comes to applying 2D metrics to 3D setting there are certain limitations. According to Rogowitz and Rushmeier, 2001 the complexities of a three dimensional object cannot be fully expressed in a flat image when shadows and light are applied. In the findings of Yixin Pan et al., 2005 it is further demonstrated that texture can be more perceptually significant than geometry of objects. Despite that, in the works of Lavoué et al., 2016

IQM have been successfully and accurately used to study the quality of an object under one constant type of distortion effect.

2.2 Plant Simulation

Botanical plant modeling remains a significant topic on modern computer graphics. Data may commonly be represented through skeletal structures, surface meshes, LiDAR point data sets or voxel modeling. (Pirk et al., 2014; Pirk et al., 2012)

The very early approaches to plant simulation analyzed fractal patterns (Aono and Kunii, 1984b; Kawaguchi, 1982; Oppenheimer, 1986; A. R. Smith, 1984) and rule based systems. (Honda, 1971) The current trend is to simulate the biological processes which guide the growth of plants in the nature. The initial representations represented plants as L-systems. (Prusinkiewicz et al., 1993; Prusinkiewicz and Lindenmayer, 1990) Subsequently, growth over time was introduced to these models. (Měch and Prusinkiewicz, 1996; Prusinkiewicz et al., 1994) as well as the influence of the surrounding environment. (Deussen et al., 1998)

The overarching theme in the environment methods is the supply of nutrients and physical restrictions. (Beneš and Millán, 2002; Greene, 1989) Multiple plants can also influence and shield each other. (Palubicki et al., 2009) These simulations have been shown to also work in real time by Hädrich et al., 2017. Additional environment factors may be introduced such as wind (Habel et al., 2009; Pirk et al., 2014) or fire (Pirk et al., 2017).

Some current approaches do not try to simulate the biology of trees and take a geometric approach. The X-frog method uses the recursive nature of certain aspects of plants to interactively generate them. (Lintermann and Deussen, 1999) Such approaches may combine possible additional inputs such as sketching. Sketching lets users draw rough outlines of the plant and uses constraints and rules to fill in the plant structure. (Anastacio et al., 2006; Ijiri et al., 2006; Okabe et al., 2007; Tan et al., 2008)

The other option for generating plant models is recreating those from the real world. Reconstruction of skeletal structures from mesh representations and point clouds remains an open topic. Point cloud approaches allow for fast gathering of substantial volumes of data, however in general are very susceptible to incomplete data. (Livny et al., 2010)

Especially the lowest tier branches appear to be near impossible to capture with LiDAR scans at conventional resolutions. General purpose skeletonization approaches do exist. However, despite significant achievements in compensating for lacking data or skewed point clouds, such methods only aid in reconstruction of larger shapes and cannot compensate for fine structures. (Huang et al., 2013; Qin et al., 2020) Lobe representation is one way to circumvent this lack of scanning resolution in botanical plants specifically. Nonetheless, it depends on knowledge of the plant specie and is not suitable in the approach presented here. (Livny et al., 2011)

Côté et al., 2009 bring up the significance of considering the influence of wind and other sources of distortion in LiDAR scan data. Since it is impossible to isolate these discrepancies, especially at the aforementioned fine branch level, reconstructive methods and artificial oversampling removal need to be employed. Furthermore, mature plants in their natural environment are often missing branches or have grown irregularly due to light conditions and obstructions in their surroundings. (Runions et al., 2007)

The most promising foundation for this experiment has proven to be the AdTree method for automatic reconstruction. Unlike many other rule-based and manual approaches, AdTree offers a stable and automatic method for reconstruction of meshes from point clouds. (Du et al., 2019)

In the subsequent step it is proposed that the state space of all the measurable metrics of a botanical tree will be searched for relationships using an analysis of variance or other more suitable approaches depending of the nature of the data gathered. As discussed in the Limitations section. There is a high chance that each tested subject will not provide a full rating for each plant pair and will therefore require an advanced statistical treatment as presented by Perez-Ortiz and Mantiuk, 2017. An example of such space exploration can be observed in the work of Liu et al., 2015.

2.3 Experiment Design

It can be challenging to measure the perception of human subjects on a numeric scale. Impressions of individuals are highly subjective and their interpretation of a rating system may vary. A commonly used way of converting subjective values to a measurable scale is a comparative experiment. Instead of requesting a numeric rating, users are presented with pairs of data and prompted to select their preferred option. Subsequently, an average normalized rating can be extracted from this data. (Rajasekaran et al., 2019) (Rubinstein et al., 2010) One such score is called just-objectionable-difference where one JOD unit indicates that 75% of the cases favored one over the other. (Perez-Ortiz and Mantiuk, 2017)

2.4 Summary

In summary, human perception is not yet fully understood and formulated, however quantification approaches do exist. Image Quality Metrics are a versatile tool for assigning ratings to comparative data, but have to be utilized carefully since their primary use case is to compare before and after scores. Processing 3D model data directly is not a straight forward tasks and extending 2D methods to renders of 3D meshes is susceptible to strong distortion due to the influence of shadows or textures.

Numerous representations of tree structures exist. In reconstruction, low tier branches prove very difficult to capture reliably, therefore it is not easy to define what constitutes pure real-world data. In the experiment step, qualitative responses can be converted to quantitative through normalization of comparative responses. Challenges associated with appropriate presentation of data are summarized within the Delimitations section.

3. METHODOLOGY

3.1 Overview

3.1.1 Summary

The research process involves several distinct steps. The overarching deliverable is a metric of appeal. Instead of guessing or mixing observable tree parameters randomly, the aim is to measure as many tree properties as possible, gather a substantial volume of data and then train a classifiers to match the opinions of the human subjects.

Before measuring any parameters, tree data is processed to a common representation using a skeletonization algorithm. A standardized TreeIO format is introduced which holds a conventional L-system representation alongside meta-data for convenience. From TreeIO parameters are calculated and stored for later. Every tree is also rendered in a realistic fashion using geometry reconstruction and soft shadows as shown in Figure 3.1.

Rendered trees are combined into all possible variations of pairs and presented to human subjects in a comparative perceptual study. Once evaluated, all the necessary information is known and classifiers can be trained as shown in Figure 3.2. In this exploration, decision forest and two deep neural networks have been selected.

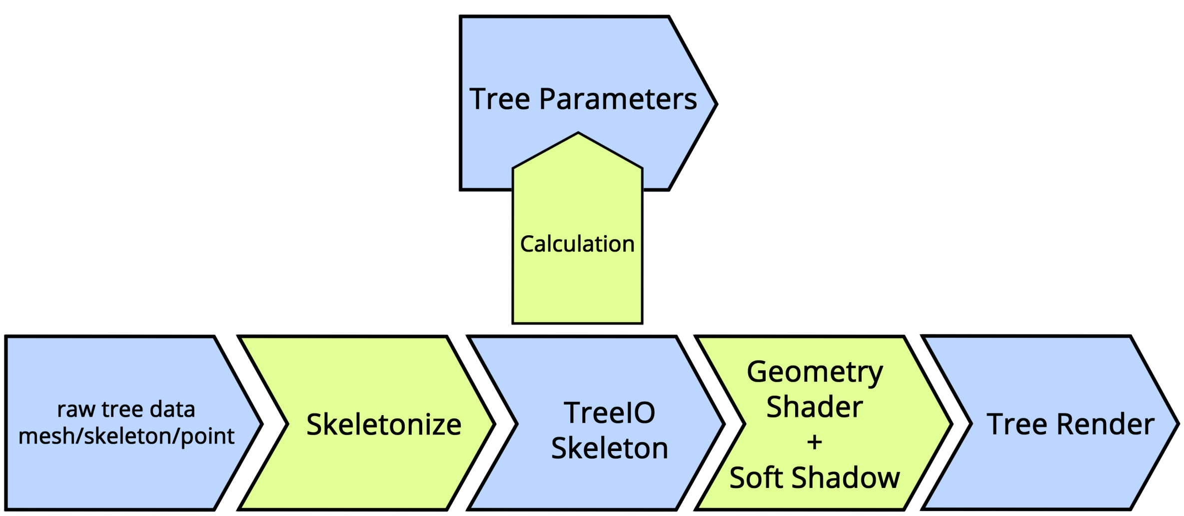

Figure 3.1. The initial step of processing takes in various tree data and yields parameters and renders for each one.

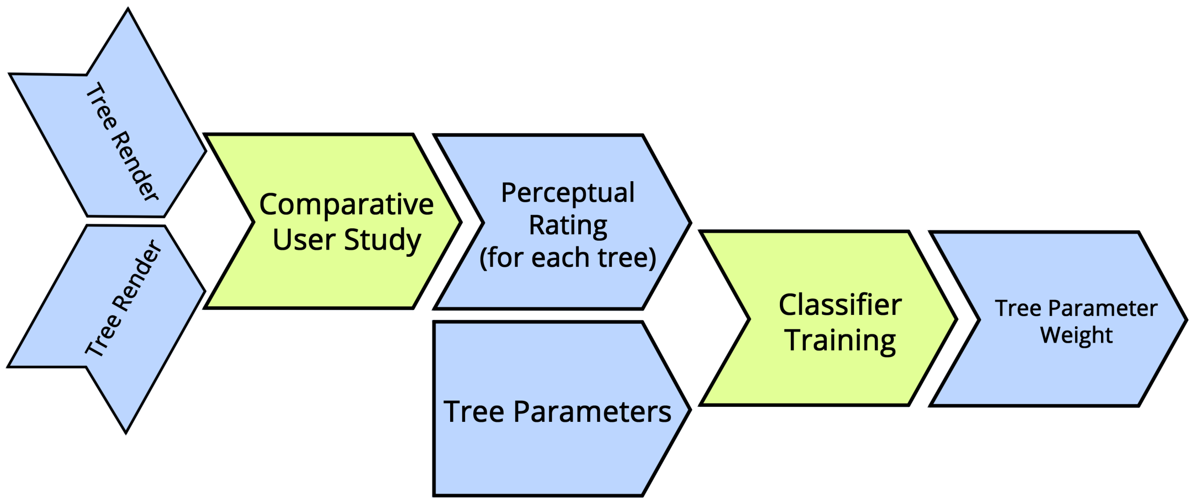

Figure 3.2. The second phase pairs two renders against each other, records the perceptual score and then matches that information with the parameters to determine the most significant ones.

Finally, from a trained classifier it is possible to deduce which properties of the tested trees play the most influential role in determining perceptual quality. If further tests or additional data support those parameters a robust predictor was found and a generalized response to the research question can be formulated.

3.2 Data Generation

The first step in conducting this study is the acquisition relevant data. Since the intent is to demonstrate a generalized method which is employable for both synthetic and organic trees, it is necessary to acquire a diverse data set.

In the case of synthetic trees, there is little difficulty. As explored in detail within the Plant Simulation section 2.2, a wide array of approaches exist to growing artificial plants with varied degrees of realism. Most of the known approaches are furthermore automatic, meaning they do not require human input and can therefore be utilized to generate near limitless amounts of testing data.

Table 3.1. Sources of real and synthetic tree geometry data. R - real, S - synthetic, 3D - 3D models, Clouds - Point Clouds

Data Source Name	Data Type	Papers
ModelNet	R 3D	N/A
LaserScan	R 3D, Clouds	Sloup et al., 2013
TreeParts	R 3D, S 3D	Xie et al., 2016
ReconTree	R 3D, Clouds	Livny et al., 2010
AdTree	R 3D, Clouds	Du et al., 2019
Shape Space	S 3D	G. Wang et al., 2018
InvTree	R 3D, S 3D	Stava et al., 2014
ShapeNet	S 3D	N/A
Multiple Views	R 3D, Source Codes	Guo et al., 2018

On the other hand, real tree data is much scarcer. Due to the inherently time and resource consuming nature of scan gathering and the aforementioned challenges of fine branch structural reconstruction, pure real tree data is difficult to gather in large quantities.

It is questionable to which standard a scanned tree can be considered an accurate scan. Any source besides a raw LiDAR scan is bound to have been reconstructed or approximated using one of the numerous algorithms. As this field is still not considered fully solved, state of the art techniques are the only option.

Since plant simulation is mainly utilized in either forest management or entertainment (Côté et al., 2009), there are currently no commercially driven efforts to gather this type of data outside of academic setting. Research papers proposing and expanding point cloud reconstruction approaches have proven as the most fruitful sources of such data sets. Lists of all included data sets along with their respective papers are included in Table 3.1.

3.2.1 Data Standardization

The collected data sets have shown large differences in modeling standards in meshes of trees. Some models were highly detailed with many segments per branch, while others were simplified and only recorded the general shape of the plant. Furthermore, certain data sets acquired were point clouds or skeletons. Skeletal representations suffered from similar discrepancies in sampling.

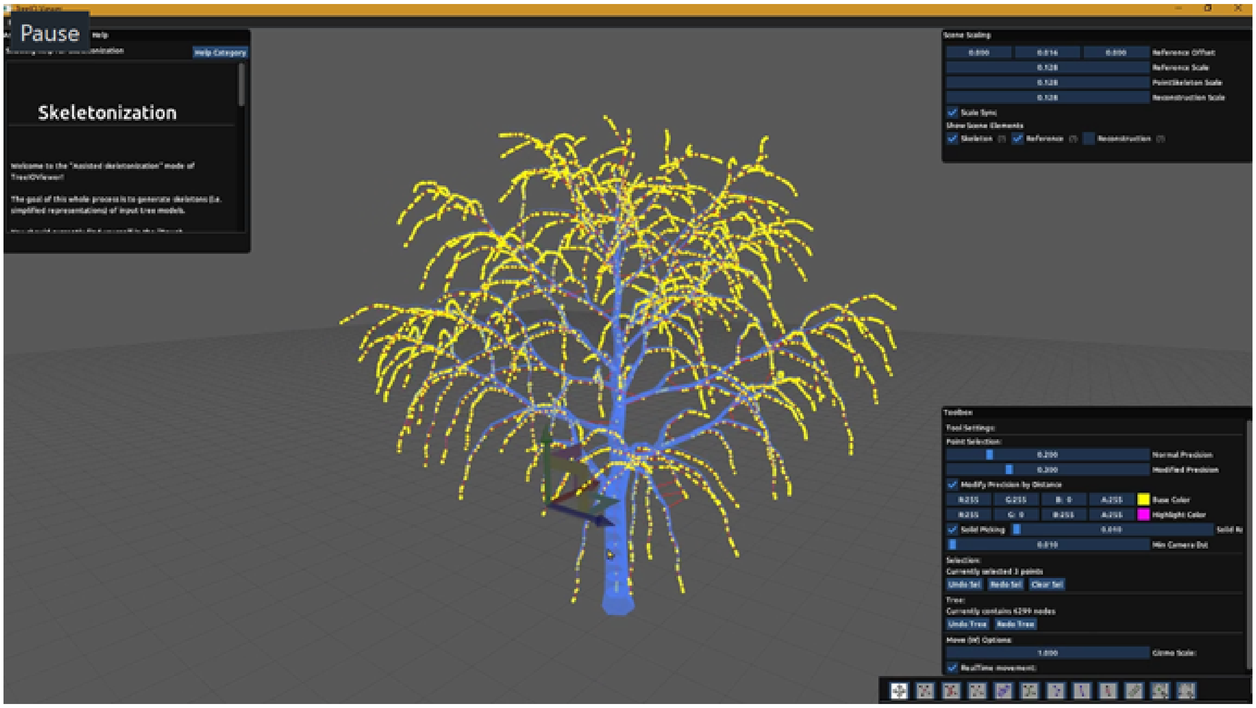

Figure 3.3. Interface of the TreeIO application used for processing the TreeIO skeletons. It provides useful to for retargeting subtree graphs and comparing the reconstructed shell (shown in blue) in real time.

Preliminary testing conducted on small samples of test renders has shown that human subjects favor more detailed trees on every occasion. In order to minimize bias stemming from uneven nature of data between real and synthetic, the sets have been filtered and trees that were either too young to show features or too old were minimized. They are still necessary points of reference, however, should not be oversampled.

Another step towards equalizing the inequalities in sampling was the re-sampling through the custom-built TreeIO application. An algorithm based on the core of AdTree (Du et al., 2019) was used to generate skeletons for all types of input data at equal granularity. Subsequently, TreeIO utilizes an improved version of cubic spline interpolation through OpenGL geometry shaders based on Stava et al., 2014.

The universal tree representation used in this paper is that of an oriented tree graph. Each node carries information regarding position, thickness and connectivity to children and parents. Thanks to the stable and definite geometry reconstruction procedure, such representation robustly expresses spatial shape of the plant without oscillations or instabilities. In cases where branch thickness is unknown, it is recalculated using the square ratio also known as ”the Da Vinci Postulate”. ”In the case where a mother branch of diameter w_1 gives rise to two daughter branches of equal diameter w_2, this postulate yields the equation $w_1^2 = 2w_2^2$, which gives a value for w_r equal to $\frac{w_2}{w_1} = \frac{1}{\sqrt{2}} \approx 0.707$” (page 57) (Prusinkiewicz and Lindenmayer, 2012)

Trees processed using the the automatic skeletonization algorithm are further manually cross-checked for inaccuracies. The algorithm reliably captures the shapes of the individual branches, but the most commonly occurring error throughout our experience has been incorrect attachment of branches in complicated or noisy parts of the mesh. Such occurrence can be quickly addressed using a graph splice operation controlled by the user. After three replacement points are selected, the faulty connection is severed. New connection is formed between the subtree and apropriate connection node within the primary graph. Finally, all connection within the subtree are reoriented to stem from the new subtree origin node.

3.2.2 Feature Generation

To support classification, features are gathered from the trees within the data set. The features are separated into several categories depending on the property which they analyze. **Local Featues** examine the structure of the tree graph. This generally means lengths of segments, angles in branching locations, straightness of individual sections and notably also the ratios of these aforementioned values between child and parent nodes within the hierarchy.

While analyzing tree graph structures the representation is grouped into segments. A segment is path in the graph connecting a pair of nodes with other than two connections. In other words, if a given branch was skeletonized into smaller segments, it is beneficial to treat them all collectively as a single unit instead of analyzing the partial sections. The curvature of the given branch also holds valuable information.

Extracting such information as a metric poses a challenge though since many scalar values have to be condensed into a reasonable number of metric values. Therefore these metrics all are considered through statistical means as a series of metrics. The maximum, minimum, ratio between the two, median, mean, variance and quantile histogram brackets. The aim is to provide as many possible insights into the nature of the tree which could contribute to the final ranking.

Besides segments, chains are also identified within the structure. Chain can either be a segment or a path starting from a non-leaf non-two-connection node and ending in a leaf node or another junction. Between these chains numerous properties can be examined, namely when two chains share a point of origin sibling angles, differences in length, full-chain lengths, number of sub-graph children, rates of branching can be either calculated for each individually, as a ratio between siblings, children and parent resulting in a plethora of potentially viable features.

Notable are topological features introduced here, such as asymetry which compares the ratio of two sibling subtrees. The branching model metric indicates how closely the distribution of the graph matches one of the models proposed by De Reffye et al., 1988 among monopodial, sympodial monochasial and sympodial dichasial. Monopodial plants extend a single long stem with many branching sub-branches. On the other hand the major thickest growth section of sympodial plants extends through secondary branches in various alternating patterns.

Global Features examine the tree as a whole. Inspiration comes from approaches by (Stava et al., 2014) where radii of the tree crown are taken at various heights to express the shape of a given tree.

Graph features are still present, but summarized into single values. Internodal length expresses the ratio between different layers of the tree structure. Age of the tree can be estimated either by examining the trunk or leaves. In this research the age is computed as the sum of the maximum depth of a leaf and the length of the trunk over mean internodal length. The tree area is computed through revolving the tree in front of a ray tracer at small increments. The result is an intensity map image which should express the shape of the tree averaged out over all possible view angles. The values can be simply summed or sampled at different heights to retrieve widths.

Figure 3.4. A tree in the study shown from five different equally spaced angles.

3.2.3 Rendering and Presentation

In relation to the other observations made while processing the tree data, it is also essential to ensure that the form of presentation does not favor a particular body plan or specie. In order to reduce variance in perception, all images are rendered in grayscale with unified lighting at 45 degrees. Besides the fact that many of the data sources outright lack texture or color information,Rogowitz and Rushmeier, 2001; Yixin Pan et al., 2005 claim that artifacts and rendering errors can strongly influence the quality of perceptual results.

First, trees chosen for sampling are rendered, then the comparison data set is generated from all possible pairs. Sides are also equally randomized, so that each image has the same chance to appear either on the left or right. In an ideal scenario, every single user would rate every single pair. However, in reality workers often quit half-way through and if compensation is withheld until full completion, the number of applying workers drops significantly. For these reasons the order is randomized and balanced over the data. Incomplete sets make statistical analysis more complex, however Perez-Ortiz and Mantiuk, 2017 provide a reliable method for processing them.

Due to a combination of issues linked to fine branch reconstruction and overall limited pool of knowledge in this area, this study presents trees without leaves. Techniques exist to construct leaf coverage realistically, however, these techniques often rely on knowledge of the plant specie which cannot easily be determined for synthetic or even interpolated organisms. Implementing such techniques is a promising area for future work, but is beyond the scope of this project. (Livny et al., 2011) Additionally, representation of leave geometry is a topic that may produce an additional range of perceptual disparities. (Rushmeier et al., 2000)

Each tree is included in the data set at 5 equally spaced views to ensure that an unfavorable angle was not picked to skew the overall perception. The height of view is deliberately placed at ground level to simulate how people usually observe trees in their daily lives. Distance was chosen such that all trees can fit into view, with the scale range being normalized and outlying oversized trees scaled down. This ensures that most trees are relatively scaled. If extremely tall trees such as Sequoya were factored into the normalization, the rest of the plants would be too small on screen and many details would be lost. The number of views can further be augmented by seeded randomization of camera properties, specifically the local angle and height offset. This simulates imprecision in how a person would hold the camera and also combats aliasing, even though our render configuration involves multi-sampling.

The rendering algorithm uses soft shadows calculated using a light map. This approach offers a reasonable balance between convincing photo-realism and real-time rendering necessary to facilitate the other uses of the application. Additional depth-based fog is added to the images to increase contrast between rear and front branches. Since the structure of the tree is the main focus of the comparative study this is done to compensate for the improved stereoscopic perception a physical observer would have in such a scenario, were it to be conducted in real life. Soft shadows are also used to cast a shadow coming towards the viewer from the light at a 45 degree angle. The shadow grounds the scene and gives the space depth. The light position is always the same across the view angles, it is the tree which changes its rotation to offer a different perspective.

As an additional aid in further analysis modality render settings are also generated. These modalities include color coded tree parts, depth, shadow maps, normals and even raytraced parameters such as depth rays and volumetric imprint of tree geometry - the projected occluded volume behind the geometry. Using histograms, such data can be stratified and included as additional inputs to the classifier.

For the purposes of machine learning the 5 angles are sampled with jittering in yaw pitch roll and distance to simulate natural variance that would be present in photos or other inputs.

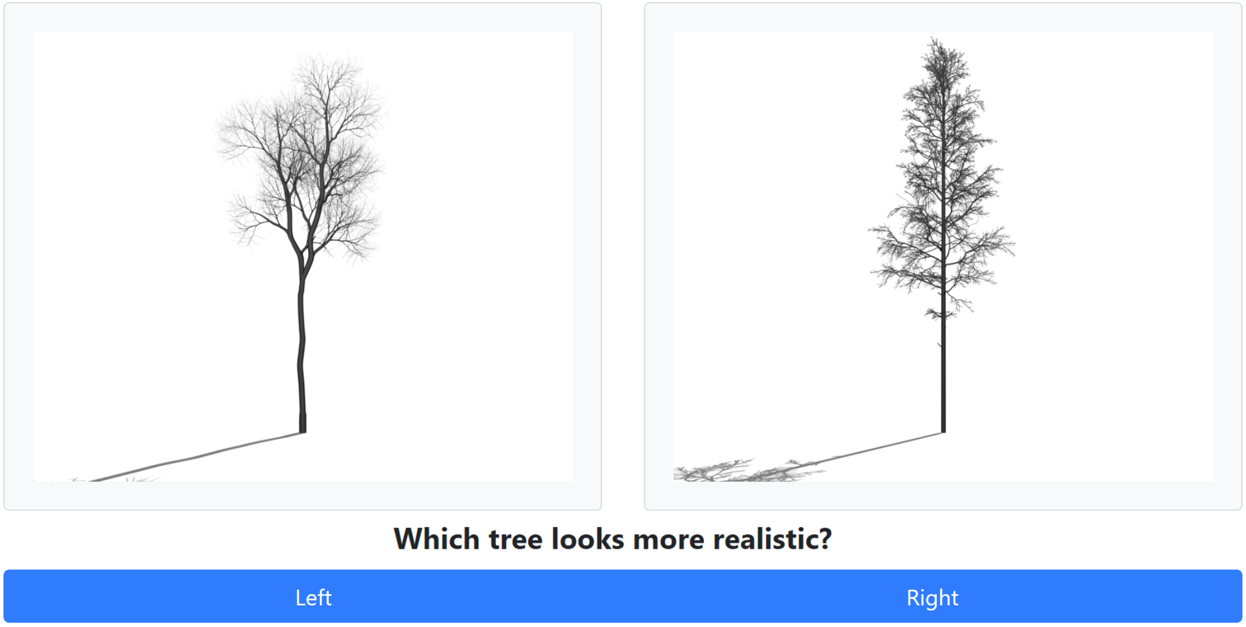

Figure 3.5. Comparison UI presented to the participants of the user study.

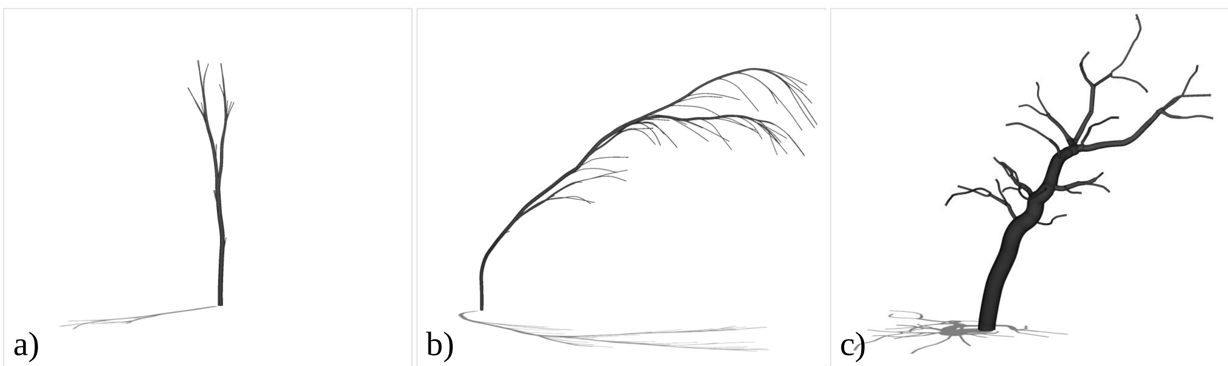

Figure 3.6. Trees showing extreme distortions or poor structure which would bias the user study. a) low amount of branching, b) strong environmental influence. c) old age.

3.2.4 Experiment Design

In order to acquire ground truth of human perception, a user study needs to be conducted. The nature of the study is comparative. The procedure used is based on the experience of Rajasekaran et al., 2019 and methodology of two-alternative forced-choice. (David, 1988) The users are be shown two trees side-by-side and asked ”Which tree looks more realistic?” In response they can press a button for either Left or Right.

Before selecting trees to use in the study, extremely distorted or poorly sampled trees are filtered out of the collection. These effects may occur as a result of extreme phototropism or other environmental influences. Since these trees have been shown to skew the perception of the viewers they have been removed.

The remaining 100 trees were matched into pairs with each other. With 5 views per each this results in 123, 750 comparison pairs to be tested. The relationship of being more realistic than another tree is not transitive and therefore each possible combination needs to be tested. This lead to 495 batches of 250 trees. Tree types were distributed equally across the batches to ensure that there would not be a batch of only conifer trees for instance. Furthermore, to ensure that trees are sampled multiple times the batches were generated ten times resulting in 4,950 unique batches.

The study itself is conducted using Amazon Mechanical Turk. Reasons for choosing this method over in-person study which would allow for better condition control are provided in the Previous Work section. In the study, subjects are presented with pairs of the aforementioned renders and asked to press a button to select which one of them appears more realistic. ”Turk Masters” were requested in the submission which should provide users who are reliable and verified. To avoid issues with reliability of results, each user is permitted to only respond to a single batch of trees.

3.2.5 Experiment Results

Upon inspecting the replies, problematic interviewees were identified, who solely chose a single option. Besides that, no serious issues arose. The resultant number of valid responses totaled 1,041,000 comparisons from 4,164 unique subjects.

Based on Amazon statistics, 50,86% were female and 49,14% make. The ages varied strongly, however 73,85% were between 20 and 40 years old.

In order to convert comparative data to a single rating number for each tree the just-objectionable-difference unit was used. (Perez-Ortiz and Mantiuk, 2017)

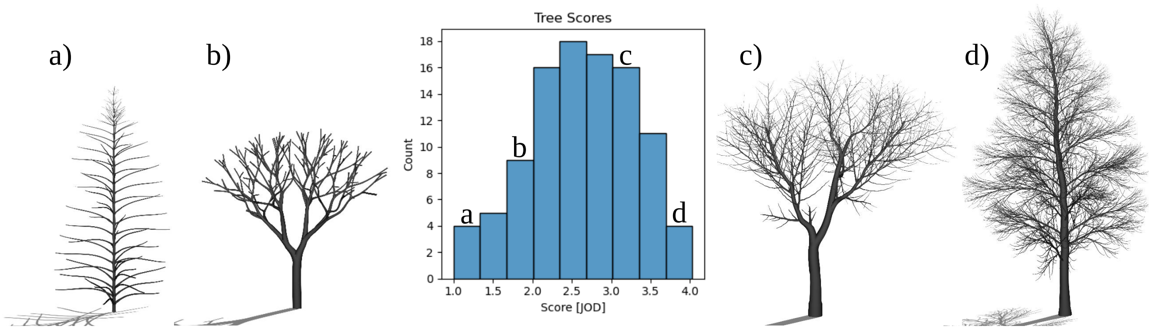

Figure 3.7. Examples of perceptually rated trees marked on a histogram of their entire set. Note how the poorly scored trees a) and b) show overly-even structure.

Just as predicted in the preliminary study, trees with complex structures scored better, however with added complexity the propensity to show abnormal or uneven growth increases which in turn decreases the quality ratings in many trees.

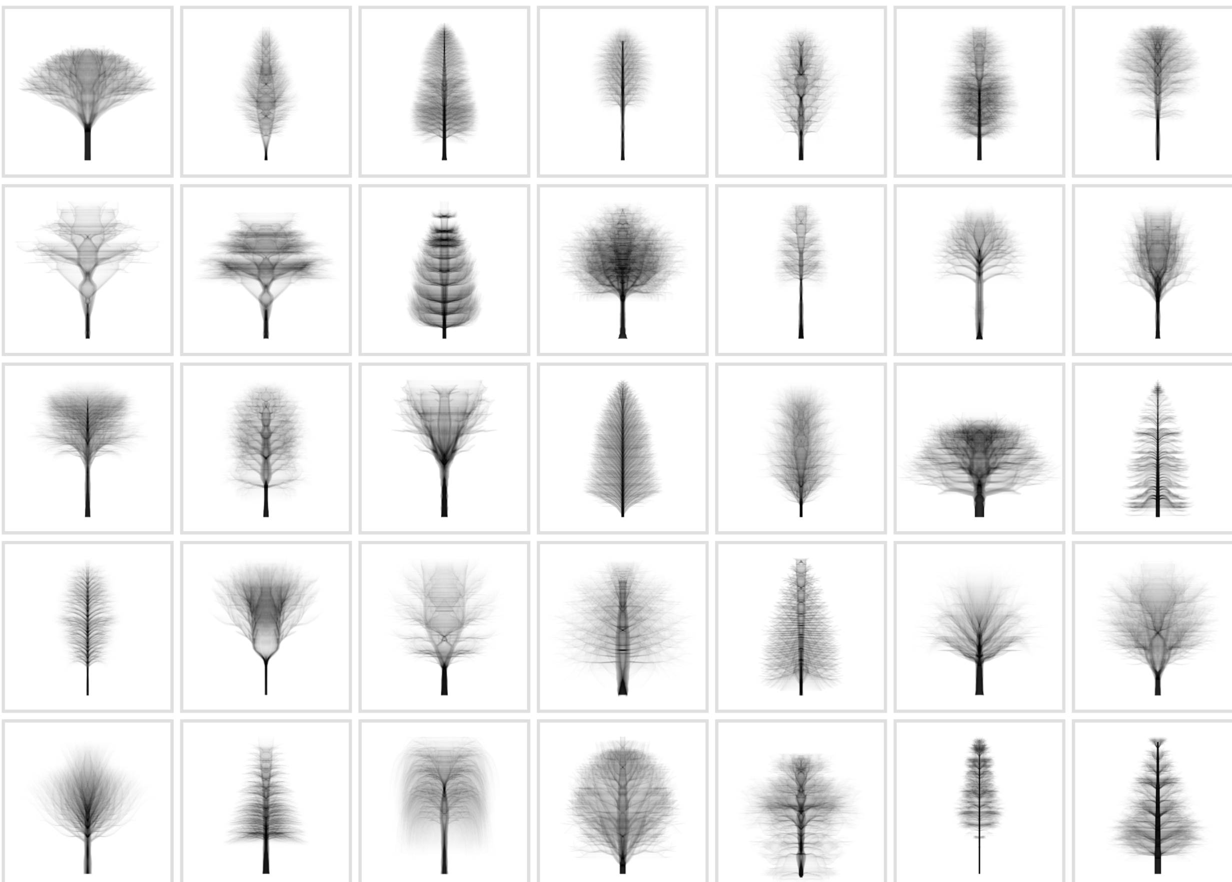

Figure 3.8. Examples of the tree area global feature being rendered by ray-tracing a revolved view of each tree.

4. DATA PROCESSING

4.1 Pattern Recognition

4.1.1 Introduction

In order to confirm the original hypothesis, a reliable rule with statistical significance and replicable outcomes needs to be formulated. Since classification algorithms train from data, it is essential to produce metrics and statistics of as many aspects of the studied plants as possible to ensure that no factor has been omitted. An exhaustive list of all statistics considered is attached via Table 4.1.

4.1.2 Decision Forest

The first implemented approach is a decision forest classifier derived from Čadík et al., 2013. A relatively traditional yet still effective method for training classification on large vol-

Table 4.1. Table of notable features factored into the proposed training. Legend - $\alpha_{|m|M}$ - represent the feature itself (α), its minimum (α_m), and maximum (α_M) values.

Type	Name	Symbol	Type	Name	Symbol
Local			**Local**		
Segment	Width	d		All Angle	$_{m\|M}$
	Volume	V		Tot. All Angle	$_{m\|M}$
Chain	Segments	S	Parent	Length Ratio	Δl
	Depth	h		Angle Sum	$\Delta\phi$
	Length	l		Continuation	$\boldsymbol{\pi}_{\|m\|M}$
	Total Length	L		Preservation	$\varsigma_{\|m\|M}$
	Deformation	σ	Branching	Asymmetry	$\kappa_{\mathrm{e}\|l\|v}$
	Straightness	τ		Branching	$\lambda_{m\|s\|d}$
	Slope	υ		Ramification	$\mu_{m\|s\|d}$
	Width	$d_{m\|M}$	**Global**		
	Tapering	Δd	Internode	Length	$\mathrm{i}_{\mu\|\sigma^2}$
	Internodes	I	Trunk	Length	l_t
	Tropism	$\zeta_{b\|h\|p}$	Leaf	Count	c_l
Sibling	Angle	α	Tree	Age	age
	Total Angle	β		Volume Imprint	$volume$

umes of parameters. A decision tree classifier is a branching tree of various simple classifiers which lead to the provided classification at a leaf. A single decision tree is not very expressive, however multiple decision trees can be averaged out to produce a probabilistic decision forest offering a much more robust classifier. (Safavian and Landgrebe, 1991) (Freund and Mason, 1999) A substantial benefit of a decision forest is the clarity with which it points towards decisive factors as opposed to the alternative classification approaches, especially the popular neural network classifiers.

The random forest regressor approach (Čadík et al., 2013) allows for these decision forests to be trained similarly to neural networks. Split of 20% testing and 80% training data is used with 1000 trees of loose depth between 9 and 12 under the mean square error function. After 30 repetitions under random splits a stable configuration is reached, verified through the Gini importance measure: (Hapfelmeier and Ulm, 2013)

$$\mathrm{i}_f = 1/n_s^{root} \sum \left(n_s^p n_\mathrm{i}^p - n_s^l n_\mathrm{i}^l - n_s^r n_\mathrm{i}^r \right)$$

For each featue f, n_f is the number of samples reaching a node, n_i is the Gini impurity and the p, l, r denote parent, left and right child. The resulting number is calculated for each ensemble and then all ensembles are averaged using mean.

4.2 Neural Network Classifiers

4.2.1 Overview

The second classifier is a convolution neural network. The pre-trained foundation used is based on ImageNet, which is an exceptionally influential database and architecture in its field. (Krizhevsky et al., 2012) (Russakovsky et al., 2015) (Deng et al., 2009) Additionally, the network in question is using so-called Siamese Neural Networks (Su et al., 2015). These networks maximize their efficiency by training for a similarity metric and treating the presented values as pairs. (Chopra et al., 2005) Such design maximizes the expressivity of smaller datasets. Since our study is restrained to 100 trees or 500 augmented tree views, such property is highly desirable.

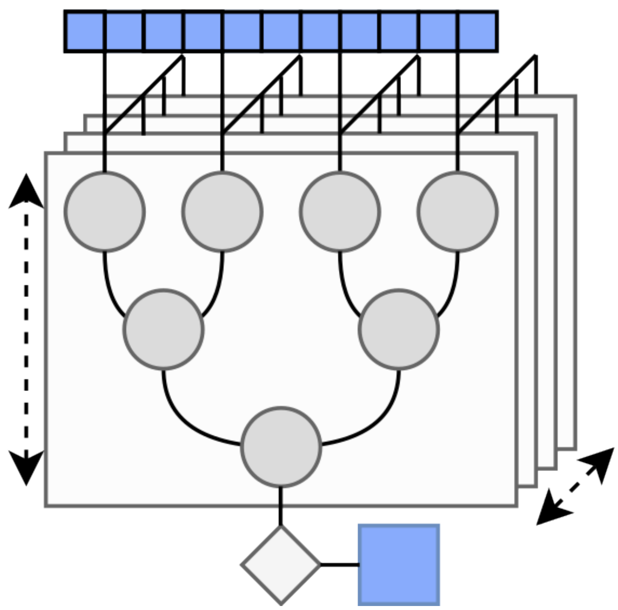

Figure 4.1. Decision forest regression model used in identifying the importance of individual features. Many different simple tree classifiers merged into a weighted forest show significantly increased expressiveness.

4.2.2 Feature-Based Predictor

The feature based predictor labeled ICTreeF uses the tree features 4.1 as its input. The structure consists of down-scaling fully connected layers with sigmoid nonlinearities inbetween leading up to a leaky rectified linear unit on the final node. (Maas et al., 2013) Meta-parameter optimization was avoided in order to avoid overfitting the data set.

Multi-layered perceptron with a hiden layer and 1D convolutional architectures were also considered, but performed worse. Pre-training has also proven to have little effect on the outcome.

The ground-truth for the training is derived from the gathered user responses to our survey alongside the full feature set expressed as a vector. The usual 80% training, 20% testing design approach is utilized alongside with the ADAM method Kingma and Ba, 2017.

The training error is evaluated through mean square $\sum(\hat{y}_i - y_i)^2$. Following the suggested approach and reduce-on-plateau technique described by Kingma and Ba, 2017.

4.2.3 Image-Based Predictor

For the image-based predictor labeled ICTreeI, the input consists of the same image used in the perceptual study and the perceptual ratings. These renders use jittering to simulare realistic data. The input images are scaled and cropped to fit a 256x256 image. Subsequently, random 224x224 subsection is cropped out of this image.

The architecture of the network is based on ResNet18 by He et al., 2015. Chosen over the other variants to increase training speed. The core of the ResNet layout is that on top of the ReLU non-linearity the raw block input is also added to the output. Every stage after the first one is down-sampled by a factor of two.

The same 80:20 split is used in training as with the feature classifier. Since assigning a direct score is not straight forward, the network is trained in incremental steps. First, the task is to select the more realistic tree out of two options. Then, the task becomes estimating the score difference between the two, until finally this advanced network is tasked with assigning a score to a singular image.

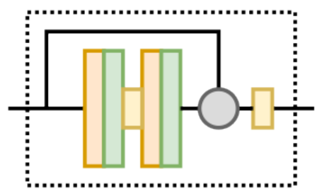

Figure 4.2. The distinct core building block of a ResNet network with the shortcut circumventing the nonlinearity.

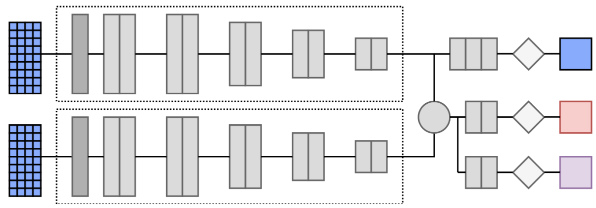

Figure 4.3. Layout of the siamese network architecture used in training binary and comparison steps of ICTreeI.

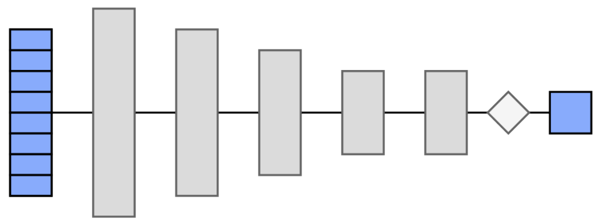

Figure 4.4. The final step of the ICTreeI training reverts back to single branch network.

The initial configuration of ResNet18 is fully random. As the first task implies, the comparison is trained on a Siamese network modification. (Su et al., 2015) The two halves of the split network share weights with each other. When the training and testing set are determined, responses to pairs across these two domains have to be removed from the set. The end of the two networks is merged, average pooled and ended with a fully connected sigmoid layer. The ADAM method is also used in training here. The error function is binary cross-entropy due to the nature of the data.

For the comparison step, the ground truth is the difference of the two perception quality values. Unlike the first step the final non-linearity is ReLU. Since the tested value has changed, the error function is mean square.

Ultimately, for finding the score the Siamese architecutre is left behind. The trained weights from the last stage carried over. Average pooling layer abd two fully connected layers of 2000 and 1 units are added at the end. ReLU is used here as well. Mean square remains the error function and reduce-on plateau is used.

5. FEATURE ANALYSIS

5.0.1 Network Implementation

Both of the networks are implemented within the PyTorch framework with CUDA support. The speeds were measured on a desktop computer with Ryzen 5 3600 processor, 16GB RAM, and NVIDIA GeForce RTX 3080 10GB GPU. *Model training:*36 hours, 5-10 minutes. *Inference:* 100-200ms, 2-5ms plus feature calculation 60ms.

5.0.2 Perception Quality Predictors

To measure the quality of the two trained networks, eight cross-validation runs were performed at the random 20:80 split over all trees in the data set. The degree of correlation of predictions with ground truth is show through Person's and Speraman's correlation coefficients cor_p and cor_s.

In the feature-based approach ICTreeF scores are more favorable at $cor_p \approx 0.735$ and $cor_s \approx 0.748$. For the image-based ICTreeI the results were $cor_p \approx 0.531$ and $cor_s \approx 0.569$. For comparison, a random classifier ranked $cor_p \approx 0.119$ and $cor_s \approx 0.086$ and a constant one $cor_p \approx 0.081$ and $cor_s \approx 0.005$.

The observations align with the hypothesis that image-based classification is a more challenging task than analyzing features. Therefore, ICTreeI could probably be trained to reach similar levels of accuracy with a larger data set.

When comparing the two metrics from different angles, ICTreeI roughly follows ICTreeF with greater variance. This indicates that even though the image metric is more varied, they both have the potential to converge.

5.0.3 Feature Interpretation

Following the user study and training of a decision forest together with a pair of neural networks, it is possible to interpret the significance of each feature's contribution to the overall perception. This is a valuable insight as it relates the number yielded by the method to practical characteristics.

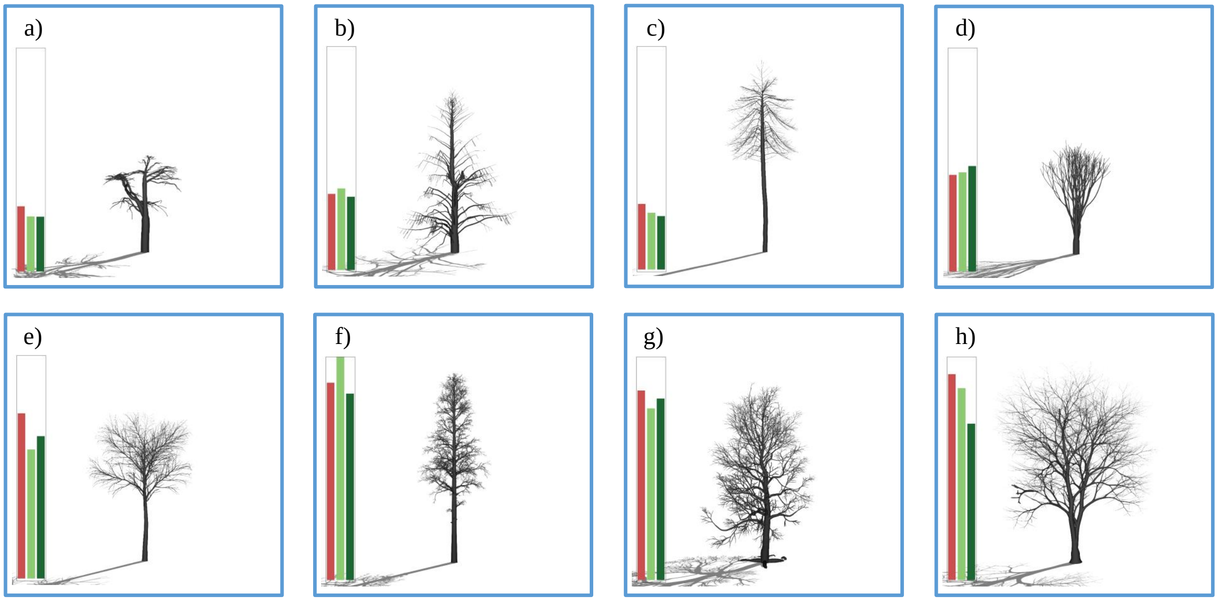

Figure 5.1. Comparison of rating for different trees sorted by ground truth quality (shown in red) alongside ICTreeF and ICtreeI (in light and dark green respectively)

The significance of individual features is determined based on the training assessment from the decision forest. Sibling angles and segment volumes have proven to be the most influential features, whereas chain segments and deformation have shown little to no effect. These observations are in line with earlier findings about the internal flow of nutrients in plants. (West et al., 1999) Likewise, Stava et al., 2014 refers to these features as important factors in inverse modeling of tree structures.

5.0.4 Result Validation

In order to apply the findings outside of the original data set, the generative method of Stava et al., 2014 was used to sample the many-dimensional vector space of features. There is no known mapping between the growth parameters of this algorithm and the space of features interpreted in this thesis, however through uniform sampling and subsequent evaluation, sufficient number of 54,944 trees across the studied feature space was found.

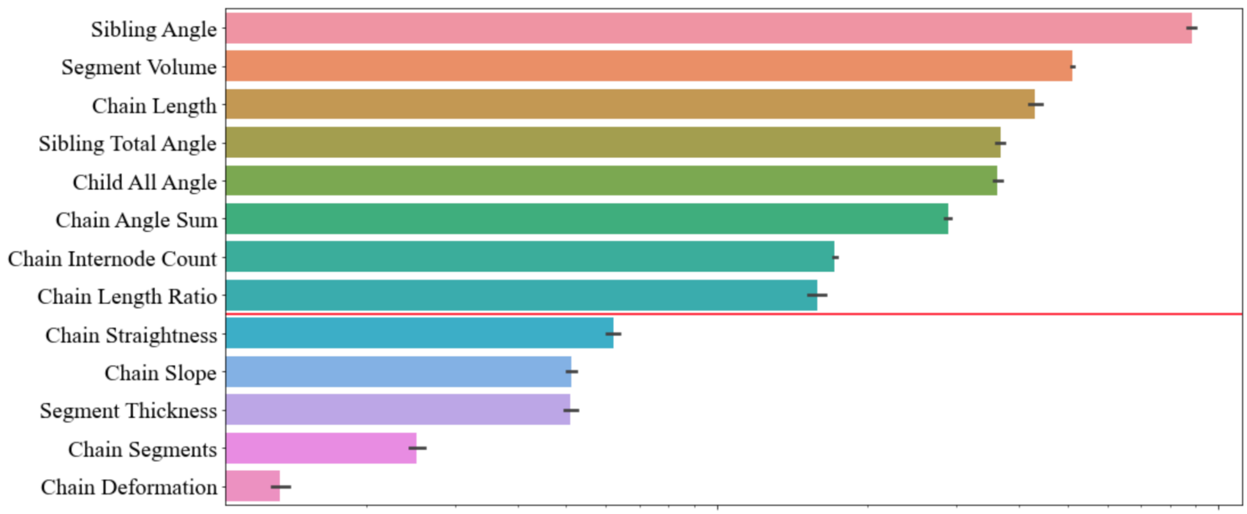

Figure 5.2. Graph illustrating the importance of features from 4.1. The width signifies importance amount on a logarithmic scale.

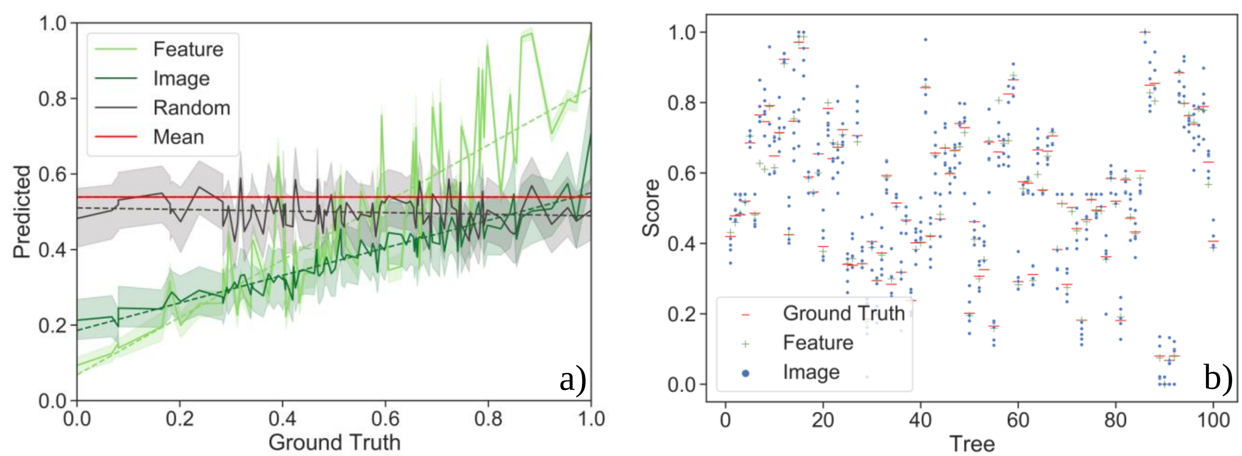

Figure 5.3. Cross validation of the ICTreeF and ICTreeI predictors alongside constant and random predictors. a) shows correlations b) plots ground truth against predictions made by ICTreeF on each tree and ICTreeI on each view.

From this set, series of trees were selected, that within a certain acceptable epsilon range varied solely in a single feature. As expected once evaluated through ICTreeF and ICTreeI it was demonstrably found that those trees had perceived quality strongly correlated with the change in an influential feature, and in a series based on a low-impact feature the differences were minimal.

To validate these perceptual observations a second smaller user study similar to the initial one was conducted through Mechanical Turk featuring 6000 pair comparisons. The perceived realism scores correlated with the two predictors. The series featuring significant features resulted in strong variations of $\sigma^2 = 0.308$ and $\sigma^2 = 0.203$, while the least significant ones only demonstrated $\sigma^2 = 0.036$ and $\sigma^2 = 0.014$.

With the ground-truth from the second study it is also possible to verify the accuracy of the predictors on new data. ICTreeF scored higher with $cor_p \approx 0.703$ and $cor_s \approx 0.517$ and ICTreeI achieved $cor_p \approx 0.430$ and $cor_s \approx 0.251$. Based on this it can be claimed that ICTreeF is a robust predictor since its scores are nearing the validation scores on the original training set. ICTreeI would require further training data to improve its generalization capabilities.

5.1 Conclusions

The research question has been scrutinized and resulted in tangeable observations. Two novel neural predictors of perceived realism in digital botanical tree models were introduced. ICTreeF uses features computed from tree structure as input while ICTreeF predicts based on an image or render of the tree in question. A large labeled data set of compared tree pairs was produced as a means of training and verifying these predictors. A secondary user study added new synthetic trees into the set and verified the assumed performance qualities of these networks.

Furthermore, it was determined which features contribute the most to those scores. This gives designers using this metric a way to know how to adjust their creations to score better or worse on the scale. This answers the proposed research question.

Figure 5.4. Table illustrating trees sampled from the vector space of features. Each row represents a different feature and each column represents an ordered development of that feature without changing other features. The top two rows are the most influential features (sibling angle and segment volume), the bottom two rows are the least significant features (chain segmants and chain deformation). The middle row is a moderately significant feature. Color bars show ground truth in red, ICTreeE in light green and ICTreeI in dark green.

5.2 Future Work

5.2.1 Limitations

The most noteworthy aspect of trees missing from this exploration is the absence of leaves which could strongly influence the outcomes of ICTreeI, if there was a way to gather or generate sufficient data. Currently, there are no large data sets of leaved trees available due to reasons related to Lidar precision. Some applications generate leaves synthetically. (Livny et al., 2011) This approach was avoided in this study to minimize data augmentation in realistic trees.

Another limiting factor is the amount of available data. Even for trees without leaves it would have been better to have access to more real tree data as well as higher number of user rated trees. Hundreds of unique tree geometries and hundreds of thousands of trials are sufficient to train a network however, the training potential is ever increasing with the size of the dataset.

As a result of varied data sources this study rendered plants in grayscale without textures. As indicated by Yixin Pan et al., 2005 this is also a very significant aspect of plant perception.

Finally, the analysis of the feature space is limited by sheer computational capabilities and number of features. It is difficult to identify gradients and formulate trends on a 30+ dimensional space. A potential future exploration could examine the behaviors of a reduced feature vector based on the proposed significance levels.

5.2.2 Future Exploration

A recurring element in tree simulation are tree species. Among others, Stava et al., 2014 based their method on knowing the specie of the trees. Since the goal of this exploration was to match trees regardless of source, meaning synthetic methods that ignore species included, the feature choice reflected that. This in practice meant that unconventional body plans such as palm trees yielded unusual feature sets. Perhaps features could also be used to label species in a follow up study.

In a similar vein, some of the metrics perform poorly on incomplete trees, young or old. Missing branches impact the interpretation of structural properties. If there was a way to measure these types of plants more reliably there could be a potential to expand this approach to more real plants. Many potential real sources of data, which would strengthen the presented perceptual claims, showcased poor sampling of fine branches just as Côté et al., 2009 notes.

REFERENCES

Anastacio, F., Sousa, M. C., Samavati, F., & Jorge, J. A. (2006). Modeling plant structures using concept sketches, In *Proc. of npar*, Annecy, France, Association for Computing Machinery. https://doi.org/10.1145/1124728.1124746

Andersson, P., Nilsson, J., Akenine-Möller, T., Oskarsson, M., Åström, K., & Fairchild, M. D. (2020). FLIP: A Difference Evaluator for Alternating Images. *ACM Trans. on Grap.*, *3*(2), 23. https://doi.org/10.1145/3406183

Aono, M., & Kunii, T. L. (1984a). Botanical Tree Image Generation. *IEEE Computer Graphics and Applications*, *4*(5), 10–34. https://doi.org/10.1109/MCG.1984.276141

Aono, M., & Kunii, T. L. (1984b). Botanical tree image generation. *IEEE Comput. Graph. Appl.*, *4(5)*, 10–34.

Aydın, T. O., Čadík, M., Myszkowski, K., & Seidel, H.-P. (2010). Video quality assessment for computer graphics applications, In *Acm trans. on graph. (proc. of siggraph asia)*, ACM.

Aydın, T. O., Smolic, A., & Gross, M. (2015). Automated aesthetic analysis of photographic images. *IEEE TVCG*, *21*(1), 31–42. https://doi.org/10.1109/TVCG.2014.2325047

Beneš, B., Andrysco, N., & Šťava, O. (2009). *Interactive Modeling of Virtual Ecosystems* (tech. rep.).

Beneš, B., & Millán, E. U. (2002). Virtual climbing plants competing for space, In *Proceedings of the computer animation*, USA, IEEE Computer Society.

Bojrab, M., Abdul-Massih, M., & Benes, B. (2013). Perceptual importance of lighting phenomena in rendering of animated water. *ACM Trans. Appl. Percept.*, *10*(1), 2:1–2:18.

Bosse, S., Maniry, D., Müller, K.-R., Wiegand, T., & Samek, W. (2018). Deep neural networks for no-reference and full-reference image quality assessment. *IEEE Transactions on Image Processing*, *27*(1), 206–219. https://doi.org/10.1109/TIP.2017.2760518

Čadík, M., Herzog, R., Mantiuk, R., Mantiuk, R., Myszkowski, K., & Seidel, H. (2013). Learning to predict localized distortions in rendered images. *Comput. Graph. Forum*, *32*(7), 401–410. https://doi.org/10.1111/cgf.12248

Cadik, M., Neumann, L., & Nemcsics, A. (2007). An Efficient Perception-based Adaptive Color to Gray Transformation. An Efficient Perception-based Adaptive Color to Gray Transformation. *Computational Aesthetics in Graphics*. https://doi.org/10.2312/COMPAESTH/COMPAESTH07/073-080

Chopra, S., Hadsell, R., & LeCun, Y. (2005). Learning a similarity metric discriminatively, with application to face verification, In *2005 ieee computer society conference on computer vision and pattern recognition (cvpr'05)*. IEEE.

Côté, J. F., Widlowski, J. L., Fournier, R. A., & Verstraete, M. M. (2009). The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial lidar. *Remote Sensing of Environment*, *113*(5), 1067–1081. https://doi.org/10.1016/j.rse.2009.01.017

David, H. A. (1988). *The method of paired comparisons*. C. Griffin. https://books.google.cz/books?id=bB21VsB%5C_GyYC

De Reffye, P., Edelin, C., Françon, J., Jaeger, M., & Puech, C. (1988). Plant models faithful to botanical structure and development. *ACM Siggraph Computer Graphics*, *22*(4), 151–158.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database, In *2009 ieee conference on computer vision and pattern recognition*. Ieee. https://doi.org/10.1109/CVPR.2009.5206848

Deussen, O., Hanrahan, P., Lintermann, B., Měch, R., Pharr, M., & Prusinkiewicz, P. (1998). Realistic modeling and rendering of plant ecosystems. *ACM Trans. on Grap.*, 275–286.

Du, S., Lindenbergh, R., Ledoux, H., Stoter, J., & Nan, L. (2019). AdTree: Accurate, Detailed, and Automatic Modelling of Laser-Scanned Trees. *Remote Sensing*, *11*(18), 2074. https://doi.org/10.3390/rs11182074

Farid, H., & Bravo, M. J. (2012). Perceptual discrimination of computer generated and photographic faces. *Digital Investigation*, *8*(3-4), 226–235. https://doi.org/10.1016/j.diin.2011.06.003

Ferwerda, J. A., Pattanaik, S. N., Shirley, P., & Greenberg, D. P. (1996). A model of visual adaptation for realistic image synthesis, In *Proceedings of the 23rd annual conference on computer graphics and interactive techniques, siggraph 1996*, Association for Computing Machinery, Inc. https://doi.org/10.1145/237170.237262

Freund, Y., & Mason, L. (1999). The alternating decision tree learning algorithm, In *Icml*.

Greene, N. (1989). Voxel space automata: Modeling with stochastic growth processes in voxel space. *SIGGRAPH Comp. Graph.*, *23*(3), 175–184. https://doi.org/10.1145/74334.74351

Grill-Spector, K. (2003). The neural basis of object perception. *Current opinion in neurobiology*, *13*(2), 159–166.

Guo, J., Xu, S., Yan, D.-M., Cheng, Z., Jaeger, M., & Zhang, X. (2018). Realistic procedural plant modeling from multiple view images. *IEEE transactions on visualization and computer graphics*, *26*(2), 1372–1384.

Guo, J., Vidal, V., Baskurt, A., & Lavoué, G. (2015). Evaluating the local visibility of geometric artifacts, In *Proc. of the acm siggraph symp. on applied perception*, Tbingen, Germany, ACM.

Guo, J., Vidal, V., Cheng, I., Basu, A., Baskurt, A., & Lavoue, G. (2016). Subjective and objective visual quality assessment of textured 3d meshes. *ACM Trans. Appl. Percept.*, *14*(2), 11:1–11:20.

Habel, R., Kusternig, A., & Wimmer, M. (2009). Physically guided animation of trees. *Comp. Graph. Forum*, *28*(2), 523–532.

Hädrich, T., Benes, B., Deussen, O., & Pirk, S. (2017). Interactive modeling and authoring of climbing plants. *Comput. Graph. Forum*, *36*(2), 49–61.

Hapfelmeier, A., & Ulm, K. (2013). A new variable selection approach using random forests. *Computational Statistics & Data Analysis*, *60*, 50–69.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition.

Herzog, R., Cadik, M., Aydcin, T. O., Kim, K. I., Myszkowski, K., & Seidel, H.-P. (2012). NoRM: No-Reference Image Quality Metric for Realistic Image Synthesis. *Computer Graphics Forum*, *31*(2pt3), 545–554. https://doi.org/10.1111/j.1467-8659.2012.03055.x

Herzog, R., Čadík, M., Aydın, T. O., Kim, K. I., Myszkowski, K., & Seidel, H.-P. (2012). NoRM: no-reference image quality metric for realistic image synthesis. *Computer Graphics Forum*, *31*(2), 545–554. https://doi.org/10.1111/j.1467-8659.2012.03055.x

Honda, H. (1971). Description of the form of trees by the parameters of the tree-like body: Effects of the branching angle and the branch length on the shape of the tree-like body. *Journal of Theoretical Biology*, *31*(2), 331–338. https://doi.org/https://doi.org/10.1016/0022-5193(71)90191-3

Huang, H., Wu, S., Cohen-Or, D., Gong, M., Zhang, H., Li, G., & Chen, B. (2013). L1-medial skeleton of point cloud. *ACM Transactions on Graphics*, *32*(4), 1–8. https://doi.org/10.1145/2461912.2461913

Ijiri, T., Owada, S., & Igarashi, T. (2006). Seamless integration of initial sketching and subsequent detail editing in flower modeling. *Comp. Graph. Forum*, *25*(3), 617–624.

Jung, M., Léger, D., & Gazalet, M. (2002). Univariant assessment of the quality of images. *J. Electronic Imaging*, *11*, 354–364.

Kawaguchi, Y. (1982). A morphological study of the form of nature. *SIGGRAPH Comp. Graph.*, *16*(3), 223–232. https://doi.org/10.1145/965145.801284

Kingma, D. P., & Ba, J. (2017). Adam: A method for stochastic optimization.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks (F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger, Eds.). In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems 25*. Curran Associates, Inc. http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

Lavoué, G., Larabi, M., & Vása, L. (2016). On the efficiency of image metrics for evaluating the visual quality of 3d models. *IEEE TVCG*, *22*(8), 1987–1999.

Lintermann, B., & Deussen, O. (1999). Interactive modeling of plants. *IEEE Comput. Graph. Appl.*, *19*(1), 56–65.

Liu, S., Wang, B., Thiagarajan, J. J., Bremer, P.-T., & Pascucci, V. (2015). Visual Exploration of High-Dimensional Data through Subspace Analysis and Dynamic Projections. *Computer Graphics Forum*, *34*(3), 271–280. https://doi.org/10.1111/cgf.12639

Livny, Y., Pirk, S., Cheng, Z., Yan, F., Deussen, O., Cohen-Or, D., & Chen, B. (2011). Texture-lobes for tree modelling. *ACM Trans. on Grap.*, *30*(4), 53:1–53:10.

Livny, Y., Yan, F., Olson, M., Chen, B., Zhang, H., & El-Sana, J. (2010). Automatic reconstruction of tree skeletal structures from point clouds, In *Acm transactions on graphics*. https://doi.org/10.1145/1866158.1866177

Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013). Rectifier nonlinearities improve neural network acoustic models, In *Proc. icml*.

Mantiuk, R., Kim, K. J., Rempel, A. G., & Heidrich, W. (2011). Hdr-vdp-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. *ACM Trans. on Grap.*, *30*(4), 40:1–40:14.

Měch, R., & Prusinkiewicz, P. (1996). Visual models of plants interacting with their environment, In *Proc. of siggraph*, ACM.

Moorthy, A. K., & Bovik, A. C. (2010). A two-step framework for constructing blind image quality indices. *IEEE Signal Processing Letters*, *17*(5), 513–516. https://doi.org/10.1109/LSP.2010.2043888

Nader, G., Wang, K., Hétroy-Wheeler, F., & Dupont, F. (2016). Visual contrast sensitivity and discrimination for 3d meshes and their applications. *Comp. Gr. Forum*, *35*(7), 497–506. https://doi.org/10.1111/cgf.13046

Okabe, M., Owada, S., & Igarashi, T. (2007). Interactive design of botanical trees using freehand sketches and example-based editing, In *Acm siggraph courses*, San Diego, California, ACM.

Oppenheimer, P. E. (1986). Real time design and animation of fractal plants and trees, In *Proceedings of the 13th annual conference on computer graphics and interactive techniques*, New York, NY, USA, Association for Computing Machinery. https://doi.org/10.1145/15922.15892

O'Sullivan, C., Howlett, S., McDonnell, R., Morvan, Y., & O'Conor, K. (2004). Perceptually Adaptive Graphics, In *Eurographics 2004 - stars*, Eurographics Association. https://doi.org/10.2312/egst.20041029

Palubicki, W., Horel, K., Longay, S., Runions, A., Lane, B., Měch, R., & Prusinkiewicz, P. (2009). Self-organizing tree models for image synthesis. *ACM Trans. on Grap.*, *28*(3), 58:1–58:10.

Perez-Ortiz, M., & Mantiuk, R. K. (2017). A practical guide and software for analysing pairwise comparison experiments. *arXiv preprint arXiv:1712.03686*.

Pirk, S., Jarząbek, M., Hädrich, T., Michels, D. L., & Palubicki, W. (2017). Interactive wood combustion for botanical tree models. *ACM Trans. on Grap.*, *36*(6), 197:1–197:12.

Pirk, S., Niese, T., Hädrich, T., Benes, B., & Deussen, O. (2014). Windy trees: Computing stress response for developmental tree models. *ACM Trans. on Grap.*, *33*(6). https://doi.org/10.1145/2661229.2661252

Pirk, S., Stava, O., Kratt, J., Said, M. A. M., Neubert, B., Měch, R., Benes, B., & Deussen, O. (2012). Plastic trees: Interactive self-adapting botanical tree models. *ACM Trans. on Grap.*, *31*(4). https://doi.org/10.1145/2185520.2185546

Prusinkiewicz, P., Hammel, M. S., & Mjolsness, E. (1993). Animation of plant development, In *Proceedings siggraph*, Anaheim, CA, ACM. https://doi.org/http://doi.acm.org/10.1145/166117.166161

Prusinkiewicz, P., James, M., & Měch, R. (1994). Synthetic topiary. *Proceedings SIGGRAPH*, 351–358. http://doi.acm.org/10.1145/192161.192254

Prusinkiewicz, P., & Lindenmayer, A. (1990). *The algorithmic beauty of plants*. Springer-Verlag New York, Inc.

Prusinkiewicz, P., & Lindenmayer, A. (2012). *The algorithmic beauty of plants*. Springer Science & Business Media.

Qin, H., Han, J., Li, N., Huang, H., & Chen, B. (2020). Mass-Driven Topology-Aware Curve Skeleton Extraction from Incomplete Point Clouds. *IEEE Transactions on Visualization and Computer Graphics*, *26*(9), 2805–2817. https://doi.org/10.1109/TVCG.2019.2903805

Rajasekaran, S. D., Kang, H., Benes, B., Čadík, M., Galin, E., Guérin, E., Peytavie, A., & Slavík, P. (2019). Ptrm: Perceived terrain realism metrics. *arXiv preprint arXiv:1909.04610*.

Ramanarayanan, G., Ferwerda, J., Walter, B., & Bala, K. (2007). Visual equivalence: Towards a new standard for image fidelity. *ACM Trans. on Grap.*, *26*(3), 76–es. https://doi.org/10.1145/1276377.1276472

Reddy, M. (1997). *Perceptually modulated level of detail for virtual environments* (Doctoral dissertation). Univ. of Edinburgh, UK.

Reddy, M. (2001). Perceptually optimized 3d graphics. *IEEE Comp. Graph. Appl.*, *21*(5), 68–75. https://doi.org/10.1109/38.946633

Reinhard, E., Ward, G., Pattanaik, S. N., Debevec, P. E., & Heidrich, W. (2010). *High dynamic range imaging - acquisition, display, and image-based lighting (2. ed.)* Academic Press. http://www.elsevierdirect.com/product.jsp?isbn=9780123749147%5C&dmnum=CWS1

Rogowitz, B. E., & Rushmeier, H. E. (2001). Are image quality metrics adequate to evaluate the quality of geometric objects? (B. E. Rogowitz & T. N. Pappas, Eds.). In B. E. Rogowitz & T. N. Pappas (Eds.), *Human vision and electronic imaging vi*, SPIE. International Society for Optics and Photonics. https://doi.org/10.1117/12.429504

Rubinstein, M., Gutierrez, D., Sorkine, O., & Shamir, A. (2010). A Comparative Study of Image Retargeting. *ACM Trans. Graph*, *29*. https://doi.org/10.1145/1866158.1866186

Runions, A., Lane, B., & Prusinkiewicz, P. (2007). Modeling trees with a space colonization algorithm. *Natural Phenomena*, *7*, 63–70.

Rushmeier, H. E., Rogowitz, B. E., & Piatko, C. (2000). Perceptual issues in substituting texture for geometry (B. E. Rogowitz & T. N. Pappas, Eds.). In B. E. Rogowitz & T. N. Pappas (Eds.), *Human vision and electronic imaging v*, SPIE. International Society for Optics and Photonics. https://doi.org/10.1117/12.387174

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Et al. (2015). Imagenet large scale visual recognition challenge. *International journal of computer vision*, *115*(3), 211–252.

Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, *21*(3), 660–674.

Saghri, J. A., Cheatham, P. S., & Habibi, A. (1989). Image Quality Measure Based On A Human Visual System Model. *Optical Engineering*, *28*(7), 813–818. https://doi.org/10.1117/12.7977038

Schwarz, M., & Stamminger, M. (2009). On predicting visual popping in dynamic scenes, In *Proc. of the apgv*, Chania, Crete, Greece, ACM. https://doi.org/10.1145/1620993.1621012

Shi, Y., Ondřej, J., Wang, H., & O'Sullivan, C. (2017). Shape up! perception based body shape variation for data-driven crowds, In *2017 ieee virtual humans and crowds for immersive environments (vhcie)*. https://doi.org/10.1109/VHCIE.2017.7935623

Sloup, P., Rebok, T., & Hanus, J. (2013). Automatic Tree Reconstruction from its Laser Scan (S. Robert, Z. Zalud, P. Cudlin, A. Farda, O. Urban, & M. Trnka, Eds.). In S. Robert, Z. Zalud, P. Cudlin, A. Farda, O. Urban, & M. Trnka (Eds.), *Global change and resilience*, Brno, Global Change Research Centre, Academy of Sciences of the Czech Republic.

Smith, A. R. (1984). Plants, fractals, and formal languages, In *Proc. of siggraph*, ACM Press. https://doi.org/http://doi.acm.org/10.1145/800031.808571

Smith, K., Landes, P.-E., Thollot, J., & Myszkowski, K. (2008). Apparent Greyscale: A Simple and Fast Conversion to Perceptually Accurate Images and Video. *Computer Graphics Forum*, *27*(2), 193–200. https://doi.org/10.1111/j.1467-8659.2008.01116.x

Spehar, B., Clifford, C. W. G., Newell, B. R., & Taylor, R. P. (2003). Universal aesthetic of fractals. *Computers & Graphics*, *27*(5), 813–820.

Stava, O., Pirk, S., Kratt, J., Chen, B., Měch, R., Deussen, O., & Benes, B. (2014). Inverse procedural modelling of trees, In *Computer graphics forum*. Wiley Online Library.

Su, H., Maji, S., Kalogerakis, E., & Learned-Miller, E. (2015). Multi-View Convolutional Neural Networks for 3D Shape Recognition, In *Proceedings of the ieee international conference on computer vision (iccv)*.

Suresh, S., Venkatesh Babu, R., & Kim, H. J. (2009). No-reference image quality assessment using modified extreme learning machine classifier. *Applied Soft Computing*, *9*(2), 541–552. https://doi.org/https://doi.org/10.1016/j.asoc.2008.07.005

Tan, P., Fang, T., Xiao, J., Zhao, P., & Quan, L. (2008). Single image tree modeling. *ACM Trans. on Grap.*, *27*(5), 108:1–108:7. https://doi.org/10.1145/1409060.1409061

Tang, H., Joshi, N., & Kapoor, A. (2011). Learning a blind measure of perceptual image quality, In *Cvpr 2011*. IEEE.

Taylor, R., Spehar, B., van Donkelaar, P., & Hägerhäll, C. (2011). Perceptual and physiological responses to jackson pollock's fractals. *Frontiers in human neuroscience*, *5*, 60. https://doi.org/10.3389/fnhum.2011.00060

Thompson, W., Fleming, R., Creem-Regehr, S., & Stefanucci, J. K. (2011). *Visual perception from a computer graphics perspective*. CRC press.

Um, K., Hu, X., & Thuerey, N. (2017). Perceptual evaluation of liquid simulation methods. *ACM Trans. on Graph. (TOG)*, *36*(4), 143.

Wang, G., Laga, H., Xie, N., Jia, J., & Tabia, H. (2018). The shape space of 3D botanical tree models. *ACM Transactions on Graphics*, *37*(1). https://doi.org/10.1145/3144456

Wang, H., Ondřej, J., & O'Sullivan, C. (2016). Path patterns: Analyzing and comparing real and simulated crowds, In *Proc. of the 20th acm siggraph symposium on interactive 3d graphics and games*, Redmond, Washington, ACM.

Wang, H., Ondřej, J., & O'Sullivan, C. (2017). Trending paths: A new semantic-level metric for comparing simulated and real crowd data. *IEEE TVCG*, *23*(5), 1454–1464. https://doi.org/10.1109/TVCG.2016.2642963

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *Trans. Img. Proc.*, *13*(4), 600–612.

Weier, M., Stengel, M., Roth, T., Didyk, P., Eisemann, E., Eisemann, M., Grogorick, S., Hinkenjann, A., Kruijff, E., Magnor, M., Myszkowski, K., & Slusallek, P. (2017). Perception-driven accelerated rendering. *Comp. Graph. Forum*, *36*(2), 611–643. https://doi.org/10.1111/cgf.13150

West, G. B., Brown, J. H., & Enquist, B. J. (1999). A general model for the structure and allometry of plant vascular systems. *Nature*, *400*(6745), 664–667.

Winkler, S., & Mohandas, P. (2008). The evolution of video quality measurement: From psnr to hybrid metrics. *IEEE Transactions on Broadcasting*, *54*(3), 660–668. https://doi.org/10.1109/TBC.2008.2000733

Wolski, K., Giunchi, D., Ye, N., Didyk, P., Myszkowski, K., Mantiuk, R., Seidel, H., Steed, A., & Mantiuk, R. K. (2018). Dataset and metrics for predicting local visible differences. *ACM Trans. Graph.*, *37*(5), 172:1–172:14. https://doi.org/10.1145/3196493

Xie, K., Yan, F., Sharf, A., Deussen, O., Huang, H., & Chen, B. (2016). Tree Modeling with Real Tree-Parts Examples. *IEEE Transactions on Visualization and Computer Graphics*, *22*(12), 2608–2618. https://doi.org/10.1109/TVCG.2015.2513409

Ye, P., Kumar, J., & Doermann, D. (2014). Beyond human opinion scores: Blind image quality assessment based on synthetic scores, In *Ieee cvpr*. https://doi.org/10.1109/CVPR.2014.540

Yixin Pan, Irene Cheng, & Basu, A. (2005). Quality metric for approximating subjective evaluation of 3-d objects. *IEEE Transactions on Multimedia*, *7*(2), 269–279. https://doi.org/10.1109/TMM.2005.843364

Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric, In *2018 ieee/cvf conference on computer vision and pattern recognition*. https://doi.org/10.1109/CVPR.2018.00068

Zhu, J.-Y., Park, T., Isola, P., Efros, A. A., & Research, B. A. (2017). *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks Monet Photos* (tech. rep.). https://github.com/junyanz/CycleGAN.