# ART-CG: Assisted Real-time Content Generation of 3D Assets through Machine Learning

## CASE FOR SUPPORT

## 1 OVERVIEW AND MOTIVATION

The production of 3D assets is a meticulous and time-consuming procedure that also requires expert knowledge. Content developers create, manipulate, and maintain high-dimensional 3D representation data ranging from order of thousands to millions.

Computer aided-design (CAD) drives engineering and manufacturing. Computer-generated image (CGI) supports the entertainment industry. 3D printing has been utilised in the field of medicine and its application is projected to expand [3]. Visualisation, simulation, and research often make use 3D assets [1]. Computer graphics (CG) is not only widely applied, according to *John Peddie Research (JPR)*, the CG market itself will exceed $235 billion by 2018 [2].

The 3D production pipeline has remained vastly static for many years. Often, an artist will plan out the asset through references or concept art. A base geometry is selected to work on- this could be entirely from scratch or modifying similar geometry already available. Traditional tools used within 3D production apply to a wide range of scenarios, but only perform steps of changes at a time. Quick methods of production such as procedural generation exist, but seldom used as the output is troublesome to control.

There exist many factors that impede production in the traditional pipeline. The pre-production and design phase begins with a medium that can visualise the concept early - such as paintings and artwork. Whilst this is a faster approach, there is a loss of data incurred when transferring from one medium to another. Art design demands consistency, but the division of labour is vital for studios to meet deadlines. Each artist can only focus on one piece of work at a time, along with projects evolving over a long span of period, stylistic consistency can be difficult to achieve.

We propose the concept of assisted real-time content generation, a novel framework that simplifies the production process of 3D assets through machine learning. Our proposition of probabilistic nonlinear dimension reduction consolidates observed variables to provide a parsimonious collection of latent variables, drastically diminishing the complexity of the production pipeline through regression.

Throughout the vast history of 3D graphical applications, An abundance of the commonly observed geometry of has been produced. The foundation established by previous work can be used to train tools that apply machine-learning to significantly reduce the overhead of repetitive tasks, allowing creators to focus more on design and refining. Content creators and consumers could enjoy the benefit of rapid prototyping and extrapolate a trained model to explore and identify novel stylistic properties. The application of machine-learning based tools could enhance the workflow of professional users and improve the experience for non-expert consumers. Such tools integrate into the production environment to improve the efficiency of acquiring initial base geometry and visually compare designs during pre-production. Non-expert users receive the ability to produce 3D geometry without requiring to learn the intrinsics of traditional 3D modelling software. The rise in popularity for augmented reality and 3D printing inspires the development of generative tools that are intuitive and simplistic to use. Applications that allow users to create their personal content could also integrate machine-learning based tools to prevent inappropriate or undesirable creation from being produced while providing options that surpass existing alternatives. An example would be avatar creation for many applications and video games. A space of reasonable options generated from predefined outputs by the developers will allow users to interpolate between sensible configurations, providing an excellent level of customisation while adhering to defined constraints.

## 2 BACKGROUND

There are numerous representations of 3D objects in computer graphics. One way to obtain 3D geometry data is to sample surfaces of the physical world with a 3D scanner. Common representations of sampled geometric data include point clouds, range maps, and voxels (figure 1.1).

3D surface representations have advantages and disadvantages depending on the use case. It is possible to convert between representations; however, conversion between representations risk incurring data loss. The CAD (Computer-Aided Design) industry often uses precise mathematical representations such as NURBS (Non-Uniform Rational Basis Spline). The most widely adopted representation for CGI (Computer-Generated Imagery) is polygonal meshes. In a production environment, it is preferred to define geometry specifically to requirements of the design as opposed to capturing examples. Polygon meshes are simple to define, yet with established techniques such as UV texturing and normal mapping, are sufficiently expressive for visualising purposes. The study of polygonal meshes is deeply rooted in computer graphics.

Elements of a polygonal mesh are vertices, edges, and faces. The topology of a mesh is concerned with the arrangement of its components, well-organised topology is required to maintain geometric qualities when performing algorithmic operations on a mesh. In practice, professionals create polygon meshes with majority quad-face topology (faces formed from four edges) during production. The rendering pipeline often automatically converts polygon meshes to triangle faces (formed from three edges) as an optimisation process. Quadrilateral mesh form what we call edge loops which can be used to define the structure of geometry, thus conform better with editing tools and preserve structure when algorithmically processed in comparison to triangle meshes.

The preferred representation of 3D objects within the industry is polygons. A polygonal object is defined by vertices, edges, and faces that form the geometry. Topology is the study of geometrical arrangement for the object. It is best practice to use quad-faces, which are faces formed from four edges asset production. The rendering pipeline often requires and automatically converts geometry to tri-faces, faces formed from three edges. Other representations include voxels which represent objects with cubes, and pixols, a 3D variant of pixels introduced by Pixelogic for their state of the art sculpting program, ZBrush.

State of the art 3D production software such as AutoDesk Maya, 3DS Max, and Blender are advanced programs with a sophisticated list of features. 3D software typically has convoluted user interfaces; even the most experienced professionals do not recognise each and every tool available. A combination of a high learning curve and overwhelming user experience causes the standard 3D software to seem intimidating for new users. In fact, most functionality is only accessible through an API (Application Program Interface) via a scripting language. Experienced users welcome this complexity as it allows creators to customise their ideal workflow, but it creates a high barrier to entry for non-experts who wish to possess the capability of such 3D programs.

Procedural generation techniques can automatically produce output that adheres to rules established by the generative model defined. Generation of terrains and city modelling sometimes employ procedural techniques to take advantage of its systematic nature to mass produce variations in agreement with specified patterns [11]. Fractals and methods such as the Lindenmayer system have been used to create patterns that resemble those observed in nature [24]. Automated techniques such as the ones discussed, however, are seldom used for modelling distinct objects with a specific design. It is an involved process to control the output of procedurally generated content without heavily restricting its capabilities. Automated methods that do not learn cannot adapt to changing demands without reimplementation, motivating for a learning-based solution.

Extensive research has been performed in graphics in the past several decades. Generative models have been used for procedural modelling and estimation. [Cite related work on traditional graphics/3D]

LAWRENCE, N. 2005. Probabilistic non-linear principal component analysis with Gaussian process latent variable models. The Journal of Machine Learning Research 6, 1783âĂŞ1816.

CAMPBELL, N. D., AND KAUTZ, J. 2014. Learning a manifold of fonts. ACM Transactions on Graphics (TOG) 33, 4, 91.

CHAI, M., SHAO, T., WU, H., WENG, Y., AND ZHOU, K. 2016. AutoHair: Fully automatic hair modeling from a single image. ACM Trans. Graph. 35, 4 (July), 116:1âĂŞ116:12.

Etc

Simplifying a procedure within the pipeline yield increased throughput and offer an opportunity for rapid prototyping. A nonlinear approach minimises loss of information; thus high dimensional data can be reduced effectively to dimensionality that is sufficiently low. Non-experts may find directing a few descriptive components within a low dimension representation easier than controlling an overwhelming number of attributes present in the original data.

## 2.1   Related Research of Machine Learning in Creative Fields

Three major paradigms divide the task of machine learning:

- Supervised learning is provided input training examples with desired outputs to learn the mapping of inputs to an output.

- Unsupervised learning seeks to learn the structure of and relation between input data.

- Reinforcement learning iteratively improve a pool of solutions by simulating an environment that applies concepts inspired by the theory of evolution.

The role that learning methods play in both manufacturing and consumer application continue to grow. However, adoption has been slow for creative fields. Obtaining more reliable data usually improves the performance of robust learning models, but example training sets of creative work are not easily attainable. Creative production values uniqueness and versatility, properties that cause difficulty in machine learning methods. Varying artistic styles in design complicate feature analysis, and ambiguity of correctness is problematic when predicting an output. In machine learning, prediction of continuous variables is a regression problem.

Style-based inverse kinematics introduced the Scaled Gaussian Process Latent Variable Model to learn the probabilistic distribution of a 3D human posture model [15]. Motion capture data represent character posture with a 42-dimensional feature vector that encapsulates joint information of a humanoid body. Learning a model of poses established the relation between joints and identified constraints exhibited in the training data. The probabilistic distribution is plot and used to select predictions in the model. There are no constraints on the joints, but sampling areas of high likelihood in the distribution models realistic motion that resembles the input data.

A latent doodle space learns latent (unobserved) variables that describe simple line drawings more concisely than the original data [7]. The motivation of a latent doodle space is to generating new drawings that are inspired by the input data. There are two key phases to derive a latent doodle space: the first challenge is to identify line strokes within drawings, the second is using a latent variable method to learn a low-dimensional latent space of input drawings.

A study by Campbell and Kautz (2014) presented a framework that learns the latent manifold of existing font styles [10]. The process involved universal parametrization of fonts to a polyline representation so that a distance measure is applicable and the generative model can interpolate between styles. Non-experts could create font styles without experience on type design by sampling points from a two-dimensional latent manifold.

Drawing assistance powered by large-scale crowd-sourcing explored the potential of data-driven drawing to prompt for correction by achieving an artistic consensus [20]. Learning a correction vector field from training drawings finds a consensus. Stroke-correction is applied using the correction vector field to adjust user input dynamically.

Chai et al. (2016) introduced AutoHair, a method for automatic modelling of 3D hair from a portrait image [12]. The approach extracts information from images and uses a database of hair meshes to construct a 3D representation of the information conveyed. A hierarchical deep neural network trained on annotated hair images learn to segment hair and estimate growth direction within portraits. Data-driven hair matching and modelling algorithm fit meshes from the database to parameters output by the neural net model to automatically produce 3D hair. The experiment developed a traversable hairstyle space of 50,000 hair models, using training images to fit segments of 3D examples obtained from the internet.

Research regarding creative content often parametrise the training data so that machine learning is applicable. There is no clearly defined solution for a problem in the creative field, effective solutions strive for versatility, employing consensus decision making or offering multiple solutions. To overcome the challenges introduced, dimensionality reduction through unsupervised learning with probabilistic latent variable models such as the Gaussian Process Latent Variable Model (GP-LVM) [18] present an opportunity to learn stylistic properties of design and predict multiple acceptable outputs by analysing the likelihood.

## 3 ACADEMIC BENEFICIARIES

AI Smart tools.

Machine Learning Small data set. Performance critical. On-line learning. Industrial application.

Graphics Generative models. Procedural modelling.

## 4 CHALLENGES

This study faces multiple challenges. Firstly, 3D meshes are difficult to compare. The training data in its raw form will have varying dimensions. We can view meshes as samples of the actual surface. Thus meshes that represent the same object could differ drastically in the number of data points depending on its level of detail. Typical feature extraction methods do not work well on meshes as artistic products are sensitive to data loss - any change could affect the perception of final result drastically.

Another problem encountered is the lack of training data. Typical machine learning solutions use huge data sets in the order of hundreds of thousands for training, but for 3D meshes, the standard size of readily available training data is much smaller. Public repositories of 3D polygonal hair ordinarily contain up to thousands of meshes [4]. Studios that store and organise past production may match the extent of public repositories, depending on the size of the company. Private repositories of independent artists will rarely exceed the order of hundreds.

The application of machine learning methods must also account for the subjectivity of evaluating artistic assets. The range of acceptable solutions is ambiguous, likened to how hair styles of characters can change drastically during the design phase.

In a production environment, the time required for a technique to return observable result directly affects throughput. For practical usage of assisted content generation, the technique should be reasonably fast in presenting observable output.

## 5 APPROACH

Our aim is to provide a framework, techniques, and methods that enable a more efficient workflow for the production of 3D geometry.

1. Parameterizing polygonal 3D models such that they are comparable for learning purposes, whether through generative models or as point clouds.

2. Applying regression models for prediction of output, comparing alternative models, deep learning, layered models for more control of output.

3. Devise techniques to repair 3D model data such that it is sensible to the context, based on semantics and constraints of the model.

4. Creating a new workflow that is beneficial for the UK digital creative economy, assisting competency of businesses by integration with state of the art tools.

5. Enable non-experts to produce personalised 3D geometry effectively based on stylistic properties of the training data set.

# 6   RESEARCH HYPOTHESIS

Success Criteria / Objectives:

1. Opportunity to use a tool that produces faster than traditional tools.

2. Complete much of the repetitive tasks instantly, then artist refine.

3. Create new assets from blending existing data in a practical manner.

4. Plausible alternative for users to have uniquely constructed 3D objects yet maintaining control over the environment to prevent inappropriate results.

Quantitative analysis: performance measurement, loss of data (residuals) from approximation Qualitative analysis: working with professional content creators, industrial partners, non-experts, compare demos to alternatives and past work

# 7   OBJECTIVES AND DELIVERABLES

## 7.1   DELIVERABLES

Success Criteria / Objectives: Opportunity to use a tool that produces faster than traditional tools. Complete much of the repetitive tasks instantly, then artist refine. Create new hairstyles from blending existing hairstyles in a practical manner. Plausible alternative for users to have uniquely constructed 3D objects yet maintaining control over the environment to prevent inappropriate results.

Quantitative analysis: performance measurement, loss of data (residuals) from approximation Qualitative analysis: working with professional content creators, industrial partners, non-experts, compare demos to alternatives and past work

1. Performing dimensionality reduction on 3D mesh data. 2. Exploring the application of nonlinear dimensionality reduction for high-dimensional mesh data such as 3D hair geometry. 3. Investigate the use of latent variables for identifying stylistic properties of 3D content. 4. Demonstrate the use of a nonlinear latent manifold to generate hair geometry. 5. Enable an intuitive method for both experienced users and non-experts to easily create 3D hair geometry. 6. Achieve performance close to real-time for practical use.

# 8   PROGRAMME AND METHODOLOGY

## 8.1   WORK PACKAGE 1 (WP1): RESOLVING THE ALIGNMENT PROBLEM OF 3D OBJECTS

Research Objective: 3D objects are defined ambiguously, it is challenging to compare structures of 3D geometry without preprocessing. This phase investigates methods of approximating object successfully so that the aligned data are comparable and minimises the data loss from approximation.

## 8.2   DELIVERABLES

1. Investigate and formulate techniques or generative model(s) that resolve the alignment problem of raw input data, allowing a sufficient approximation of generative parameters with the same dimension.

2. Implementation of said approximators, integrated within the production pipeline (extending state of the art tools such as 3DS Max, Maya, and Blender).

A challenge presented when attempting to learn the variation of 3D objects is the alignment problem. Data produced by creators defined in the form of vertices, edges and faces have varying dimensionality. Representations of the same object can vary, due to the level of detail and the pose of the object. Approximation of the object should be: translation, rotation, scale, and level of detail invariant, demanding a volumetric or an alternative generic approach.

Voxels have had success in estimating 3D objects but typically require closed shapes, this is not enough for production as planes, and unfilled shapes are often used to optimise by reducing the polycount.

## 8.3   Work Package 2 (WP2): Learning the Generation of 3D Objects

Research Objective: Investigate methods that are ideal for learning the generation of 3D objects and how to best visualise it so that it is useful. The scope of investigation encapsulates comparison of different regression models, extending viable models through layered hierarchy or applying deep learning. When applicable to latent variables, determining whether it is possible to infer semantic meaning from common latent variables for classification. Considerations include the limitations of learning models, evaluating the trend of performance as training data set scales.

### 8.3.1   Deliverables

1. Investigate viable learning models that are suitable for generation of 3D objects.

2. Critically evaluating the advantages and disadvantages of each model.

3. Applying research of complex models such as layered models and deep learning.

## 8.4   Work Package 3 (WP3): Generation and Reparation

Research Objective: The output of trained models contain a level of uncertainty. In the context of 3D assets, it is often flexible with approximation, but some scenarios outline specific constraints. An example would be hair object that should not grow inwards into a head object. Engineering design and architecture also apply strict limitations. This phase investigates methods of generation that conform with specified constraints, performing reparations appropriately so that the output is optimised and practical.

### 8.4.1   Deliverables

1. Investigate and formulate methods of generation and reparation for the uncertain output of a trained model.

## 8.5   Work Package (WP4): Integration and Applications of Machine-Trained Tools

Research Objective: Present a framework or specification of training-based tools that are fitting of integration to production and consumer applications. Discover factors that determine performance attained and identify the performance differences of scenarios, such as real-time programs.

### 8.5.1   Deliverables

1. Evaluation of online (continually) learning models against offline learning models.

2. Developing a framework of ideal workflow that integrates machine-trained tools into state of the art production software.

3. Measure the efficacy for both expert and consumer use cases.

## REFERENCES

[1] Y. Gao, P. Liu, D. Liu, and L. Zhang. 3D simulation research on hazard of fuel system leakage of LNG bus. pages 1–4, 2011.

[2] Jon Peddie Research. The CG market will exceed \$235 billion by 2018. `http://jonpeddie.com/press-releases/details/the-cg-market-will-exceed-235-billion-by-2018`, 2017. [Online; accessed 11-May-2017].

[3] C. L. Ventola. Medical applications for 3D printing: Current and projected uses. 39(10):704–711, 2014.

## BUDGET

Employment Machines & monitors, accessories Travelling Conference Communication

# IMPACT STATEMENT

The production of 3D assets play a pivotal role in business, engineering, and entertainment. [Cite market surveys of CAD, business, visualisation, entertainment, other benefits]

More 3D in future, Microsoft 3D Paint, augmented reality on tablets and smartphones, etc

Related topics: Graphics and visualisation, artificial intelligence, machine-learning

Dimensionality reduction on enormously high dimensional data.