LATENT PETROGLYPHS: PATTERN EXTRACTION FROM PREHISTORIC ROCK ART THROUGH GENERATIVE WORKFLOWS FOR A DESIGN PROJECT IN GREECE

SMARO KATSANGELOU[1], PARKER WILMOTH[2], DIMITRIS A. PADOS[3] and EMMANOUIL VERMISSO[4]
[1,4] School of Architecture, Florida Atlantic University
[2,3] Department of Electrical Engineering & Computer Science, Florida Atlantic University
1skatsangelou2023@fau.edu 0000-0003-1135-4667
2pwilmoth2023@fau.edu, 0000-0003-1731-4690
3dpados@fau.edu, 0000-0001-8959-6450
[4]evermiss@fau.edu, 0000-0001-9116-8877

Abstract. This paper regards the translation of indigenous rock art (petroglyphs) into training data for deep learning algorithms. Vis-à-vis the recent popularity of pre-trained AI models, the authors examine the potential of domain-specific search procedures to inform the process for a building design in Greece. Petroglyphs are a primitive form of artistic expression which has survived through the ages due to the medium upon which it was engraved. The practical aspect of this art was navigating through nature. The “Rock Art Center” aims to exhibit the narratives and culture behind rock art scattered in the mountains. Considering the adoption of generative adversarial networks (GANs) in the architectural workflow, the landscape and local prehistoric graffiti are viewed as datasets for tackling different design decisions, formally and conceptually interrogating the project’s scope. The existing rock art sites serve as the primary dataset to explore the building’s form, by accessing the 'latent' space of prehistoric rock art and its interpolation with the demands of the project. A number of algorithms and digital tools is employed to interpret the data in question.

Keywords. Artificial Intelligence, Deep Learning, Architectural Ideation, Design Workflow, Image Generation.

1. Introduction

The intersection of prehistoric art, architecture and deep learning (DL) is explored to navigate a design method for the creation of a building for a 'Rock Art Center' in Northern Greece. The authors draw inspiration by leveraging the gestural simplicity present in prehistoric carvings and the robustness of deep learning to extract features from large amounts of data. Architecture is often viewed as system, which comes as response to an existing set of conditions (Ching, 2007) In addition to this, "Architecture

ACCELERATED DESIGN, Proceedings of the 29th International Conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA) 2024, Volume 1, 149-158.

is a gesture" according to (Wittgenstein, 1980); in that sense, the satisfaction of functional requirements in a building is inadequate to render this "Architecture". True architectural achievement relies on communicating a notion, regardless -or in addition to- other (utilitarian) criteria. The need to communicate ideas and experiences in a non-verbal manner is inscribed in humanity's course, transcending time. Gesture in architecture and visual arts plays a crucial role in conveying ideas in a non-verbal manner. When "gesture" as a movement comes into play with the specific demands of professional architectural projects, a design negotiation occurs between the functional and aesthetic aspects of design.

One of the first, primitive efforts of man to create artistic gestures survives to this day in the form of petroglyphs, geoglyphs, rock, and cave paintings. The archaic is one of the greatest inventions of the 20th century (Davenport, 1981). The clarity and simplicity of prehistoric art has been the basis of many 20th-century artists' work such as Picasso and Modigliani. Romanticizing and misinterpreting prehistoric or indigenous heritage has been at times criticized as cultural appropriation of "primitive" art forms (Lock, 1994). Petroglyphs can only be experienced in-situ, "en plein air". The point of their existence is not to be carried and exhibited in the neutrality of a museum space and maybe not even to be viewed by the greater public (Anati, 1994).

Rock etchings are found across a span of 20km in the mountainside of Mount Paggaion, telling a story of non-verbal, prehistoric communication shared amongst groups of shepherds to mark and navigate through the natural landscape (Chatzilazaridis, 2000). This graffiti, dating from the Early Bronze Age up until the Middle Ages, depicts deer, horse riders, archers, hunters, standing figures, tools, and boats. Due to the inaccessibility of the locations where they were discovered (Moutsopoulos, 1969) they survive to this day. Some of the positions are impossible to access, so the creation of a building, which narrates the story of prehistoric art in Greece was necessary. The “Rock Art Center” Project is developed to inform the public about prehistoric civilization in the area, drawing visitors and attention to this yet unexplored and unknown part of Greek cultural heritage and history.

Prior related work exploring petroglyphs through generative design methods includes artificial petroglyphs generation through Hough Transform (Zhu, 2009). In this instance, the authors considered the use of deep neural networks (DNNs), in order to broaden the search in the ideation process drawing inspiration from these gestural depictions. Reinterpreting and exploring the design solution ("latent") space of the petroglyphs becomes a driving force for designing features of the building. These "latent" design opportunities explored in the proposed experimental workflow spark a debate about design agency and the idiosyncrasy of design as 'layered' process. Architectural design is a complex, multi-step process, which cannot be replicated entirely by AI, but can rather be infused with artificially generated outputs. Singular AI models cannot adequately capture a global configuration of complexities and multimodality of architectural production (Bolojan, Vermisso & Yousif, 2022). This has typically been addressed by multiple models, a sort of 'ecosystem' of algorithms designed to tackle different architectural tasks using the Deep Himmelblau network (Bolojan, 2022). Along these lines, the authors used custom-made Deep Learning algorithms to address the multimodality of architecture when considering the workflow for this design project.

2. Design Stages and Relevant Datasets

The case of using a piece of art or an everyday object as a starting point for lining out the shape of a building has unravelled manifold in the course of architectural history. From Claude-Nicolas Ledoux's Oikema to modern interpretations of mimetic architecture, the case for drawing inspiration from everyday life and embedding it into the formal configuration of buildings has been explored in various manners.

This project tries to intuitively link the gestural and artistic qualities of primitive carvings to those of a contemporary building attributes (façade; massing) through the designer's agency and personal design sensibilities. Our proposed workflow examined an AI-assisted method for addressing two architectural layers (1) Building Façade and (2) Building Massing, featuring two deep neural networks (GAN1; GAN2 respectively) which are informed by features from the petroglyphs data.

2.1. DEEP NEURAL NETWORK 1: PETROGLYPH CREATOR (GAN1)

The use of two-dimensional characters as design space driver has been previously explored in projects like Zaha Hadid's seminal design for Monsoon restaurant in Sapporo, Japan (Pavitt & Howe, 2023), where the Arabic writing of "Zaha" and "Monsoon", Hadid were used to kick-start the design process.

Inspired by the architecture of a StyleGAN, this network (DNN1) creates 'deep-fake' petroglyphs. The artificially generated petroglyphs are used as a reference to create a metal clad façade for the building, while also serving as serve as synthetic dataset for running a second network for massing. The objective is to explore the latent space of prehistoric rock art; by interpolating the different forms encountered within the dataset, new, artificial, yet untapped formations are navigated. Retrieved from archaeological documentation studies, sketches depicting the different prehistoric etchings found in the Mountain are used as a second datasets involved in the design of this building. This dataset initially included 122 images of 1024x1024 pixels depicting animals, bows, chariots, deer, hunters, people, ships, and miscellaneous objects.

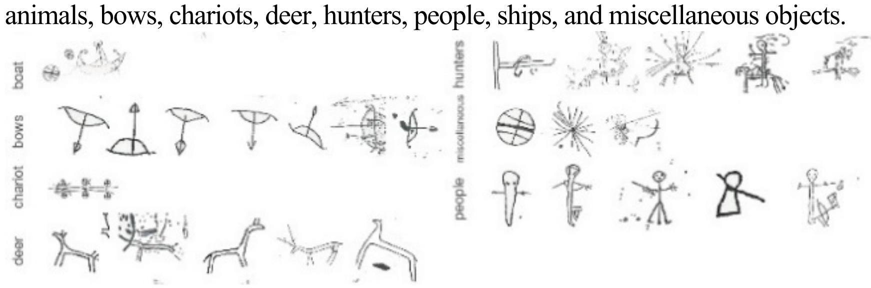

Figure 1 Part of the Initial Petroglyph Dataset.

As this number of instances in inadequate for training a DNN, the dataset was augmented tenfold. During augmentation, there were no restrictions on orientation, scaling, or other distortion of the images, as their semantic meaning was not altered in any way. These 8 classes of instances are inequilaterally represented in the initial dataset, as some appear in larger numbers than others. The augmentation was done in a manner that does not distort the quota of each class in the overall dataset. The results of this DNN trained on petroglyphs extracts features encountered in all eight classes.

Harun Farocki concocted the term "operative images" to describe images which are not meant for human viewing, but rather to be interpreted by machine vision (Farocki, 2010). In that sense, even though the images were initially cropped and formatted into 1024x1024 pixels, the team found that the semantic information could be conveyed in far lesser pixels, reducing the trainable parameters of the network and hence the training time. Consequently, the dataset comprises of 10k images of petroglyphs with a resolution of 64x64 pixels. The idea of operative image consideration therefore interrogates the threshold of size, aspect ratio needed to find optimize our computation needs for this experiment. The objective is to pinpoint the minimum resolution threshold, through which semantically meaningful results, which useful to the progression of the project can be produced.

2.2. DEEP NEURAL NETWORK 2: BUILDING PROGRAM (GAN2)

The use of artificial intelligence in architecture has been criticized for the lack of domain-specific datasets and agency in the design process (Bernstein, 2022). In this project, the creation of a dataset, which caters to a specific need for the design process, is used to inform design decisions regarding the spread of the square feet within the plot and the building program. Exploring the formal configuration of the project was of particular importance and interest, due to the lack of adjacent buildings in the plot. Found in a mountainous area, with no built but rather only natural environment surrounding it, this plot measures 5000 m^2, whereas one can construct up to 800 m^2.

A holistic approach to designing both the shell of the building, the exhibition and figuring out the building program was followed. The building program is structured around four fundamental pillars linked to the function of the spaces. Firstly, a robotic arm petroglyph carver is set as the main focus of the entrance of the building. Secondly displays of a DNN, which extracts and compares features encountered in petroglyphs from different regions of the world structures the space. Thirdly, a space for physical models either of the mountain, the rocks or the building itself and 3D printed real size petroglyph rocks, produced through photogrammetry models, are showcased. Lastly, a space where a shop, café and a gathering space is envisioned.

Using a Pix2Pix network architecture, corresponding massing iterations are created by the designer, based off 100 different petroglyphs which were created using DNN1, as explained in section 2.1. Even though the correlation between the provided petroglyphs and the formal configurations is at times visible, a direct correlation is avoided. In that sense, the objective is not to generate a building, which emulates the shape of a deer or a hunter, but rather to transcribe the fundamental rules, which connect the two systems. To encode the four different modules of the building program and corresponding square footage, four colours -Green, Blue, Peach and Grey- are used to distinguish the function of the different building volumes. The order of the system components come together to make up a whole based on the principles of hierarchy, rhythm, datum and transformation (Ching, 2007). From the resulting 3D iterations, a top view snapshot is captured to describe the building's shape. To work with the DNN, the produced dataset is subsequently augmented to 2000 images of building massings, initially formatted to 1024x1024 pixels. To make this experimental workflow semantically meaningful and useful to the design process, every module of the floorplan for each iteration has specific square footage, representing the exact space, which can be built on the plot. Split between $250m^2$ for the petroglyph carver, $250m^2$ for the showcase of petroglyphs from around the world, $200m^2$ for the physical models and $100m^2$ for the café and shop, for each iteration, the massings coalesce into a cohesive ensemble, whose geometry could potentially evolve into the final building.

Even though most GANs are trained on 8 GPUs, at a resolution of 1024x1024 pixels, this magnitude of computational power is not attainable for most designers and architects. To make this experimental workflow easily replicable by other architects, the team chose to run it on a single GPU, rather than multiple ones. For time-efficiency during training, images were converted to Grayscale, decreasing parameters by 66.6% and consequently, the necessary training time and computational power. The democratization and proliferation of an architectural workflow which includes custom, domain-specific datasets were driving forces for designing the specific network. Within architectural workflows, most AI tools commonly used are either pre-trained and readily available online, trained on a general dataset such as ImageNet (which is not catered to designers' needs) or developed with specific foci and large-scale computational resources, which are hard to replicate by small practices or teams of designers. To this end, bridging the gap between small-scale architectural production and existing AI generative models was considered paramount in this study.

3. Workflow - Methodology

Design is viewed as the accumulation of data and information from various sources and their attribution to inform the multimodality of architectural design (Ching, 2007). Before computational tools were invented, this process was carried out through extensive amounts of travelling by young, aspiring architects and artists, who roamed through Europe to experience and observe antiquities (Frampton, 1980). In this process, the petroglyphs found in the vicinity of the site are used as the conceptual basis for tackling two distinct tasks interwoven with the design of the building. Architectural

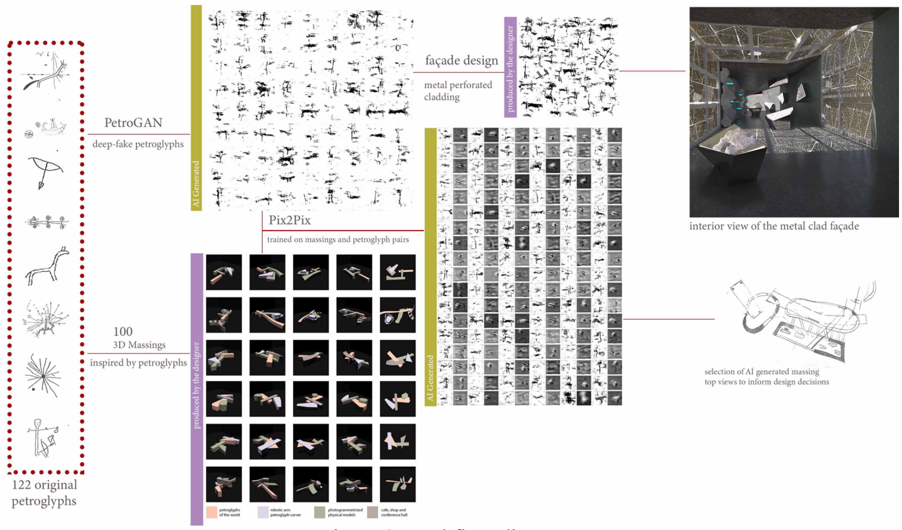

Figure 2 Workflow diagram

objects embody a thought process, which is repeated mentally by each person experiencing the space (Heinrich de Fries). Composing a building is often a negotiation between utilitarian and aesthetic considerations (Pehnt, 1973). The two realms can at times be conflicting, walking the line between what 'could' and what 'would' be. The term 'Expressionist Architecture' was coined in the brink of the 20th century to describe a style of architecture, which creates a certain mood or atmosphere (Behne,1915).

3.1. TECHNICAL DETAILS: GENERATIVE MODELS' ARCHITECTURE

This project used two DNNs (GAN1; GAN2) run on a virtual environment using Jupyter Notebook with Tensorflow, due to the obsolescence in many libraries used to train GANs. GANs gradually turn initial Gaussian noise into artificially generated images (Karras et al., 2021) through a minimax game, where a generator and a discriminator network compete against each other. The activation function used for both networks is LeakyReLU. The training progression of the first network showcases that more training time outputs decreasingly semantically meaningful results, whereas the second network moves from abstract interpretations of fake petroglyphs towards more defined geometries until it overfits, making a direct translation of the petroglyphs.

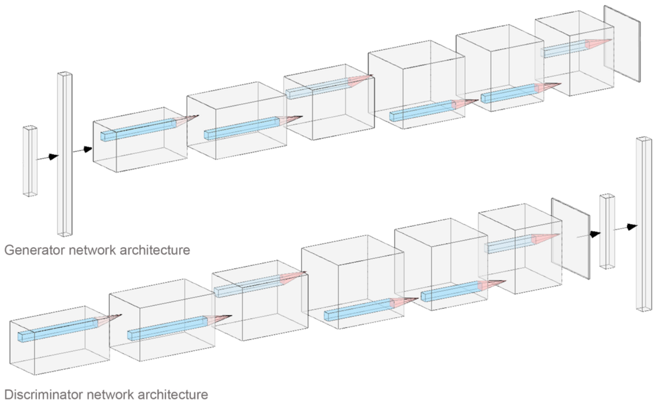

Figure 3 Employed GAN Network Architecture for GAN1

GAN1, a StyleGAN-type network, was trained on original petroglyphs, and comprises of a discriminator which inlcudes a LeakyReLU activation function, a dropout rate to regularize the network as well as a final dense layer with sigmoid activation for one-hot encoding. The discriminator has a 64x64x1 pixel image as input and outputs a value from 0.00 to 1.00. The Generator includes multiple convolutional layers, batch normalization and LeakyReLU activation. The output layer uses linear activation function to cater to a wider range of value. The generator initially takes random Gaussian noise vector of size 100 and outputs an image of 64x64 pixels.

GAN2, a Pix2Pix type network is trained to map images from one domain (petroglyphs) to another (massings). Like GAN1, its architecture consists of a "generator" and a "discriminator". The generator tackles transforming the petroglyphs from 2D representations to 3D top-down massing iterations. A series of convolutional layers extract features from the input (fake petroglyphs). The transposition of convolutional layers (UpSampling2D) up-samples the features of the images to gradually reconstruct the target image through Gaussian Noise. Finally, the last layer uses linear activation to produce the predicted massing. The discriminator distinguishes between fake and real using convolutional layers to process the input and extract features.

Other experiments conducted by the team include the use of different angles of the representations of the 3D iterations to train a third GAN. The network did not seem to produce semantically meaningful results to inform the ideation process of the architectural design. Consequently, this network was not used further in our proposed workflow. Potentially, the use of more high-resolution images of the iterations could yield more useful results. Nonetheless, the team chose to keep the resolution low to reduce the training time and required computational power.

4. Results

GAN1 trained on 10k images of Petroglyphs, solely from the area of Paggaion. The network trained for more than 1000 epochs. The results of the network can be identified as the initial different classes. Hunters and animals occur, as well as standing figures. Moreover, hybridized forms, which interpolate between the forms of people and animals. When we attempt a one-on-one comparison of the provided original petroglyphs and the results emulate the structure and compositional values of the training set. The line weight, as well as the clarity of the lines is not of the same consistency as the training dataset. The outputs of GAN1 were used to create a new composition, which turned to be a laser-cut metal clad façade for the 'Rock-Art Center'. Anthropomorph and animal-like figures can be distinguished from the outputs. While training, the team decided to omit some of the samples belonging to the 'Miscellaneous' class, as their unstructured composition seemed to affect the network negatively.

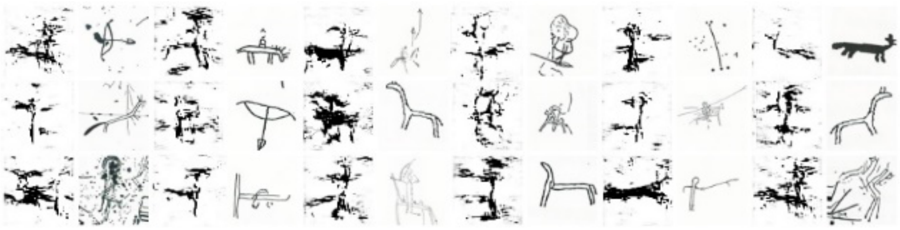

Figure 2 Fake and Real Petroglyph Comparison

The second network (GAN2) for this experimental workflow was based on the Pix2Pix architecture, combining the top view of the massings created by the designer as well as the original augmented petroglyphs. The network was trained on original petroglyphs as well as the original massings created by the team. For the final use of the artificially produced massings, the input of the network was 'deep-fake' petroglyphs. In that sense, the concept of exploring the latent space of prehistoric forms was expanded in the overall configuration of the building.

From a total of 1000 massings based on fake petroglyphs, the team chose 10 which were considered for the process of finalizing the formal aspect of the building. The network, as it is pretrained on the provided original petroglyphs as well as the massings, can potentially output new massings provided any two dimensional black and white drawing. The designer's agency and personal style is passed onto the network through the manual creation of the initial massing dataset. This network caters to a very specific design problem, bridging the gap between function and aesthetics. The ambiguity of the results is beneficial to the development of the architectural project, as it leaves room for further interpretation and curation based on the designer's intuition and aesthetics.

It is observed that most of the produced massing diagrams include two, three or four predominant volumes, which are correlated with the initial dataset upon which the network was trained. The initial conception for the dataset was to be structured around four color-coded modules which make up a whole. So, it is accessed that some of the principles followed to create the dataset were learned successfully by the network, even with a resolution as low as 64x64. There are various instances in the training progression of the second network where the results create stimulating results for the aforementioned purpose. It is nonetheless obvious when the network starts to overfit, as the outputs start to mimic the petroglyphs as a one-to-one translation rather than a loose interpretation. Even when subtracting color and resolution from the dataset, the outputs still show variation with 3 or 4 differentiated volumes, which follows the rationale of the dataset of massings produced by the designer. This shows that the network has learnt some of the compositional characteristics of the input, whilst definitely lacking definition. What's more, the output images take up similar square footage as the one's provided as training material.

5. Conclusions and Future Work.

In the 20th century, the use of material from the deeper past of human civilization is equivalent to the use of design and aesthetic ideals from Hellenistic Rome during the European Renaissance (Davenport, 1981). Prehistoric art has been explored and interpreted in the context of modernity, lacking at times context and deeper understanding of the circumstances upon which these creations came to be. The ideation process as well as the final output of the building is informed through generative algorithms. Interpreting and reiterating the petroglyphs of Paggaion gave ground to the conception of new ideas, as well as formal configurations for the building. Moreover, the accessibility of this method, rather than using multiple GPUs, has potential of broader use either commercially or pedagogically.

The intersection of Architecture and Artificial Intelligence has been explored in multiple ways spanning from GANs to diffusion models. This year, the debate has been monopolized by the rise of powerful text-to-image models which heuristically kick-start the design process, serving as a tool to boost brainstorming. In our study, a domain-specific dataset, created manually is used to aid in exploring the latent space of Petroglyph forms and building configurations. The designer's agency or copyrights to the produced building is not compromised as the network serves as a tool for specific sub-modalities of what makes up architecture.

Based off selected produced massings, the designer can use the synthetic output as a basis for tracing the floorplan of the building. Going back and forth between analogue, digital and artificially generated content is proposed as an alternative workflow, which could be easily, and (time-wise) efficiently replicated by small scale design teams of 5 to 10 people. The team is working towards higher resolution outcomes, which can capture the expressiveness of BIM modelled building massings.

For these experiments, the dataset's resolution was reduced to make this workflow easily accessible and replicable for potential educational purposes. The training time for each model is estimated at 2 minutes with a GForce RTX 3070 GPU. The reduced size of the dataset results in reduced training time and computational power. The network was trained locally, rather than on GoogleCollab, making it more resilient to updates and easily accessible to architects and students. This study acts as an early version of an architectural workflow ran on a very small initial dataset. This experiment is to be used as a proof of concept for the team's workflow.

In future experiments, the authors will focus on using three-dimensional geometry to produce artificially generated massings in 3D. The use of actual three-dimensional geometry, rather than two-dimensional representations of it, is envisioned as a basis for the design process in a more sculptural manner. Experiments of this sort have been conducted by NVIDIA Lab. The generation of high-quality 3D textured shapes has been achieved through training a network with 442 shapes of animals, 563 houses of shapes and 500 shapes of people (Gao, 2022). The non domain-specific nature of the datasets and their random collection from Turbosquid may yield high-quality results but cannot be utilized for particular design needs, so the authors intend to curate further bespoke datasets for training the fine-tuned current models (GAN1; GAN2).

Acknowledgements

The project was commissioned and funded by the Kavala Municipality in Northern Greece. The team also acknowledges the contribution of the resources and know-how of the "Center for Connected Autonomy & Artificial Intelligence" at FAU.

References

Anati, E. (1994). Archetypes, constants, and universal paradigms in prehistoric art. *Semiotica 100* (2-4):125-140. https://doi.org/10.1515/semi.1994.100.2-4.125

Bolojan, D., Vermisso, E., & Yousif, S. (2022). Is Language All We Need? A Query Into Architectural Semantics Using a Multimodal Generative Workflow. *CAADRIA 2022: Post-Carbon*. Sydney: CAADRIA.

Behne, A, (1915). *Biologie und Kubismus. Der Sturm* 5(11/12), 68–71.

Bernstein, P. (2022). *Machine Learning: Architecture in the age of Artificial Intelligence*. RIBA publishing, London.

Bolojan, D.: Creative AI: Augmenting design potency. *Archit. Des*. 92(3), 22–27 (2022). https:/ /doi.org/10.1002/ad.2809

Chatzilazaridis, L. (2000). *Prehistoric rock engravings of Northern Greece*. Doctoral Dissertation, Aristotle University of Thessaloniki. Thessaloniki, Greece. https://thesis.ekt.gr/thesisBookReader/id/17938#page/1/mode/2up

Ching, F. (2007). *Architecture: form, space and order.* John Wiley & Sons, Inc. New Jersey. https://rsd2-alert-durden-reading-room.weebly.com/uploads/6/7/1/6/6716949/ching-architecture_form_space_order.pdf

Davenport, G. (1974). The symbol of the archaic. *The Georgia Review,* 28(4), 642–657. http://www.jstor.org/stable/41397160

Davenport, G. (1981). *The Geography of the Imagination: Forty Essays.* North Point Press, San Franscisco. https://issuu.com/polinecia/docs/the_geography_of_the_imagination.

Farocki, H. (2010). *Installations. 2010. Re-pouring. 2010*. Retrieved November 17, 2023, from http://www. harunfarocki.de/installations/2010s/2010/re-pouring.html.

Frampton, K. (1980). *Modern Architecture: A Critical History.* Thames & Hudson, London. DOI:10.1080/10331867.2021.1989898

Gao J., Shen T., Wang Z., Chen W., Yin K., Li D., Litany O., Gojcic Z., and Fidler S. (2022). GET3D: A Generative Model of High Quality 3D Textured Shapes Learned from Images. In *Advances in Neural Information Processing Systems.* Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2022/file/cebbd24f1e50bcb63d015611fe0fe767-Paper-Conference.pdf

GET3D: A Generative Model of High Quality 3D Textured Shapes Learned from Images https://research.nvidia.com/labs/toronto-ai/GET3D/assets/paper.pdf

Harding, J. (2016). Dimensionality reduction for parametric design exploration. In: *Advances in Architectural Geometry 2016*, September 9th - 13th, 2016, Zurich, Switzerland. DOI:10.3218/3778-4_19

Karras T., Aittala M., Laine S., Harkonen E, Hellsten J., Lehtinen J., and Aila T. (2021). Alias-free generative adversarial networks. *Advances in NeurIPS*, 34, 2021.

Moutsopoulos, N. (1969). *Ta akidografimata tou Paggaiou (The graffiti of Paggaion).* Aristotle University of Thessaloniki Publications. Thessaloniki, Greece. http://apothesis.teicm.gr/xmlui/handle/123456789/3031

Pavitt, J. and Howe, C. (2023) Zaha Hadid Foundation: Interdisciplinarity and Intersectionality. In *annual conference of the Jaap Bakema Study Centre Architecture Archives of the Future,* TU Delft – Nieuwe Instituut in Rotterdam, November 2023.

Pehnt, W. (1973). *Expressionist Architecture.* Thames & Hudson, London. https://pdfcoffee.com/expressionist-architecture-2-pdf-free.html

Zhu Q., Wang X., Keogh E., and Lee S. (2009). Augmenting the generalized hough transform to enable the mining of petroglyphs. In *Proceedings of the 15th ACM SIGKDD International conference on Knowledge discovery and data mining (KDD '09).* Association for Computing Machinery, New York, 1057–1066. https://doi.org/10.1145/1557019.1557133