Interlacing Latent Features: Synthesis of Past and Present in Architectural Design through Artificial Intelligence in a Case Study of Japanese Houses

Citation

Permanent link
https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37378316

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
Interlacing Latent Features: Synthesis of Past and Present in Architectural Design through Artificial Intelligence in a Case Study of Japanese Houses

A Thesis Submitted to the Department of Architecture
Harvard University Graduate School of Design,

by

Rio Kobayashi

In Partial Fulfillment of the Requirements for the Degree of Master of Architecture

December 2023

Copyright © 2023 by Rio Kobayashi

“The author hereby grants Harvard University permission to reproduce and distribute copies of this thesis, in whole or in part, for educational purposes.”

(signature of author) (signature of advisor)

Rio Kobayashi Andrew Witt
Interlacing Latent Features:
Synthesis of Past and Present in Architectural Design through Artificial Intelligence
in a Case Study of Japanese Houses

Harvard University
Graduate School of Design
Master of Architecture I

Rio Kobayashi

Advisor: Andrew Witt
Tapestry of Architectural Data
Throughout history, humans have created and accumulated rich stores of architectural knowledge and data. This vast repository (of drawings, images, writings, and other media representations) includes architecture with and without architects.
While some forms of knowledge and data fade into oblivion, others persist and continue to evolve within contemporary design landscapes.
Knowledge and data inherited from the past can be deeply embedded in our conscious and subconscious thinking. It stretches across the temporal and spatial expanse of human civilization like a grand tapestry.
My project likens the web of knowledge to such a tapestry consisting of diverse threads each representing a distinct architectural environment: a line for form, a cross-stitch for function, delicate silk filament for aesthetics, sturdy twine for structure.
Threads of information manifest in various forms, from text to 3D model, while the manner of weaving can be likened to how every design practitioner revisits and reinterprets each thread and tool they come across.
Threads are interlaced to create patterns and motifs. Those that are bold and pronounced remind me of grand edifices and monuments, while others that are more subtle or imperceptible might represent the nuances of vernacular dwellings and modest charm of utilitarian spaces. Various movements have taken place in time, and the weaving manner has changed many times in the past.
<table>
<thead>
<tr>
<th>Contextual Architecture</th>
<th>Vernacular Architecture</th>
<th>Indigenous Architecture</th>
<th>Traditional Architecture</th>
<th>Regional Architecture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contextual architecture, also known as Contextualism is a philosophical approach in architectural theory that refers to the designing of a structure in response to the literal and abstract characteristics of the environment in which it is built. Contextual architecture contrasts modernist architecture, which values the imposition of their own characteristics and values upon the built environment. Contextual architecture is usually divided into three categories: vernacular architecture, regional architecture, and critical regionalism, all of which also inform the complementary architecture movement. [1]</td>
<td>Vernacular architecture is building done outside any academic tradition and without professional guidance. This category encompasses a wide range of variety of building types, with differing methods of construction, from around the world, both historical and extant, representing the majority of buildings and settlements created in pre-industrial societies. It constitutes 95% of the world’s built environment, as estimated in 1995 by Amos Rapoport, as measured against the small percentage of new buildings every year designed by architects and built by engineers. It usually serves immediate local needs; is constrained by the materials available in its particular region; and reflects local traditions and cultural practices. Traditionally, the study of vernacular architecture did not examine formally schooled architects, but instead, that of the design skills and tradition of local builders, who were rarely given any attribution for the work. [2]</td>
<td>The field of Indigenous architecture refers to the study and practice of architecture of, for and by Indigenous people. It is a field of study and practice in the United States, Australia, Aotearoa/New Zealand, Canada, Arctic area of Sápmi and many other countries where Indigenous people have a built tradition or aspire translate or have their cultures translated in the built environment. This has been extended to landscape architecture, urban design, planning, public art, placemaking and other ways of contributing to the design of built environments. [3]</td>
<td>Architecture based on a way of thinking, behaving, or doing something that has been used by the people in a particular group, family, society, etc., for a long time: following the tradition of a certain group or culture. [4]</td>
<td>The main idea inherent in the concept of regional architecture/regionalism is context-specific architecture. This, in turn, is based on knowledge of the history of a place, climatic conditions, concerns, materiality, topology, ecology, environmental conditions, culture and traditions, skills, tools, and technology available in a particular area. The driving idea behind Critical Regionalism is resistance to the standardization of Architecture. The increasing standardization is a modern phenomenon caused by globalization. [5]</td>
</tr>
</tbody>
</table>
Today, one can argue that much of our knowledge/data and access to it has been largely overwritten by capitalist motives and urbanist biases thus resulting in the homogenization of design and deemphasizing sustainability and diversity in design, a homogeneity that benefits from anonymity.
We question; will our society ever see a residential building like Cappadocia again?
As designers, we learn and digest information both individually and collectively. It is intimate and cognitively demanding work, and there is a limit to our computing capacity. Hence the creation from our own learning is inevitably exposed as biased and limited as well.
This period of exhaustion is likely where the neglect and erasure of data/novelty happens. Can there be a way to balance our current architectural tapestry from the encroachments of globalization and homogenization?
Machines have served humans in various ways for a long time and as predicted by experts decades ago, it has achieved an intelligence via neural networks that now closely mirrors the cumulative nature of collective human intelligence.
Architecture is hardly exempt, as we have come to appreciate and acknowledge the added value these machine learning technologies and digital innovations bring to design as a discipline and practice.
"With the advent of computer vision, for the first time, art and design can be quantified. Never before have we had the power to attribute artwork with the support of AI to confirm artist technique with data. We are in an exciting period in technological design with wide implications of this innovation in a variety of fields." [6]
Machine intelligence can take varied forms: Classification, evaluation, optimization, generation.
One of the more interesting capacities of machine intelligence is its ability to identify hidden features and fragments that are typically imperceptible to humans, and then transmit them into their creations.
Today researchers leverage machine intelligence in order to augment design perceptions and thinking.
Language and text is a way of storing, retrieving, and generating meaning in architecture and is an essential input into machine learning as well.
Using machine learning as tool and inspiration, my project seeks to track and archive AI’s ability to read hidden features in order to synthesize the past and the present fabric of architecture across various places and times, and in doing so, produce intriguing and evocative forms.
My project features three case studies in Tokyo which aim to reimagine the modern, westernized home...
through the introduction of machine leaning models, classification, and image generation.
My design process began with collecting data fabrics and defining synthetic fabric qualities. I followed this up with the classification and generation phases – reading between threads and synthesizing the fabrics.
My initial data collection focused on plans and facades in order to highlight spatial features that face the local context.
Japanese Houses and Metadata
Japan’s postwar economic growth is emblematic in several of that period’s archetypes. Housing, offices, and commercial spaces were deeply influenced by mass production and standardization.
Spaces began emphasizing contemporary values of living such as speed and privacy. Interior spaces became reclusive and more enclosed, with smaller windows. This rapid transition can be owed to the predominance of wooden construction and market preferences for new construction over second-hand properties, as the latter’s value drops significantly as soon as it goes on sale.
Japan’s year calendar system aligns with its imperial lineage, beginning a new era with each emperor’s ascent. Each era thus represents one human’s life span, and consequently corresponds to the architectural style of that time.
Elevations and plans of Japanese houses can be traced back to the Edo period which started 420 years ago, covering 5 periods until now.
A thick straw roof and shoji doors are typical in the Edo period.
Whereas white and gray exterior colors with small windows begin to proliferate in the modern period.
A key indicator of the shift from traditional to modern living can be seen with the introduction of the "small house," the first prefabricated building in Japan which was marketed as a solution to the shortage of schools and housing due to the baby boom era of the fifties and sixties. This pre-fabricated system would go on to significantly influence Japan’s architectural landscape thereafter.
This project uses distinct historic eras as labels and prompts that teach AI how to distinguish the features from one period to the next.
<table>
<thead>
<tr>
<th>Region</th>
<th>Period</th>
<th>Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daido</td>
<td>806 (Oldest Existing House)</td>
<td>117</td>
</tr>
<tr>
<td>Edo</td>
<td>1603 - 1867</td>
<td>245</td>
</tr>
<tr>
<td>Taisho, Meiji</td>
<td>1868 - 1926</td>
<td>177</td>
</tr>
<tr>
<td>Showa</td>
<td>1926 - 1989</td>
<td>153</td>
</tr>
<tr>
<td>Heisei</td>
<td>1989 - 2019</td>
<td>171</td>
</tr>
<tr>
<td>Reiwa</td>
<td>2019 ~</td>
<td></td>
</tr>
</tbody>
</table>

Plans and facade images are likewise collected from the internet and input into the machine.
Some random architecture images were also added to the dataset to augment machine understanding of houses in the Japanese context in general.
During the data curation phase, image styles were optimized to highlight different expressions of architectural elements.
This dataset and period labels allowed machine learning to nuance the features of Japanese architecture.
Model Finding Journey
“Read between Threads”

Image Classification
(Discriminative AI)

Classifier based image generation model
Based on classification model provided in SCI 6487: Machine Aesthetics: The Binary and the Spectrum
By Panagiotis Michalatos
at Harvard GSD Spring 2023

→

Classification assessment model
Based on classification model provided in SCI 6485: Introduction to Generative Artificial Intelligence
By Sabrina Osmany
at Harvard GSD Fall 2023

Here it begins to read what is between the threads – revealing hidden features
Classification assessment model

Based on classification model

SCI 6485: Introduction to Generative Artificial Intelligence
By Sabrina Osmany

To identify the features of "fabrics" and "threads" contains in the scraped dataset, a supervised machine learning called classification is used here.
The model tries to predict the correct label of a given input data, in this case periodic information. By optimizing some parameters in the code that controls the machine's learning capabilities and the influence weight of the dataset, the classification model provides more accurate predictions.
The percentages in the images provide a more detailed quantitative prediction, indicating the likelihood of the houses belonging to each historical period.
My project took experimental steps for generation process to find the most effective methods including popular generative AI tools, classification-based modeling, CycleGAN, and stable diffusion models.
In the experimental iterations that used AI tools like DALL-E2 and Midjourney, the only way to synthesize data is mainly through texts in natural language, which forces one to give up most of design control.
The next model I explored is classification based generative model, which transforms features from one label period to another.

Edo

Reiwa

Other

*GIF
The results highlighted the predominant features of each period, thereby validating the quality of this dataset and machine's understanding of architectural features.
This project landed on stable diffusion modeling, which has been the most popular image generative model over the last few years.
This model is known for its high-quality and diverse output compared to GAN and autoencoder.
Stable Diffusion (SD) model
(text2image, image2image)

Text2image SD model
↓
Fine-tuned SD model with a pre-trained Pokemon dataset
↓
Fine-tuned SD model with my trained dataset
↓
Transformation SD model between prompts - GIF
↓
Image2image SD model
↓
Fine-tuned image2image SD model with my dataset

Starting from text to image generative model (.. which is basically an armature version of the AI tools mentioned earlier)
Stable Diffusion text-to-image fine-tuning

The `train_text_to_image.py` script shows how to fine-tune the stable diffusion model on your own dataset.

The text-to-image fine-tuning script is experimental. It’s easy to overfit and run into issues like catastrophic forgetting. We recommend to explore different hyperparameters to get the best results on your dataset.

Running locally

Installing the dependencies

Before running the scripts, make sure to install the library’s training dependencies:

```
$ pip install git+https://github.com/huggingface/diffusers.git
$ pip install -r requirements.txt
```

And initialize an `Accelerate` environment with:

```
accelerate config
```

And to make machine understand more specific subject and style, fine-tune the stable diffusion model, first with Pokemon image dataset as a test
Successfully created new Pokémon through text as an experiment.
Stable Diffusion text-to-image fine-tuning

```python
from diffusers import StableDiffusionPipeline
from PIL import Image
import torch

# Text2image SD model
Text2image SD model
↓
Fine-tuned SD model with a pre-trained Pokemon dataset
↓
Fine-tuned SD model with my trained dataset
↓
Transformation SD model between prompts - GIF
↓
Image2image SD model
↓
Fine-tuned image2image SD model with my dataset
```

Referencing the methods, built a text to image stable diffusion model extra-trained with my own Japanese housing dataset with periodic labels.
Stable Diffusion (SD) model
(text2image, image2image)

Text2image SD model
↓ Fine-tuned SD model
with a pre-trained Pokemon dataset
↓ Fine-tuned SD model
with my trained dataset
↓ Transformation SD model
between prompts - GIF
↓ Image2image SD model
↓ Fine-tuned image2image SD model with my dataset

Simultaneously created a text-to-gif model
To see a walk the model is taking, transformation process that reflects machine’s understanding and suggestions.
Machine walks around between the information threads and intersections pondering where to go and taking a detour – called latent walk.

GIF

Rio Kobayashi (MArch I 2024), Advisor: Andrew Witt

Harvard University Graduate School of Design
I further finetuned the model to make it image-to-image so that it can be effectively applied to the case study.

Stable Diffusion (SD) model
(text2image, image2image)

Text2image SD model
↓
Fine-tuned SD model
with a pre-trained Pokemon dataset
↓
Fine-tuned SD model
with my trained dataset
↓
Transformation SD model
between prompts - GIF
↓
Image2image SD model
↓
Fine-tuned image2image SD model
with my dataset

Make sure to check out the Stable Diffusion Tips section to learn how to explore the tradeoff between scheduler speed and quality, and how to reuse pipeline components efficiently!
The stable diffusion model transforms images by referencing the dataset of Japanese houses and capturing the nature of the hidden qualities.
Case Study:
Transforming Japanese Modern Houses
I return to my case study of three houses which are assumed to be located at the intersection of new central areas and an old residential town in northern Tokyo. Responding to Tokyo’s high population density and the limited space, residential areas near central Tokyo often consist of compact apartments and homes jostling for space.
Each house has two or three stories designed for a single family - a couple and children. This minimalistic façade indicates that they were newly built in the Reiwa period, during the last several years.
Their interiors use generic materials, adhering to modern housing design conventions.
Their assumed plans and façades exhibit typical patterns seen in modern Japanese housing.
When taking architectural plans and façade images and feeding them into a fine-tuned stable diffusion model...
the model manipulates period-specific design elements in the images by blending, amplifying, or transforming them and then optimizing them for a coherent and relatively seamless outcome.
Each façade is generated by different period-based prompts which demonstrate new hybridized possibilities of form, composition, and material use from eras like the Reiwa, Taisho, and Edo.
Transformed plans suggest how the spatial organization can blended by periodic essences and suggest new conditions of living.
In greater detail, some conventional plans absorbing hints of period based spatial logics.
<table>
<thead>
<tr>
<th></th>
<th>Edo</th>
<th>Meiji/Taisho</th>
<th>Showa</th>
<th>Heisei</th>
<th>Reiwa</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 420</td>
<td>now</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

House 1
Synthesis of Edo, Meiji/Taisho, and Reiwa
My process begins with an original building image that is fed into a stable diffusion model in order to generate suggestive images. I then curate and post-process those images to make them more legible while maintaining the suggestion of period influence. For the plans I sorted them into categories such as: wall structure, texture, spatial properties which were then extruded. Finally, I merged the hybridized façade onto the interior spaces.
In this case study, a distinct feature that is maintained throughout the diffusion transformation process is the courtyard space.
Interior spaces arranged around a central courtyard are typical in traditional Japanese homes and are known as 'tsubo-niwa'.
Transformed façade propose materiality palette mixed from the selected eras. Sugi-ban and wooden louver doors, ceramic tiles.
Further demonstrations of the transformation process: Here machine learning suggestions were translated and then realized in more specific and functional architectural language.
Here I further render machine-generated elements into detailed spaces and facades emphasizing an aesthetic of functionality and comfort.
The presence of an ‘engawa’, transitional space, further mediates between the personal and common areas, reinforcing the concept of semi-private zones that are open yet secluded. This approach allows the residents to experience the exterior environment without stepping outside, maintaining a delicate balance between exposure and privacy.
It also creates intimate outdoor spaces that enhance privacy in dense neighborhood settings. Vertical voids above the garden facilitate natural light and ventilation, contributing to environmental comfort.
House 2
Synthesis of Showa, Heisei, and Reiwa

Edo - Meiji/Taisho - Showa - Heisei - Reiwa

- 420 - now
It follows the same process with the first case.
The organization of rooms in these images showcases a 'box-in-the-box' situation where spaces are defined yet retain a sense of openness.
The use of wooden louvers allows for a visual connection between spaces while still providing separation.
Rio Kobayashi (MArch I 2024'), Advisor: Andrew Witt

Harvard University Graduate School of Design
Rio Kobayashi (MArch I 2024), Advisor: Andrew Witt

Harvard University Graduate School of Design
This arrangement is reflective of the Japanese principle of interconnectedness, where boundaries are suggested rather than explicitly stated,
allowing for flexibility and flow within the interior environment.
House 3
Synthesis of Meiji/Taisho, Showa and Heisei

- 420
now
Original data of House 3
It follows the same process with the first case
In the process of translation of the suggested plan, this case exhibits the concept of layered walls,
The composition is linked to traditional Japanese architectural technique often seen in castles that creates depth and a dynamic effect through the use of Fusuma doors.
The sliding panels not only serve as functional room dividers but also contribute to the visual narrative of the space. They create a sense of progression as one moves through the layers, offering varying degrees of privacy and interaction.
This arrangement can be a reflection of the Japanese spatial concept of 'ma', emphasizing the importance of voids, or negative space, which are just as significant as the solid, or positive space.
Each image represents a nuanced interpretation of Japanese design principles through collaboration with machine intelligence and collaboration with the past. The vast repository of architecture data embodies the diverse qualities of human existence that are closely intertwined with time and place, allowing us to weave obsolete features and knowledge into contemporary design thinking and practice, and in doing so, further enrich and support pluralistic projections far into the future.

This is just a humble introduction to prototypical workflow in the coming world of AI.
The integration of AI in the design process has the theoretical potential to bypass the 2D phase completely, allowing for the direct generation of 3D data from 3D datasets. 

However, these case studies indicate that 2D data-based generation process could be beneficial for designers within this AI-enhanced design framework.

These cases illustrate the possibility of maintaining authorial control without compromising the dynamic interplay between design elements. Machine intelligence is intended to enhance human creativity, not undermine it, as long as we learn how to effectively merge it with our thinking mind.
AI in architecture functions much like an intricate loom networked to the vast archives of design knowledge. It processes and proposes patterns that might elude the human eye, weaving new materials, technologies, and theories into the existing fabric. This integration does not disrupt the continuity of the tapestry but rather enhances it, introducing novel textures and colors that augment depth and nuance to the architectural landscape.
Dedicated to

my parents, Mari and Naonori
my aunt and sister, Miki and Nanao
my grandparents, Yoshiaki, Atsumi, and Noriko
and all of my friends and advisors

Thank you;)