Lego Mosaics

Algorithms, Software, and Other Fun Stuff

Deb Banerji

Last updated: 2022-9-16

Contents

- Intro
- Color Distance + Variations
- The Remix Problem
- 3D Mosaics
- Pixelization (If We Have Time)
- Code Links
- Q & A

Intro

- Who am I
- Who is this presentation for
- Disclaimer
  - Everything shared here is open source - please see the license for details
  - Feel free to contact me if you have questions, but I may not always have time to respond

What Is a Lego Mosaic

- Also known as ‘Lego Art’
- Brief history

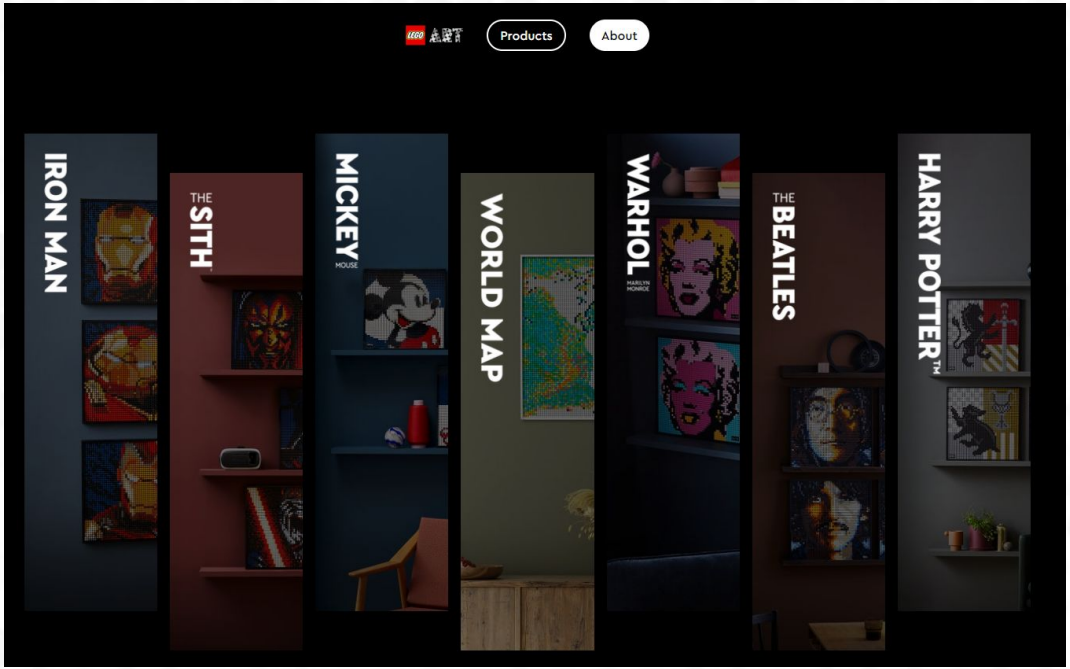
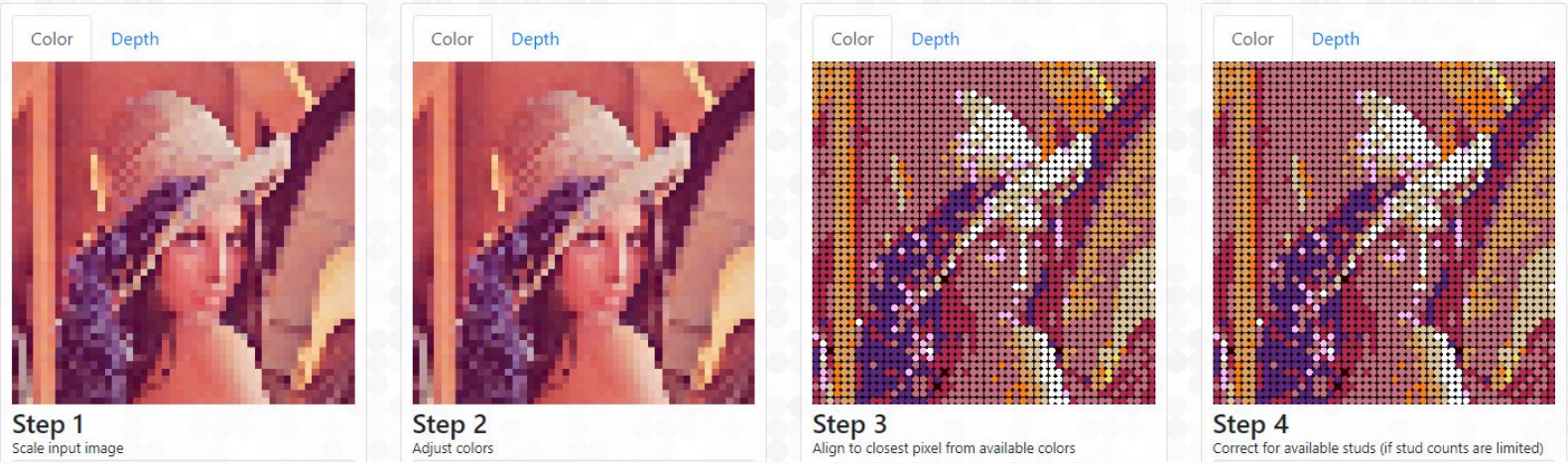


Image source: Lego.com

Lego Mosaic Software

- Has also existed for a while
- Some even created by Lego themselves!
- Designed to assist humans
- What we're going to explore today



Example: Image of Lena Forsén put through Lego Art Remix

What I Won’t Speak About (Much)

- Project architecture
- Deployment
- Security
- Performance optimizations
- Feature management
- Memory management
- HTML canvas usage
- Interactive pixel editing
- PDF generation
- Neural network deployment + details
- Other software engineering stuff
- Generative model integration possibilities, perf optimizations

Don’t Panic!

- Don’t worry if you don’t understand the details of the math
- It’s way more useful to focus on understanding the ideas the math is trying to capture

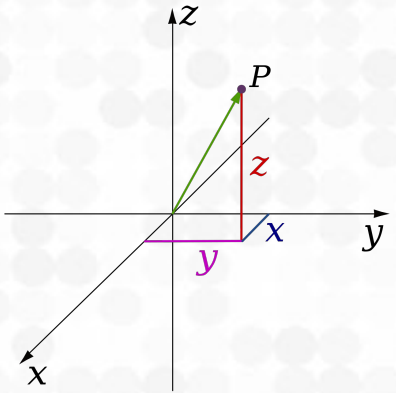
Color Distance

- Computer representation of colors (RGB, Hex Codes)
- What is color distance?

colorDist(#FC33FF, #8033FF) < colorDist(#FC33FF, #176C14)

An Example Color Distance Function

- Euclidean RGB is relatively simple to code + interpret



Source: https://en.wikipedia.org/wiki/Three-dimensional_space

- Other functions include CIE94, CIEDE2000, Euclidean LAB, etc.

Mosaic Creation

- Lego mosaic creation as minimizing a color distance function
- Scaling down + matching available Lego colors to pixels
  - For every pixel, find the ‘closest’ Lego color
- Mathematically, we’re minimizing the sum of the color distances between corresponding pixels (quantization error)

$$\sum \text{colorDist}(\text{original}_{i,j}, \text{replacement}_{i,j})$$

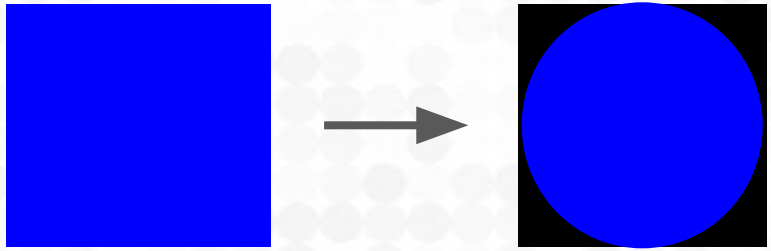
i = 1 to width, j = 1 to height

Examples

Source: ‘It Doesn't Have to Make Sense’ Album Art,
Ingrid Michaelson, 2016

Bleedthrough

- Pictures with round studs as pixels may be darker than expected
- Solution: multiply brightness of studs by $\pi/4$ before alignment



Transparency

- Image formats such as .PNG contain transparency channel for each pixel
- We can account for this by adding a transparency factor to the result of each color distance check
  - colorDist(original$_{i,j}$,replacement$_{i,j}$) + transparencyFactor
- ‘transparencyFactor’ is lower the closer the original is to the replacement in transparency
- Also need to change backing plate!

Brick image source: Bricklink.com

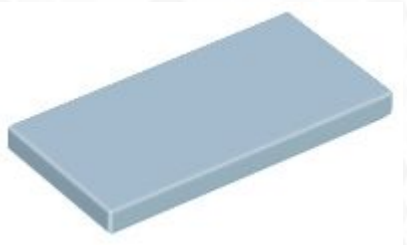
Metallicity

- Much trickier than transparency
- Isn’t stored as numbers in images
- Predict for each pixel with convolutional neural net
- Use same transparency trick afterwards
  - colorDist(original$_{i,j}$,replacement$_{i,j}$) + metallicityFactor

Brick image source: Bricklink.com

Complex Pixels

- What if pixels aren’t points?
- Asymmetrical pixels
- Pixel pieces with prints
- Multi pixel pieces (see ‘efficient tiling’)
- Raised pixels (see ‘3D mosaics’)



Brick image source: Bricklink.com

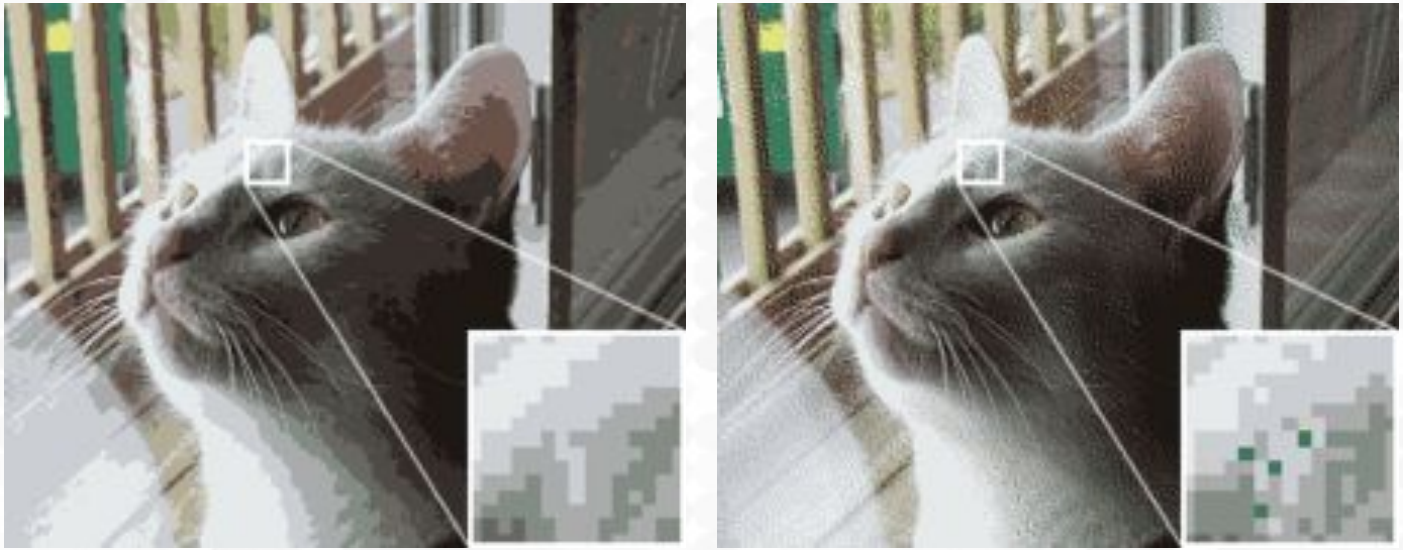
Brick image source: Lego.com

Price Optimization

- Some colors can be significantly more expensive
- May still be worth it!
- Quality vs. price tradeoff
- We can use the transparency trick again:
  - colorDist(original$_{i,j}$,replacement$_{i,j}$) + k*price(replacement$_{i,j}$)
- The higher ‘k’ is, the more we care about price
- This is a very efficient way to account for price

Dithering

- Useful for large outputs

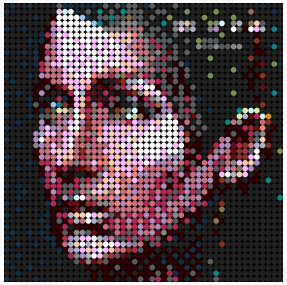


Source: https://en.wikipedia.org/wiki/Dither

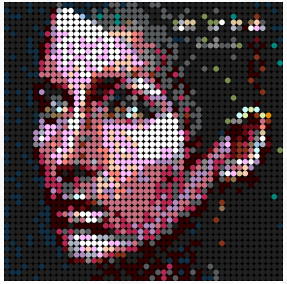
Dithering Strategies

- Caveat: Spreading error in color spaces not always accurate
- Most are linear algorithms - can’t handle remix problem, but are cleaner than greedy gaussian

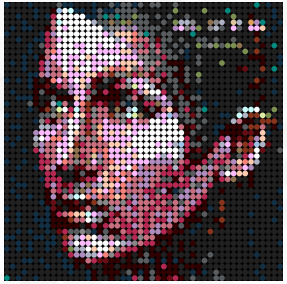
No Dithering



Floyd-Steinberg Dithering



Sierra Dithering



Greedy Gaussian Dithering

The Remix Problem

- What if we have limited numbers of studs for each color?
- Same goal, different constraints
  - $\sum$ colorDist(original$_{i,j}$,replacement$_{i,j}$)
- Much harder to minimize this now

Use Cases

- You want to design something with your existing pieces
- You want to design a MOC from an existing set
- You want to design something that’s easy to acquire
- You think it would be cool
- You work for Lego!

Assignment Problem/Linear Programming

Assignment problem

From Wikipedia, the free encyclopedia

The **assignment problem** is a fundamental combinatorial optimization problem. In its most general form, the problem is as follows:

> The problem instance has a number of *agents* and a number of *tasks*. Any agent can be assigned to perform any task, incurring some *cost* that may vary depending on the agent-task assignment. It is required to perform as many tasks as possible by assigning at most one agent to each task and at most one task to each agent, in such a way that the *total cost* of the assignment is minimized.

Alternatively, describing the problem using graph theory:

> The assignment problem consists of finding, in a weighted bipartite graph, a matching of a given size, in which the sum of weights of the edges is a minimum.

Solution by linear programming [edit]

The assignment problem can be solved by presenting it as a linear program. For convenience we will present the maximization problem. Each edge (i,j), where i is in A and j is in T, has a weight w_{ij}. For each edge (i,j) we have a variable x_{ij}. The variable is 1 if the edge is contained in the matching and 0 otherwise, so we set the domain constraints:

$$0 \leq x_{ij} \leq 1 \text{ for } i,j \in A,T,$$

$$x_{ij} \in \mathbb{Z} \text{ for } i,j \in A,T.$$

The total weight of the matching is: $\sum_{(i,j)\in A\times T} w_{ij}x_{ij}$. The goal is to find a maximum-weight perfect matching.

To guarantee that the variables indeed represent a perfect matching, we add constraints saying that each vertex is adjacent to exactly one edge in the matching, i.e,

$$\sum_{j\in T} x_{ij} = 1 \text{ for } i \in A, \quad \sum_{i\in A} x_{ij} = 1 \text{ for } j \in T,$$

.

All in all we have the following LP:

$$\text{maximize} \sum_{(i,j)\subset A\times T} w_{ij}x_{ij}$$

$$\text{subject to} \sum_{j\in T} x_{ij} = 1 \text{ for } i \in A, \quad \sum_{i\in A} x_{ij} = 1 \text{ for } j \in T$$

$$0 \leq x_{ij} \leq 1 \text{ for } i,j \in A,T,$$

$$x_{ij} \in \mathbb{Z} \text{ for } i,j \in A,T.$$

Solving the Remix Problem

- The remix problem reduces to the assignment problem in polynomial time
  - Instead of agents and tasks, we have studs and pixels
- The assignment problem similarly reduces to linear programming
- Many libraries exist for solving linear programs
- The catch: it’s slow, and memory usage can be bad
  - Can get around this by reducing problem size, but this can be unreliable

Solving the Remix Problem (Faster)

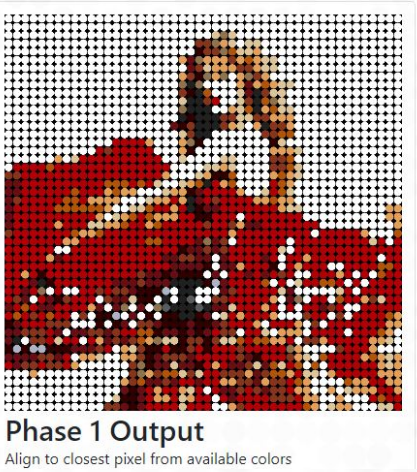
- To get around, this, we use a much faster 2 phase approximation algorithm
- The algorithm first fills in the ‘best matches’ for each available color as best as it can
- Then, it fills in the gaps using a similarly greedy approach
- It’s not perfect, but in practice, it works well under some reasonable assumptions:
  - Ex: If a user has at least one stud of a given color, then they can acquire more in the worst case (or remove this color entirely)
- In many cases, recreating the input image accurately just isn’t possible, so things will look quite weird

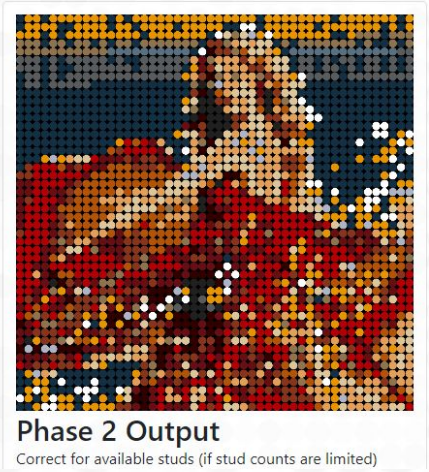
Solving the Remix Problem (Faster)

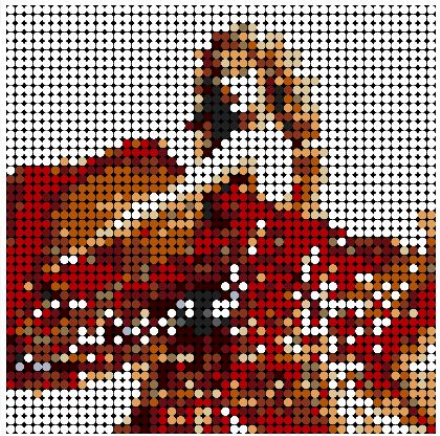
- Alternative to 2 phase: heap based algorithm
  - Pair each pixel with closest available color
  - Put each pair in a priority queue with priority = color distance
  - Keep dequeuing/assigning while decrementing color counts
  - If a color runs out, remove it and rebuild the queue
- Usually slower than 2 phase, but can handle dithering
  - We can spread the error to neighboring unassigned pixels
    - I use a gaussian pattern for this, but it is flexible
  - This is why a heap is more useful than sorting
  - Often looks weird at low resolutions
- Once again, a good result may not always be possible

Debugging

- We can get the difference to see what extra studs would help
- Ex: Trying to make an image from the iron man set
- Still need to match set and image for reasonable results





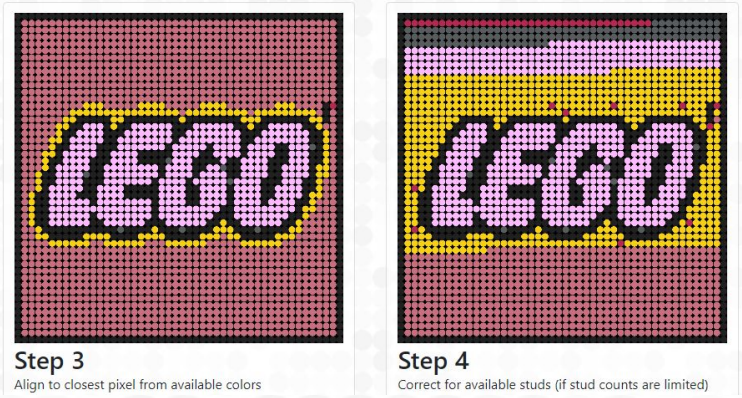


Easy to see that biggest problem is missing white studs, but other imperfections aren't too bad

Source: 'Speak Now' Album Art, Taylor Swift, 2010

Segmented Backgrounds

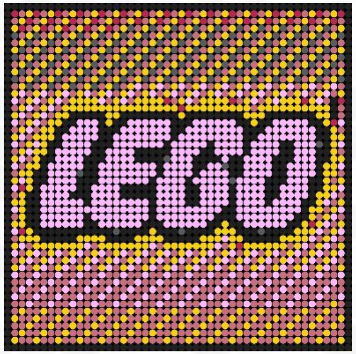
- For backgrounds and other monochromatic patches, we may fall short with color counts
  - Ex: previous slide
- May lead to weird/unpredictable segmentation

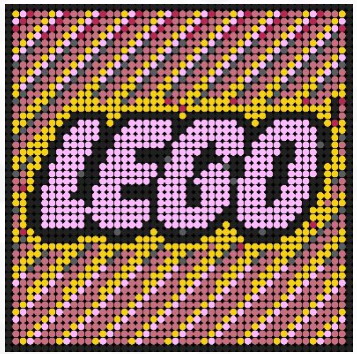


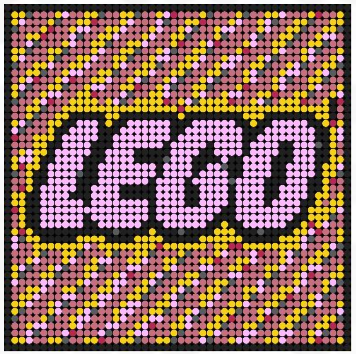
- We could detect, then manually process these
  - Can be difficult, may not be reliable

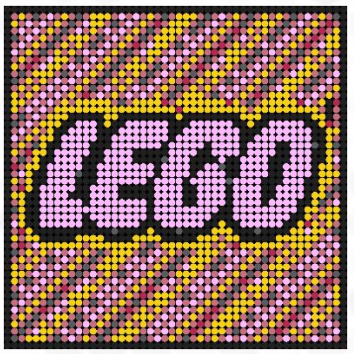
Pseudo ‘Dithering’

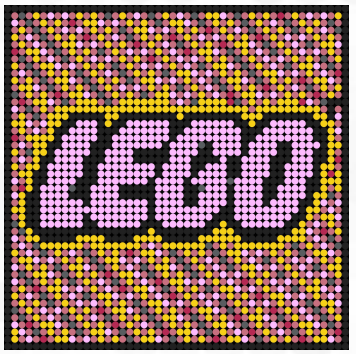
- Reuse our transparency color distance strategy
- This time, the extra factor is very small
  - Tiebreaking only
- colorDist(original$_{i,j}$,replacement$_{i,j}$) + ϵ((i+j) mod k)
  - Where ϵ and k are small, and k is an integer
  - Can also use cascading/alternating mod
  - Can also add even smaller noise after


Mod 3


Mod 5

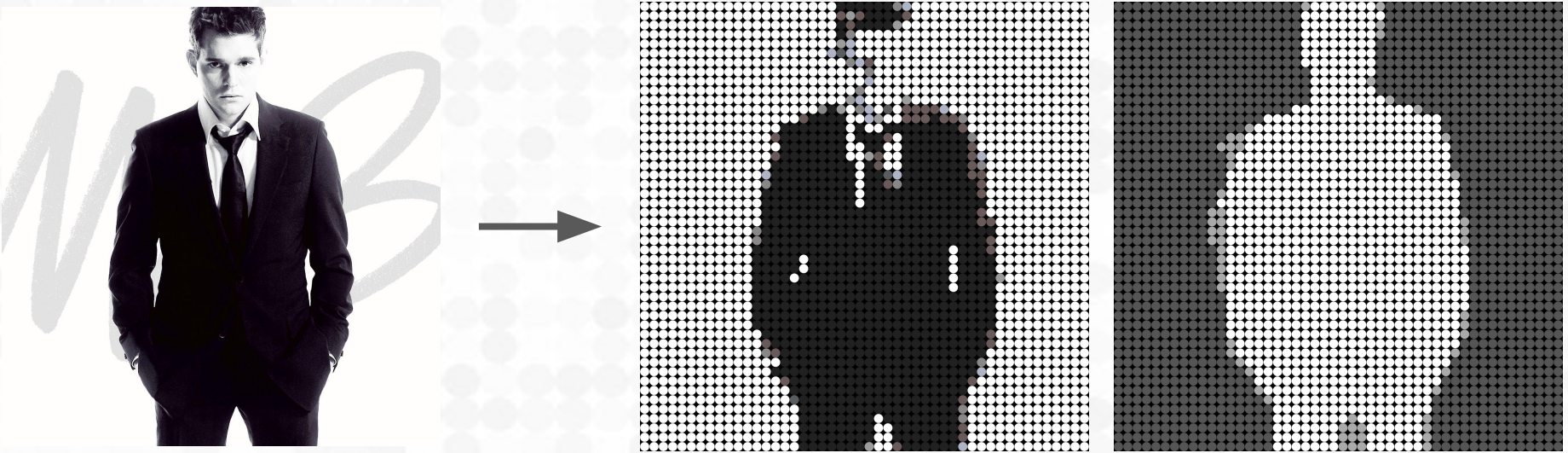

Noisy Mod 5


Cascading Noisy Mod


Alternating Noisy Mod

3D Mosaics

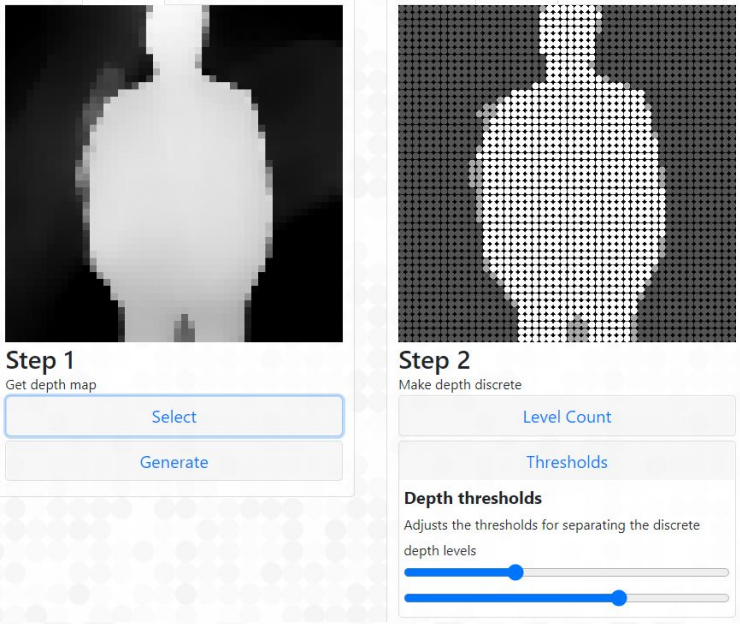
- Designing mosaics with depth
- Adding plates under certain pixels to raise them up



Source: `It's Time` Album Art,
Michael Bublé, 2005

Depth Maps

- Usually used for portrait mode, etc.
- Usually contain a depth value from 0-255 per pixel
- Reduced from 256 to 2-5 levels for 3D Lego mosaics

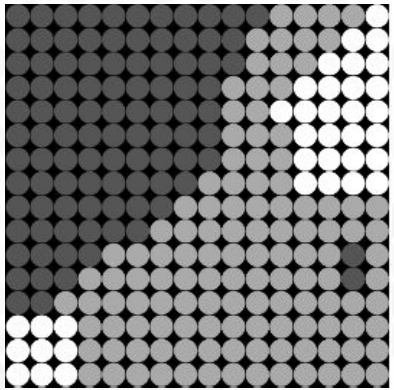


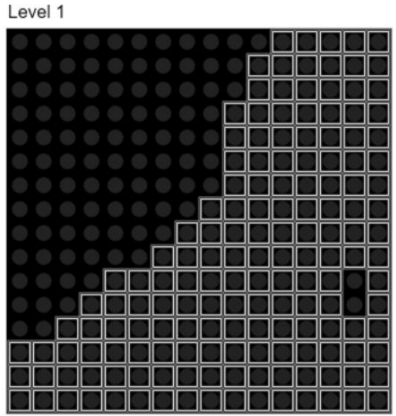
Inferring Depth Maps

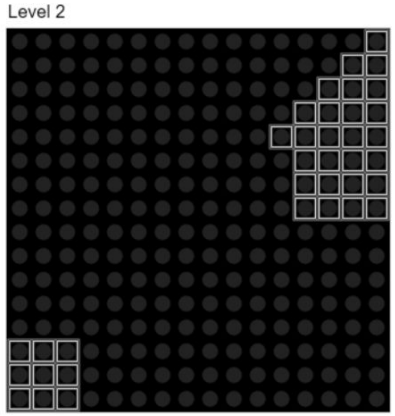
- What if a depth map for the image isn't provided?
- We can infer one using monocular depth estimation
- Many algorithms for this exist
  - A bunch of these are neural network based
  - The one I use in production is MiDaS
    - Ranftl, René, Katrin Lasinger, David Hafner, Konrad Schindler, and Vladlen Koltun. "Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer." (2020). *IEEE Transactions on Pattern Analysis and Machine Intelligence*

Backing Plates

- Using plates in order to raise up the pixels to create depth
- Generally one layer of plates per layer of depth
- Can also fill in more space with bricks (for >3 layers)
  - Can be annoying in many cases, affect reusability
- Best to avoid plates with large areas
  - Difficult to remove

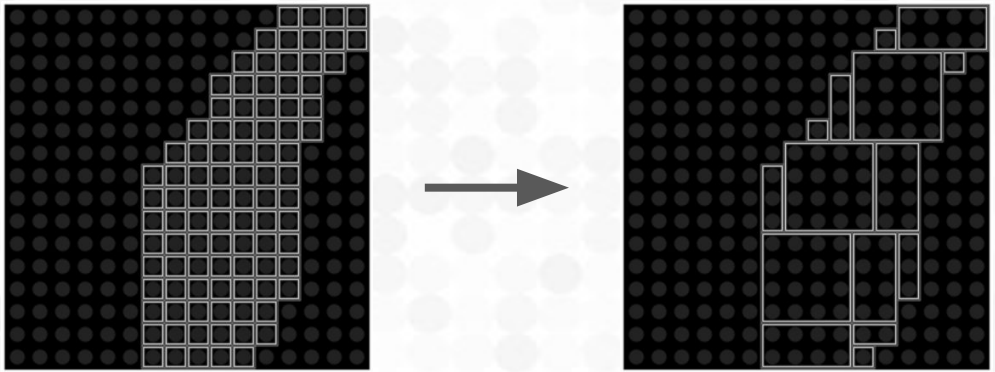






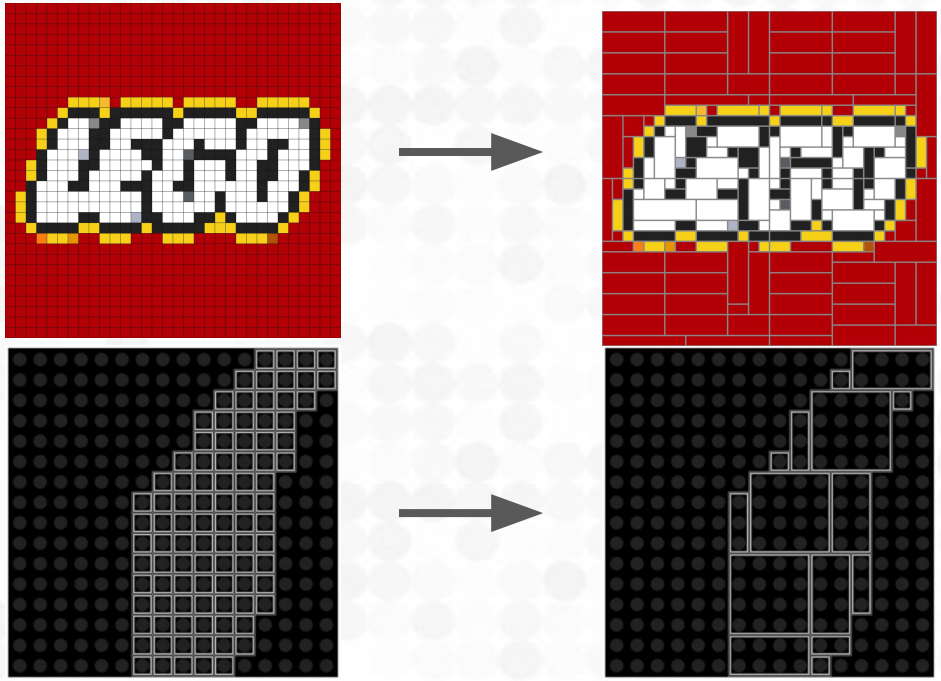
Efficient Tiling

- Want to minimize the number of plates used
- Given a set of possible plate dimensions
  - Might want to exclude large plates
- Solution: Sort by area, then fill from edges (‘stones + sand’)
  - Not perfect, but optimal in many cases
  - Guarantees adjacent small plates can’t be ‘merged’
  - Very fast



Efficient Tiling for Colors

- We apply the previous algorithm to each color individually
- Our remix algorithms are no longer applicable
- Can conflict with depth



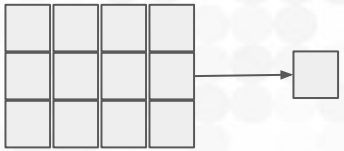
Pixelization

- Pixelization/pixelation/retargeting/downsampling/interpolation
- Low resolution displays vs. pixel art, Lego mosaics
- ‘Quality’ can be subjective
- Not as relevant for larger resolutions
- Blurred edges can be problematic in Lego mosaics
- Challenging to handle edges and gradients simultaneously

Source: ‘Red’ Album Art, Taylor Swift, 2012

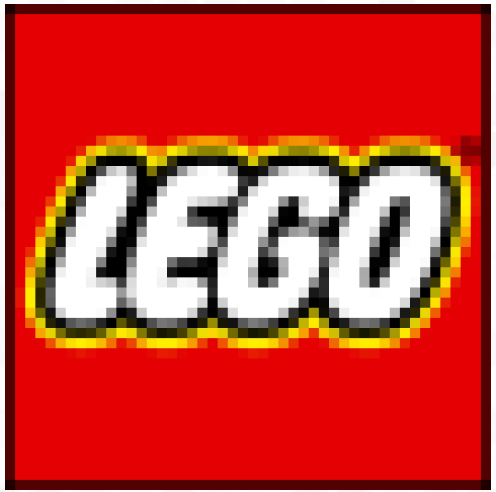
Pooling

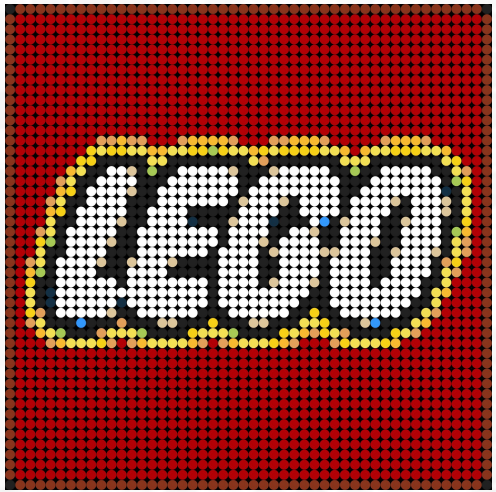
- Process a pool of pixels from original image and combine them somehow to create a pixel in the downsampled image
  - Ex: Average pooling, Max pooling
- Popular when downsampling in CNNs, because we're usually trying to extract information
- Browsers generally don't do pooling because they want to go for smoothness when seen from afar
  - Usually use bilinear/trilinear interpolation or something similar instead
- For our purposes, we can do per-channel RGB pooling



Average Pooling

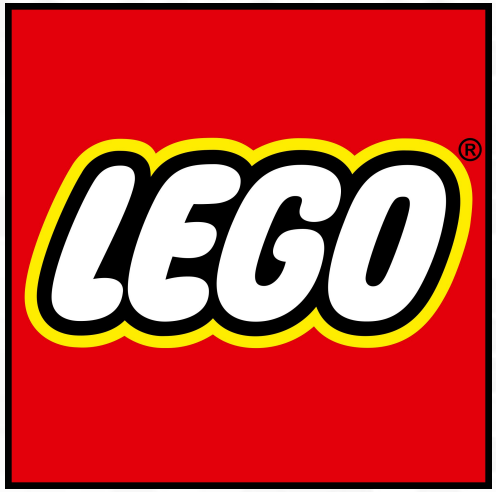
- Average out the value in the source pool to get the pixel
- Can lead to blurry edges

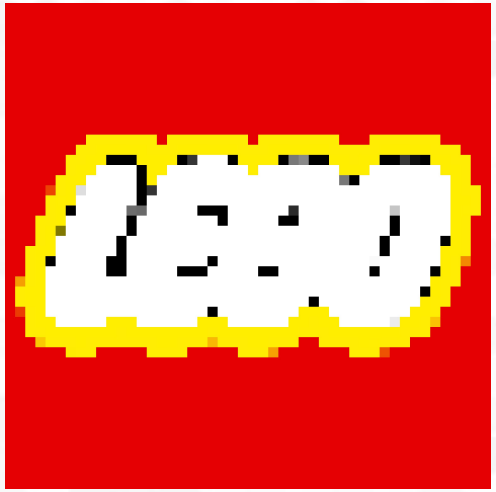


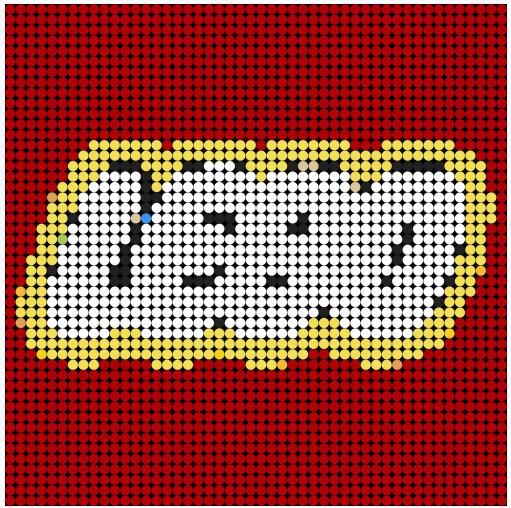


Max Pooling

- Fewer blurry edges
- Biased towards brightness
- Not suited for accurately recreating images

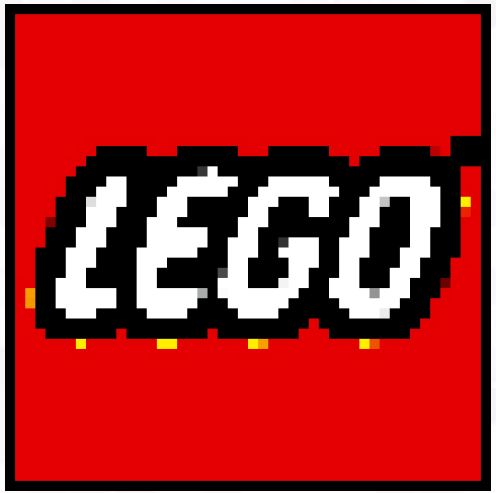


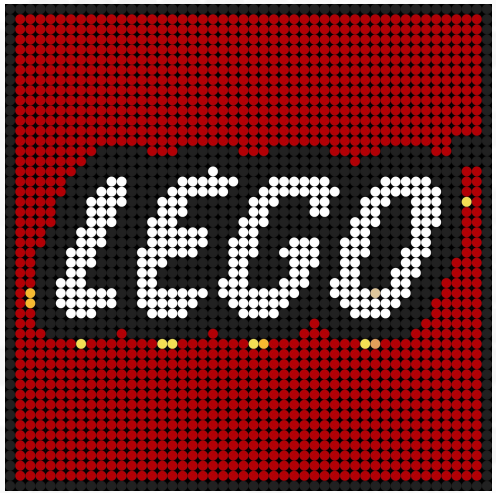




Min Pooling

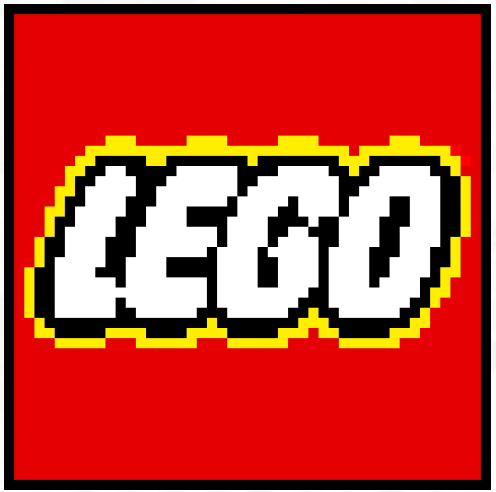
- Opposite of min pooling
- Opposite problem - resulting images are darkened
- Not as common in deep learning
- Once again, Not suited for accurately recreating images

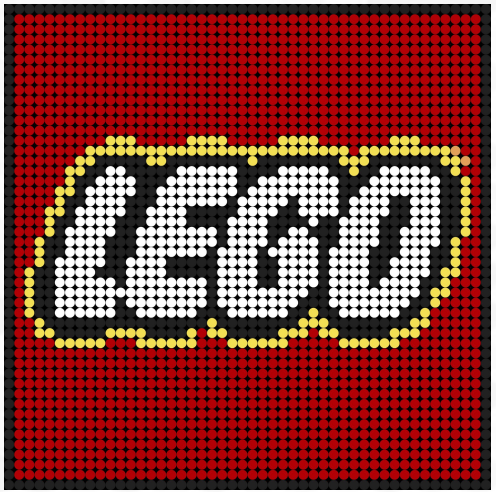




Dual Min-Max Pooling

- Designed to clean up edges without messing up total brightness
- Calculate min, max, and average, and then choose either min or max based on what is closer to average
- Works quite well in many situations
- Bilinear interpolation, etc. still works better in many situations





Segmentation Based Pooling

- General idea
  - Segment the image before pooling
  - Find majority category within each pool
  - Average only pixels in the majority category
- Doesn’t use semantic properties of segmentation model
  - Theoretically, a flexible non semantic model has speed/accuracy advantages here
- Not implemented yet, but theoretically addresses some biases in previously designed techniques
- Other idea: Super resolution -> Segmentation -> Pooling

Superpixels

- Gerstner, Timothy, et al. "Pixelated image abstraction." *NPAR@ Expressive*. 2012.
- Designed for pixel art, so good for Lego mosaics
- Can be expensive (many iterations usually used)

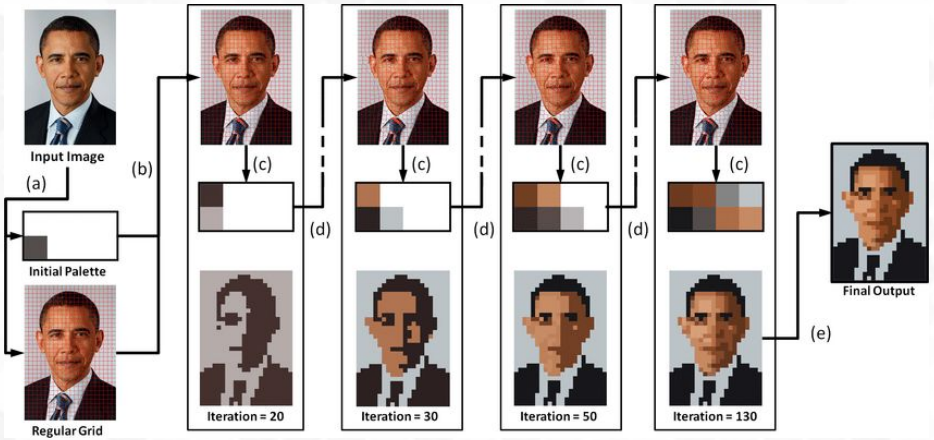


Image source: Same as citation

Learning Based Approaches

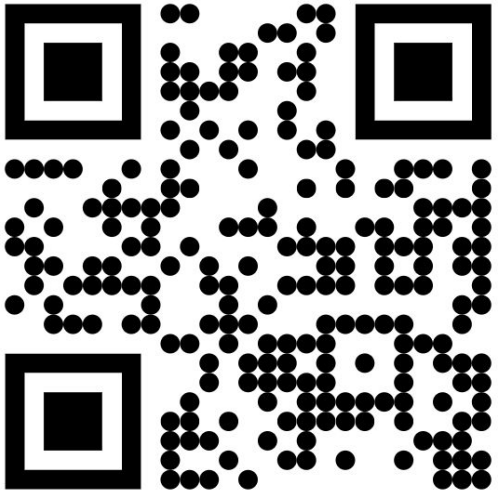
- Can we learn how to downsample + convert to Lego colors end to end?
- In theory, knowing which Lego colors are available in the downsampling step can help us more intelligently downsample
- A CNN based approach may be able to do so, but wouldn’t be able to address the remix problem out of the box
- Difficult to teach ‘good pixelization’ - i.e. getting a good dataset that in itself hasn’t been built from previous approaches
- Not implemented, but there are a few ideas for approaching this (ex: VAE that compares to pooled images for error)

Semantic Information

- Ex: “Each eye should be 1 pixel”, “Number of pixels in reflection should match”
- Hard to define rules around fine details and enforce them in code
- Today, humans are much better at this than computers - these are subjective, so tricky to decide what rules to enforce
- This is one of the big reasons the ability to edit mosaics in software is an important feature

Code Links

- github.com/debkbanerji/lego-art-remix
- See citations for non Lego art specific algorithm code
- See lego-art-remix.com if you don’t care about code and just want to use a Lego mosaic tool



Q & A

- If you have a feature request for Lego Art Remix specifically, you can also open an issue on github.com/debkbanerji/lego-art-remix/issues