An Automated Framework for Designing Subpixel Layouts with Gradient Descent Optimization

Kun-Yuan Lin*, Chieh-En Lee, Chung-Hao Tien
***i-want-the-degr.ee13@nycu.edu.tw**
Department of Photonics, National Yang Ming Chiao Tung University, Hsinchu, 300093 Taiwan

Abstract
In this study, we propose an optimization method to refine subpixel layouts by minimizing the Learned Perceptual Image Patch Similarity (LPIPS) loss. This approach ensures the final subpixel layouts can display images that closely resemble the original ones, achieving high visual fidelity and image quality.

Author Keywords
Subpixel Layout Design; Display Layout Design; Optimization; Gradient Descent.

1. Introduction
Innovations in display technology continually seek to enhance image clarity and visual fidelity. A critical area of development is subpixel rendering (SPR) [1]. Unlike traditional methods that treat display pixels as single units displaying one color, SPR introduces additional color detail within each pixel, improving overall image quality.

Traditionally, SPR techniques have been employed in various down-sampling methods to adjust images before display, transforming low-resolution images into high-resolution perceptions for the human eye [2]. Despite these advancements, traditional design methods [3, 4] remain limited by the existing subpixel configurations, such as RGB PenTile and PenTile diamond shape [5], constraining their performance and potential.

In this work, a design method based on gradient descent optimization is proposed, which aims to design subpixel layouts that can display clear images perceived by human eyes, meaning the perceived image should closely resemble the original image. The proposed method involves simulating how human eyes perceive images displayed on a layout formed by repeating a 7×7 unit layout pattern. The unit layout is optimized by minimizing the perceptual differences between the simulated and original images. Learned Perceptual Image Patch Similarity (LPIPS) is used as the loss function in this study, as it is less affected by minor pixel value changes that do not significantly impact overall image perception [6], making it a robust metric for evaluating image quality based on perceptual similarity.

While the optimized layouts discovered in this research have the potential to be further enhanced by applying SPR techniques in the future, the primary focus remains on achieving excellent display quality independently of SPR. This study aspires to contribute significant insights and practical solutions to the ongoing evolution of display technologies by improving image quality through optimized subpixel layouts.

The rest of the document is organized as follows. The proposed method and its corresponding simulation techniques are introduced in Sec. 2. The experimental results and discussion are presented in Sec. 3. Finally, a conclusion will be given in Sec. 4.

2. Method
(a) Dataset: The optimization uses 2000 images randomly chosen from the Stanford Dogs Dataset [7]. Each image is resized to 300×300 for consistency during processing and evaluation.

(b) Simulation: Simulating the image perceived by human eyes involves two key steps:

1. Simulate how images are displayed on the screen.
2. Simulate how human eyes perceive these images.

To simulate the image displayed on the screen $\left(\mathbf{I}_d^{R,G,B}\right)$, we compute Kronecker products between the original image $\left(\mathbf{I}_0^{R,G,B}\right)$ and the unit layout $(\mathbf{L}^{R,G,B})$ for each color channel C. The relation is expressed as:

$$\mathbf{I}_d^C = \mathbf{I}_0^C \otimes \mathbf{L}^C \tag{1}$$

where $\otimes$ denotes the Kronecker product operation. This process effectively models how the original image is "expanded" by the repeating unit layout, simulating the output of the display. Given the small size of the subpixels, the light they emit tends to spread out as it travels from the panel to the eyes, causing a blending effect. In this study, we apply a Gaussian kernel as the Point Spread Function (PSF) in the spatial domain to simulate this spreading effect. Different kernel sizes simulate various observation distances; when the display is viewed from a greater distance, the light spreads out more, resulting in a blurrier image. Thus, a larger kernel size is chosen to simulate this effect. For our purposes, the kernel size is set to 19 to ensure an appropriate level of blurring. This PSF convolution is necessary to accurately replicating how human eyes perceive images on the display. The simulation process is shown in Figure 1.

(c) Optimization Process: A gradient-based approach is proposed to optimize the layout. Figure 2 illustrates the optimization process. The simulation is applied to both the objective layout $\tilde{\mathbf{L}}^{R,G,B}$ and the target layout $\mathbf{L}_0^{R,G,B}$. We then calculate the LPIPS loss between the two perceived images and optimize the objective layout via gradient descent. The objective is to discover an $\tilde{\mathbf{L}}^{R,G,B}$ that achieves a comparable level of clarity to $\mathbf{L}_0^{R,G,B}$. A desired layout should have the following properties:

- Every pixel value is either zero or one.
- RGB channels do not overlap on the same position.

However, these properties are often violated during the optimization process without any constraints. To address this, we process each pixel value of $\tilde{\mathbf{L}}^{R,G,B}$ at location i, denoted as $x_i^{R,G,B}$, with three procedures before applying the simulation:

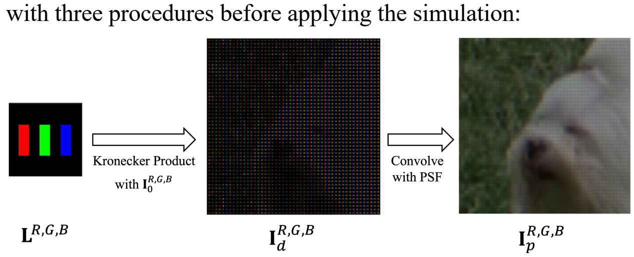

Figure 1. The simulation process.

ISSN 0097-996X/25/5603-1838-$1.00 © 2025

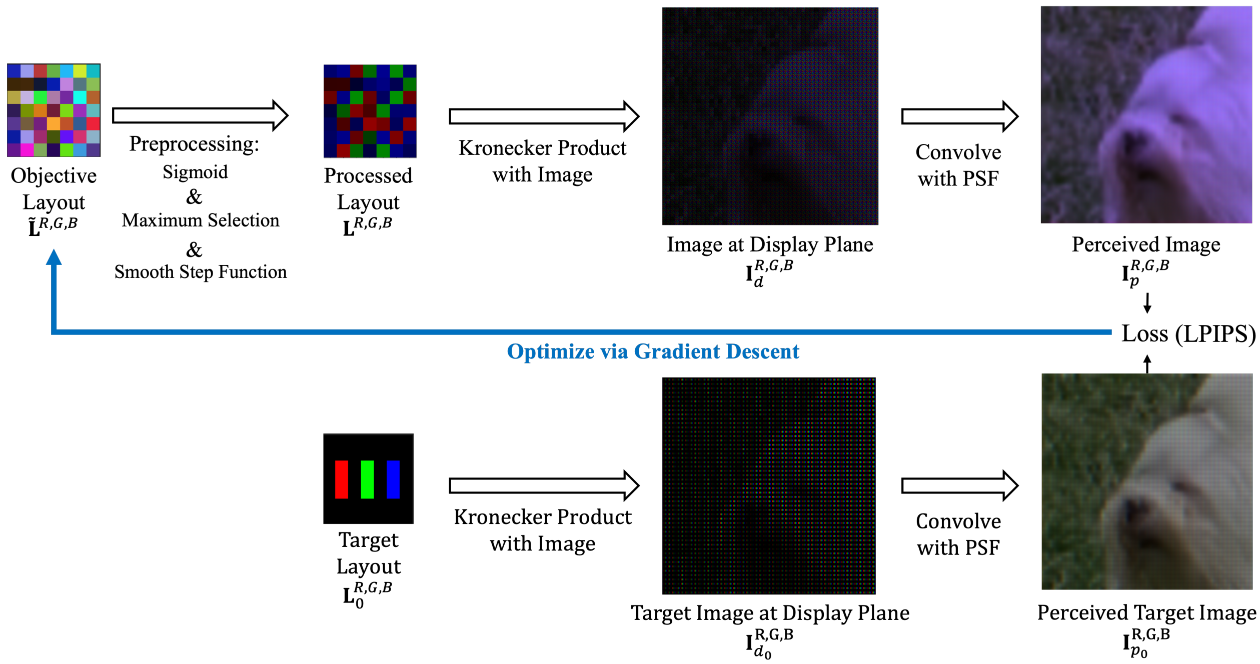

Figure 2. The flowchart of the optimization process. It starts with a randomly initialized layout $\tilde{\mathbf{L}}^{R,G,B}$ and a target layout $\mathbf{L}_0^{R,G,B}$ (for example, an RGB stripes layout). The goal is to optimize $\tilde{\mathbf{L}}^{R,G,B}$ to create an $\mathbf{L}^{R,G,B}$ capable of displaying images with clarity comparable to that achieved by $\mathbf{L}_0^{R,G,B}$.

1. Sigmoid
2. Maximum Selection
3. Smooth Step Function

Sigmoid: The definition of the sigmoid function is given in equation 2:

$$\sigma(x_i^C) = \frac{1}{1+e^{-x_i^C}} \tag{2}$$

Any value that's been transformed by the sigmoid function would be ranging from zero to one. The Sigmoid function is applied to ensure that every value in the layout is within the range of zero and one. This is essential because it limits the pixel values to a valid range for display purposes, preventing unrealistic values that could distort the image representation.

Maximum Selection: After applying the Sigmoid function, the channel with the highest value for each pixel is retained, while the others are set to zero. This ensures that each pixel represents a single-color component without interference from other channels. The selection process is mathematically described in Equation 3.

$$M(x_i^C) = \begin{cases} x_i^C, & if\ x_i^C = max\ (x_i^R,\ x_i^G,\ x_i^B) \\ 0, & otherwise \end{cases} \tag{3}$$

Smooth Step Function: The values in the layout must be binarized, which can be achieved by applying a step function:

$$u(x_i^C) = \begin{cases} 1, & if\ x_i^C > threshold \\ 0, & if\ x_i^C \le threshold \end{cases} \tag{4}$$

However, step functions are not differentiable, making them unsuitable for gradient descent optimization. To address this, we approximate the step function using a modified sigmoid function:

$$s(x_i^C) = \frac{1}{1+e^{-\left(scale\cdot\left(x_i^C - threshold\right)\right)}} \tag{5}$$

This "smooth step function" behaves similarly to the step function but is differentiable, enabling its use in optimization. Elements greater than the threshold are approximate to one, and others to zero, with a larger scale value pulls the values further away from the threshold.

We experimented with different thresholds and scales for equation 5 to determine the optimal parameter. The optimization process was run for 50 iterations, and the results are shown in Figure 3, where invalid pixel values of $\tilde{\mathbf{L}}^{R,G,B}$ are clipped. Each layout begins with an identical random initial objective layout and was optimized using an RGB stripes $\mathbf{L}_0^{R,G,B}$. After evaluating the convergence results, we found that the configuration with a threshold of 0.7 and a scale of 10 yields the best $\mathbf{L}^{R,G,B}$. Therefore, this configuration would be used for subsequent optimizations.

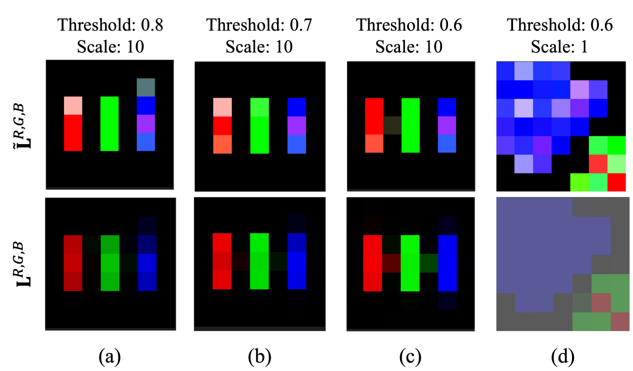

Figure 3. Results of different thresholds and scales.

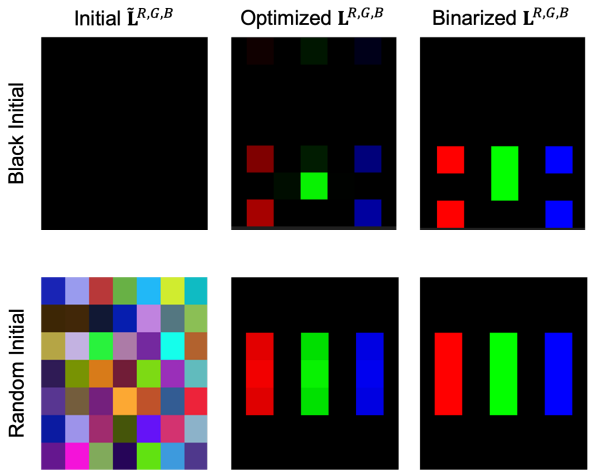

Figure 4. Optimized layouts from an RGB stripes $\mathbf{L}_0^{R,G,B}$.

(d) Achieving the Final Layout:

We obtain the outcome once the $\mathbf{L}^{R,G,B}$ converges. However, it's important to note that $\mathbf{L}^{R,G,B}$ may still contain non-binary values, meaning there could be values other than one and zero. To address this, we further binarize into a binary layout where values above a certain threshold are set to 1 (indicating activation) and values below the threshold are set to 0 (indicating non-activation). This binary layout is then used as the desired layout, representing the final outcome of the optimization process, ensuring that only the desired channels are activated in the final layout.

3. Results

To evaluate the effectiveness of the proposed optimization method, we conducted experiments using both RGB stripes and white target layouts. For each target layout, we used both completely black masks and random masks as initial objective masks for optimization. All optimized layouts were obtained after 1000 iterations.

The results of using an RGB stripes layout as the target layout are shown in Figure 4. The optimized layouts converge to a configuration that resembles the RGB stripes, demonstrating that the optimization process can closely match the target layout. Next, we used a white mask as the target layout. The white mask target layout is expected to yield a layout that can produce images with clarity closely resembling the original images since a white mask implies no color filtering. The resulting layouts and the simulated perceived images can be seen in Figure 5. The optimized layouts display clear and perceptually accurate images, validating the efficacy of our approach.

Figure 5 also shows the comparison between the optimized layouts and some existing layouts, including the traditional RGB layout, RGB PenTile, and PenTile diamond shape. A randomly generated layout that has 7 subpixels for R, G, and B respectively is also included as a reference. We also calculate the LPIPS of the simulated perceived images within the dataset. The average LPIPS loss is presented in Table. 1. Lower LPIPS loss suggests higher perceptual similarity between two images. The LPIPS values demonstrate that our method consistently produces layouts with high perceptual quality.

In summary, the experiments show that our optimization method effectively produces subpixel layouts that achieve high perceptual similarity to the target layouts. This is evidenced by the convergence to layouts that visually and quantitatively (LPIPS values) match the target layouts. The ability to optimize towards white target layouts highlights the versatility and robustness of our approach in generating high-quality subpixel layouts suitable for enhancing display technologies.

4. Conclusion

In this paper, we proposed an optimization method for designing subpixel layouts capable of displaying images with high clarity, closely resembling the original images perceived by human eyes. By simulating the perceived images using Point Spread Function (PSF) and minimizing the differences between the perceived and original images with the LPIPS loss function, the proposed method successfully converged to an optimized subpixel layout.

Our experiments demonstrated that the optimized layouts, when evaluated against both RGB and white target layouts, produced images with high perceptual similarity, validated through LPIPS metrics. The method proved effective regardless of the initial objective mask, whether completely black or randomly initialized, consistently converging to high-quality layouts.

This research underscores the potential of our optimized layouts in enhancing display technologies, with applications that could extend to subpixel rendering techniques in the future, paving the way for display systems.

5. Impact and Significance

Our findings demonstrate that the proposed method produces layouts with superior perceptual quality compared to existing layouts, including traditional RGB layouts. By effectively addressing limitations in subpixel configurations and utilizing the LPIPS metric to prioritize perceptual similarity, our approach introduces an innovative strategy for optimizing layout designs, achieving enhanced visual fidelity.

Unlike conventional subpixel rendering, which relies on pre-defined layouts and focuses on post-rendering adjustments, our method optimizes layouts at the design level, ensuring the perceived image quality is inherently higher. Furthermore, the use of a differentiable optimization framework allows for greater adaptability to diverse target layouts, showcasing its versatility.

These advancements have implications for a wide range of applications in display technologies, particularly in improving resolution, color accuracy, and visual fidelity. By enhancing the visual clarity, our work sets a strong foundation for integrating optimized subpixel layouts into modern display manufacturing, ultimately benefiting consumer electronics, virtual reality, and other visual media industries.

Table. 1 Average LPIPS Between the Perceived Images Simulated from the White Mask and Different Layouts

Layout	RGB	RGB PenTile	PenTile Diamond	Processed (black init.)	Binarized (black init.)	Processed (rand. init.)	Binarized (rand. init.)	Random Layout
Avg. LPIPS (10^{-3})	310.37	50.52	72.84	0.09	39.11	0.02	15.26	107.10

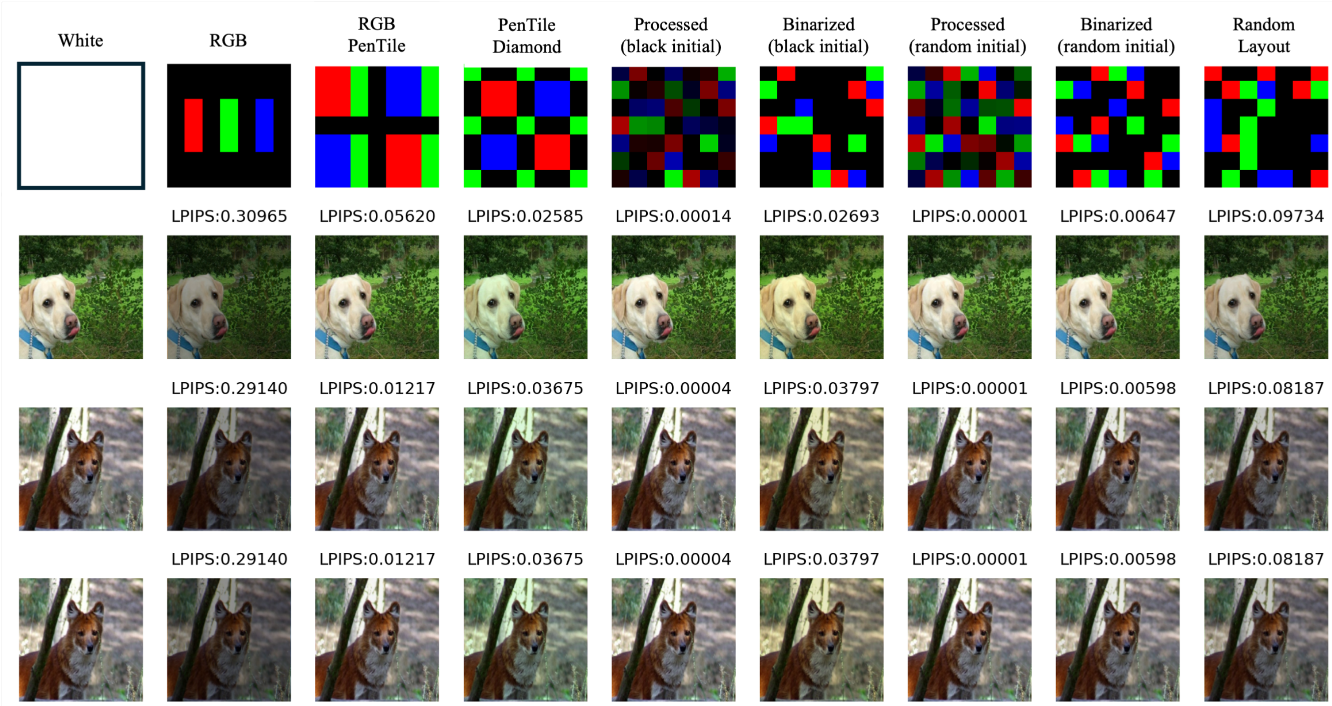

Figure 5. Results comparison for existing layouts and the layouts optimized using a white mask target layout.

6. Acknowledgements

This work is supported by the NSTC, Taiwan under the contracts: NSTC 113-2218-E-A49-019-MBK and NSTC 113-2622-E-A49-031-.

7. References

[1] L. Fang, O. C. Au, K. Tang, and X. Wen, "Increasing image resolution on portable displays by subpixel rendering – a systematic overview," *APSIPA Transactions on Signal and Information Processing,* vol. 1, p. e1, 2012, Art no. e1, doi: 10.1017/ATSIP.2012.3.

[2] S.-H. Chae, C.-H. Yoo, J.-Y. Sun, M.-C. Kang, and S.-J. Ko, "Subpixel rendering for the pentile display based on the human visual system," *IEEE Transactions on Consumer Electronics,* vol. 63, pp. 401-409, 11/01 2017, doi: 10.1109/TCE.2017.015103.

[3] S. K. Hong *et al.*, "New Pixel Design on Emitting Area for High Resolution Active-Matrix Organic Light-Emitting Diode Displays," *Journal of Display Technology,* vol. 6, no. 12, pp. 601-606, 2010, doi: 10.1109/JDT.2010.2063694.

[4] B. Shi *et al.*, *Sub-pixel Layout for Super-Resolution with Images in the Octic Group*. 2015.

[5] C. Brown Elliott, T. Credelle, S. Han, M. Im, M. Higgins, and P. Higgins, "Development of the PenTile Matrix™ color AMLCD subpixel architecture and rendering algorithms," *Journal of The Society for Information Display - J SOC INF DISP,* vol. 11, 01/01 2003, doi: 10.1889/1.1831725.

[6] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 18-23 June 2018 2018, pp. 586-595, doi: 10.1109/CVPR.2018.00068.

[7] E. Dataset, "Novel datasets for fine-grained image categorization," in *First Workshop on Fine Grained Visual Categorization, CVPR. Citeseer. Citeseer. Citeseer*, 2011, vol. 5, no. 1: Citeseer, p. 2.