Creating Enhanced Autostereograms Using Disparity Convolution and CNNs

1. Main Project Idea:

The goal of this project is to create visually compelling autostereograms, commonly known as "Magic Eye" images, by leveraging advanced deep learning techniques like disparity convolution. Traditional autostereograms are limited by their reliance on simple repeating patterns to encode depth information. In contrast, my approach uses neural networks, particularly CNNs and disparity convolution, to generate richer depth maps and more dynamic textures, improving both depth perception and the overall aesthetic quality of the images. The core concept is to incorporate cutting-edge computational techniques to enhance traditional autostereogram design, making the depth information clearer and more visually engaging.

2. Proposed Approach:

Our approach involves the following key steps:

A. Depth Map Generation (Alternative to 3D Model Rendering):

Instead of using a complex 3D-to-2D rendering pipeline, I have simplified the process by directly working with depth maps that can be easily generated from 3D models. This eliminates unnecessary steps and streamlines the process. Depth maps are generated using publicly available random images. The depth information represents how far or near different parts of the object are in space.

- **Input:** Images from user
- **Process:** Use Python libraries like Pyrender or other depth-rendering tools to create depth maps from the images. This step directly outputs depth maps, eliminating the need for intermediate stereo-image pairings.

This approach cuts down on the unnecessary complexity of going from stereo to images to depth map, as was previously planned.

B. Autostereogram Generator:

Once we have the depth maps, the next step is to encode the depth information into an autostereogram. The depth map values are used to shift the pixels of a texture horizontally—closer objects are represented by larger shifts, while distant objects are represented by smaller shifts.

- **Input:** Depth maps generated in the previous step.

- **Process:** Select either randomly generated textures or CNN-generated textures (which could be trained to match real-world scene textures). The horizontal shifts in the texture pixels are mapped based on the depth information, creating the illusion of 3D objects when viewed correctly.

By combining these shifted textures, the final stereogram image is created, which appears as a flat pattern but reveals hidden 3D shapes when the viewer focuses their eyes appropriately.

C. Disparity Convolution Network:

This step focuses on improving the depth perception and texture clarity in the stereogram by using disparity convolution within a neural network framework. Here, we simulate the concept of stereopsis, which is the human ability to perceive depth from the visual information provided by two eyes, using horizontal disparity between pixels.

- **Disparity Convolution:** Instead of comparing a pixel's feature vector with its adjacent neighbor (which was a mistake in the original proposal), disparity convolution involves learning the disparity map based on feature differences across the image. This process estimates depth more accurately by learning the depth relationships across the entire image.
- **Network Architecture:** We use a CNN architecture such as ResNet or U-Net, where disparity is encoded into feature maps that guide the neural network to adjust texture representations for better depth transitions and clarity. The features are refined through multiple convolutional layers to enhance the depth representation.
- **Process:**
    1. The depth map features are encoded into feature maps using disparity convolution, which learns the depth information through horizontal shifts of pixel values across the image.
    2. The CNN then refines these depth representations to make the textures more realistic, ensuring smoother transitions between near and far objects in the stereogram.

The use of CNNs allows for dynamic texture generation, adding variety and complexity to the stereogram beyond traditional methods.

3. Input/Output Data:

- **Input Data:** Images from user
- **Output Data:**
    - **Enhanced Autostereograms:** The final images will be the autostereograms where viewers can perceive hidden 3D shapes when viewed correctly. These images will incorporate optimized textures for improved depth perception and aesthetic clarity.

4. Training Data:

- **3D Object Models:** ShapeNetCore dataset with over 50,000 3D models, which includes 55 categories of objects. These models are rendered into depth maps directly using tools like Pyrender.
- **Textures:** Textures will be collected from online repositories or procedurally generated. To simulate real-world conditions, textures will be augmented with transformations like blurring, compression, and color adjustments to add variety and realism.
- **Augmented Data:** To improve robustness, the dataset will be augmented using common image transformations like cropping, flipping, rotation, and scaling. This will help train a network that is resilient to different input conditions and improves the generalization of the model.

5. Evaluation Plan:

- **Robustness Testing:** We will test the stereograms under various degradation scenarios, such as blurring or compression, to assess their robustness and the network's ability to recover accurate depth information under real-world conditions.
- **Comparison:** The final results will be compared to traditional stereogram methods to evaluate improvements in terms of clarity, depth perception, and aesthetic value.

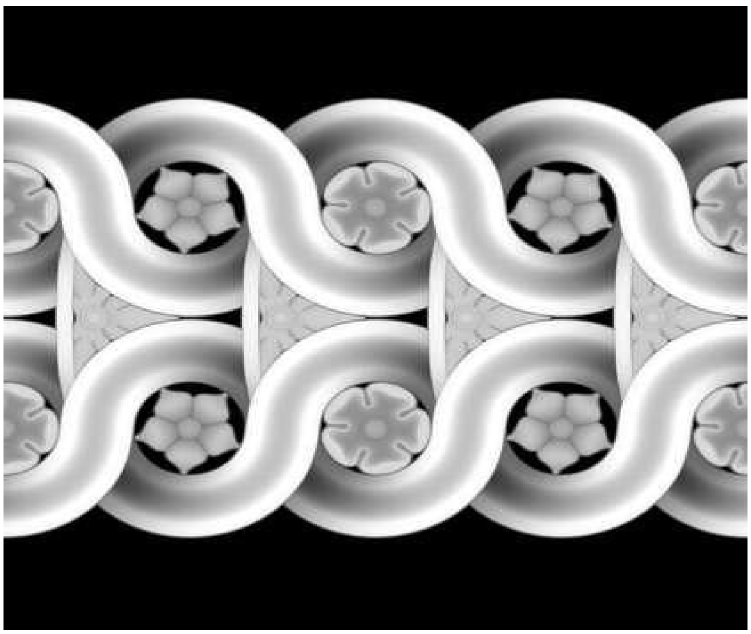
6. Impact:

This project has significant implications across several fields:

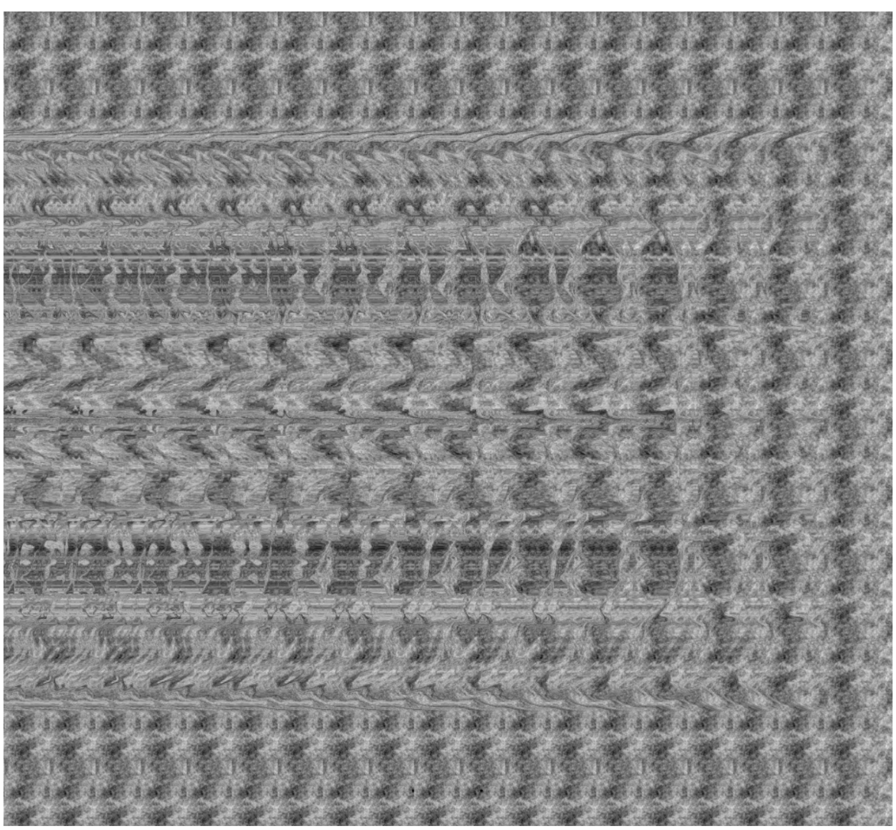
- **Art and Design:** Provides new creative possibilities for generating complex and interactive 3D illusions, offering more visually engaging stereograms for both artistic and educational purposes.
- **Education:** Offers a practical application of vision science, allowing students to explore human stereopsis and depth perception in a computational context. This could be useful in teaching computer graphics, vision science, or machine learning.
- **Accessibility:** The enhanced depth perception could aid individuals who find traditional stereograms difficult to interpret. The disparity convolution process mimics human depth perception more effectively, potentially offering a more accessible version of autostereograms.
- **Technology and Research:** This project demonstrates how advanced deep learning techniques, such as disparity convolution, can be applied to unconventional image synthesis problems, expanding the boundaries of what's possible in both artistic and scientific image processing.

Depth Map: Shows the grayscale image where brightness represents depth (closer is brighter).
Output autostereogram: Appears as a repeating texture with hidden 3D information. When viewed correctly (diverging eyes), it reveals the 3D object encoded from the depth map.

1. Depthmap 1
   Input:



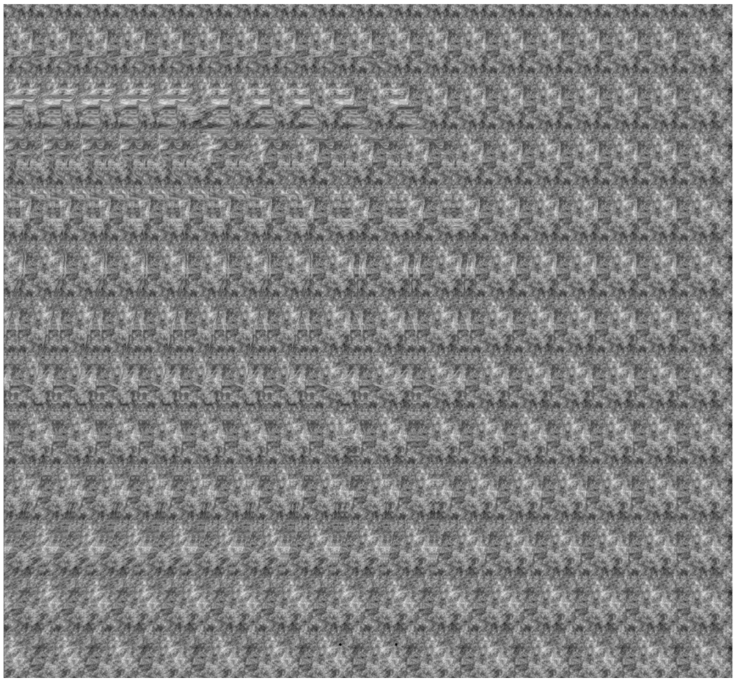
   Output autostereogram:



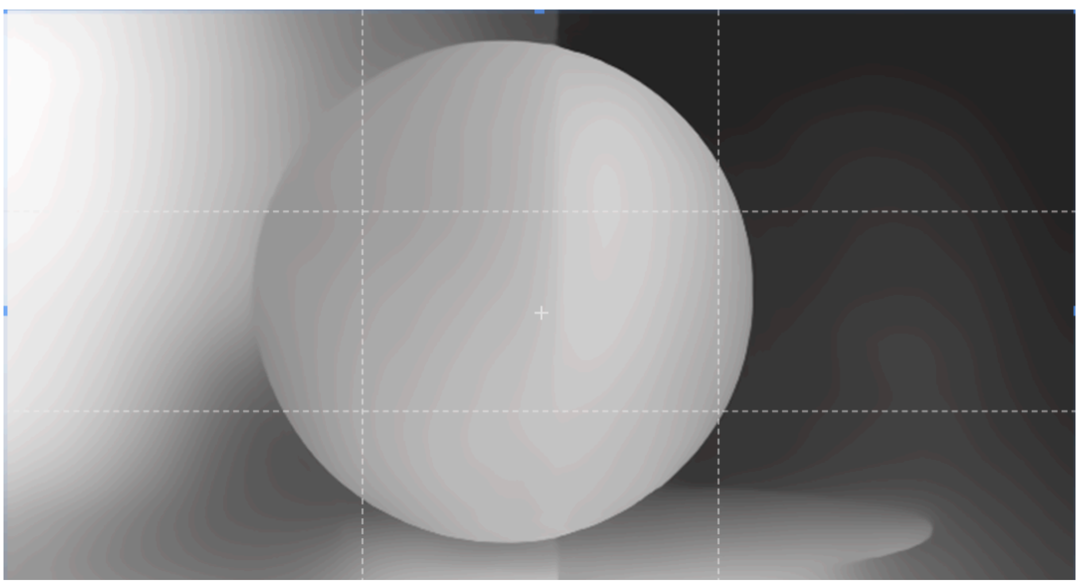
2. Ironman
   Input:

   Converted to a grayscale depth map :

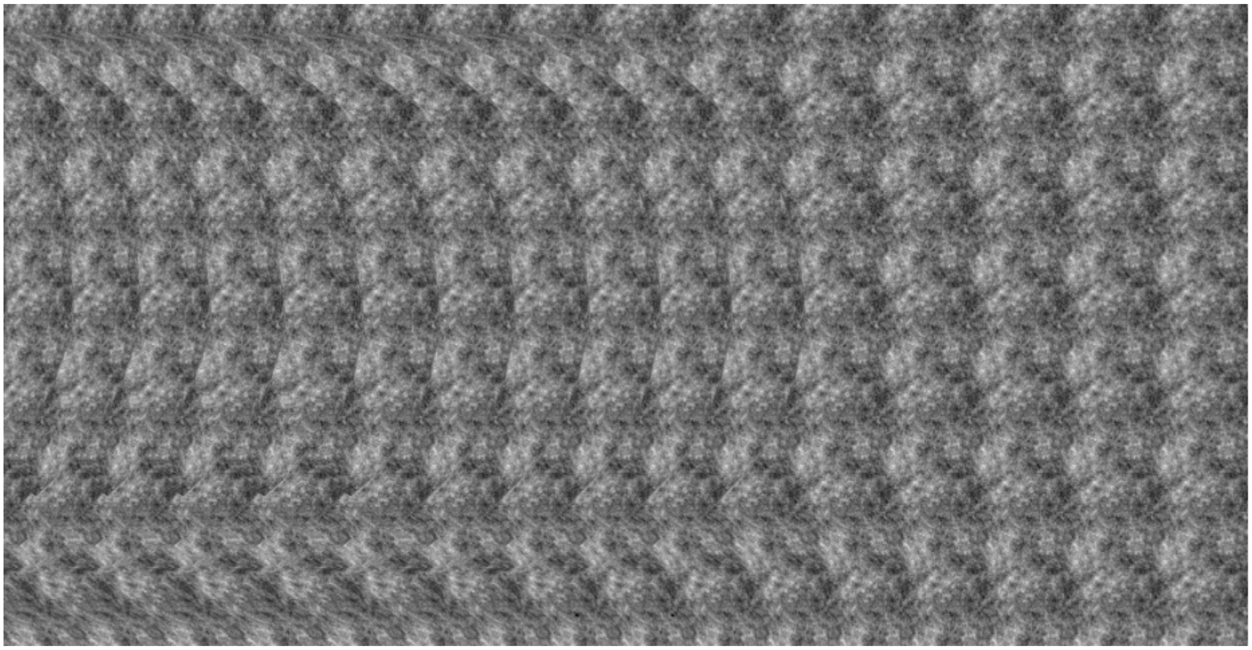
Output autostereogram



3. Depth map 2:
   Input: is a grayscale image



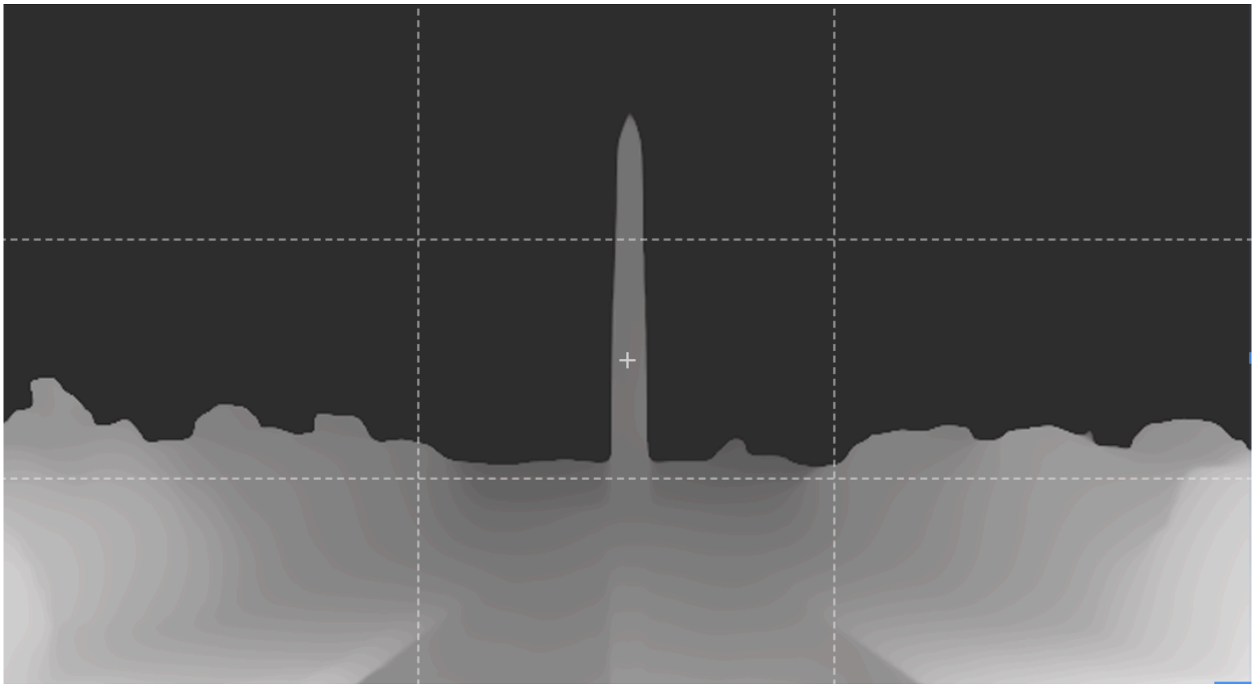
Output autostereogram :



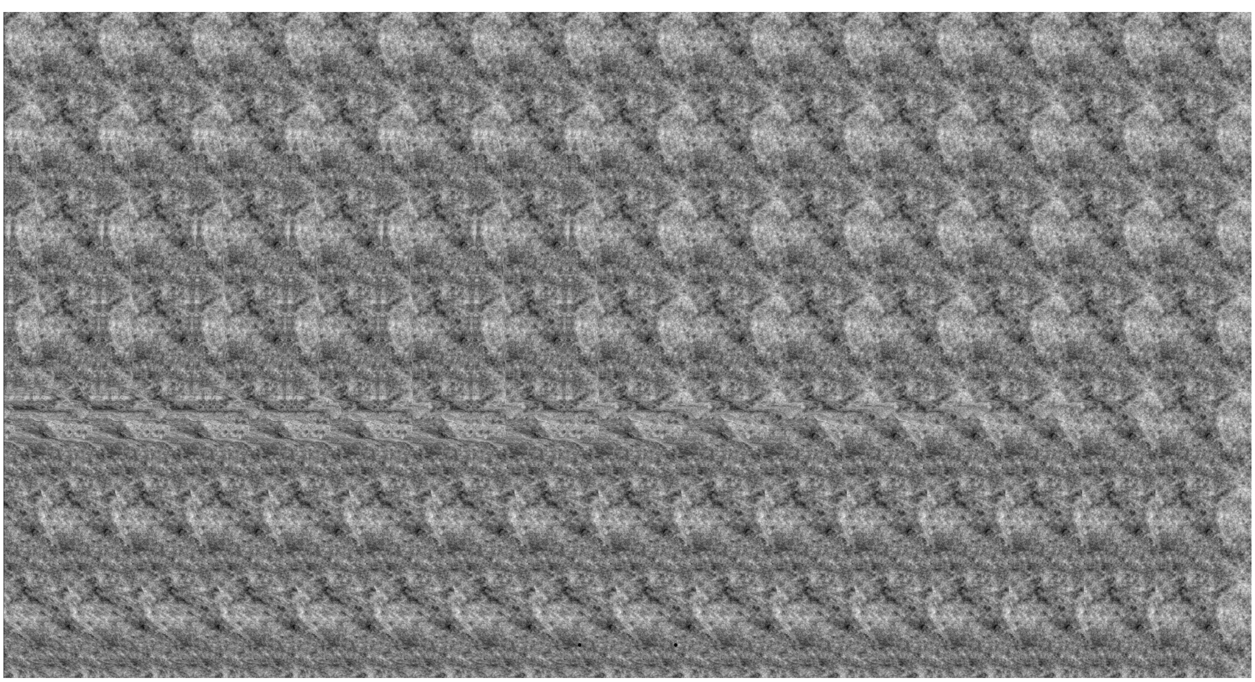
4. Washington Monument:

Input:

Converted to grayscale depth map:



Output autostereogram:



Conclusion: This implementation demonstrates an innovative approach to creating autostereograms. By combining disparity convolution with CNNs, it achieves enhanced depth representation, making the 3D effect more vivid. Dynamic textures, adding artistic and aesthetic value. This methodology can be extended to real-world applications in education, art, and human vision research, emphasizing the synergy between computational techniques and visual design.

Google Colab Link:
https://colab.research.google.com/drive/1EcxGkhXvj1eoSKN3eiaeCDa5zyp68ejj?usp=sharing