Style-Transfer via Texture-Synthesis

Fares Atef, Ghaith Mohammed, Amr Mohammed, Amr Magdy

Cairo University, Computer Engineering Department

***Abstract*— Style transfer has emerged as a pivotal technique in image processing and computer vision, drawing considerable attention for its capacity to blend the artistic essence of one image (the style image) with the content of another (the content image). This transformative process intricately marries the aesthetic appeal of the style image while preserving the foundational structure of the content image. This paper presents an in-depth exploration into the pioneering work of Michael Elad and Peyman Milanfar in the realm of style transfer. Building upon their foundations, our study delves into experimental modifications that demonstrably refine and augment specific aspects of the established methodologies.**

I. Introduction

Style transfer, a transformative process in the realm of image manipulation, revolves around imbuing the visual style of one image, termed the style image, onto the content of another, the content image. However, the transfer of stylistic attributes is inherently abstract, allowing for varied interpretations and approaches. The evolution of style transfer techniques has witnessed significant strides, especially with the advent of Convolutional Neural Networks (CNNs), catalyzing unprecedented advancements in this domain.

The seminal work by Gatys et al. laid the groundwork for neural style transfer, harnessing CNNs to capture and reproduce artistic styles, resulting in compelling image transformations. While CNN-based methodologies revolutionized style transfer, recent explorations have rekindled interest in classical approaches. Michael Elad and Peyman Milanfar reinvigorated the classical methodologies by revisiting Kwatra's seminal paper, innovatively augmenting its framework to achieve aesthetically enriched results.

The resurgence of interest in classical methods stands as a testament to the quest for holistic and refined style transfer techniques. Elad and Milanfar's enhancements, rooted in the foundations laid by Kwatra, transcend the confines of traditional paradigms, offering a nuanced amalgamation of classical wisdom and contemporary computational prowess.

This paper embarks on an exploratory journey, traversing the evolution of style transfer from its neural network-driven advancements to the reinvigoration of classical methodologies. Delving into the intricacies of these approaches, we aim to elucidate the nuances and divergences, thereby contributing to a comprehensive understanding of the dynamic landscape of style transfer in image processing.

II. The Proposed Algorithm

The essence of the style transfer algorithm lies in the minimization of an Energy function (1). This function encapsulates pivotal parameters crucial to the style transfer process, each contributing significantly to the augmentation of the content image based on specific criteria. The nuanced interplay of these parameters forms the bedrock of the algorithm, orchestrating a sophisticated interweaving of content and style.

$$\frac{1}{c}\sum_{(i,j)\in\Omega_{L,n}} Min\left\|R_{ij}^{n}X - Q_{kl}^{n}D_{L}^{S}S\right\| + \|D_{L}^{C}C - X\| \qquad (1)$$

s.t:

X represents the estimated image
C represents the content image
S represents the style image
R_{ij}^{n} represents extraction of the i,j-th patch of size nxn
Q_{kl}^{n} represents extraction of the k,j-th patch of size nxn
L represents the working scale
c is for normalization

The inclusion of these parameters within the Energy function is integral to the algorithm's functionality. Our study aims to meticulously dissect these parameters, exploring their individual contributions to the criteria governing style transfer. This detailed analysis seeks to provide a comprehensive understanding of how these parameters collectively impact the augmentation of content images in the style transfer process.

The Algorithm consists of eight main components which are the following:

1) *Segmentation*
2) *Building Gaussian Pyramids*
3) *Initialization*
4) *Color Transfer*
5) *Patch Matching*
6) *Robust Aggregation*
7) *Content Fusion*
8) *Denoise*

Each step will be discussed in details.

A. Segmentation

Before initiating the core algorithm, a segmentation map serves as a crucial input, delineating the salient regions within the content image. This map assigns importance measures to each patch, enabling the preservation of significant content elements to a certain degree.

Our exploration encompassed the implementation of different segmentation techniques. Initially, we employed GrabCut, as advocated in Elad's paper. However, GrabCut's reliance on user-defined foreground and background annotations often resulted in suboptimal performance, particularly when these annotations were absent, leading to degraded outcomes.

Subsequently, an alternative edge-based segmentation technique was adopted. This technique involved a series of steps, starting with the application of the Canny edge detection algorithm, followed by dilation, contour extraction, and subsequent filling. Notably, this approach exhibited superior performance in multiple scenarios, showcasing improved outcomes compared to GrabCut in diverse contexts.

It is noteworthy that while GrabCut excelled, particularly in well-defined portrait scenarios characterized by distinctly defined foregrounds, the edge-based technique demonstrated greater versatility and consistency across various image contexts.

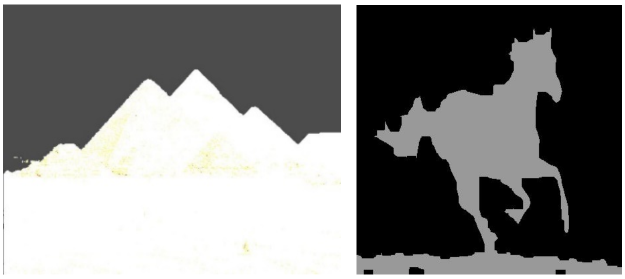

Fig. 1. **Left**: Example of segmentation mask using GrabCut algorithm **Right**: Example of segmentation mask using edge-Based Segmentation.

B. Building Gaussian Pyramids

The process of constructing Gaussian pyramids involves downscaling the resolution of the content, style and segmentation mask images, creating multiple layers. Operating the style transfer algorithm on each of these layers is instrumental in enhancing the content image's capacity to assimilate broader stylistic features from the style image.

By iteratively reducing the resolution, the content image progressively incorporates higher-level stylistic elements present in the style image. This hierarchical approach facilitates the inheritance of larger and more encompassing features from the style image, enabling a richer fusion of stylistic attributes within the content image.

The utilization of Gaussian pyramids, with their multi-layered representations, serves as a pivotal strategy to impart an encompassing and comprehensive style essence from the style image to the content image, enriching the overall stylistic transformation process.

C. Initialization

Following the completion of pre-processing functions and pyramid construction, the algorithm initiates X, the primary image subjected to algorithmic operations. Initially set as the content image, X undergoes an additional procedural step wherein a significant amount of Gaussian noise is applied.

The intentional introduction of substantial Gaussian noise serves a pivotal purpose in facilitating the subsequent patch matching phase. This deliberate perturbation amplifies the divergence within patch matching, preventing multiple segments of the content image from excessively converging onto the same patch from the style image. This deliberate diversification mitigates the tendency for pattern repetition, ensuring a more nuanced and diversified style transfer effect.

The deliberate injection of Gaussian noise into the content image, serving as the basis for subsequent operations, fosters a more pronounced and authentic style transfer experience, minimizing repetitive patterns and enhancing the overall stylistic coherence within the transformed image.

Fig. 2. **Left**: Example of patch matching on uniform white image with no noise **Right**: Example of patch matching on uniform image with added noise.

D. Color Transfer

The process of color transfer involves transposing the color palette from the style image onto the content image. Several methodologies were explored to accomplish this task, commencing with histogram matching, as advocated in Elad's paper

While histogram matching showcased efficacy in many instances, it occasionally exhibited limitations, leading to instances of extreme color transfer. To mitigate these challenges and ensure a more nuanced and controlled color transformation, an alternative approach was adopted.

Subsequently, color transfer in LAB channels was implemented as an alternative methodology. This method leverages the LAB color space to perform the color transformation, offering greater flexibility and control over the color transfer process.

The adoption of color transfer in LAB channels overcomes the limitations encountered with histogram matching, enabling a more refined and controlled approach to color palette application from the style image to the content image.

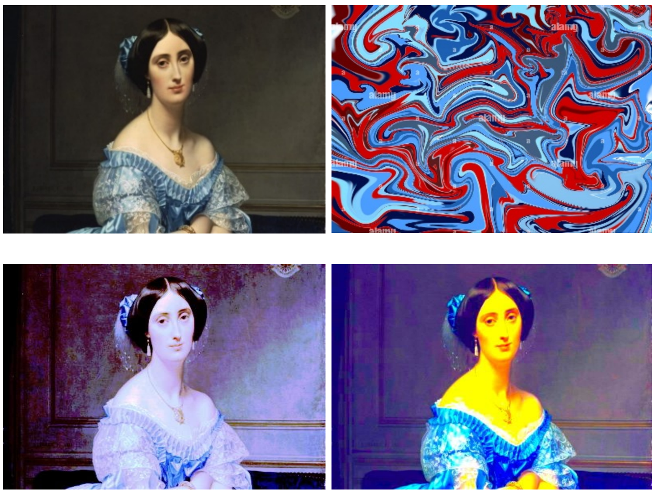

Fig. 3. **Left**: Example of color transfer using histogram matching **Right**: Example of color transfer using LAB channels.

E. Patch Matching

The crux of patch matching revolves around identifying the nearest patch in the style image that corresponds to the current patch in the content image. This fundamental process ensures that the changes introduced post-style transfer remain coherent with the original content. For instance, in a scenario where the content image features a sky, after style transfer, the resultant image's sky retains a semblance to that depicted in the painting, such as Van Gogh's "Starry Night." This preservation of content is pivotal, underscoring the significance of patch matching in achieving a harmonious blend of style transfer while upholding content integrity.

$$Min\left\|R_{ij}^{n}X - Q_{kl}^{n}D_{L}^{S}S\right\| \qquad (2)$$

The minimization task, denoted by equation (2) aims to determine the most fitting patch in the style image, corresponding to each patch in the content image, using a Nearest Neighbor approach. However, the exhaustive nature of Nearest Neighbor for each patch incurs substantial computational overheads.

To address the computational complexity while retaining fidelity in patch matching, Principal Component Analysis (PCA) was employed as an alternative. PCA's principle lies in reducing the image's dimensionality to its essential components, enabling effective differentiation between patches. This reduction in dimensionality streamlines the matching process, offering computational efficiency without compromising the accuracy of patch correspondence.

Fig. 4. **Left**: Example of patch matching with patch sizes = [40, 30] **Right**: Example of patch matching with patch sizes = [20, 10] (using Van Gogh "starry night" style image).

F. Robust Aggregation

Robust Aggregation, a crucial procedure within the style transfer algorithm, involves the systematic adjustment of each patch in the content image. This adjustment aims to iteratively align each content patch closer to its matched counterpart in the style image.

The methodology employed for this purpose, as elucidated in Kwatra's paper, relies on Iteratively Reweighted Least Squares (IRLS). This technique systematically refines the content patches by iteratively adjusting their attributes to converge towards the corresponding style patches.

IRLS operates by iteratively recalculating and reweighting the contributions of individual patches in the content image. Through this iterative reweighting process, patches are gradually modified to exhibit closer alignment with their matched style patches. This adaptive adjustment ensures a robust and refined aggregation of stylistic attributes from the style image onto the content image, contributing significantly to the coherent application of style transfer.

G. Content Fusion

Content Fusion serves as the pivotal stage in the style transfer process, focusing on reapplying the content image in alignment with the segmentation outcome. This crucial step aims to ensure the preservation of significant content elements within the final stylized output.

The segmentation map, guiding this fusion process, plays a decisive role in delineating the essential areas of the content image. By leveraging the segmentation results, the algorithm selectively integrates the content image, emphasizing the preservation of these marked crucial segments.

This meticulous fusion mechanism allows the retention of vital content details, ensuring that the style transfer process prioritizes and upholds the integrity of these identified important areas. As a result, the final stylized output maintains the essence of the original content image, strategically blending it with the stylistic attributes borrowed from the style image.

H. Denoise

Denoising stands as a crucial step within the style transfer process, pivotal in guaranteeing seamless transitions within the resultant image. It plays a pivotal role in refining the overall visual quality by minimizing unwanted artifacts or noise.

Throughout our experimentation, various denoising methodologies were rigorously assessed for their efficacy in achieving smooth transitions within the output image. Among the methodologies scrutinized, the Bilateral Filter emerged as the most proficient performer.

The Bilateral Filter, distinguished by its ability to preserve edges while effectively reducing noise, showcased superior performance in refining the image, ensuring smooth transitions without compromising on the fidelity of structural details. This denoising technique significantly contributed to the enhancement of the final stylized output, achieving a desirable balance between noise reduction and preservation of image details.

III. Experiment Results

We managed to get good results from our implementation, and also got some failure cases. In this section, we will show some of our results.

Fig. 5 displays some of our successful results. The left images represent the content, the middle ones depict the style, and the right ones show the results. In each case, patch matching worked perfectly. For instance, in the horse example, the dark part of the style image appears around its legs. Additionally, the edges were well-preserved with a smooth transition between the segmented part and the background, thanks to the denoising process.

Fig. 6 demonstrates an additional feature where we preserve geometric shapes in the background.

Fig. 7, 8. show some failure cases due to bad styles or segmentation.

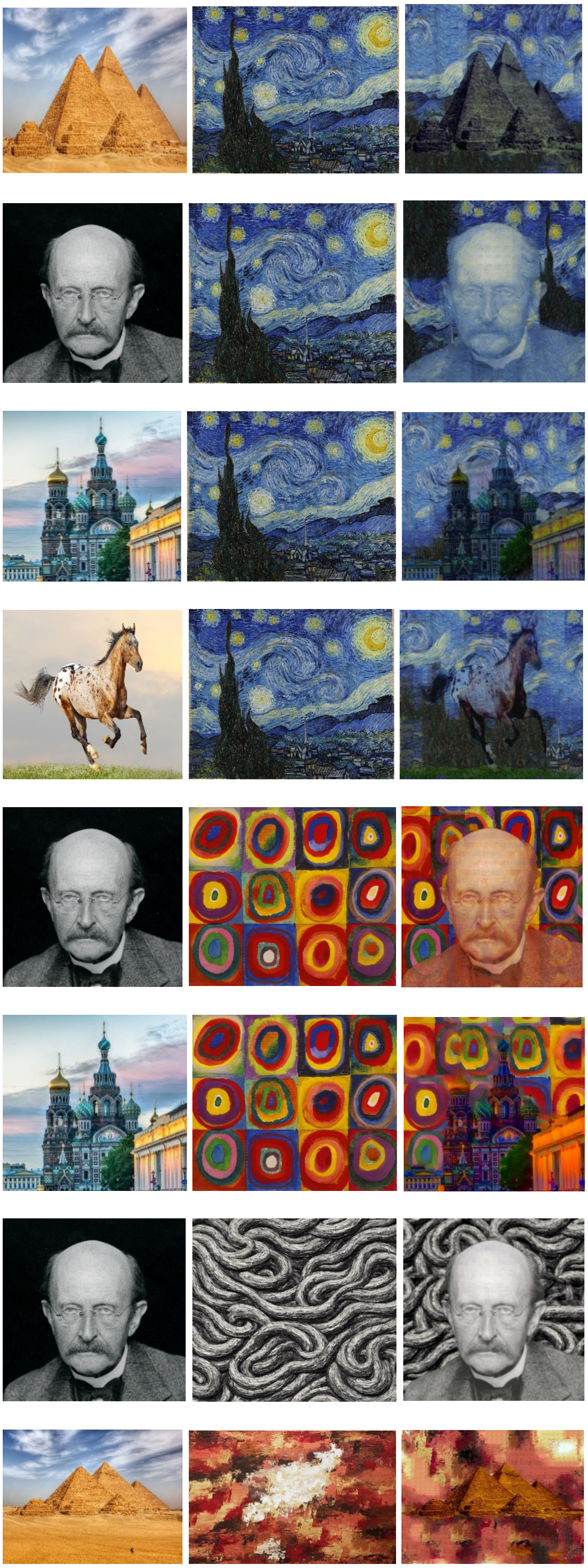

Fig. 5. Good results.

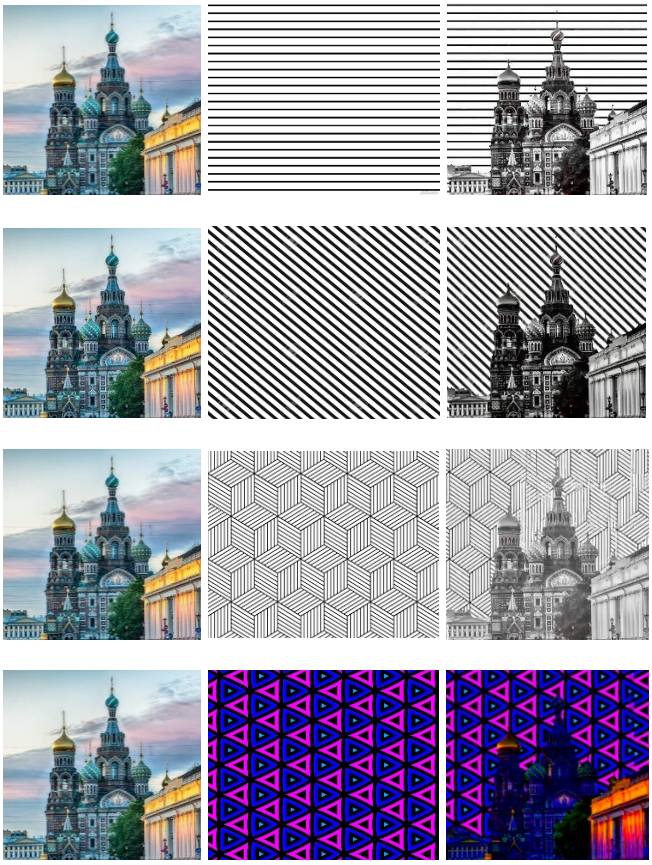

Fig. 6. Good results with geometric shapes in the style images.

Fig. 7. Bad results due to poor style.

Fig. 8. Bad result due to bad segmentation.

IV. PERFORMANCE

Our algorithm takes about 30 seconds with patch sizes [40, 30] and 60 seconds with patch sizes [20, 10], using a content image of size 400x400.

V. CONCLUSION

In this project, we engaged with a diverse array of image processing techniques, exploring their nuances and applications. In conjunction with these methods, we seamlessly integrated classical machine learning approaches, including Principal Component Analysis (PCA), Iteratively Reweighted Least Squares (IRLS), and K-Nearest Neighbors (KNN). Each of these methodologies played a crucial role in our overarching objective. Through a thorough exploration of various techniques, our project achieved a harmonious fusion of image processing and classical machine learning, demonstrating the efficacy of this combined approach in accomplishing the intricate task of style transfer.

REFERENCES

[1] Style Transfer via Texture-Synthesis "Micahael Elad and Peyman Milanfar".

[2] Artisitc Style Transfer "Elias Wang and Nicholas Tan".

[3] Domain Transfer for Edge-Aware Image and Video Processing "Eduardo S. L. Gastal and Manuel M. Oliveira".