Unveiling Color Dynamics in Andy Warhol's "Shot Marilyns": A Study on Visual Variations and Perception

A Preprint

Erick S. Arenas V
Department of Statistics
University of California, Davis
Davis, CA 95616
esarenas@ucdavis.edu

Weilin Cheng
Department of Statistics
University of California, Davis
Davis, CA 95616
wncheng@ucdavis.edu

Hengyuan Liu
Department of Statistics
University of California, Los Angeles
Los Angeles, CA 90095
hengyuanliu@g.ucla.edu

Xinhui Luo
Department of Statistics
Tufts University
Boston, MA 02155
xinhui.luo@tufts.edu

Kathy Mo
Department of Statistics
University of California, Los Angeles
Los Angeles, CA 90095
kathymo24@g.ucla.edu

Li Yuan
Department of Statistics
University of Michigan, Ann Arbor
Ann Arbor, MI 48109
leeyuan@umich.edu

July 15, 2024

Abstract

This study delves into the comparative analysis of five distinct versions of Andy Warhol's "Shot Marilyns," focusing on the intricacies of their color composition and distribution. Employing a range of analytical methods, including relative conditional entropy, this research investigates the unique color distributions and interrelations present in each artwork. Through the clustering of the artworks and the meticulous examination of specified regions of interest (ROIs)—namely, the backgrounds, hair, eyeshadow, and faces—we have unearthed profound insights into the constructional variances and similarities among the images. Our findings reveal that the presupposed uniformity in the coloration of certain elements stands contradicted, thereby underscoring the complexity and illusionary nature of color perception in visual art.

***K*eywords** shot marilyns · marilyn monroe · andy warhol · region of interest · python

1 Introduction

"Shot Marilyns," an illustrious artwork by the American artist Andy Warhol, stands as a testament to his exploration of celebrity culture and the commodification of images. Created in 1964, this series features multiple depictions of Marilyn Monroe, crafted in the wake of her untimely demise, thereby immortalizing her status as a cultural icon. Warhol's fascination with fame, consumerism, and the media's influence on societal perceptions is evident in his innovative use of silkscreen printing. This technique, adopted in August 1962, facilitated a production-line approach in art-making, enabling Warhol to replicate images with minor variations. The "Shot Marilyns" series, inspired by Monroe's death in the same year, utilized this method to produce screenprints of her image, employing photo stencils and a palette of vibrant inks to transfer her likeness onto canvas. This body of work underscores Warhol's endeavor to dissect Monroe's persona through repetitive imagery, each variation rendered in distinct colors and compositions, allowing for diverse visual interpretations.

The series' repetitive nature serves as a critique of the pervasive commodification of celebrities, transforming them into ubiquitous symbols within popular culture. Warhol's strategic use of bright, contrasting colors aims to encapsulate Monroe's dynamic celebrity essence, with the vivid hues underscoring her societal allure and influence. This study seeks to analyze the pixel color distribution in RGB space across the five "Shot Marilyns" images, examining the interplay between primary colors through relative conditional entropy. Such an analysis aims to reveal the underlying emotional and symbolic connotations Warhol might have intended, with color distributions potentially reflecting varied moods and themes.

In 1964, an intriguing incident further contributed to the series' lore when Dorothy Podber, a performance artist, mistakenly received permission from Warhol to "shoot" the Marilyns, leading to an act of literal gunfire that damaged four of the five canvases. This event birthed the "Shot Marilyns," adding a layer of physical and conceptual depth to the artwork. Part of our project involves digitally restoring the "Blue Marilyn," focusing on the gunshot damage, by analyzing and replicating the surrounding area's color distribution.

This paper is intended for publication in the Journal on Computing and Cultural Heritage and aims to shed light on the technological intersections of art restoration and analysis, using Warhol's "Shot Marilyns" as a pivotal study case.

(a) Orange Marilyn

(b) Red Marilyn

(c) Turquoise Marilyn

(d) Blue Marilyn

(e) Eggblue Marilyn

Figure 1: Different color variations of Marilyn Monroe

2 Methods

An image comprises pixels, with each pixel containing three color components: Red (R), Green (G), and Blue (B), denoted as (R, G, B) respectively. These components define the intensity of the respective colors, with each component represented by an integer value within the range of [0, 255] in the RGB color space. Therefore, each color component is a discrete variable capable of assuming 256 distinct values. In the equations below, $Y = y$ or $X = x$ can be selected from any of the three color components, R, G, or B.

2.1 Entropy Calculation

The probability of a specific color component, $P(Y = y)$, is determined by dividing the number of pixels with color coordinates corresponding to that component by the total number of color components in the entire image. The following equations illustrate the calculation of entropy, conditional entropy, and relative conditional entropy

The entropy of a color component Y is defined as:

$$H(Y) = -\sum_{y=0}^{255} P(Y = y) \cdot \log(P(Y = y)) \tag{1}$$

The conditional entropy of Y given X is given by:

$$H(Y|X) = \sum_{x=0}^{255} P(X = x) \cdot H(Y|X = x) = -\sum_{x=0}^{255}\sum_{y=0}^{255} P(X = x, Y = y) \log_2 \left(\frac{P(X = x, Y = y)}{P(X = x)} \right) \tag{2}$$

The relative conditional entropy is calculated using the following formula:

$$HR(X|Y) = \frac{H(X|Y)}{H(X)} \tag{3}$$

2.2 K-Means Clustering Analysis

Start this section from here ...

2.3 ROI Extraction

Start this section from here ...

2.4 K-Nearest Neighbors Repair

Start this section from here ...

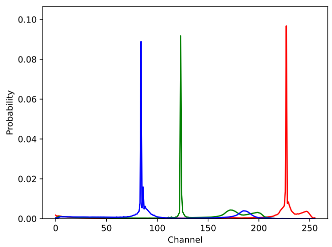
3 Data Description

Our study utilized a dataset comprising five images, each retrieved via URL links provided by "The Interior Review" website [1]. The images, titled Orange Marilyn, Red Marilyn, Turquoise Marilyn, Blue Marilyn, and Eggblue Marilyn, are digitally encoded in the RGB color model. This model synthesizes a spectrum of colors through the additive mixing of Red, Green, and Blue light. The intensity of each primary color is quantized into discrete levels, ranging from 0 to 255, offering a finite palette within this cubic color space. The dimensions of each image are 960 by 960 pixels, totaling 921,600 unique data points per image, each specified by a distinct location and its chromatic composition.

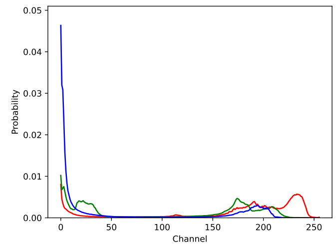
4 Data Exploration and Visualization Analysis

Our initial analytical procedure involved scrutinizing the RGB channels of the images and assessing their distribution profiles. Figure 2 delineates the variation in the red, green, and blue distributions across the five images. Notably, Figure 2a and Figure 2d display pronounced disparities when contrasted with the others. For instance, in Figure 2a, the blue channel's most significant probability density is localized within the [50, 100] interval, reaching an estimated 27%. The green channel's peak probability lies between [120, 130], at around 28%. Similarly, a notable concentration for blue is observed in the [220, 240] range with a likelihood of approximately 29%.

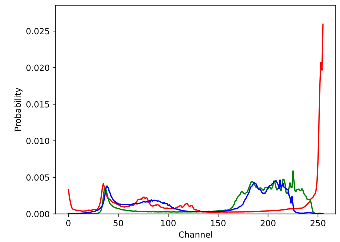
Intriguingly, Figure 2d mirrors the Figure 2a in terms of green channel probabilities, predominantly in the [120, 130] range. However, it diverges markedly in the red and blue spectra. Figure 2d features a red channel probability apex in the [70, 80] interval, representing a 32% likelihood, while its blue channel's highest probability is within the [190, 210] range, also accounting for a 32% probability. The divergence in the red and blue channel distributions between Figure 2d and Figure 2a accentuates their unique attributes in the collective color distribution.



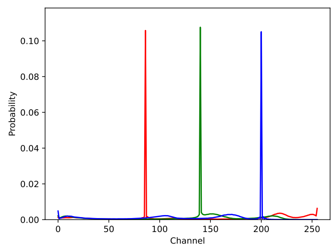
(a) Orange Marilyn RGB Distributions



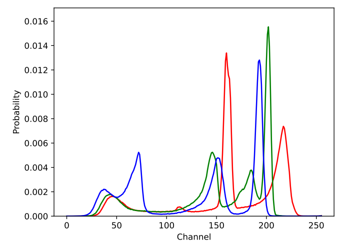
(b) Red Marilyn RGB Distributions



(c) Turquoise Marilyn RGB Distributions



(d) Blue Marilyn RGB Distributions



(e) Eggblue Marilyn RGB Distributions

Figure 2: Distributions of values of Red, Green, and Blue channels for five images with all pixels

5 Clustering based on Whole Images

To analyze the distribution of RGB colors in a 2D space, we plotted each color channel against the others for the entire image. Figures 5 to 9 display a series of scatter plots, with each dot representing a single pixel. The top three scatter plots in each figure show the distribution of pixel colors in the 2D RGB space, using combinations of Red, Green, and Blue coordinates. The bottom three scatter plots use grayscale to represent the intensity of each color component. In these grayscale plots, lower intensity values appear as white, while higher intensity values appear as black. Specifically, the leftmost panel illustrates the relationship between red and green, with blue represented in grayscale. The middle panel shows the interplay between red and blue, with green in grayscale. The rightmost panel depicts the interaction between green and blue, with red in grayscale.

The following is a reference to the whole figure above, type this in .Rmd:

`Figure \ref{fig:orange_marilyn_scatter}`

It will show as:

Figure 4

To reference a subplot:

`Figure \ref{fig:4_2_orange_marilyn_original_scatter}`

It will show as:

Figure 4b

The following is a reference to the whole figure above, type this in .Rmd:

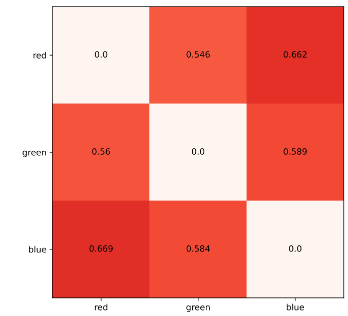
`Figure \ref{fig:red_marilyn_scatter}`

It will show as:

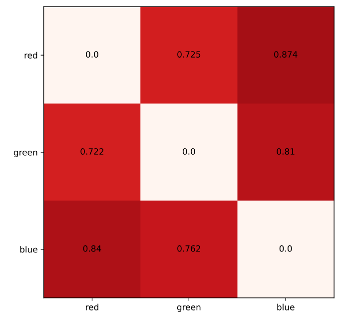
Figure 5

To reference a subplot:

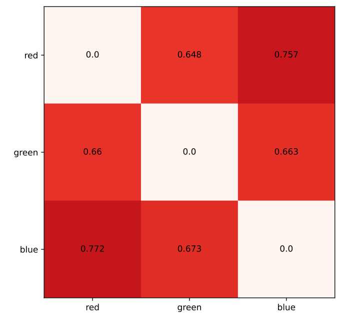
`Figure \ref{fig:4_7_red_marilyn_original_scatter}`



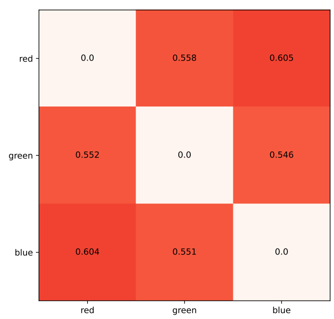
(a) Orange Marilyn Relative Conditional Entropy



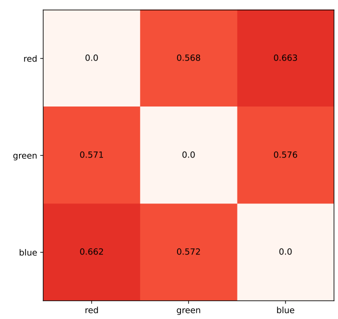
(b) Red Marilyn Relative Conditional Entropy



(c) Turquoise Marilyn Relative Conditional Entropy

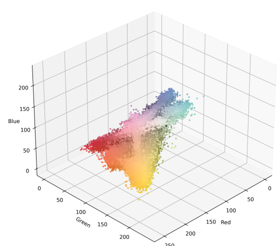


(d) Blue Marilyn Relative Conditional Entropy

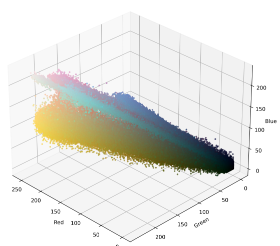


(e) Eggblue Marilyn Relative Conditional Entropy

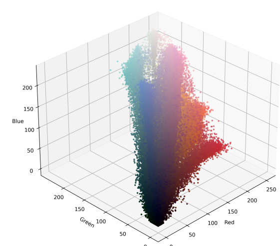
Figure 3: The relative conditional entropy values among the red, green, and blue coordinates of pixels



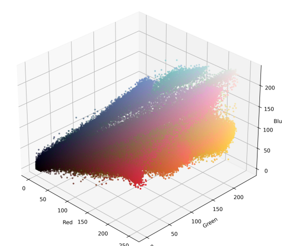
(a) Orange Marilyn RGB Space Angle 1



(b) Orange Marilyn RGB Space Angle 2



(c) Orange Marilyn RGB Space Angle 3



(d) Orange Marilyn RGB Space Angle 4

Figure 4: The RGB space occupied by the pixels for the entire image of Orange Marilyn

It will show as:

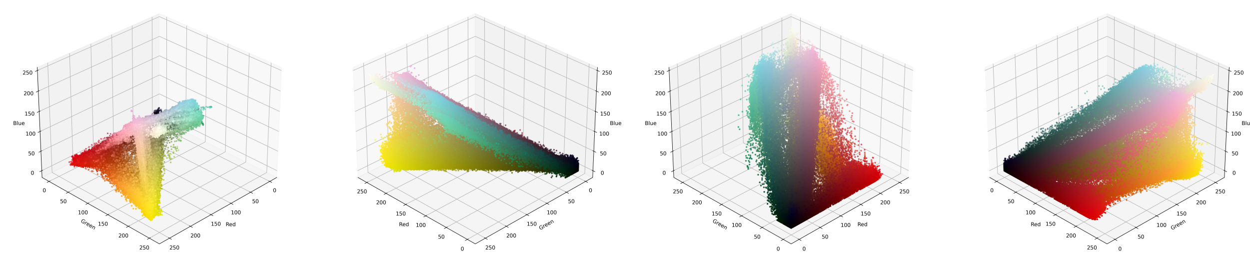
Figure 5c

6 Clustering based on Region of Interest (ROI)

7 Repair Gunshot of Image

8 Disuccsion

9 Conclusion and Future Work



(a) Red Marilyn RGB Space Angle 1 (b) Red Marilyn RGB Space Angle 2 (c) Red Marilyn RGB Space Angle 3 (d) Red Marilyn RGB Space Angle 4

Figure 5: The RGB space occupied by the pixels for the entire image of Red Marilyn