OmniFace: Skin Tone and Faces in Artworks

Olga Konstantinova
Student number: 13094653

Leonard Herold
Student number: 12854743

Lucien Martijn
Student number: 12969168

Stijn Meijerink
Student number: 12970247

Jim Cuijpers
Student number: 10999418

***Abstract*—The way faces are portrayed in art changes throughout history, as throughout time, the view on how people should be portrayed also changes. To magnify the notable differences in skin tone and facial structure of each period, and how different age or gender groups were expressed, this work introduces OmniFace. Through OmniFace, users can facilitate an exploration tool to understand how skin tone in artworks change over time and to what extent the individual faces are similar to the average face of that particular time period.**

***Index Terms*—Design Study, Art, Skin Tone, Facial Landmarks, Face Morphing**

I. Introduction

Color is an integral part of our everyday life. Whether it is the choice of clothing in the morning, the color of the wall paint in one's apartment, or the selection of the appropriate color palette of an artist: color is not only functional, but also transports feelings or other inconspicuous signals for an individual and, indeed, society. Especially in paintings and photography, color is often an essential part of the art. In some cases, the color should tell us something; in other cases, the combination of colors should be as harmonic as possible. Research regarding the use of color in paintings have been conducted in the past. For example, Sari C. et al. [1] have studied the color of clothing in art. Besides the clothing present in a portrait, the color of the skin can also be an interesting component, especially when the color is aggregated to allow analysis over a period of time or any other dimension relevant to the user. Therefore, light will be shed on the functions and aspects an application must cover to enable users to examine large corpora of art in terms of the color of skin and other facial characteristics used by the artist. The underlying research question of this work is as follows: *How can the characteristics of a face, such as color of skin, used by artists in paintings visualized against multiple dimensions, enabling people to analyze art and discover the data to gain further insights?*

Based on this question, it has been chosen to extract and visualize the skin color together with an average face and other characteristics from portraits of the OmniArt data set. This data set is a combination of multiple museums' collections amounting 432,217 artworks with additional meta-data [2]. The data set and associated images have been processed with regard to the visualization and the obtained insights have been made available to users via an appropriate user interface. The visualization is expected to be used on similar corpora of portraits that, for example, could be part of an art exhibition or similar enabling visitors to gain insights about the collections skin tone characteristics.

The remainder of this work is structured as follows. First, the problem area is specified in Section II and enriched with relevant details. Subsequently, the Section III presents the literature which is relevant in this context and introduces important foundations required for this work. Based on this, a detailed description follows the process of data extraction in Section IV and the visualization based on this in Section V. The paper will be finalized with a discussion of this work and finally a short conclusion.

II. Background

Skin tone and average face are the main topics of interest in this paper, which means that the extraction of features such as skin color, age, gender and facial landmarks are important aspects of this project. Therefore, several techniques will be described below that are used for data extraction.

A. Skin color detection

Skin color detection has been studied by many scientists. It is a process of splitting the image on skin and non-skin pixels. Since people from different regions have different skin tone, it is difficult to develop a uniform method for skin color detection. However, there is a general process which is used to detect the skin color, which has the following structure [3]:

1) Representing image pixels into the proper color space;
2) Detecting the skin and non-skin pixels;

Choosing the appropriate color space is one of the problems that exists in skin detection algorithms. Vezhnevets V. et al. in their research discuss the color space choice problem. [4] According to them, many scientists tend to believe that color space selection is important for the skin model. However, others doubt the fact that the final skin detection result has significant influence of color space selection. Nevertheless, many scientists, who have been studying this problem, such as Shaik K. et al. [5], Buza E. et al. [6], Vezhnevets V. et al. [4] and others, are certain that RGB color space is not preferred as a choice for color-based recognition algorithms. The reason for this is the fact that RBG color space is mixing color (chrominance) and intensity (luminance) information, has high correlation between channels and non-uniform characteristics.

The most popular choices for color space are HSV, YCrCb and YIQ. It has been decided to make use of HSV (Hue, Saturation, Value), since different researchers make use of different color spaces.

According to Shaik K. et al., skin detection using HSV color space is based on the method using a threshold value of the individual component of corresponding color space. [5] The algorithm for the detection of human skin color in color images contains a few steps, which are as follows: [5]

1) Obtaining an input image;
2) Converting the image into HSV color space;
3) Determining the threshold value for three components HSV;
4) Creating single channel mask, that denotes the presence of colors in the threshold;
5) Applying threshold to the masked image.

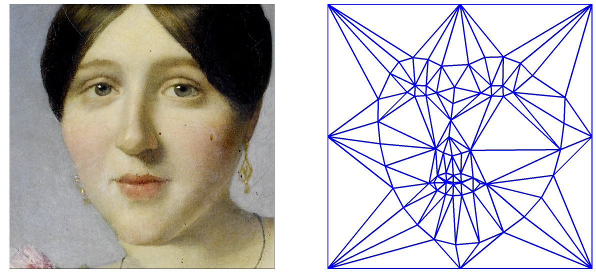
The output of this algorithm contains only skin pixels.

B. *Age and gender classification*

The increasing popularity of social platforms and social media is one of the reasons why age and gender classification becomes more important, due to their wide use in different applications. According to Levi G. and Hassner T., “performance of existing methods on real-world images is significantly lacking”. [7] In their research “Age and Gender Classification using Convolutional Neural Networks”, they proposed a method for classifying age and gender by using a deep Convolutional Neural Network (CNN) which can be used even in cases of limited data. The result of this research shows an improved performance of age and gender classification. [7] o

C. *Average face*

Representing an averaged face for a certain group of people is used in different areas. For example, in 1993 the TIME magazine published the *The New Face of America* representing the influence of immigration on the average face of US citizens [8]. In order to aggregate and blend multiple faces together, different techniques have been proposed. In general, one can distinguish between simple averaging of faces by overlaying pictures without any adjustments or instead more complex methods that morph faces based on predefined points (e.g. the eyes) called landmarks before blending them together [9]. According to Clement J. G. and Marks M. K. [9], morphing faces based on predefined landmarks yields better warped faces, as it ensures that the same parts of a face are blended together. An established method is to determine so-called Delaunay triangles for a set of landmarks of a specific face [10]. The triangles shown in Figure 1 are calculated based on the facial landmarks the machine learning library dlib detects. Dlib provides a pretrained model using the method from Kazemi and Sullivan [11], [12] to detect the location of 68 facial landmarks. Generally, there can be any number of facial landmarks, however, dlib uses the frequently used landmarks defined by Sagonas et al. [13].



(a) Transformed Image (b) Delauny Triangulation

Fig. 1: Blending Faces

Based on the landmarks detected, a set of faces can be transformed such that certain points are ensured to overlay. In the shown example, the eyes were fixed while other parts of the face were transformed using the similarity transform method of OpenCV [14]. Based on this and the calculated triangles shown in Figure 1, faces can be blend together by blending the same regions of each face. Based on this, any number of faces can be averaged effectively and yielding appealing faces despite different postures.

D. *Similarities*

In addition to warping faces, facial landmarks can also be used to calculate the difference between two faces. The fact that the facial landmarks represent a set of points in a coordinate system allows applying common distance metrics, such as the Euclidean distance. Thus, it is possible using the landmarks of the average face and the landmarks of every individual face to determine the distances between them. As the landmarks have a predetermined ordering [15], the sum of Euclidean distances between every two coordinates of the facial landmarks is used in this work to determine a relative error between the average face and an arbitrary face.

III. RELATED WORK

Data visualization is a powerful tool which is used in a variety of applications and has already been applied to the field of art. Art based data has been used at museums such as the Museum of American Art and the San Francisco Museum of Modern Art. [16] This section represents an overview on the projects in the field of visualization of skin tone and the average face in art.

A. *Visualization of skin tone in art*

Lots of articles can be found related to skin tone extraction techniques. However, only a small amount of papers refer to the visualization of skin tone. Moreover, in these researches and projects the visualization of skin tone in art is not often mentioned. In the papers regarding the skin tone extraction, visualization is not the main objective, since it is used with the purpose of showing the result of the used approach or technique. Examples of these papers have been already mentioned: Shaik K. et al. [5], Buza E. et al. [6], Vezhnevets V. et al. [4]. One example of such a visualization is shown in Figure 2.

Fig. 2: Results of the fuzzy skin detector [4]

Regarding the information visualization projects about the skin tone visualization, John Nami Choi created two interesting projects. One of them was made in collaboration with Inpyo Chang and Woonyung Choi, which is called "Narrow minded in color". In this project, the authors started with the question: "How different are we?". They built a program for skin tone detection and its visualization, in which users had to go through few steps. One of these steps was to let the user make a selection of what he or she believes to be his or her own skin tone. In the end, users were shown their average skin tone, followed by how its placed on the spectrum of all skin tones. The goal of the project was to show to users that our skin belongs to a narrow spectrum of colors, and that there is no need for discrimination in this narrow spectrum [17]. The interface of the project is shown in Figure 3.

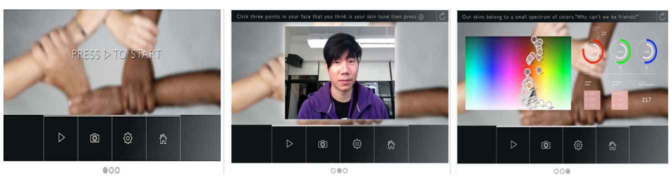


Fig. 3: Narrow minded in color [17]

The other project was made in collaboration with Inpyo Chang: "Self Perception". Following the previously mentioned project, John Nami Choi decided to prove that people choose their own skin to be brighter than their actual skin color [18]. The project has shown that about 70 % of people choose brighter color of their skin than their original skin tone [18]. The interface of the project is shown in Figure 4.

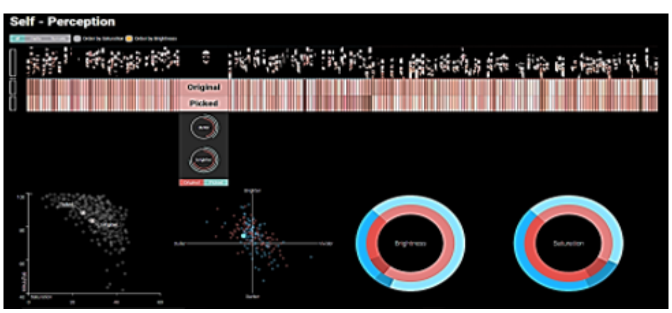


Fig. 4: Self perception [18]

The paper by Viegas F.B and Wattenberg M. does mention information visualization of skin tone in art [16]. In the paper, the authors studied the "implications of artistic interest in visualization". They describe the projects that have been successful in the artistic community and that make use of data visualization. One example the authors refer to is Jason Salavon's art piece *Every Playboy Centerfold, The Decades*. This work is shown in Figure 5 and depicts the result of averaging the pixel colors for all Playboy centerfold collection of images between 1960 and 1999 [16]. This evolution of portraiture indicates a change of skin and hair color over the years, particularly that colors get lighter as time passes.

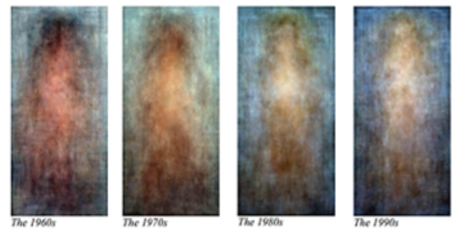


Fig. 5: Every Playboy Centerfold, The Decades [16]

No further projects have been found that visualize data of skin tone related to art. Therefore, a new way of data visualization will be introduced in this paper. The introduced way of visualization allows the user to explore the changes in skin tone of people on portraits, based on time, gender and age. The visualization also shows the average face found in portraits for the chosen parameters.

B. *Visualization of average face in art*

Many articles can be found that describe the algorithms of obtaining the average face. The visualization in such researches is mainly used to show the result of these implemented algorithms. However, there are few examples of data visualization tools, related to average faces, that allows the user explore the data. One of them is shown in the research "AverageExplorer: Interactive Exploration and Alignment of Visual Data Collections". The authors created a tool called "AverageExplorer" that helps users to explore and visualize a large image collection using the medium average images. Figure 6 shows a snapshot of the tool. According to Zhu J.-Y. et al., this interactive framework has a set of brush tools that can be used to edit the average image in order to interactively explore the data and each user interaction provides a new constraint to update the average [19].

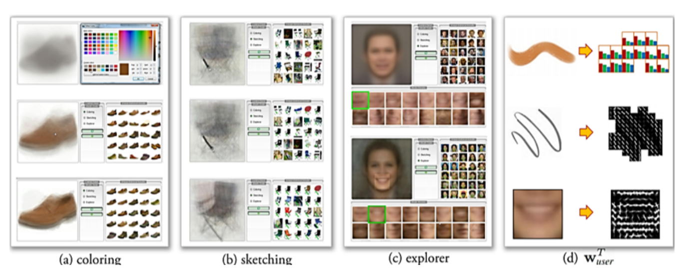


Fig. 6: AverageExplorer framework [19]

Two other projects are "Face Of Tomorrow" and "The World's Most Average Faces" and are similar in nature. The first project represent a series of photos which shows the effect of globalization on identity. It shows how a person's appearance might look in the future through the average faces from the different countries for both genders [20]. One of the examples from this project is shown in Figure 7. The second project follows the same logic, but shows the average face that represents a whole region, for example, Europe, Africa etc [21].

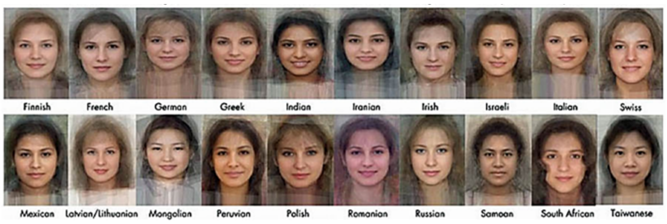


Fig. 7: Average women's faces [20]

More projects exist that make use of the visualization of an average face, for example the average face of a selection of celebrities [22]. However, not many projects describe the visualization of the average face in art. One of the examples is the paper by Yaniv J. et al.

The authors have presented a facial landmark detection framework for artistic portraits. They studied the differences between the natural and artistic faces with the purpose to use these observations for proposing a method for artistic augmentation. According to Yaniv J. et al., this method brings the natural face domain closer to the artistic one [23]. One example of their result is shown in Figure 8.

Fig. 8: Geometric Stylization [23]

No further projects have been found that make use of average faces in the field of art. Therefore, the tool created and described in this paper brings something new to the field of information visualization of art. The tool allows the user in such a way that the user can set parameters and obtains an average face based on historic artistic portraits.

IV. Data Extraction

In addition to the visualization of the data, processing was necessary to extract the required features from the images. In this section, a comprehensive overview of the entire processing is presented first. Afterward, single aspects are explained in more detail.

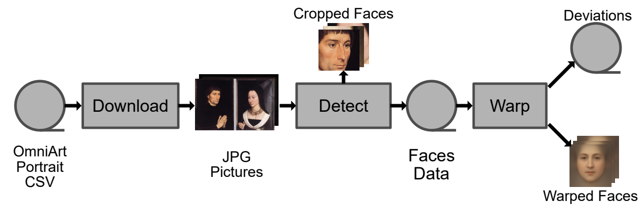


Fig. 9: Data Extraction Pipeline

A. *Overview*

The entire data processing pipeline, as displayed in Figure 9, consists of three major steps, namely download, detect and warp, and produce four outputs including information about the portraits relevant for the visualization.

The input data set for the pipeline required some pre-processing, due to the fact that the OmniArt data set consists of a myriad of different artworks and artworks types. As is focused on extracting the skin color of human beings, it was decided that only portraits are relevant. It is expected that this would keep the noise, due to erroneous processing (i.e., classification errors) in the pipeline, low. As displayed in Table I, this filtering reduced the size to 14,970 artworks.

In a first step all portraits were downloaded and transformed into a single image format. Then these images were processed in the *detect* step of the pipeline determining the age group, gender, facial landmarks, dominant skin color and a group mapping for the dominant skin color. This resulted in the main part of the data, which the user should be able to explore. However, it was also required that the user can explore an aggregated view of the data and the faces detected. The pipeline post processes the extracted faces data and combines the faces along the predefined dimensions age group, gender, group mapping and time (based on year, decade and century). Furthermore, the deviations between the aggregated face and incorporated faces were calculated allowing to display the most similar face according to the landmarks. In the subsequent subsections, the most challenging and important features that are extracted will be elaborated upon.

TABLE I: Data Properties

OmniArt data set Artwork Count	432,217
OmniArt Portrait Count	14,970
Amount of Portraits in Data Product	11,436
Faces Extracted From Portraits	12,897

B. *Detection of Age, Gender, Landmarks and Skin Color*

The detection step processes each image individually and determines four features as shown in Figure 10. First, the pipeline step detects the number of faces and their location in the image. Based on this, the face's age and gender is extracted first, then the landmarks for each face are determined. This allows to define a narrow bounding box for the face, which means that the dominant skin color can be extracted. Lastly, the skin color is clustered into a set of nine groups.

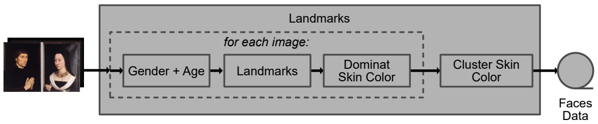
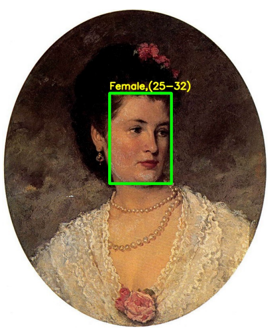
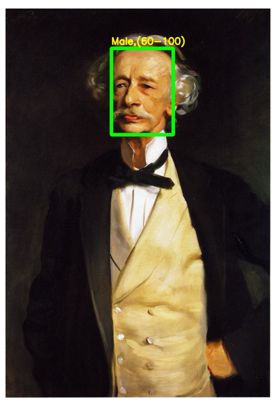


Fig. 10: Detection Step

Regarding the age and gender classification, age and gender is extracted from the image using the method by Levi and Hassner [7]. The derived implementation of Chauhan [24] was partly followed. This yields a gender that is either male or female and an age class from the following classes: 0-2, 4-6, 8-12, 15-20, 25-32, 38-43, 48-53 and 60-100. Figure 11 shows two example portraits from the obtained results of gender and age classification.



(a) Example 1



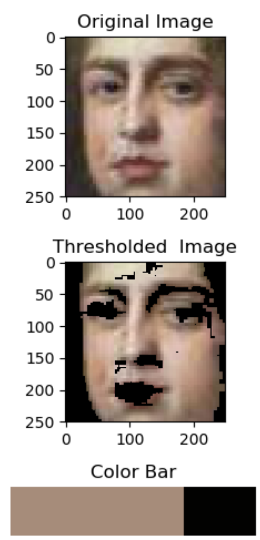
(b) Example 2

Fig. 11: Age and Gender classification

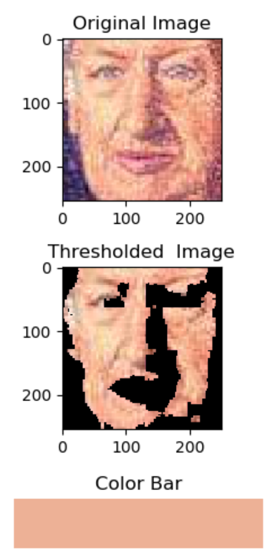
Next, the facial landmarks of each face are extracted by using the Kazemi and Sullivan's method that is implemented in the dlib machine learning library [11], [12]. The model predicts the location of the 68 facial landmarks defined by Sagonas et al. [13] as shown in Figure 12. The figure displays two faces identified in an artwork and the detected facial landmarks in red. Next, the skin color is extracted. However, since the data set does mainly consists of artworks rather than normal photos, the color of the skin is close to the color of the clothing or other objects in the painting. In order to prevent the number of cases with wrong skin detection, it has been decided to implement a few data enhancing steps, which are: cropping the detected faces and applying the proposed algorithm for the skin detection on the area of the face. Thereafter, the pixels were clustered through a simple K-means clustering in order to obtain the dominant skin color [6].

As shown in Figure 13, applying the skin color detection algorithm combined with extraction of dominant skin color provides reasonable results. In Figure 13, the *original image* represents the cropped image from the portrait and the *thresholded image* is the image that contains only skin pixels and the *color bar* contains two dominant skin colors. However, in some cases portraits consists of non-natural skin color resulting in the dominant color of black. As a fallback solution, a simple averaged skin color has been used to circumvent the problem in such cases. Finally, after all images are processed the extracted colors are further clustered into nine color groups such that the images can be grouped by dominant color within the visualization. Overall, the detection step outputs the cropped faces and the data extracted for each face as shown in the pipeline overview.

Fig. 12: Facial Landmarks



(a) Example 1

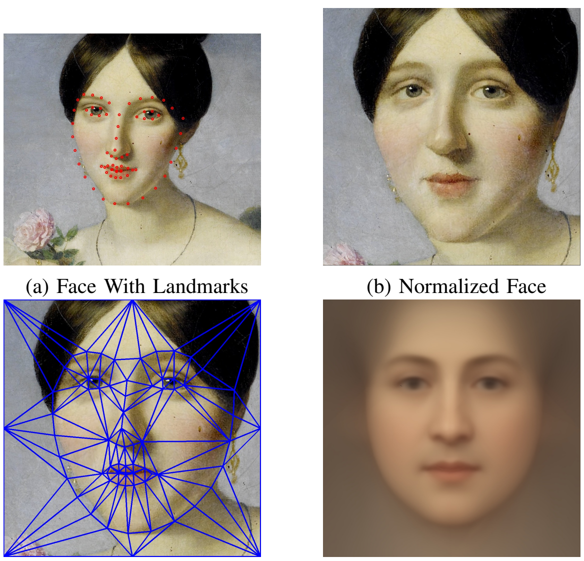


(b) Example 2

Fig. 13: Skin detection

C. Average Face Warping and Similarity

In order to display an average face for certain dimensions, as described in detail in Section V, the final processing step of the pipeline leverages the previously extracted data, such as the landmarks of each face, the associated metadata from the OmniArt data set. The pipeline groups the data set along the following dimensions: all images, per century, per decade and yearly. Based on these four time groups, the data set is further grouped based on age, gender and color group information previously extracted. Each of these groupings are then used to determine an average face over the group and the time period. In a first step the facial landmarks, that were detect in the previous step, are read. These landmarks are then used, more specifically the outer corners of both eyes, to transform image into a similar shape. This is important, as faces have different shapes or postures and, additionally, the images vary in their resolution. Therefore, the pipelines transforms an image as shown in Figure 14 from the raw image in a) to a normalized face shown in b). As the normalization transforms all images into a common coordinate system, the previously described Delaunay triangles can be used to warp faces into one another. The implemented pipeline uses the triangles as shown in Figure 14c) to average any number of faces. An example that combined an overall of 12,879 faces is displayed in Figure 14d).



(a) Face With Landmarks

(b) Normalized Face

(c) Normalized Face with Delaunay Triangles

(d) Averaged Face

Fig. 14: Skin detection

Furthermore, besides the blending faces to obtain the average faces, the distance between the average face and the added face is calculated for each face. The distance corresponds to the pairwise Euclidean distance of the extracted and the determined (in the case of the average face) facial landmarks. This is a sensible approach as it mitigates the differences between the various resolutions and the different postures. Overall, the main objective was to show the most similar facial shapes and not postures. In summary, the data extraction already prepares valuable artifacts, which are used in the visualization described in the following Section and should support the user in his intention.

V. Art Visualization

In total, four main dimensions can be identified in the data set after extracting the previously described pipeline. Namely: age, gender, time and skin color. These main dimensions can be used to deduce aggregated information, such as frequency and averages. To create a concise visualization that can conceive the main research topic, choices on visualizations had to be made. Combining all dimensions in one visualization would not lead to the intended insight. Therefore, multiple visualizations are created to give different perspectives on the topic. This section will describe the following views and what insight it should bring to the user(s): Multi-view analyzer, Heat map and Scroll View.

All of the following views brings some insight to the user. The views have one or more user goals underlying it, which is based on The science of interaction, published by Pike et al. [25]. Every user goal is annotated in an *italic* font style.

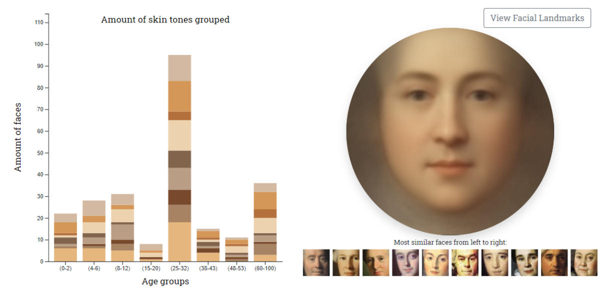


Fig. 15: Multi-view analyzer

A. *Multi-view analyzer*

The multi-view analyzer consist of three views, see Figure 15. Most prominently an averaged face, over a user-defined time period is displayed. This is done in combination with a similarity view. Lastly, there is a view showing the skin color distribution of the selected averaged face. The main goal in this case is to support the investigation of the single concept, skin tone. Multi-view has been implemented according to the guidelines as described by Wang and Kuchinsky [26].

1) Average face view: The Average face view is displayed on the right hand side of the Multi-view analyzer. The user is presented with an averaged face that is constructed from multiple faces of portraits from the OmniArt data set as described in subsection IV-C. As the user can select different time periods in a time slider (described later in subsection V-B1), the face changes in a smooth fade in / fade out transition manner, it feels like the average face morphs into the average face of the new selected time period. The more faces available for the selected time period, the smoother the averaged face is.

2) Similarity view: The Similarity view is displayed right under the Face View. This gives the user the ability to *understand* how the average face is constructed and *compare* similarity of faces. As discussed in the previous point, the average face consists of multiple faces that, when morphed together, matches the average face. To give the average face a layer of abstraction, an overview of the top 10 most similar faces is provided. By determining the relative distance between the landmarks of the faces against the landmarks of the average face using the sum of Euclidean distances between every two facial landmark points, it is possible to make a statement about how similar or not similar a face is to the average face, based on how far facial landmarks of two faces are apart from each other.

3) Color distribution: To help the user understand which skin colors are most frequently portrayed, a color distribution view has been implemented. This view displays for multiple skin color groups, per age group the amount of faces that are extracted within the selected time period. It has been decided, with the help of tools used in the field [27], that for these dimensions a stacked bar graph was the most fitting solution. To help the user get a better understanding of the exact amount of faces inside a color group, a tool tip is shown when hovering over the group. The main goal of this part of the view is to help the user *understand* and *compare* the skin colors. When extracting distinct skin colors for every face, a problem arose. It became apparent that there are 9970 distinct skin colors in the prepared data set. Such a broad color space is difficult or even impossible to visualize in current visualization techniques, while respecting the other dimensions. Therefore, data reduction needed to be performed on the broad color space. A frequent used technique is K-means clustering which results in a specified amount of color groups where similar colors are being grouped together [28]. Based on these groups, a center of the group can be determined. This loses information and these centers will only reflect the middle point in this group, which might not be a complete accurate color. However, this significantly reduces the color space and enables for a visualization that is easy to grasp.

Fig. 16: Camille Daurelle in the Park at Yerres, By Gustave Caillebotte (1877)

A problem this method brings up, can be found when looking at the portrait in Figure 16. One can see that the skin color is mainly blue which is, of course, possible when expressing artistic freedom. This image, however, is grouped in a pink (#D3B9A2) color group, which to the user might seem odd. However, it can be argued that this presentation of color is too far away from the ‘’normal” skin colors. Therefore, it has been deemed too time-consuming and irrelevant to the project to improve.

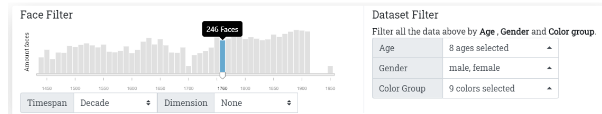


Fig. 17: Filter box

B. Interaction

The controls are split up in two sections, see Figure 17. One section which enables the user to gain more insight in the way the average face is constructed and another section which is used to explore the full breadth of the data set. In both sections, it is possible to adjust the age, color and gender parameters. However, only one time period slider is present.

1) Time slider: The time slider can show multiple granularities of time periods, which are: yearly, decades, centuries and overall(15th - 20th century). The time slider visualizes a distribution of the amount of faces per time period, in a bar-plot fashion. This design is called a scented widget and in this case supports the user to navigate to the most interesting data points [29]. In this case, the more interesting data points can be characterized by the higher amount of faces, which will result in a smoother averaged face. Furthermore, the color distribution will show more age and color categories, which brings forward a better comparison between average faces. The main user goal of this view can be defined as *selecting*.

2) Filtering: Next to the time slider, the user can filter the age groups, color groups and gender. These parameters have an influence on multiple aspects, which are the average face, similarity view, color distribution and scroll view. However, in the case of the average face, a user can only choose a single dimension to drill down. This is constrained in the User Interface by the left side of the filter box. A user first selects the dimension it wants to *explore*, and then selects a value. For example, a user selects that it want to explore on age and selects the (60-100) age group. Next, the user gets presented the average face corresponding to this selection and the corresponding color distribution. If the user wants to explore more in-depth between gender and/or color groups, it can use the controls on the right side. These controls only update the color distribution and scroll view visualization. Updating the faces according to all possible filter options would result in many distinct average faces, because for every average face, multiple gender and age options exist, with multiple color group options (which would result in calculating the Cartesian product). It had been decided, this would require too much pre-processing, furthermore it was not deemed useful for gaining extra insights.

3) Facial landmarks: The facial landmarks of an averaged face have been calculated by using the average face image. The facial landmarks show the contour shape of the average face, as well as the eye, nose and lip shape. The facial landmarks are shown by clicking on the button top-right of the face view in Figure 15. The text changes from “View Facial Landmarks” to “View Average Face” when clicking the button to show the Facial landmarks. The same happens vice versa when clicking again. This helps the user understand which action will be performed, when clicking the button.

C. Heat map

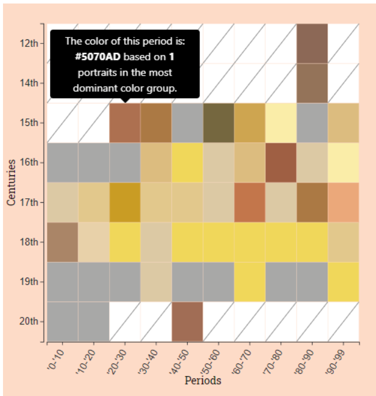


Fig. 18: Heat map

The heat map, as displayed in Figure 18, is not part of the multi-view analyzer. This visualization is to show the most frequent skin color over time. The centuries and periods are chosen in such a way that they align with the time-period slider. In this way, it is possible to have a high-level view in the heat map and get more insight in the other views. To gain this insight, K-means clustering has once more been utilized. This time, nine color groups were assumed to not give enough insight. Next to that, this visualization's purpose is not the same as the color distribution's, because the heat map should show a more fine grained view on the most prominent color used in this time period. After experimenting with multiple settings for K (in K-means clustering algorithm), 200 has been chosen as the final value. This view therefore presents the user with the option to *assimilate* the skin color spectrum per period.

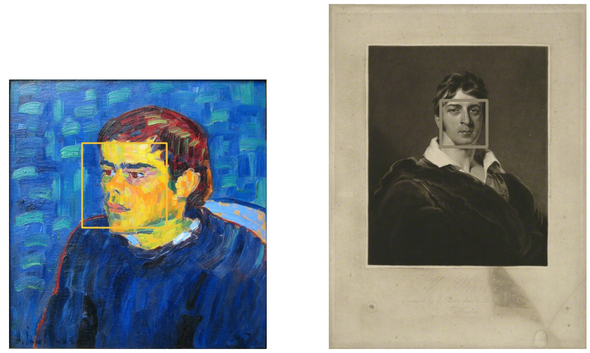
One could extract a pattern from the visualization, namely that darker colors were more frequent occurring during the 14th until 16th century, whereas the 19th and 20th century are characterized by lighter more gray looking skin colors. The pattern, however, might seem odd by the average user. However, taking into account artistic freedom, one can find many examples of "strange" colored faces in the data set, examples in Figure 19. Furthermore, because of color group clustering, the centers of the groups can be positioned in the middle, which might not reflect accurately the grouped colors.

D. Scroll view

This view lets the user *explore* all the portraits based on the set filters. In this view, the user is able to browse the portraits one by one and gain in-depth knowledge about the (sub)set of portraits. The user is given the possibility to *understand* which painter painted the art-work and how the art-work is named.

E. General layout

To finish the visualization, a layout had to be chosen which was up to current times web-page standards. In addition, the multi-view components had to be positioned such that they are all visible at the same time. In this way, a user should understand more easily that all views are interconnected. A current website design trick is Parallax scrolling. This brings the illusion of depth on a web-page [30]. This effect has been chosen because the design philosophy is based on drilling down further the more the user scrolls, by placing the high-level visualizations at the top of the page and the most detailed views at the bottom. Some user studies also suggested that people find Parallax scrolling a fun experience. Another aspect is color, which as earlier stated in the intro, is a very important aspect of communicating stories and feelings. Therefore, Colorbrewer 2.0 [31] was used to compose a color scheme. This color scheme is based on a skin tone. Moreover, the color scheme also takes into account people that suffer from colorblindness.



(a) Yellow colored face (b) Grey colored face

Fig. 19: Portraits of unnatural face colors

VI. Discussion

This paper presents a tool for skin tone and average face visualization in art. The research on the related works, has shown that the authors have more focused on the algorithms and data extraction. In addition, the existing data visualization projects are aiming to visualize skin tone or average face with different purposes. The created tool visualizes the characteristics of a face, such as color of skin, facial landmarks used by artists in paintings, average faces and associated metadata. This data is visualized against multiple dimensions, allowing people, that are keen to analyze art and discover the data, to gain further insights.

The data set behind the visualization is the OmniArt data set, filtered such that only portraits are used. The visualized features are skin tone, age, gender, time, facial landmarks, average face and the distance between the faces. In order to extract these features, the data extraction pipeline has been built, which is described in Section *Data Extraction*. However, some techniques that were used in the pipeline have their limitations.

First, the technique of obtaining the dominant color of the skin uses the color grouping with K-means clustering. This means that the dominant color is the center of the cluster. The same approach of grouping colors has been applied to the created color groups in the filters and heat map. Although this technique aggregates all skin color information to one color, the result loses accuracy. In addition, the skin color has been extracted with a method, which uses the threshold value of the individual component of the corresponding color space. As a result, skin colors are closer to reality. However, this approach might counteract the artistic freedom a painter might have wanted to express.

Second, the model for extracting age and gender information from the portraits has limited accuracy. Existing models show a quite high performance for correctly predicting one's gender and age. However, these techniques make use of actual faces, rather than faces found in portraits. This results in a lower accuracy for correctly predicting age and gender for the faces found in a portrait. Similarly, the OmniEye data set revealed the problem that the algorithm lacks accuracy for face detection in the OmniArt data set. Therefore, as was found for the OmniEye data set, new models should be created to make the predictions more accurate [32].

Third, when computing the relative distance between two sets of facial landmarks it has to be taken into account that not every image is of the same size, this would lead to large errors between two sets of facial landmarks as every set of two points are far apart from each other although the original faces have the same structure. This issue is resolved by scaling sets of facial landmarks according to a universal coordinate system. This universal coordinate system is created based on the common location of the eyes in faces.

By doing this the original set of facial landmarks that gets scaled against the facial landmarks of the average face are slightly warped, which results in a slightly different error than originally would have been the case (i.e., the facial landmarks shrink or expand on the horizontal axis based on how far the eyes of the original image are apart from each other). However, by scaling the sets of landmarks to the same size we acquire more accurate errors than with the original sets.

The visualization that is described in this paper is not the only way of information visualization in art and allows room for improvement. One example of this is the filter box, which could be less complicated. The current design is reasoned by the pre-calculations required for the warped faces. Providing pre-calculated warped faces for all combinations that are possible with the data set filters is was not feasible.

Besides the mentioned limitations, the tool brings several features that make the tool unique, easy-to-use and valuable and helps a user to explore the data. First, the concept of having multiple views, which uses three distinct views to support the exploration of a single conceptual entity, which is skin tone. Second, the interaction in the visualization that uses elements such as a time slider, filters (based on age and color groups and gender) and a button that allows the switch between the average face and facial landmarks. This helps the user to analyze the data from different perspectives. Furthermore, a view that helps the user assimilate the data set on a higher level, namely heat map, has been constructed. Lastly, an in-depth view has been created to explore the fine grained details of each art-work.

VII. Conclusion

In summary, this paper describes the creation of a skin tone and average face visualization tool that enables users to explore multiple perspectives of artistic works created between the 15th and 20th century. A data processing pipeline has been built in order to extract information about the age, the gender and the skin color of the chosen portrait, including average face, facial landmarks and the distance between faces. Several limitations have been identified regarding this project, which leaves room for improvement. However, the unique features of the tool allows a user to explore the data, through interactive components, such as a time slider and several filters. Moreover, a multi-view analyzer has been built that shows multiple aspects of an average face at once. Finally, a scroll view has been built which allows the user to explore all the portraits and their fine grained details, which are filtered on the time slider and other several filters mentioned above.

Acknowledgments

A sincere gratitude is due to the people who have assisted us during the past five weeks. We would like to offer a specific thanks to both Gjorgji Strezoski for pointing us in the right direction and Inske Groenen for providing us with valuable feedback and discussion.

References

[1] C. Sarı, A. A. Salah, and A. Akdag Salah, "Tracing the colors of clothing in paintings with image analysis," 2017.

[2] G. Strezoski and M. Worring, "Omniart: multi-task deep learning for artistic data analysis," *arXiv preprint arXiv:1708.00684*, 2017.

[3] H. K. Al-Mohair, J. Mohamad-Saleh, and S. A. Suandi, "Color space selection for human skin detection using color-texture features and neural networks," in *2014 International Conference on Computer and Information Sciences (ICCOINS)*. IEEE, 2014, pp. 1–6.

[4] V. Vezhnevets, V. Sazonov, and A. Andreeva, "A survey on pixel-based skin color detection techniques," in *Proc. Graphicon*, vol. 3. Moscow, Russia, 2003, pp. 85–92.

[5] K. B. Shaik, P. Ganesan, V. Kalist, B. Sathish, and J. M. M. Jenitha, "Comparative study of skin color detection and segmentation in hsv and ycbcr color space," *Procedia Computer Science*, vol. 57, no. 12, pp. 41–48, 2015.

[6] E. Buza, A. Akagic, and S. Omanovic, "Skin detection based on image color segmentation with histogram and k-means clustering," in *2017 10th International Conference on Electrical and Electronics Engineering (ELECO)*. IEEE, 2017, pp. 1181–1186.

[7] G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2015, pp. 34–42.

[8] (1993) Time magazine cover: The new face of america - nov. 18, 1993 - ethnicity - race - demographics - immigration. [Online]. Available: http://content.time.com/time/covers/0,16641,19931118,00.html

[9] J. G. Clement and M. K. Marks, *Computer-graphic Facial reconstruction*. Elsevier, 2005.

[10] Z.-C. Li, J. Y. Chiang, and C. Y. Suen, "Face transformation with harmonic models by the finite-volume method with delaunay triangulation," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 6, pp. 1543–1554, 2010.

[11] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1867–1874.

[12] D. King. (2014) Real-time face pose estimation. [Online]. Available: http://blog.dlib.net/2014/08/real-time-face-pose-estimation.html

[13] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "300 faces in-the-wild challenge: The first facial landmark localization challenge," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2013, pp. 397–403.

[14] S. Mallick. (2016) Average face : Opencv (c++ / python) tutorial. [Online]. Available: https://www.learnopencv.com/average-face-opencv-c-python-tutorial/

[15] A. Kaynar. (2019) Extract eyes in face images with dlib and face landmark points. [Online]. Available: https://medium.com/@aanilkayy/extract-eyes-in-face-i%CC%87mages-with-dlib-and-face-landmark-points-c45ef480c1

[16] F. B. Viégas and M. Wattenberg, "Artistic data visualization: Beyond visual analytics," in *International Conference on Online Communities and Social Computing*. Springer, 2007, pp. 182–191.

[17] J. N. Choi, I. Chang, and W. Choi. (2013) Narrow minded in color. [Online]. Available: http://cargocollective.com/namutree/Narrow-Minded-in-Color

[18] J. N. Choi and I. Chang. (2015) Self perception. [Online]. Available: http://cargocollective.com/namutree/Self-Perception

[19] J.-Y. Zhu, Y. J. Lee, and A. Efros, "Averageexplorer: Interactive exploration and alignment of visual data collections," *ACM Transactions on Graphics (SIGGRAPH 2014)*, vol. 33, 07 2014.

[20] Mike. (2013) Face of tomorrow – how we will look like in the future. [Online]. Available: https://ultrafeel.tv/face-future/

[21] Chris. (2011) The world's most average faces. [Online]. Available: https://www.manmadediy.com/873-the-world-s-most-average-faces

[22] R. Kramer, R. Jenkins, and A. Burton, "Interface: A software package for face image warping, averaging, and principal components analysis," *Behavior Research Methods*, vol. 49, 12 2016.

[23] J. Yaniv, Y. Newman, and A. Shamir, "The face of art: Landmark detection and geometric style in portraits," *ACM Trans. Graph.*, vol. 38, no. 4, Jul. 2019. [Online]. Available: https://doi.org/10.1145/3306346.3322984

[24] N. S. Chauhan. (2018) Predict age and gender using convolutional neural network and opencv. [Online]. Available: https://towardsdatascience.com/predict-age-and-gender-using-convolutional-neural-network-and-opencv-fd90390e3ce6

[25] W. A. Pike, J. Stasko, R. Chang, and T. A. O'connell, "The science of interaction," *Information visualization*, vol. 8, no. 4, pp. 263–274, 2009.

[26] M. Q. Wang Baldonado, A. Woodruff, and A. Kuchinsky, "Guidelines for using multiple views in information visualization," in *Proceedings of the working conference on Advanced visual interfaces*, 2000, pp. 110–119.

[27] S. Ribecca. (2020) The data visualisation catalogue. [Online]. Available: datavizcatalogue.com/

[28] M. Sahu and D. K. Parvathi, "Segmentation of colour data base image by implementing k-means clustering," *International Journal of Innovations in Engineering and Technology*, vol. 2, no. 4, pp. 229–234, 2013.

[29] W. Willett, J. Heer, and M. Agrawala, "Scented widgets: Improving navigation cues with embedded visualizations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1129–1136, 2007.

[30] D. Frederick, "The effects of parallax scrolling on user experience in web design," *Journal of Usability Studies*, vol. 10, no. 2, 2015.

[31] M. H. Cynthia Brewer. (2020) Colorbrewer 2.0, color advice for cartography. [Online]. Available: https://colorbrewer2.org/

[32] G. Strezoski, R. Knoester, N. van Noord, and M. Worring, "Omnieyes: Analysis and synthesis of artistically painted eyes," in *International Conference on Multimedia Modeling*. Springer, 2020, pp. 628–641.