1 Introduction

The history of color is full of instances of how and why certain colors become to be associated with certain concepts, ideas, politics, status and power. Sometimes the connotations occur arbitrarily, like in the instance when pink was assigned to baby girls, and blue started to be associated with baby boys at the turn of 19th Century [Paoletti, 1987]. Sometimes though, color associations have very tangible reasons, such as in the case of Marian blue and why over the centuries it was reserved only for painting Virgin Mary. The reason is to be found in the scarcity of the rock lapis lazuli -even more valuable than gold-, from which the blue pigments were extracted. Individual colors have convoluted and contested histories, since they have been attached to many symbols at any given time. John Gage, an art historian who has devoted 30 years of research on the topic of color, explains the conundrum of what he terms as “politics of color” in a simple fashion: “The same colors, or combinations of colors can, for example, be shown to have held quite antithetical connotations in different periods and cultures, and even at the same time and in the same place.” [Gage, 1990].

The purpose of the present study is to introduce a method for automatically extracting color distributions and main colors of paintings, as well as color schemes of people in paintings. By visualizing these over time for cross-referencing with historical data will reveal changes of how particular colors were used in a given time period and culture. In this study, we will look at artworks to find out whether certain colors or tones are associated with a specific gender, and if these connotations change over time. To that end, we apply algorithmic tools to process very large datasets automatically, and information visualization tools to depict the findings.

2 Related Work

Today, major cultural heritage collections are available online. Digitization and preservation of artworks is an important occupation of museums and cultural heritage institutions, as well as many Digital Humanities projects. Part of such
digitized collections are made available to further computer vision research to scrutinize art historical questions. Such collections are usually enriched with meticulously tagged meta data about the origins of each artwork. However, these datasets do not provide comprehensive gender annotations. For example, Rijksmuseum’s arts database has a wide selection of categories with rich meta-data that is primarily about the art objects themselves (see Table 1), but without any reference to what these artworks hold [Mensink and Van Gemert, 2014]. Automatically determining whether a sitter of a portrait is female or male in a painting is not an easy task.

### Table 1: Sample from Rijksmuseum meta data

<table>
<thead>
<tr>
<th>filename</th>
<th>format</th>
<th>date</th>
<th>type</th>
<th>title</th>
<th>subject</th>
<th>contributor</th>
<th>relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>009071</td>
<td>brons</td>
<td>1475</td>
<td>beeldhouwwerk</td>
<td>De gerechtigheid...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>008450</td>
<td>brons</td>
<td>1476</td>
<td>graffiguur</td>
<td>Pseudo Philips van...</td>
<td>41D2</td>
<td>Bruikleen...</td>
<td>BK-AM-33</td>
</tr>
<tr>
<td>008550</td>
<td>polychromie</td>
<td>1480</td>
<td>beeldhouwwerk</td>
<td>Bewening van...</td>
<td>73D7211(+5)</td>
<td>Aankoop...</td>
<td></td>
</tr>
<tr>
<td>104894</td>
<td>boekdruk</td>
<td>1706</td>
<td>inhoudsopgave</td>
<td>Byvoegsel tot de...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>091654</td>
<td>polychromie</td>
<td>1735</td>
<td>schaal (objectnaam)</td>
<td>Hartvormige schaal...</td>
<td>25G3</td>
<td>Schenking...</td>
<td></td>
</tr>
<tr>
<td>060116</td>
<td>etsen</td>
<td>1764</td>
<td>prent</td>
<td>Portret van P. Born</td>
<td>31A5331</td>
<td></td>
<td></td>
</tr>
<tr>
<td>069114</td>
<td>hout</td>
<td>1780</td>
<td>werfmodel</td>
<td>Halfmodel van een...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Several publications have appeared in recent years with the aim of automatic gender recognition. The survey on [Ng et al., 2012] describes a variety of approaches on gender recognition in natural images. [Xiong and De la Torre, 2013] proposed a practical and effective method for automatically detecting faces in natural or man-made images. Once the face is detected, a supervised classifier is used to determine whether it belongs to a male or female.

There has been focused studies to address face recognition tasks on artistic images [Srinivasan et al., 2015]. For the purposes of face detection, mainstream algorithms perform sufficiently well on paintings that are of interest for this study.

### 3 Methodology

In this study, the aim is to analyze the trends of clothing color in different periods, for each gender. For this purpose, we work on a database of paintings, for which the era (or date) is provided, and we seek to annotate each image with the gender of the depicted person, as well as a rough segmentation of the area of the clothing. The general workflow of the proposed approach is depicted in Figure 1.

#### 3.1 Database

The Rijksmuseum is a Dutch national museum dedicated to arts and history in Amsterdam. The Rijksmuseum database contains 112,039 high-resolution images with extended meta data [Mensink and Van Gemert, 2014]. However, as mentioned previously in Section 2, Rijksmuseum database has neither gender, nor clothing color information embedded into its meta data. We describe briefly how we determine the missing information.
3.2 Gender classification

Following an approach similar to Jia’s work [Jia and Cristianini, 2015], we collected images using the Google Image search tool. 5600 male and 5300 female faces were downloaded from the IMDB website, using actor and actress names. These images were used to train a supervised gender classifier. For each image, we find facial landmarks using the approach in [Xiong and De la Torre, 2013], align the faces to a mean shape [Gower, 1975], and extract features that are resistant to illumination effects [Ojala et al., 2002]. We then train a classifier using the sequential minimal optimization (SMO) method [Platt et al., 1998].

The biggest challenge for evaluating gender recognition performance on the paintings is to make sure the ground-truth gender data are actually correct [Mathias et al., 2014]. From our experience, this demanding task requires a full view of the painting, rather than just the detected face. We could reach around 80% accuracy on paintings, just by using photographs of actors and actresses in the training of the system.

3.3 Clothing color information

Portrait paintings that are completely focused on the sitter’s face have still a lot of background noise that disrupt the color representation of the paintings (see Figure 2). Our hypothesis is that color representation, when focused on the clothing of the model, will still reflect the color scheme that is associated with the gender of the sitter.

In order to extract color information of a garb we need to do a coarse segmentation of the clothing. We used the GrabCut approach [Rother et al., 2004]. In this method, a user defines an area of interest, as well as foreground and background seeds for the segmentation. In our study, background and foreground seeds are initialized based on the detected face landmarks.
Fig. 2: Four paintings from the Rijksmuseum collection, classified and segmented in terms of colors.

Figure 3 provides an initial visualization of the dominant color distributions for each era, for males and females. Concentric circles have thickness associated
Fig. 3: Clothing colors over time. Females wear significantly lighter colors than males. Best viewed in color.

with the frequency of the color. Bright colors are relatively rare, as the paintings in our tagged collection are generally dark, with the occasional shaft of light illuminating part of the painting. But a very distinct pattern can be observed in that females wear lighter colors compared to males, and show higher variance over the years. Some painting examples are given in Figure 4.

4 Conclusions

Every period and location has certain dominant color associations and symbolism. To investigate hundreds of thousands paintings in a single sweep requires automatic analysis tools. Our main objective in this work in progress is to perform an analysis on the usage of color for different genders along the centuries, and to develop tools for establishing semantic associations of colors for each particular period of study. With the increased popularity of open-art, this study can be extended significantly by introducing more databases alongside Rijksmuseum, for example, drawing on the Europeana collection [Doerr et al., 2010].

References


[Xiong and De la Torre, 2013] Xiong, X. and De la Torre, F. (2013). Supervised descent method and its applications to face alignment. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.