

The Geometry of Colors in Van Gogh's Sunflowers

Shuting Liao

University of California Davis

Patrice Koehl

University of California Davis

Jennifer Schultens

University of California Davis

Fushing Hsieh (✉ fhsieh@ucdavis.edu)

University of California Davis

Research article

Keywords: van Gogh, Sunflowers, color geometry, color transfer

Posted Date: July 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-676336/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

RESEARCH

The geometry of colors in van Gogh's Sunflowers

Shuing Liao¹, Patrice Koehl², Jennifer Schultens³ and Fushing Hsieh^{4*}

*Correspondence:

fhsieh@ucdavis.edu

⁴Department of Statistics,
University of California, Davis,
One Shields Avenue, 95616 Davis,
USA

Full list of author information is
available at the end of the article

Abstract

“Paintings fade like flowers”: van Gogh's prediction on the impact of age on paintings came true for most of his paintings. We have studied the consequences of this aging on the *Sunflowers in a vase with a yellow background* series, namely its original, F454, and two replicates, F457 and F458, at the van Gogh's museum in Amsterdam, which van Gogh painted using the original as a model. The background and flower renditions in those paintings have faded and turned brown, making them less vibrant that van Gogh had most likely intended. We have attempted to restore van Gogh's intent using a computational approach based on data science. After identifications of regions of interest (ROI) within the three paintings F454, F457, and F458 that capture the flowers, stems of the flowers, and background, respectively, we studied the geometry of the color space (in RGB representation) occupied by those ROIs. By comparing those color spaces with those occupied by similar ROIs in photographs of real sunflowers, we identified shifts in all three color coordinates, R, G, and B, with the positive shift in the blue coordinate being the more salient. We have proposed two algorithms, PCR-1 and PCR-2, for correcting that shift in blue and generate representations of the paintings that aim to restore their original conditions. The reduction of the blue component in the yellow hues has lead to more vibrant and less brownish digital rendition of the three *Sunflowers in a vase with a yellow background*.

Keywords: van Gogh; Sunflowers; color geometry; color transfer

1 Introduction

Vincent van Gogh's series of paintings of *Sunflowers* are among his most famous creations. There are two such series, the first one executed in Paris in 1887 that depicts sunflowers lying on the ground, while the second series was made later in Arles and represents sunflowers in a vase. Van Gogh made four versions of the Arles series in 1888, with different flower arrangements and different backgrounds: turquoise for the first version, F453, royal-blue for the second version, F459, blue-green for the third version, F456, and yellow for the fourth version, F454. Note that we use here the recognized classification of van Gogh's paintings, where F stands for the De La Faille catalogue of these paintings. A year later van Gogh made one repeat of the third version (F455), and two repeats of the 4th version (F458 and F457). The authenticity of the second repetition of the 4th version, F457, has been questioned, although experts now believe it is authentic [1]. These three repetitions and their two originals are usually referred to as the *Sunflowers in a vase*, where the number of sunflowers vary even in van Gogh's own recollection as described in his correspondence [2]. They are all currently housed in different museums: the third version, F456, is in the Neue Pinakothek museum in Munich, Germany, and its repetition, F455, is the the Philadelphia Museum of

Art, United States; the fourth version, F454, is at the National Gallery in London, England, while its repetitions, F457 and F458 are at the Seiji Togo memorial Sompō Japan Museum of Art, Tokyo, Japan, and at the van Gogh Museum, Amsterdam, Netherlands, respectively. Those paintings are rarely reunited, with possibly the latest occurrence being an exhibit at the National Gallery in London in 2014 where the original and one repetition of the 4th version were shown side to side. Interestingly, however, it is possible to see them virtually in the same room in a virtual 360° exhibition commented by Vincent van Gogh's grandson (<https://www.facebook.com/VanGoghMuseum/videos/10159187334010597/>). In this paper we focus on the fourth version, F454, and its two repetitions, F457 and F458, in which the vase with sunflowers is portrayed against a yellow background. This version is considered as a study of variations on the theme of yellow, with the aim of achieving a light-on-light effect [1, 3]. Pictures of those three paintings are shown in Fig 1.

Van Gogh's *Sunflowers* appeal to many. The fascination exerted by van Gogh's stay in the south of France in 1888-1889, his friendship and fallout with Paul Gauguin, his mental health, and the seven sunflowers canvases he painted during that time remain a tremendous source of inspiration in many forms of popular entertainment as well as for artists, museum curators, and scientists in general. Biologists have studied how bees interact with those paintings [4], the genetic fabric of the sunflowers [5], doctors have attempted to associate his medical conditions and related treatments to his perception of colors [6, 7, 8]. Of direct relevance to the paintings themselves, scientists and museum curators have used a broad array of traditional and state-of-the-art techniques to look at and below the paint surface itself [9, 10, 11, 12, 13, 14, 15, 16, 17]. Of special interest, we advise the reader to consider the recent book, *Van Gogh's Sunflowers Illuminated* published by the Amsterdam van Gogh museum that summarizes the results of the research undertaken by an international team of scientists, curators, and art historians aimed at comparing the F454 *Sunflower* from the London National gallery with one of its repetitions, F458 from the van Gogh museum [18]. In this paper, we approach the same problem of comparing F454 with its repetitions F457 and F458 from a very different perspective. Instead of analyzing directly the canvases and the paint surfaces, we study high resolution images of the paintings using a data science approach. The availability of large collection of digital images of paintings has opened the door to the use of state-of-the-art supervised machine learning techniques in art, for understanding the artist style [19], for classifying paintings [20], to detect forgeries [21], and possibly for even creating art [22]. Our approach to analyzing one of the *Sunflowers* painting and two of its repeats differs as it is unsupervised. We consider a painting based on a high-resolution image of it. This image is a collection of pixels, with each pixel characterized by its location and color. The color is quantified based on the RGB color model. This is an additive color model in which Red, Green, and Blue lights are added together to reproduce any colors. In a digital image, the amount of each R, G, and B color is discrete, usually an integer in the range [0, 255]. As such, a pixel in an image belongs to a discrete color space of size $[0, 255]^3$. Our analysis of F454, F457, and F458 amounts to comparing their representations in this discrete color space.

We acknowledge that our analyses may be somewhat subjective. We only use indirect representations of the three paintings by van Gogh, namely digital images of those paintings. While those images are high resolution, it is not impossible that there are some color distortion, although those are expected to be small with modern digital cameras. More importantly, the paintings themselves are more than 130 years old and the modern images we have of those paintings certainly differ from their original renderings. It is interesting that van Gogh himself was well aware that paintings age, as he wrote to his brother on April 30th 1889 that “paintings fade like flowers” [2]. Many studies have focused on the chemistry of aging of van Gogh’s paintings [9, 10, 11, 12, 13, 16, 15, 14], as well as attempts to predict the evolution of the appearance of those paintings [23, 24, 25]. With this as a background, there are two main directions that we have followed in our analyses of the digital images of F454, F457, and F458:

- i) *Identify markers of aging from the color distributions in the digital images.*
We were able to identify shifts in the red, blue, and green components of the images that are most likely a result of aging. The most significant shift is observed for the blue component of yellow hues, leading to a browning and fading of those colors. Using this shift, we propose pseudo-color reconstruction schemes that enable us to generate model images of the original paintings. We note that the same procedure can be used to extrapolate to models of the paintings in the future.
- ii) *Compare and contrast the use of colors in the original version, F454, and the two repetitions, F457 and F458, of the Sunflowers in a vase with a yellow background.* In particular, we identify differences between the repetitions that most likely reflect different intents by van Gogh.

We organize the paper as follows. Major findings are shown in Section 2. In particular, we provide an in depth analysis of the RGB color spaces of the digital images corresponding to the three paintings, F454, F457, and F458. By comparison with recent images of actual sunflowers, we identify significant shifts in the blue components, B , of the images of those paintings and use those shifts to derive pseudo-color reconstruction (PCR) schemes to attempt to restore the original images. We also propose a color transfer scheme that enables us to compare the three paintings. In the conclusion section 2 we briefly discuss the benefits of our approach as well as future directions of research. Finally, in Section 4 we provide technical details on the images we have used, on the statistical methods that we implemented, as well as on the algorithms we have implemented.

2 Results and Discussion

2.1 Basic statistics on the color content of the three digital images of the *Sunflowers* paintings

A digital image is composed of basic image elements, i.e. pixels. Each pixel is characterized by its position in the image and its color. In the RGB model, a color is described by 3 components, also called coordinates, representing the amounts r , g , and b , of red, green, and blue that need to be combined to generate that specific color. Each coordinate is an integer value taken in the range $[0, 255]$. There are therefore 256^3 (approx. 17 millions) possible colors. Interestingly, in the digital images of F454, F457, and F458 we observe only 2.56 %, 4.64 %, and 3.24 % of those

256³ colors, respectively, indicative of relatively low diversities of colors in the three paintings. While this may not be surprising as all three paintings are predominantly yellow (see Figure 1), it is noteworthy that the color diversity differ significantly between the painting. Indeed, F457 and F458 are defined as repetitions of F454 and therefore we could expect more similarities.

In Figure 2, we plot the distributions of the values for the red, green, and blue coordinates over the three digital images (see Section 4 for details on how to generate those distributions). While the distributions of red and green values are relatively similar over the three images, with values that concentrate in the range [100, 255], the distributions of values of the blue coordinate differ significantly. In F457, there are basically no pixels whose blue intensity is below 60 or otherwise stated nearly all its pixels have some blue component. In contrast, in F458 there are more than 5 % of the pixels that have a blue coordinate close to zero, namely no blue component. Clearly, something is different with the contributions of blue in those paintings. This will be discussed in detail below.

We looked also at the association between pairs of fundamental colors in all three paintings, by computing their relative conditional entropy HR (see Section 4 for details). The conditional entropy HR measures the amount of information shared between two colors. Our definition of HR places it in the interval [0, 1]; it is equal to zero if the two colors are fully determined by each other, while conversely it is equal to 1 if the two colors are completely independent. In Figure 3 we display the values for the nine pairs of colors in the form of tables for the three images corresponding to F454, F457, and F458. As expected, a color compared to itself leads to a conditional entropy of zero. In contrast, the HR values for pairs of different colors reveal very low dependence between them. There are, however, noticeable differences between the three images.

2.2 Color distributions in the background, the sunflower crowns, and the sunflower stems of the *Sunflowers in a vase with a yellow background*

The sunflowers, with their yellow crowns and stems are at the heart of the three paintings *Sunflowers in a vase with yellow background* that we are analyzing, namely F454, the original, and F457 and F458, two copies that Van Gogh made a year after painting F454, using the latter as a model. Visual inspections show differences between those three paintings, especially when it comes to their colors. We see that the green color of the stems is similar to the expected green for stems of real sunflowers. On the other hand, the yellow flowers are darker than the expected vibrant yellow of the petals of a real sunflower, especially in F454 (see Figure 1). There are many possible reasons for this darker color of the flowers. First, it could have been Van Gogh's intent. It is known that Van Gogh painted his series of paintings on sunflowers using real sunflowers as models. He painted F454 in late august 1888, when sunflowers usually start to fade [26]. Second, van Gogh used the yellow chrome pigment to color the sunflowers and it is known that this pigment darkened as it ages [9, 10, 27]. Finally, we analyze the colors of the paintings using digital images of those paintings. The images are influenced by the lighting that was used at the time that they were taken, which may influence our perception of their colors [25]. It is therefore difficult to isolate a specific origin, especially as those

mentioned above are ultimately non-exclusive. We propose a data-driven approach based on the analysis of the color space of the three paintings to provide some quantitative elements that describe the differences between the three paintings, and their differences with real-life sunflowers.

The *Sunflowers in a vase with a yellow background* paintings include 15 sunflowers with their stem in a vase standing on a table, with a yellowish background. For each painting, we identify three regions of interest, or ROI. A ROI is defined according to its position in the image as well as from its color consistency. The first two ROIs relates to the flower crowns of the sunflowers (mainly yellow, yf-ROI) and the stems of the sunflowers (green, g-ROI), while the third ROI refers to the background (bg-ROI), which includes the table, vase, and the region behind the sunflowers. In parallel, we analyzed two recent photographs of sunflowers (see Figure 12), from which we extracted the corresponding flower ROIs and stem ROIs. Extractions of the ROI was performed as described in the Material and Methods section. Each ROI is defined as a set of pixels, characterized by their locations and colors, with the latter given in the RGB space with three discrete coordinates with values in the range $[0, 255]$.

2.2.1 The yellow flowers of the Sunflowers in a vase with a yellow background

We compared first the ROIs associated with the yellow parts of the sunflowers of the three paintings. Figure 4 shows the distribution of colors of the pixels associated with those ROIs in the red-green, red-blue, and green-blue planes. In the red-green plane, most pixels are found to follow the first diagonal, consistent with the general yellow color of those ROIs. Interestingly, the original, F454, and the repetition F458 are found to be very similar. Both differ from the repetition F457. The differences with F457 are striking in the red-blue and blue-green planes, as blue appears with a wide range of values in F454 and F458, from 0 to 120, while it is always present in F457, appearing in a much smaller range centered around 60.

While there are differences between the three paintings in the diversity of color of their yellow sunflowers, there are even bigger differences when we compare those sunflowers with real sunflowers, as observed by comparing Figure 4 with Figure 5. Those differences can be summarized as follows:

- i) The yellow of real sunflowers contain a minimal amount of blue, as illustrated in the blue-red and blue-green planes for the two photographs SF1 and SF2. In contrast, as indicated above, the yellow observed in the sunflowers of the paintings contains a significant amount of blue.
- ii) The yellow of the flowers of real sunflowers contains a wide range of green covering nearly the whole spectrum of possible values from 0 to 250. In contrast, the green in the yellow of the flowers in the paintings is limited to the range $[50, 250]$ in all three paintings.
- iii) Similarly, the yellow of the flowers of real sunflowers contains a wide range of red with values between 20 and 255. In contrast, the red in the yellow of the flowers in the paintings is limited to the range $[100, 250]$ in all three paintings.

2.2.2 The stems of the Sunflowers in a vase with a yellow background

We repeated the analysis described above on the ROIs corresponding to the green stems of the sunflowers, both for the 3 paintings of sunflowers and for the two pho-

tographs of real sunflowers. Results are shown in Figure 6 and Figure 7, respectively. Comparisons of those two figures lead to the following observations:

- i) Among the two copies, F457 and F458, F458 is the closest to the original F454. This result is the same as what was observed for the yellow parts of the petals and crowns of the sunflowers. Differences associated to F457 again come from a smaller range of blue coordinates for the pixels representing the stems.
- ii) On average, the pixels in the green ROI of F457 contain more blue than the corresponding pixels in F454 and F458. This is visible in the paintings themselves, as the stems in F457 appear with darker green (a consequence of the addition of blue)
- iii) There is a stronger linear relationship between red and green coordinates for the stems of the real sunflowers, compared to the stems in the three paintings
- iv) Besides the stronger relationship indicated above, the colors of the stems of the painted sunflowers resemble qualitatively with the colors of the stems of real sunflowers, as captured by photographs.

2.2.3 The backgrounds of the Sunflowers in a vase with a yellow background

Finally, we compared the backgrounds of the three paintings in Figure 8. The background ROIs are expected to be the most diverse as they include multiple parts of the paintings. In addition, visual inspections of those paintings clearly indicate significant differences (see Figure 1): the background of F454 is globally pale yellow, while those of F457 and F458 are more green-yellow and solid yellow, respectively. Those differences are reflected in the differences between the RGB space occupied by the pixels of the background, as illustrated in Figure 8. The background ROI within F454 shows a large number of pixels with large R, G, and B coordinates; those pixels will show as close to white. In contrast, a significant number of pixels in the background of F457 have large G coordinates, consistent with a green coloration. Finally, the background of F458 includes many pixels with high R and G coordinates, and low B coordinate, consistent with a solid yellow.

2.2.4 Why those differences between the paintings, and between the painted Sunflowers and real sunflowers?

As mentioned in the introduction of this section, there are possibly three main reasons for differences between the *Sunflowers* paintings themselves, and between the paintings and real sunflowers. Those reasons are associated to the painter's intent, to the aging of the paintings, and to artifacts associated with the digital images we consider. We exclude the latter as we believe that such artifacts are minor compared to elements of the two other reasons [25]. The analyses we have presented above provide elements that highlight the importance of aging.

The three paintings F454, F457, and F458 are more than 130 year old and as such they have been aging. They have been subject to color degradation and deterioration as the color pigments are constantly exposed to light, humidity, pollution, and microbial contamination [28, 13]. It is known that Van Gogh used commercial oil paints for his paintings. In a letter to Arnold Koning, dated January 1889 and believed to refer to the the *Sunflowers* F454, van Gogh described them as being

”painted with the three chrome yellows, yellow ochre and Veronese green and nothing else” (letter 740 [2]). Recent X-ray fluorescence spectrometry analyses of F454 [29] and F458 [14] revealed that indeed van Gogh was using chrome yellows to render yellow in those paintings; it is very likely that he was using the same pigments for F457. The aging of chrome yellow pigments used in paintings has been studied in details [9, 10, 11, 13, 27], including studies on van Gogh’s *Sunflowers* [15]. Those studies have highlighted that the lightest hues in the chrome yellow family contain sulfate groups, which reduce the pigments’ stability under light: bright yellow on canvases then turns to brownish green. Our comparisons of the yellow observed for real sunflowers and the yellow visible in the sunflowers in F454, F457, and F458 are consistent with those observations. In particular, we observe an increase in the amount of blue for the yellow pixels associated with the flowers in those paintings, leading to a less vibrant yellow that may even look brown. In addition, we observe a shift to increased amount of green for those pixels (as none of them have green coordinates below 50, while the distribution of green in the yellow of real sunflowers covers the whole spectrum. This increase in blue and green is consistent with the yellow colors appearing more brownish green.

While the differences in the yellow hues of the painted *Sunflowers* compared to the yellow hues in real sunflowers are explicit for all three paintings, highlighting aging of those paintings, we have also observed significant differences between the repetition F457 and the two other paintings, F454, the original, and F458, the second repetition. This is most likely the intent of van Gogh, as proposed by Bakker and Ripoele [26]. Indeed, van Gogh painted F457 not based on real sunflowers but based on another work of art, F454 (van Gogh painted F457 in January when there were no sunflowers available). In a letter to his brother Theo (letter 736, [2]), van Gogh mentioned that this repetition was meant to be “equivalent and identical”, although it was clear that this referred to the subject (the vase and sunflowers), and not in details and colors. Van Gogh pushed chromatic intensity even further, with the aim of achieving a radical light-on-light effect, such that the green stalks of the flowers contrast even more strikingly with the various yellows than in the original, F454. Those differences remain despite the aging of the paintings. In contrast, our analyses show that F458 appears closer to F454.

2.3 Can we remediate aging of colors in paintings digitally?

Paintings age and consequently do not look today as they were originally designed by the artists. This is not evident at the level of the colors that undergo fading and sometimes changes in hue, such as yellow turning into brown. While this aging is inevitable, there is great interest among curators to recreate the artists’ original colors to enhance the experience of museum visitors. Paintings, however, cannot be physically restored to their original colors and only reconstructions offer the possibility of recreating their appearance as intended by the artist. Over the past few years, digital reconstructions of several paintings by van Gogh and other post-impressionist artists have been published, including investigation of van Gogh’s series of *The Bedroom* [30], *The Starry Night* [31], *Undergrowth with Two Figures* [32], *Irises* [33], *Roses* [33], and *Fields with Irises near Arles* [25]. Some of those reconstructions rely on identifications of regions within the paintings that contain pigments that

may have aged, using X-ray fluorescence, followed by digital reconstruction using software that can manipulate images (see for example the digital reconstruction of van Gogh’s “Undergrowth with two figures” [32], or applications of software for optimizing the display of images on mobile devices [34]. The analyses of the RGB space occupied by the pixels of the paintings *Sunflowers in a vase with a yellow background* provided above suggest two other methods for digital reconstruction, namely color correction and color transfer, which we illustrate below.

2.3.1 Restoring colors in van Gogh’s *Sunflowers* using color transfer

One approach to correcting aging in a painting is to transfer color from a recent representation of an object onto the region representing that object in the painting. For example, we can transfer the yellow color from life sunflowers observed in photographs to the regions representing sunflowers in one of the paintings. We illustrate this process in Figure 9 for F454, the original *Sunflowers in a vase with a yellow background*. We start with the ROI corresponding to the yellow flowers of F454 as well as the corresponding ROI of the yellow flowers in the photograph SF1. We apply hierarchical clustering on the pixels of each of those ROIs, using the Euclidean distance between their RGB coordinates as a metric, and complete linkage (see Methods). The hierarchical clustering generates a tree; we represent the leaves of that tree with vertical bars whose colors are the colors of the corresponding pixels. Figure 9 A and Figure 9 B illustrates the corresponding complete color bars for the yellow flower ROI of F454, and for the yellow flower ROI of SF1, respectively. In the color bar associated to F454, we identify the cluster with the darkest hues of yellow and label the corresponding pixels as F454_A. The positions of those pixels within F454 are highlighted in red in Figure 9 C. Similarly, we isolate the region with the brightest hues of yellow in SF1 and label corresponding pixels as SF1_A. Finally we transfer the color associated with SF1_A onto F454_A using the transfer algorithm described in the Method section. Result of the transfer is shown in Figure 9 E, to be compared with the original F454 shown in Figure 9 D. As expected, the sunflowers in the modified F454 are much brighter.

The color transfer strategy presented above can be expanded to play with color contrasts within the painting. We illustrate this concept by modifying the background of the flowers in F454, as illustrated in Figure 10. van Gogh painted many variations of the sunflowers, with background varying from pale to deep blue and yellow as he explored chromatic effects in the juxtaposition of those backgrounds with the yellow flowers of the sunflowers [26]. We decided to modify at least part of this background to see how it visually impact our perception of the paintings, using F454 as a support. We first identified pixels in the background of F454 that are located mostly behind the sunflowers and have light yellow hues. Those pixels are referred to as F454_B (see Figure 10 C). We then selected pixels in the real sunflowers depicted in the photograph SF1, which correspond to the dark yellow part of the flower crown. Those pixels are referred to as SF1_B. Finally we transfer the color associated with SF1_B onto F454_B using the transfer algorithm described in the Method section. Result of the transfer is shown in Figure 10 E, to be compared with the original F454 shown in Figure 10 D. As expected, the sunflowers in the modified F454 appear darker than in the original, as they are now put in a context of a darker background.

2.3.2 Restoring colors in van Gogh's Sunflowers using color correction

Possibly the most striking difference we observed when comparing the hues of yellow in the flower depiction of the three version of *Sunflowers in a vase with a yellow background* with the hues of yellow in real sunflowers observed in modern photographs is an increase in the amount of blue coordinates in those yellow hues in the paintings, leading to a less vibrant and brownish yellow, in agreement with chemical analyses of the aging of chrome yellow [9, 10, 11, 13, 27]. This observation hints at an opportunity to restore algorithmically the original colors of the *Sunflowers*: reducing the amount of blue for all pixels in the images of the paintings. We implemented two versions of such an algorithm as follows. We first note that the images of the paintings have been divided into three ROIs, namely the flowers, yf-ROI, the stems of the flowers, g-ROI, and the background, bg-ROI. The first two ROIs are relatively homogeneous in color, yellow and green, respectively, while the latter includes a more diverse spectrum of colors, as it includes the table, vase, and flower background. Each ROI is then processed independently. All pixels within a ROI are clustered first, using the differences in their color as a metric, and hierarchical clustering (see methods). The leaves of the corresponding tree are represented as vertical color lines, where the color of the line is the color of the leaf. Those color lines are organized as a color bar, which is then segmented into clusters. Figures 9 A and 10 A provide illustrations of such a color bar with its clusters for the yf-ROI and bg-ROI of F454, respectively. Each cluster regroups pixels with similar color within an ROI. Those clusters are then processed separately. For a given cluster k within an ROI, we compute first the minimal blue coordinate, m_k , over all pixels in the cluster. Let i be one such pixel, and let b_i be its blue coordinate. We correct this blue coordinate using one of the two following pseudo color restoration (PCR) schemes:

$$\text{PCR-1: } b'_i = b_i - m_k$$

$$\text{PCR-2: } b'_i = b_i - \min(m_k, 60)$$

PCR-1 was designed to remove as much blue as possible from the color coordinates of the pixels in the images of the paintings. This correction is based on the observation that the yellow colors of real sunflowers contain nearly no blue, as seen in Figure 5. It is expected to work well for all yellow hues that have been tarnished. PCR-2 is a more gentle correction as it limits the amount of blue that can be deducted. The upper limit of 60 for this deduction comes from the observation that the blue shift detected in the flower regions of the three paintings is close to this value (this is especially clear for F457, see Figure 4 B).

In Figure 11 we show the results of applying the two strategies PCR-1 and PCR-2 described above. There are a few observations associated with those reconstructions. First, visually the differences between PCR-1 and PCR-2 are small. This is expected as those two schemes are expected to act similarly on yellow hues, and yellow dominates in the three paintings. Second, there is a clear difference between the restored images and the original images as the colors appear much brighter in the reconstructed images, especially the yellow in the background (see for example the effects on F454). While we need to be cautious as to ascertaining that the reconstructions represent the paintings as originally intended by van Gogh, we can safely say that they are most likely closer to his intent, as van Gogh was playing with

the contrasts between the sunflowers and the background (as he was experimenting with different colors and intensities for those background) and that darker colors reduce this contrast (see Figures 10 D and E). Finally, the reconstructed models of the paintings highlight differences in the contrast between flowers and background between the original, F454, and the two repetitions, F457 and F458, with the original based on a brighter background. This observation reemphasize the importance given by van Gogh to the capture of light, colors, and contrast for the *Sunflowers in a vase*.

3 Conclusion

Vincent van Gogh was by no means a chemist to his own admission (letter 889 to his brother Theo [2]). However, he was well aware that colors in paintings evolve due to changes in the chemical nature of their pigments: "...paintings fade like flowers" (letter 765 to Theo, [2]). In the case of the *Sunflowers in a vase with a yellow background*, namely an original, F454, and two replicates, F457 and F458 that van Gogh painted using the original as a model, this sentiment came true, as the background and flower rendition in those paintings have faded and turned brown, making them less vibrant that van Gogh had most likely aimed at. His intents with those paintings were to "push chromatic intensity ... with the aim of achieving a radical light-on-light effect" [26]. We have attempted to restore van Gogh's intent using a computational approach based on data science. After identifications of regions of interest (ROI) within the three paintings F454, F457, and F458 that capture the flowers, stems of the flowers, and background, respectively, we studied the geometry of the color space (in RGB representation) occupied by those ROIs. By comparing those color spaces with those occupied by similar ROIs in photographs of real sunflowers, we identified shifts in all three color coordinates, R, G, and B, with the shift in the blue coordinate being the more salient. This shift in blue that leads to hues of yellow that are faded and even brownish are consistent with the fading of the chrome yellow, the pigments used by van Gogh for representing yellow in *Sunflowers*, observed by chemical spectroscopic methods [9, 10, 11, 13, 27]. We have proposed two algorithms, PCR-1 and PCR-2, for correcting that shift in blue and generate representations of the paintings that aim to restore their original conditions (see Figure 11). While we acknowledge that these are models, the reduction of the blue component in the yellow hues has lead to more vibrant and less brownish digital rendition of the three *Sunflowers in a vase with a yellow background*.

While we believe that the models we have generated for the *Sunflowers* take the viewers closer to the paintings created by van Gogh, we acknowledge that the current state-of-the-art techniques for digital reconstructions, including our own, are still limited in number and far from actually restoring the original version of the painter. Progress will ultimately come from combinations of techniques. We did not have access to spectroscopic fluorescence data on the *Sunflowers*; we believe that such data would have helped us delineate the regions of interest in the paintings, in particular those regions where chrome yellow dominates as those regions are more susceptible of fading due to aging of the chrome yellow pigments. In return, the computational methods proposed here for analyzing the RGB space occupied by those regions should help identify the impact in color space of aging, as well

as methods for reverting those effects. In addition, those methods enable manipulations of the images of the paintings of interest and therefore the assessment of hypotheses on how the painter was experimenting with color juxtapositions within a painting. Based on our preliminary studies of the *Sunflowers*, it is our intent to develop a general computational framework for digital restoration and manipulation of paintings and make this framework available as a tool for enhancing painting viewing experience.

4 Materials and methods

4.1 Material: the digital images used in this study

F454 is the fourth version of the *Sunflowers* painted by van Gogh while he was living in Arles, with the specificity of having a yellow background in contrast with the turquoise (F453), royal-blue (F459), and blue-green (F456) for the other versions. It is currently owned by the National Gallery in London, UK, and stored under the inventory number NG3863. The National Gallery only provides a low resolution image of this painting on its website; we found a high resolution version on Wikimedia commons, from the URL https://upload.wikimedia.org/wikipedia/commons/4/46/Vincent_Willem_van_Gogh_127.jpg. This high resolution image is provided in jpeg format, with a resolution of $3,349 \times 4,226$ pixels.

F458 is one of the two repetitions of F454 painted by van Gogh, owned by and displayed at the Van Gogh Museum, Amsterdam. The van Gogh museum only provides a low resolution image of this painting on its web site (<https://www.vangoghmuseum.nl/en/collection/s0031V1962>). However, just like for F454, we found a high resolution version on Wikimedia commons, from the URL https://upload.wikimedia.org/wikipedia/commons/9/9d/Vincent_van_Gogh_-_Sunflowers_-_VGM_F458.jpg. This high resolution image is provided in Jpeg format, with a resolution of $3,224 \times 4,226$ pixels.

The high resolution image of F457 was generously provided by the Seiji Togo memorial Sampo Japan Museum of Art, Tokyo, Japan. It was also provided in jpeg format, with a resolution of 3626×4829 pixels.

In addition to the three high resolution images of the three paintings F454, F457, and F458, we used two recent digital images of sunflowers in a field, to capture actual colors of sunflowers. There two pictures, which we label as SF1 and SF2, were obtained from:

- SF1) A picture of sunflowers in Provence (Shuttersandsunflowers.com), downloaded with permission of the author.
- SF2) A picture of sunflowers in a field, by Leo Adamchuk, available within Adobe Stock and downloaded as part of their free trial.

4.2 Color statistics in digital images

A digital image is an image composed of picture elements, or *pixels*. Each pixel is characterized by spatial coordinates, (x, y) , that define its positions along the x-axis and y-axis within the image, as well as by color components, (r, g, b) , that define the amount of red (R), green (G), and blue (B) which, when combined, describe the color at the pixel. Note that we rely here on the additive RGB color model; this model is not universally accepted and other models such as CMYK (a subtractive

model) or Lab are possible. We have used RGB as it is the color model that was available with the images we recovered. In this RGB model, each color component r , g , or b is an integer in the range $[0, 255]$. Namely, each fundamental color R , G , or B is described by a discrete variable that can take 256 distinct values, while a composite color is described by its 3 coordinates along those fundamental colors, and therefore belongs to a discrete space (in this case a cube) with $[0, 255]^3$ possible values.

Let us define a digital image I as the set $S(I)$ of its pixels. Let N be the cardinality of that set, namely the total number of pixels in the image. Let C be one of the fundamental colors, namely C can be R , G , or B . As we have seen above, C can take 256 possible discrete values, each in the range $[0, 255]$. We can compute the distributions of these values over a given image as follows. Let $S(c)$ be the set of pixels such that its color coordinate for the color C is c :

$$S(c) = \{p \in P(I) \mid C(p) = c\}$$

The probability of observing color C with intensity c within the image is then given by:

$$P(C = c) = \frac{|S(c)|}{N} \quad (1)$$

where $|S(c)|$ stands for the number of elements of set $S(c)$, i.e., its number of elements. The entropy of the color C within the image is then given by:

$$H(C) = - \sum_{c=0}^{255} P(C = c) \log(P(C = c))$$

The entropy measures the amount of “information” associated with the color C in the image. If the color C is always represented with the same value over each pixel, the entropy is zero, while if the possible values for the color are evenly distributed, the entropy is at its maximum with a value of $\log(256)$.

We can also measure the association of fundamental colors within the image. Let C and D be two of the three fundamental colors. Let $P(C = c, D = d)$ be the joint probability of those colors C and D taking the values c and d , respectively. If we define the set $S(c, d)$ as

$$S(c, d) = \{p \in P(I) \mid C(p) = c \ \& \ D(p) = d\}$$

then

$$P(C = c, D = d) = \frac{|S(c, d)|}{N} \quad (2)$$

From this joint probability, we can compute the conditional probability that $C = c$, knowing that $D = d$, as

$$P(C = c \mid D = d) = \frac{P(C = c, D = d)}{P(d)}$$

Let us define now $H(C | D = d)$ as the entropy of the color C conditioned on the color D taking the value d . This entropy is computed as:

$$\begin{aligned} H(C | D = d) &= - \sum_{c=0}^{255} P(C = c | D = d) \log(P(C = c | D = d)) \\ &= - \sum_{c=0}^{255} \frac{P(C = c, D = d)}{P(D = d)} \log \frac{P(C = c, D = d)}{P(D = d)} \end{aligned}$$

where the sums extend over all 256 possible values for the coordinates associated to the color C . The entropy of the color C conditioned on the color D is then computed as the average of those numbers over all possible values for d :

$$\begin{aligned} H(C | D) &= \sum_{d=0}^{255} P(D = d) H(C | D = d) \\ &= - \sum_{d=0}^{255} P(D = d) \sum_{c=0}^{255} \frac{P(C = c, D = d)}{P(D = d)} \log \frac{P(C = c, D = d)}{P(D = d)} \\ &= - \sum_{d=0}^{255} \sum_{c=0}^{255} P(C = c, D = d) \log \frac{P(C = c, D = d)}{P(D = d)} \end{aligned}$$

where the probabilities $P(D = d)$ and $P(C = c, D = d)$ are computed using Equations 1 and 2, respectively. Finally, we compute the relative conditional entropy as

$$HR(C | D) = \frac{H(C | D)}{H(C)}$$

$HR(C | D)$ takes values between 0 and 1. $HR(C | D) = 0$ if and only if the values of C are completely determined by the values of D . Conversely, $HR(C | D) = 1$ if and only if the values of C and D are independent of each other.

4.3 Hierarchical Clustering Analysis (HCA)

Clustering is the task of regrouping objects such that those that belong to the same group, called a cluster, are more similar to each other than to those in other groups. In our analyses, the objects are pixels within an image. A pixel $P(i)$ is characterized by its color, $c(i)$, given as a combination of its components along the fundamental colors R , G , and B , i.e. $c(i) = (r(i), g(i), b(i))$. As described above, $c(i)$ belongs to the 3D discrete space $[0, 255]^3$. The similarity between two pixels $P(i)$ and $P(j)$ is set to be the Euclidean distance between their colors:

$$d(P(i), P(j)) = \sqrt{(r(i) - r(j))^2 + (g(i) - g(j))^2 + (b(i) - b(j))^2} \quad (3)$$

The clustering of the pixels is then performed using the agglomerative hierarchical clustering analysis, or HCA. This is a bottom-up approach in which each pixel starts in its own cluster, and pairs of clusters are merged iteratively until all pixels

belong to the same cluster. The whole procedure defines a hierarchy of clusters, also referred as a clustering tree. A key element to this procedure is to define the distance between two clusters, also referred to as the linkage criterion. When the two clusters contain a single element, this distance is simply the distance between those elements, as defined in equation 3. When the two clusters A and B are sets of elements, the distance is then defined as a function of the pairwise distances between those elements. Two common choices are the single linkage:

$$d(A, B) = \min\{d(a, b), a \in A, b \in B\}$$

and the complete linkage:

$$d(A, B) = \max\{d(a, b), a \in A, b \in B\}$$

We have used the complete linkage in all our analyses.

4.4 Identifying regions of interest in the paintings *Sunflowers in a vase with a yellow background*

A region of interest (ROI) is defined according to its position in the image as well as on its color consistency. In the *Sunflowers in a vase with a yellow background*, we identify three such ROIs, namely the flower crowns of the sunflowers (yellow), yf-ROI, the stems of the sunflowers (green), g-ROI, and the background, bg-ROI. Automated segmentation is one approach for selecting the pixels that belong to each ROI. We have used instead a combination of manual and automated processing, using the algorithm described below.

This algorithm is based on the fact that the yf-ROI and g-ROI are chosen based on color consistency. We first divide the whole discrete RGB space $[0, 255] \times [0, 255] \times [0, 255]$ into a collection of $5 \times 5 \times 5$ color cubes. All pixels of an image are then attached to those cubes based on their colors. Cubes that are occupied are then assigned a representative, whose color is identified as the center of the cube. All those representatives are then clustered using the HC technique described above, using the Euclidean distance between their colors as a similarity measure, and complete linkage. At the bottom of the HC tree, we draw a color bar with each leaf of the tree represented with a vertical line whose color corresponds to the cube associated with that leaf. The HC tree is then cut to form 70 clusters; we chose 70 somewhat arbitrarily, such that we would have enough cluster. Most of those clusters have consistent colors along the color bar described above. . Two groups of clusters are selected visually, those that are predominantly yellow, with a yellow that resembles the yellow of the flower in the paintings, and those that are predominantly green. Each of this group is then processed separately.

Let G be the group of clusters of green colors identified above. Collect all the pixels in G , and repeat the procedure described above, but with the RGB space divided into a collection of $4 \times 4 \times 4$ cubes. Again, identify all cubes that are not empty, cluster their representatives, and select only those clusters that contain pixels whose colors are consistently green. The remaining pixels are processed one last time, with the RGB space divided into a collection of $3 \times 3 \times 3$ cubes. All the

corresponding occupied cubes are ordered with respect to the number of pixels they contain. Cubes are retained by going down on the ordered list, starting with the most populated cube, until 90% of the remaining pixels are selected. Finally, those pixels are identified in the paintings, and those that do not correspond to a flower stem are discarded. The remaining pixels form the green ROI, g-ROI, associated with the stems of the sunflowers.

Let Y be the group of clusters of yellow colors identified in the first step of the algorithm. This group of clusters is processed in the same way as G was processed, leading to a group of pixels that are mostly yellow and that belong to the crowns of the sunflowers. This group of pixels forms divided the yf-ROI, associated with the crowns of the sunflowers.

Finally, all pixels that do not belong to the g-ROI and yf-ROI are deemed to belong to the background (this includes the table in the front, the vase, and the background behind the flowers. Those pixels form the bg-ROI.

4.5 Region specific Color transfer between two images

Assume that we have identified a set of pixels A in one picture and that those pixels have similar (also not necessarily equal) colors. Similarly, we have identified a set of pixels B in another picture, with similar colors. A and B may be of different size and may contain different colors. Our goal is to transfer the colors from A to B . Let $A = \{\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_N\}$ and $B = \{\mathbf{b}_1, \dots, \mathbf{b}_j, \dots, \mathbf{b}_M\}$ where the \mathbf{a} s and \mathbf{b} s are 3D vectors containing the R, G, and B coordinates of the colors of the pixels in A and B , respectively. We apply the following algorithm:

- i) Compute the centers of mass of A and B : $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$;
- ii) Translate B to B^* so that A and B^* have the same center of mass: $B^* = \{\mathbf{b}_1^*, \dots, \mathbf{b}_j^*, \dots, \mathbf{b}_M^*\}$ with $\mathbf{b}_j^* = \mathbf{b}_j - \bar{\mathbf{b}} + \bar{\mathbf{a}}$;
- iii) Build a HC-tree on the union of A and B^* , using the Euclidean distance as a metric and complete linkage, and cut at a tree level to define clusters. If a cluster only contains elements originally from B^* , merge it with its closest cluster (in terms of tree distance) that contains elements of A . At the end of this procedure, we have a set of clusters, with each cluster containing either elements of A and B^* , or elements of A only. The latter are discarded.
- iv) Let C be one of the clusters from step iii), and let $AC = \{\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_{NC}\}$ and $BC = \{\mathbf{b}_1^*, \dots, \mathbf{b}_j^*, \dots, \mathbf{b}_{MC}^*\}$ be the subsets of A and B^* that belong C , where NC and MC are their sizes, respectively. For each j in $[1, MC]$, replace \mathbf{b}_j^* with a color picked randomly in AC . Repeat over all clusters C .

At the end of this procedure, each element j of B has been assigned a new color b_j^* that is inherited from A . The pixels in the second picture are then assigned those new colors.

Availability of data and material

Digital copies of the three paintings F454, F457, and F458, of the two photographs SF1 and SF2, as well as all R scripts used for this paper are included in the submission as a compressed zip file, attached as electronic supplementary material

Competing interests

The authors have no competing financial interests

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors

Author's contributions

S. L. performed all the data analyses as well as wrote the code needed for those analyses. F. H and P. K supervised the study. All authors designed the study, wrote and edited the manuscript.

Acknowledgements

We thank Mr Akihiro Murata, the Exhibition General Manager of the Seiji Togo memorial Sompo Japan Museum of Art, Tokyo, Japan, for providing us with the high-resolution image of F457 as well as for informative discussions.

Author details

¹Graduate Group in Biostatistics, University of California, Davis, One Shields Avenue, 95616 Davis, USA.

²Department of Computer Science and Genome Center, University of California, Davis, 451 East Health Sciences Drive, 95616 Davis, USA. ³Department of Mathematics, University of California, Davis, One Shields Avenue, 95616 Davis, USA. ⁴Department of Statistics, University of California, Davis, One Shields Avenue, 95616 Davis, USA.

References

- van Tilborgh, L., Hendricks, E.: The Tokyo *Sunflowers*: a genuine repetition by Van Gogh or a Schuffenecker forgery? *Van Gogh Museum Journal*, 16 (2001)
- Bakker, N., Jansen, L., Luijten, H.: *Vincent van Gogh: A Life in Letters*. Thames and Hudson, London, UK (2020)
- van Tilborgh, L.: *Van Gogh and the Sunflowers*. Van Gogh Museum, Amsterdam, NL (2008)
- Chittka, L., Walker, J.: Do bee like Van Gogh's *Sunflowers*? *Optics and Laser Technology* **38**, 323–328 (2006)
- Chapman, M.A., Tang, S., Draeger, D., Nambeesan, S., Shaffer, H., Barb, J.G., Knapp, S.J., Burke, J.M.: Genetic analysis of floral symmetry in Van Gogh's sunflowers reveals independent recruitment of CYCLOIDEA genes in the Asteraceae. *PLoS Genetics* **8**, 1002628 (2012)
- Arnold, W.N., Loftus, L.S.: Xanthopsia and van Gogh's yellow palette. *Eye (Lond.)* **5**, 503–510 (1991)
- Blumer, D.: The illness of Vincent van Gogh. *The American journal of psychiatry* **159**, 519–526 (2002)
- Gruener, A.: Vincent van Gogh's yellow vision. *British Journal of General Practice* **63**, 370–371 (2013). doi:10.3399/bjgp13X669266. <https://bjgp.org/content/63/612/370.full.pdf>
- Monico, L., Van der Snickt, G., Janssens, K., De Nolf, W., Miliani, C., Verbeeck, J., Tian, H., Tan, H., Dik, J., Radepon, M., Cotte, M.: Degradation process of lead chromate in paintings by Vincent van Gogh studied by means of synchrotron x-ray spectromicroscopy and related methods. 1. artificially aged model samples. *Analytical Chemistry* **83**, 1214–1223 (2011). doi:10.1021/ac102424h. PMID: 21314201
- Monico, L., Janssens, K., Miliani, C., Van der Snickt, G., Brunetti, B.G., Cestelli Guidi, M., Radepon, M., Cotte, M.: Degradation process of lead chromate in paintings by Vincent van Gogh studied by means of spectromicroscopic methods. 4. artificial aging of model samples of co-precipitates of lead chromate and lead sulfate. *Analytical Chemistry* **85**, 860–867 (2013). doi:10.1021/ac3021592. PMID: 23051631
- Monico, L., Janssens, K., Vanmeert, F., Cotte, M., Brunetti, B.G., Van der Snickt, G., Leeuwestein, M., Salvant Plisson, J., Menu, M., Miliani, C.: Degradation process of lead chromate in paintings by Vincent van Gogh studied by means of spectromicroscopic methods. part 5. effects of nonoriginal surface coatings into the nature and distribution of chromium and sulfur species in chrome yellow paints. *Analytical Chemistry* **86**, 10804–10811 (2014). doi:10.1021/ac502841g. PMID: 25301536
- Monico, L., Janssens, K., Hendriks, E., Vanmeert, F., Van der Snickt, G., Cotte, M., Falkenberg, G., Brunetti, B.G., Miliani, C.: Evidence for degradation of the chrome yellows in Van Gogh's *Sunflowers*: A study using noninvasive in situ methods and synchrotron-radiation-based X-ray techniques. *Angew. Chem. Int. Ed.* **54**, 13923–13927 (2015)
- Vanmeert, F., Hendriks, E., Van der Snickt, G., Monico, L., Dik, J., Janssens, K.: Chemical mapping by macroscopic X-ray powder diffraction (MA-XRPD) of Van Gogh's sunflowers: Identification of areas with higher degradation risk. *Angewandte Chemie (International ed. in English)* **57**, 7418–7422 (2018). doi:10.1002/anie.201713293
- Hendriks, E., Geldof, M., Monico, L., Johnson, D.H., Miliani, C., Romani, A., Grazia, C., Buti, D., Brunetti, B.G., Janssens, K., Van der Snickt, G., Vanmeert, F.: *Methods and Materials of the Amsterdam Sunflowers*, pp. 85–123. Amsterdam University Press, ??? (2019). <http://www.jstor.org/stable/j.ctvx8b758.7>
- Monico, L., Hendriks, E., Geldof, M., Miliani, C., Janssens, K., Brunetti, B.G., Cotte, M., Vanmeert, F., Chieli, A., Van der Snickt, G., Romani, A., Melo, M.J.: *Chemical Alteration and Colour Changes in the Amsterdam Sunflowers: A Focus on Geranium Lakes and Chrome Yellows*, pp. 125–157. Amsterdam University Press, ??? (2019). <http://www.jstor.org/stable/j.ctvx8b758.8>
- van den Berg, K.J., Hendriks, E., Geldof, M., de Groot, S., van der Werf, I., Miliani, C., Moretti, P., Cartechini, L., Monico, L., Iwanicka, M., Targowski, P., Sylwestrzak, M., Genuit, W.: *Structure and Chemical Composition of the Surface Layers in the Amsterdam Sunflowers*, pp. 159–173. Amsterdam University Press, ??? (2019). <http://www.jstor.org/stable/j.ctvx8b758.9>
- Iwanicka, M., Sylwestrzak, M., Szkulmowska, A., Targowski, P.: *Methods and Techniques*, pp. 207–227. Amsterdam University Press, ??? (2019). <http://www.jstor.org/stable/j.ctvx8b758.11>
- Hendriks, E., Vellekoop, M.: *Van Gogh's Sunflowers Illuminated: Art Meets Science*. Amsterdam University Press, ??? (2019). <http://www.jstor.org/stable/j.ctvx8b758>
- van Noord, N., Hendriks, E., Postma, E.: Toward discovery of the artist's style: learning to recognize artists by their artworks. *IEEE Signal Processing Magazine* **32**, 46–54 (2015)
- Sandoval, C., Pirogova, E., Lech, M.: Two-stage deep learning approach to the classification of fine-art paintings. *IEEE Access* **7**, 41770–41781 (2019)
- Polatkan, G., Jafarpour, S., Brasoveanu, A., Hughes, S., Daubechies, I.: Detection of forgery in paintings using supervised learning. In: 2009 16th IEEE International Conference on Image Processing (ICIP), pp. 2921–2924 (2009). doi:10.1109/ICIP.2009.5413338
- Coeckelbergh, M.: Can machines create art? *Philosophy and Technology* **30**, 285–303 (2017)

23. Hendriks, E., Brokerhof, A.W., van den Meiracker, K.: Valuing van Gogh's colours: From the past to the future. In: ICOM Committee for Conservation 18th Triennial Meeting (2017)
24. Geldof, M., Proaño Gaibor, A.N., Ligterink, F., Hendriks, E., Kirchner, E.: Reconstructing van Gogh's palette to determine the optical characteristics of his paints. *Heritage Science* **6**, 17 (2018)
25. Kirchner, E., Geldof, M., Hendriks, E., Proaño Gaibor, A.N., Janssens, K., Delaney, J., van der Lans, I., Ligterink, F., Megens, L., Meedendorp, T., Pilz, K.: Recreating van Gogh's original colors on museum displays. In: Eschbach, R., Marcu, G.G., Rizzi, A. (eds.) *Color Imaging XXIV: Displaying, Processing, Hardcopy, and Applications*, Burlingame, CA, USA, 13-17 January 2019. Ingenta, ??? (2019)
26. Bakker, N., Riopelle, C.: The Sunflowers in Perspective, pp. 21–47. Amsterdam University Press, ??? (2019). <http://www.jstor.org/stable/j.ctvx8b758.5>
27. Barberio, M., Skantzakis, E., Sorieul, S., Antici, P.: Pigment darkening as case study of in-air plasma-induced luminescence. *Science Advances* **5** (2019). doi:[10.1126/sciadv.aar6228](https://doi.org/10.1126/sciadv.aar6228). <https://advances.sciencemag.org/content/5/6/eaar6228.full.pdf>
28. Korenberg, C.: The photo-ageing behaviour of selected watercolour paints under anoxic conditions. *British Museum Technical Research Bulletin* **2**, 49–57 (2008)
29. Higgitt, C., Macaro, G., Spring, M.: Methods, Materials and Condition of the London Sunflowers, pp. 49–83. Amsterdam University Press, ??? (2019). <http://www.jstor.org/stable/j.ctvx8b758.6>
30. Fiedler, I., Hendriks, E., Meedendorp, T., Menu, M., Salvant, J.: *Materials, Intention, and Evolution*. Yale University Press, ??? (2016)
31. Zhao, Y., Berns, R.S., Taplin, L., Coddington, J.: An investigation of multispectral imaging for the mapping of pigments in paintings. In: Stork, D.G., Coddington, J. (eds.) *Computer Image Analysis in the Study of Art*, vol. 6810, pp. 65–73. SPIE, ??? (2008). International Society for Optics and Photonics
32. Fieberg, J.E., Knutås, P., Hostettler, K., Smith, G.D.: "paintings fade like flowers": Pigment analysis and digital reconstruction of a faded pink lake pigment in Vincent van Gogh's undergrowth with two figures. *Applied Spectroscopy* **71**, 794–808 (2017)
33. Centeno, S.A., Hale, C., Carò, F., Cesarotto, A., Shibayama, N., Delaney, J., Dooley, K., van der Snickt, G., Janssens, K., Stein, S.A.: Van Gogh's irises and roses: the contribution of chemical analyses and imaging to the assessment of color changes in the red lake pigments. *Herit. Sci.* **5**, 18 (2017)
34. Kirchner, E., van der Lans, I., Martínez-Verdú, F.M., Perales, E.: Improving color reproduction accuracy of a mobile liquid crystal display. *J. Opt. Soc. Am. A* **34**, 101–110 (2017)

Figures

Figure 1 Pictures of the three paintings of *Sunflowers in a vase with a yellow background* from Vincent van Gogh: (A) the fourth version, F454, painted in 1888 in Arles and currently at the National Gallery in London, and its two repetitions, also painted in Arles in 1889, F457, currently in the Seiji Tojo memorial Sampo Museum in Tokyo (B), and F458, at the van Gogh museum in Amsterdam (C).

Figure 2 Distributions of the values of the color coordinates for Red (in red), Green (in green) and Blue (in blue) for all pixels of the images of the three *Sunflowers* paintings (A) F454, (B) F457, and (C) F458.

Figure 3 Relative conditional entropy values between the red, green, and blue coordinates of all pixels of (A) F454, (B) F457, and (C) F458 *Sunflowers in a vase with yellow background*; values range from 0 to 1 where 0 indicates dependence and 1 independence between the two coordinates.

Figure 4 The RGB space occupied by the pixels of the yellow flower ROIs of the original *Sunflowers in a vase with yellow background*, F454 (top row), compared with the two duplicates F457 (middle row), and F458 (bottom row) (A) red vs green; each dot represents one pixel; its color (grayscale) suggests the intensity of its blue coordinate, ranging from low (white) to high (black); (B) red vs blue; (C) green vs blue

Figure 5 The RGB space occupied by the pixels of the yellow flower ROIs of the photographs SF1 (top row) and SF2 (bottom row) of real sunflowers. (A) red vs green; each dot represents one pixel; its color (grayscale) suggests the intensity of its blue coordinate, ranging from low (white) to high (black); (B) red vs blue; (C) green vs blue

Figure 6 The RGB space occupied by the pixels in the green stem ROIs of the original *Sunflowers in a vase with a yellow background*, F454 (top row), compared with the two duplicates F457 (middle row), and F458 (bottom row) (A) red vs green; each dot represents one pixel; its color (grayscale) suggests the intensity of its blue coordinate, ranging from low (white) to high (black); (B) red vs blue; (C) green vs blue

Figure 7 The RGB space occupied by the pixels of the green stem ROIs of the photographs SF1 (top row) and SF2 (bottom row) of real sunflowers. (A) red vs green; each dot represents one pixel; its color (grayscale) suggests the intensity of its blue coordinate, ranging from low (white) to high (black); (B) red vs blue; (C) green vs blue

Figure 8 The RGB space occupied by the pixels in the background ROIs of the original *Sunflowers in a vase with a yellow background*, F454 (top row), compared with the two duplicates F457 (middle row), and F458 (bottom row). Note that the background includes the table, the vase, and the region behind the flowers. (A) red vs green; each dot represents one pixel; its color (grayscale) suggests the intensity of its blue coordinate, ranging from low (white) to high (black); (B) red vs blue; (C) green vs blue

Figure 9 Color transfer from a real sunflower to the sunflowers in F454 (A) The color bar representing all hues of yellow in the ROI corresponding to the flowers in F454. This color corresponds to the leaves of the HC tree computed from the colors of all pixels in this ROI. We select the region (cluster) 454.yE in this color bar which contains some of the darkest yellow pixels. The pixels associated to this region are referred to as F454.A. (B) The color bar representing all hues of yellow in the flower ROI of SF1. We identify the cluster SF1.yG in this color bar that contains some of the brightest yellow. The pixels associated to this region are referred to as SF1.A (C) Pixels in F454 that belong to F454.A are highlighted in red (D) F454: before color transfer (E) F454: after color transfer.

Figure 10 Color transfer from a real sunflower to the yellow part of the background in F454 (A) The color bar representing all hues of colors in the ROI corresponding to the background in F454. This color corresponds to the leaves of the HC tree computed from the colors of all pixels in this ROI. We select the region (cluster) 454.bG in this color bar which contains some of the pixels that are directly located behind the sunflowers in the vase. The pixels associated to this region are referred to as F454.B. (B) The color bar representing all hues of yellow in the flower ROI of SF1. We identify the cluster SF1.yK in this color bar that contains some of the darkest yellow. The pixels associated to this region are referred to as SF1.B (C) Pixels in F454 that belong to F454.B are highlighted in red (D) F454: before color transfer (E) F454: after color transfer.

Figure 11 Color restoration of the *Sunflowers in a vase with a yellow background* For the original sunflower in a vase with yellow background, F454 currently in London (top row), and the two repetitions F457, currently in Tokyo, middle row, and F548, currently in Amsterdam, bottom row, we show a picture of the current painting (left column), as well as two possible models of the painting intended by van Gogh, using the color reconstruction methods PCR-1 (middle column) and PCR-2 (right column). See text for details.

Figure 12 Two images of natural live sunflowers (see text for details): (A) SF1; (B) SF2.

Figures

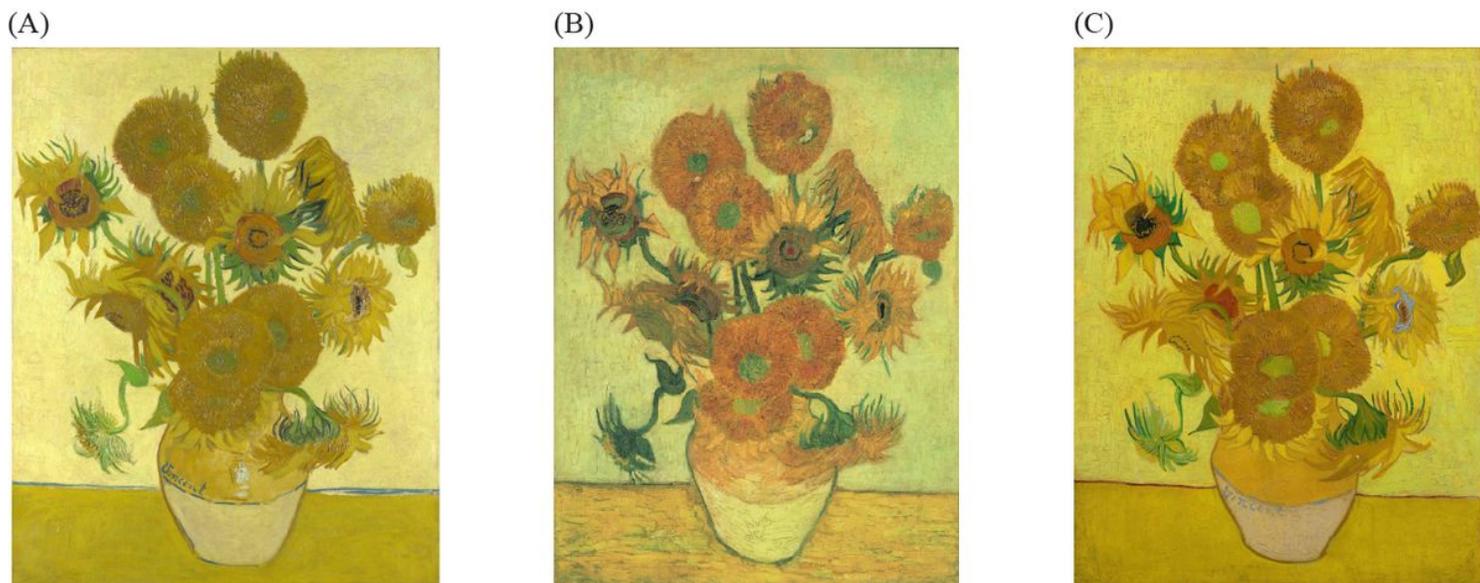


Figure 1

Pictures of the three paintings of Sunflowers in a vase with a yellow background from Vincent van Gogh: (A) the fourth version, F454, painted in 1888 in Arles and currently at the National Gallery in London, and its two repetitions, also painted in Arles in 1889, F457, currently in the Seiji Tojo memorial Sampo Museum in Tokyo (B), and F458, at the van Gogh museum in Amsterdam (C).

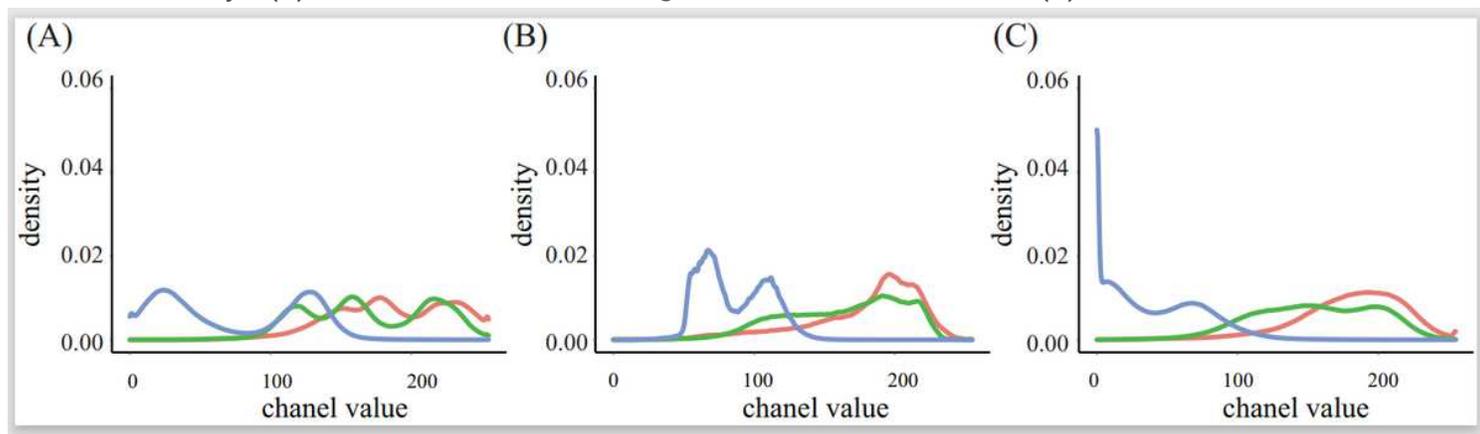


Figure 2

Distributions of the values of the color coordinates for Red (in red), Green (in green) and Blue (in blue) for all pixels of the images of the three Sunflowers paintings (A) F454, (B) F457, and (C) F458.

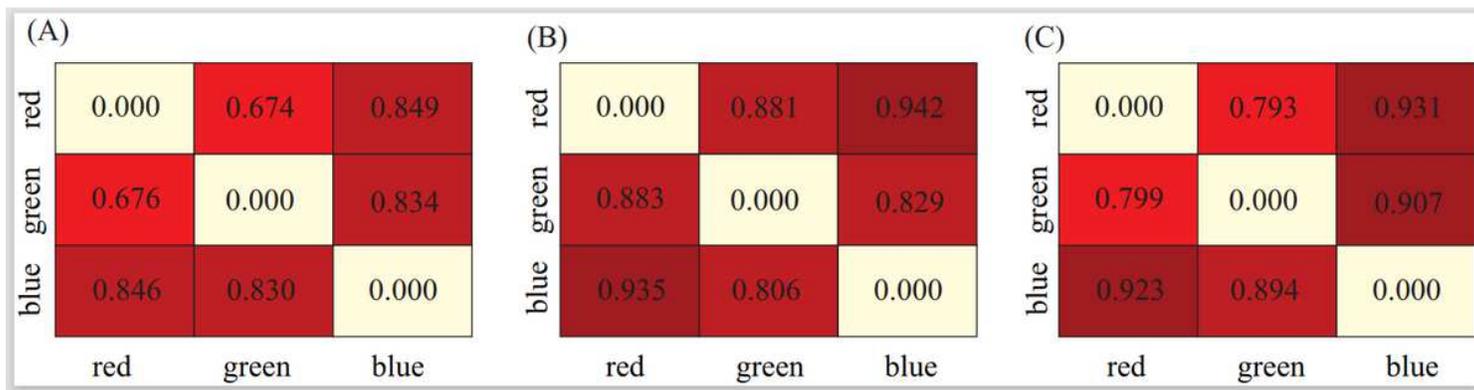


Figure 3

Relative conditional entropy values between the red, green, and blue coordinates of all pixels of (A) F454, (B) F457, and (C) F458 Sunflowers in a vase with yellow background; values range from 0 to 1 where 0 indicates dependence and 1 independence between the two coordinates.

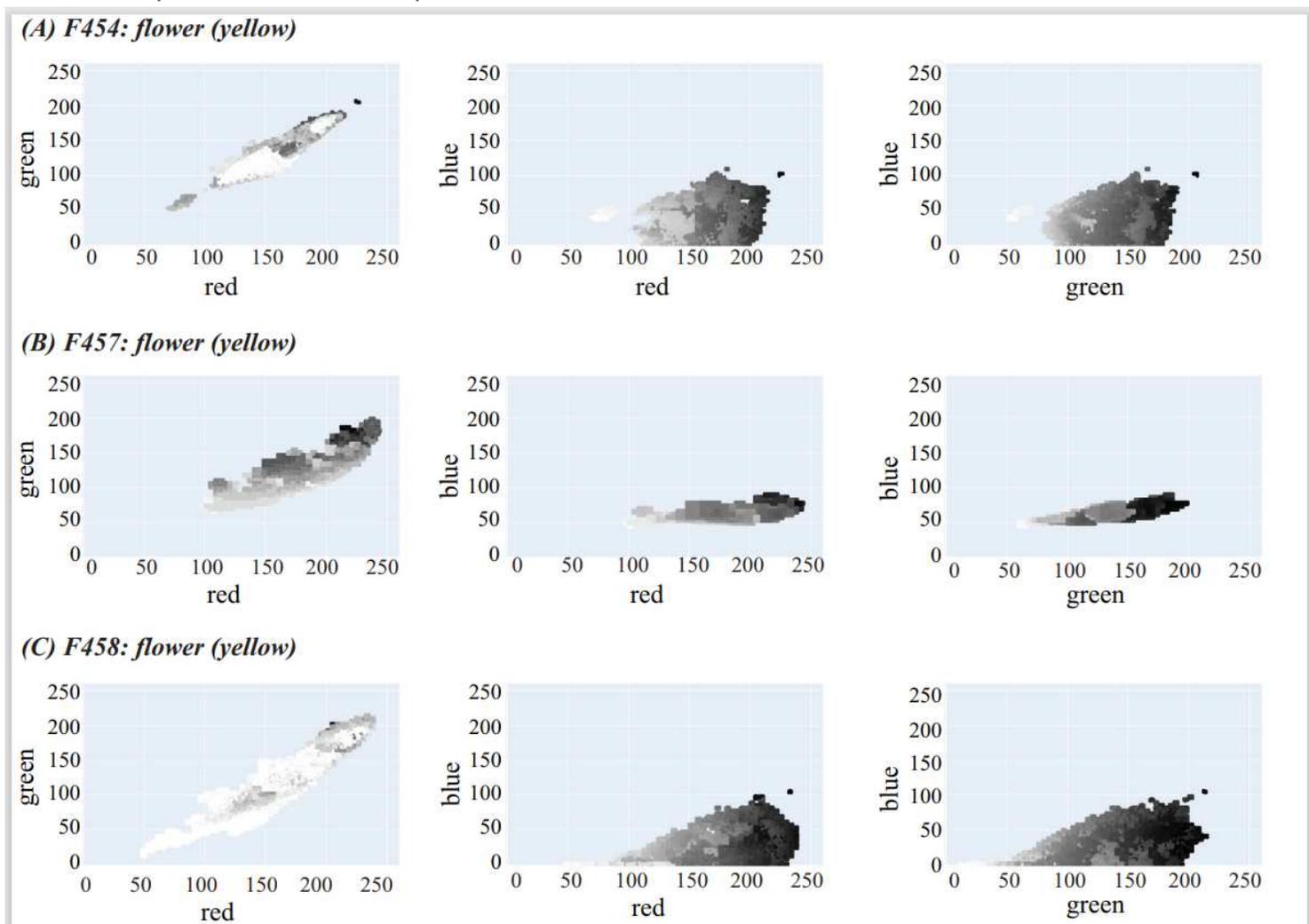


Figure 4

The RGB space occupied by the pixels of the yellow flower ROIs of the original Sunflowers in a vase with yellow background, F454 (top row), compared with the two duplicates F457 (middle row), and F458 (bottom row) (A) red vs green; each dot represents one pixel; its color (grayscale) suggests the intensity of its blue coordinate, ranging from low (white) to high (black); (B) red vs blue; (C) green vs blue

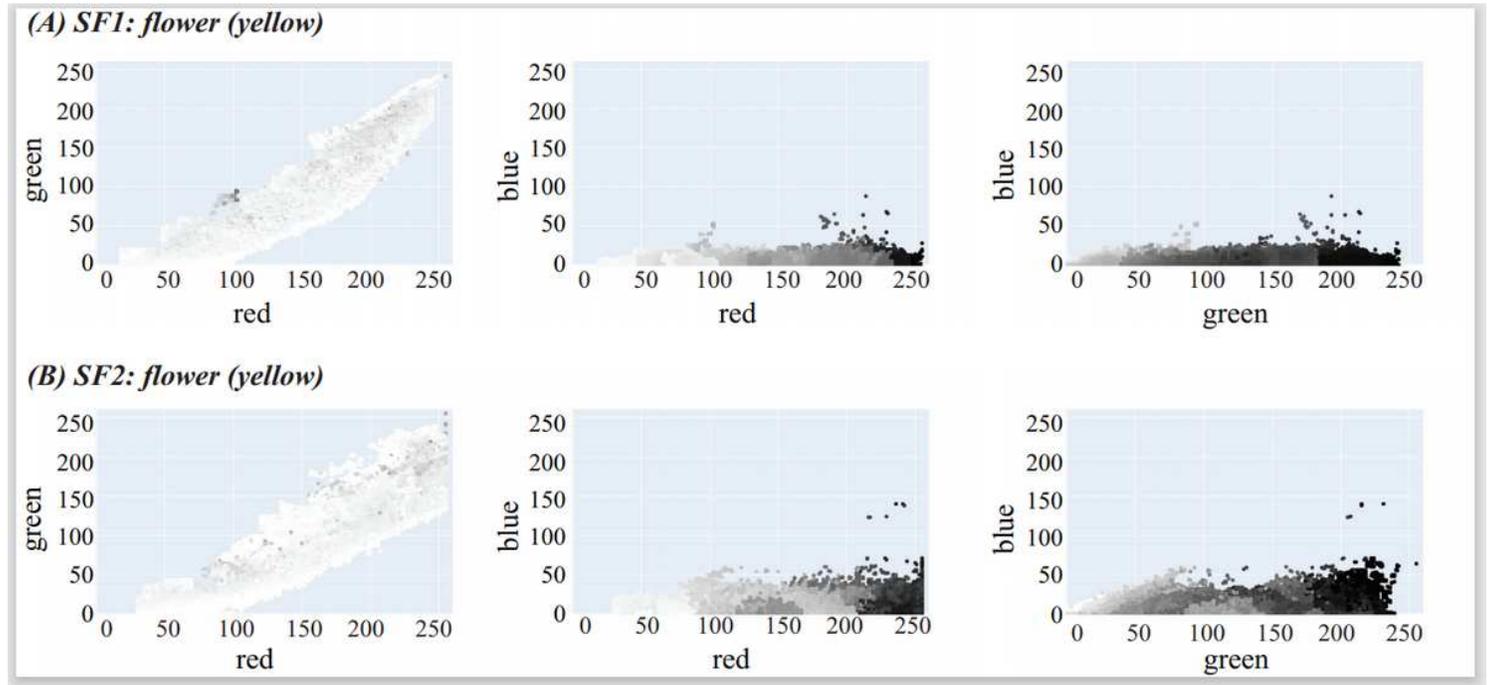


Figure 5

The RGB space occupied by the pixels of the yellow flower ROIs of the photographs SF1 (top row) and SF2 (bottom row) of real sunflowers. (A) red vs green; each dot represents one pixel; its color (grayscale) suggests the intensity of its blue coordinate, ranging from low (white) to high (black); (B) red vs blue; (C) green vs blue

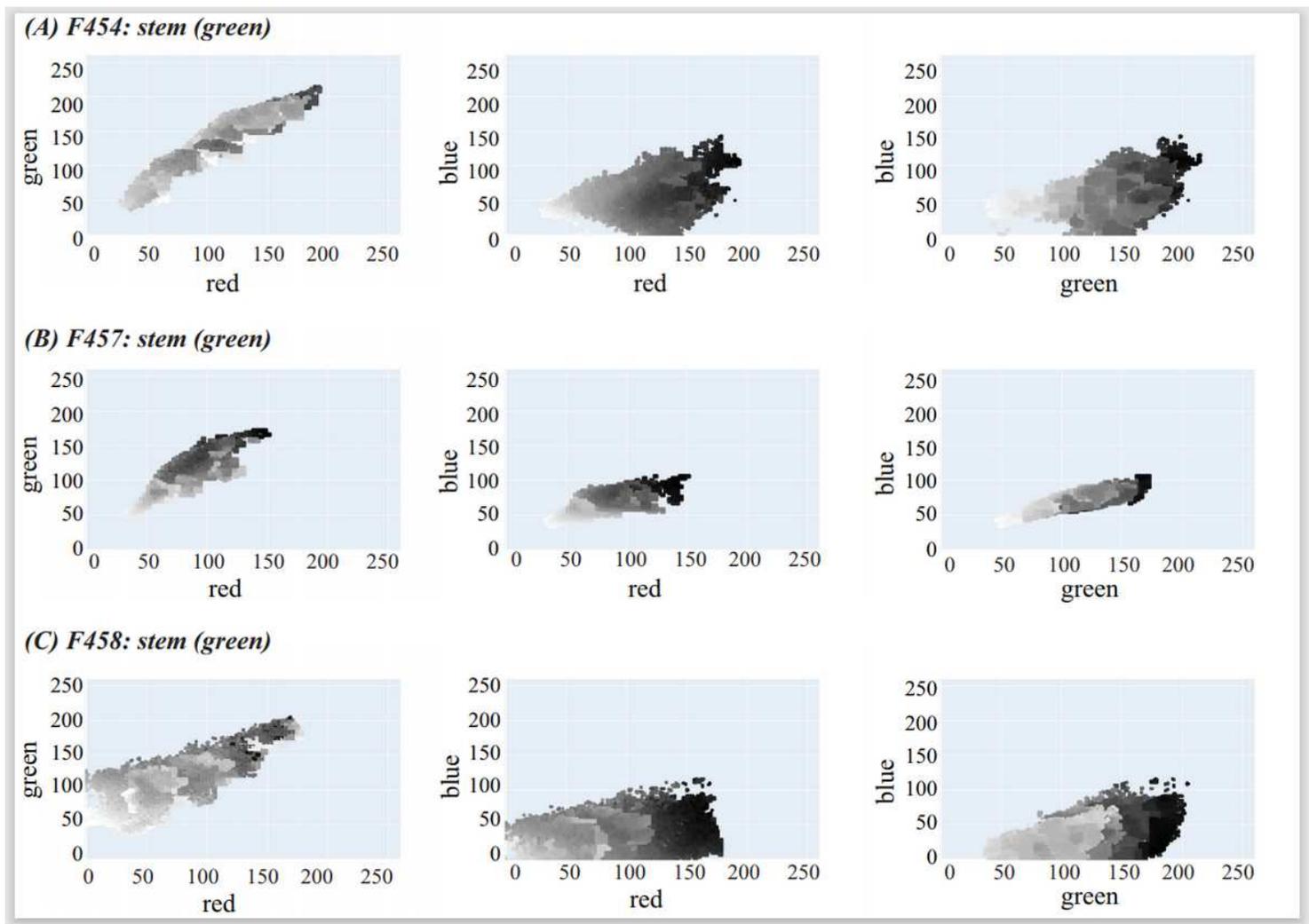


Figure 6

The RGB space occupied by the pixels in the green stem ROIs of the original Sunflowers in a vase with a yellow background, F454 (top row), compared with the two duplicates F457 (middle row), and F458 (bottom row) (A) red vs green; each dot represents one pixel; its color (grayscale) suggests the intensity of its blue coordinate, ranging from low (white) to high (black); (B) red vs blue; (C) green vs blue

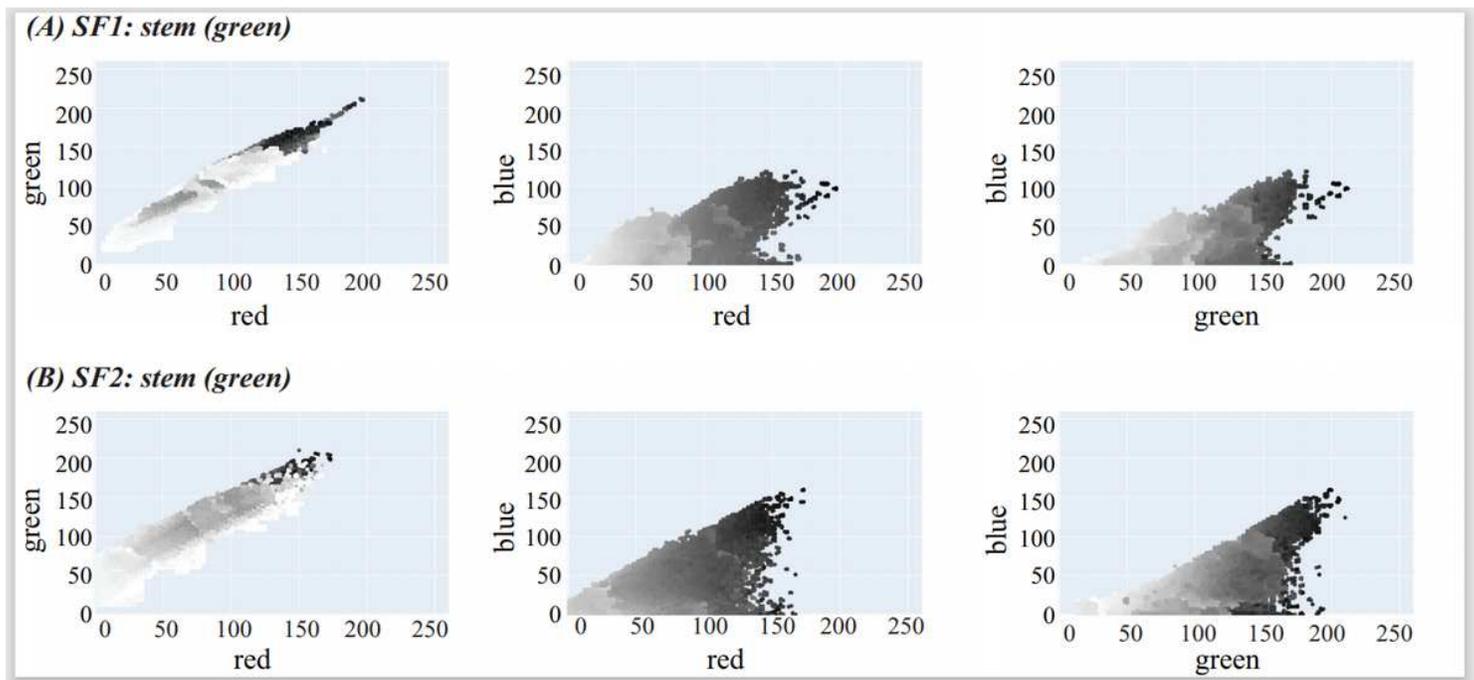


Figure 7

The RGB space occupied by the pixels of the green stem ROIs of the photographs SF1 (top row) and SF2 (bottom row) of real sunflowers. (A) red vs green; each dot represents one pixel; its color (grayscale) suggests the intensity of its blue coordinate, ranging from low (white) to high (black); (B) red vs blue; (C) green vs blue

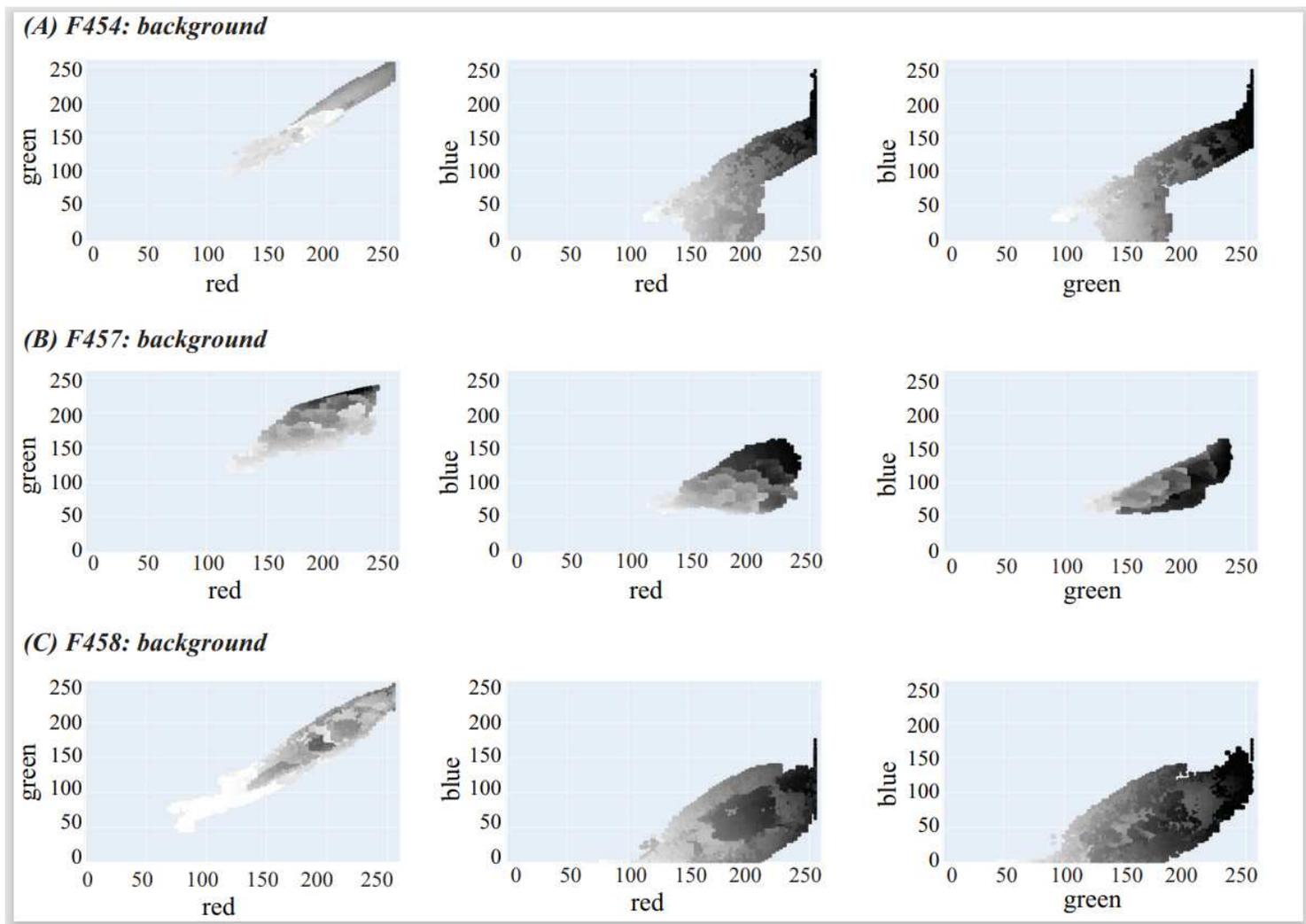


Figure 8

The RGB space occupied by the pixels in the background ROIs of the original Sunflowers in a vase with a yellow background, F454 (top row), compared with the two duplicates F457 (middle row), and F458 (bottom row). Note that the background includes the table, the vase, and the region behind the flowers. (A) red vs green; each dot represents one pixel; its color (grayscale) suggests the intensity of its blue coordinate, ranging from low (white) to high (black); (B) red vs blue; (C) green vs blue

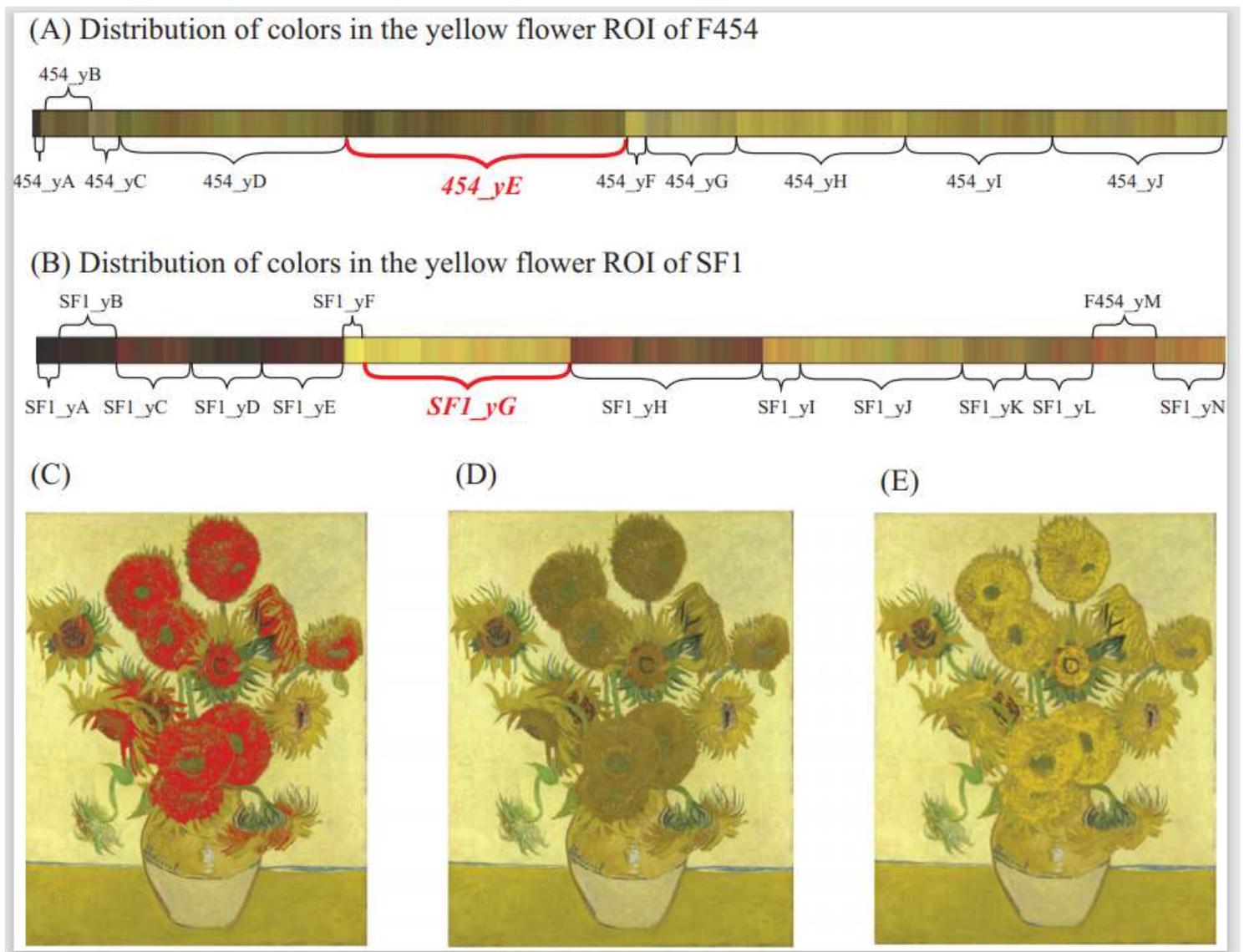


Figure 9

Color transfer from a real sunflower to the sunflowers in F454 (A) The color bar representing all hues of yellow in the ROI corresponding to the flowers in F454. This color corresponds to the leaves of the HC tree computed from the colors of all pixels in this ROI. We select the region (cluster) 454_yE in this color bar which contains some of the darkest yellow pixels. The pixels associated to this region are referred to as F454 A. (B) The color bar representing all hues of yellow in the flower ROI of SF1. We identify the cluster SF1_yG in this color bar that contains some of the brightest yellow. The pixels associated to this region are referred to as SF1 A (C) Pixels in F454 that belong to F454 A are highlighted in red (D) F454: before color transfer (E) F454: after color transfer.

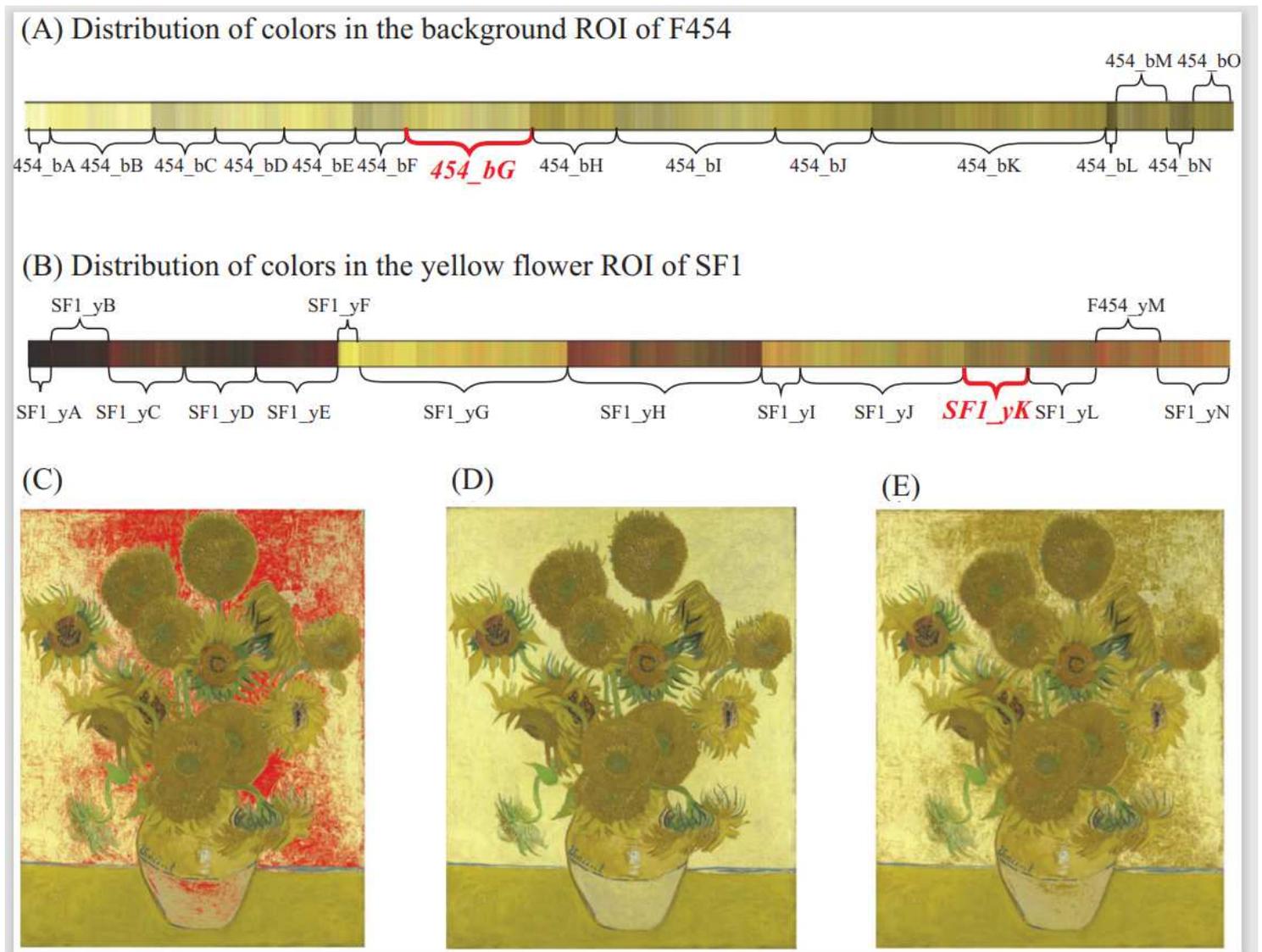


Figure 10

Color transfer from a real sunflower to the yellow part of the background in F454 (A) The color bar representing all hues of colors in the ROI corresponding to the background in F454. This color corresponds to the leaves of the HC tree computed from the colors of all pixels in this ROI. We select the region (cluster) 454_bG in this color bar which contains some of the pixels that are directly located behind the sunflowers in the vase. The pixels associated to this region are referred to as F454 B. (B) The color bar representing all hues of yellow in the flower ROI of SF1. We identify the cluster SF1_yK in this color bar that contains some of the darkest yellow. The pixels associated to this region are referred to as SF1 B (C) Pixels in F454 that belong to F454 B are highlighted in red (D) F454: before color transfer (E) F454: after color transfer.

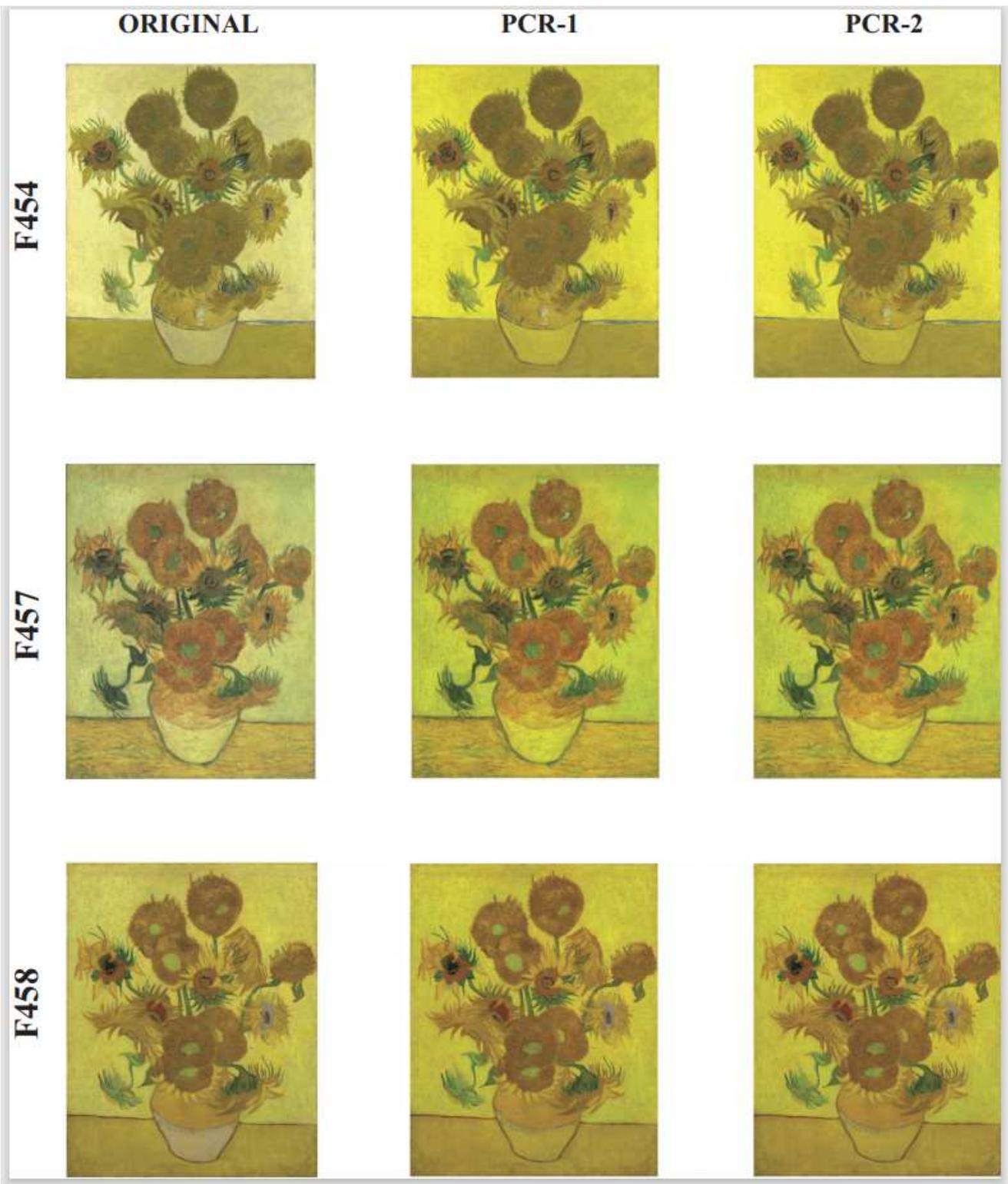


Figure 11

Color restoration of the Sunflowers in a vase with a yellow background For the original sunflower in a vase with yellow background, F454 currently in London (top row), and the two repetitions F457, currently in Tokyo, middle row, and F548, currently in Amsterdam, bottom row, we show a picture of the current painting (left column), as well as two possible models of the painting intended by van Gogh, using the color reconstruction methods PCR-1 (middle column) and PCR-2 (right column). See text for details.



(A)



(B)

Figure 12

Two images of natural live sunflowers (see text for details): (A) SF1; (B) SF2.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [vanSFF454.jpeg](#)
- [vanSFF458.jpeg](#)
- [vanSFF457re.tif](#)
- [SF1.jpg](#)
- [SF3.jpg](#)
- [SF1analysis.r](#)
- [colortransfer.r](#)
- [SF2analysis.r](#)
- [funccolorchangeupdate121220.r](#)
- [PCR.r](#)
- [F457sunflowerVGanalysis.r](#)
- [F458sunflowerVGanalysis.r](#)
- [F454sunflowerVGanalysis.r](#)