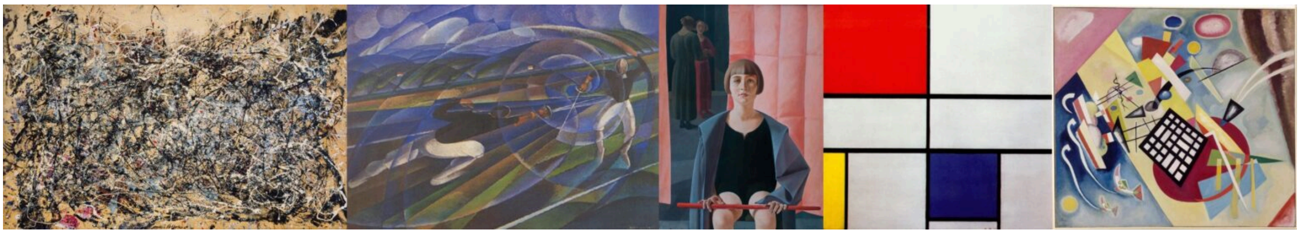
Centre de Sémiotique et Rhétorique

Un /carnet/ de « recherches »

Advances on the F.R.S.-FNRS research project "Towards a Genealogy of Visual Forms" : On edge directions and dynamicity in paintings

by Adrien Deliege



How to visualize and quantify the dynamicity of these paintings? *Human observers can "feel" that some of these images are more "dynamic" than others, and can sometimes "see" dominant directions emerging from these paintings. In this paper, we propose an edge detection-based method to compute tangible characteristics that model these feelings and allow further large-scale analyses.*

This post summarizes some works carried out on Maria Giulia Dondero's F.R.S.-FNRS research project "Towards a Genealogy of Visual Forms". More details will appear in a dedicated paper.

Background of the project. The project aims at analyzing, quantifying, clustering, characterizing the evolution of "forms" in the broad sense, that have been represented in large corpuses of paintings throughout centuries, up to modern art such as in fashion photography. For that purpose, the project builds upon recent computer vision techniques powered by deep learning-based models trained on generic large-scale datasets. The interaction between these two worlds, computer vision on one hand, art analysis on the other hand, benefits them both, by producing insights on the capabilities and limitations of the former in the artistic domain, and by resulting in novel ways of revisiting the latter.

Contributions. In this post, we present a method for computing and comparing the dynamicity in digitized paintings through edge detection. Besides, we demonstrate the applicability of our method in three main ways: 1) we show how some different artistic styles can be distinguished and how the production of an artist (examplified for Kandinsky) evolves through time, from a “dynamicity” perspective, 2) we show how to visualize hundreds of images and cluster them by dynamic similarity through the PixPlot software (as also done in our previous post here), 3) we retrieve similarly dynamic images in an image corpus with respect to a query image. By providing a quantitative and scalable approach to analyzing dynamicity in art, we aim to offer new insights into the study of visual culture in the growing field of digital humanities.

Method

Preprocessing. We first preprocess images with the 3 following steps.

a) Image rescaling. Digitized paintings are saved numerically as regular 3-channel RGB images with pixel values ranging from 0 to 255 and whose dimensions may vary from one digitization tool to another. The resolution/dimension of the images might thus unnecessarily influence our numerical analysis. Therefore, we first resize all the images while keeping their original aspect ratio, such that they roughly have the same number of pixels as a square 512 x 512 image.

b) Conversion to grayscale. Edge detection is typically performed on grayscale images because it simplifies the process, improves computational efficiency, strengthens edge detection accuracy, and reduces noise interference. Grayscale images have a single channel representing luminance intensity, making it easier to analyze and detect edges compared to color images with multiple channels. We thus convert our images to grayscale images.

c) Pixel values rescaling. Finally, the pixel values of the grayscale image are divided by 255 such that they range from 0 to 1, for the sake of faster processing and interpretation in our process.

Computing edges. For the edges computation itself, we proceed as follows.

a) Use of Sobel filters. We use Sobel filters to compute the edges in an image. For details, see here. First, the grayscale image is convolved with the Sobel filters to get the horizontal and vertical gradient images, from which the gradient magnitude and direction are then computed.

b) Magnitude normalizing factors. To better compare and aggregate edge magnitudes per direction, we rescale the computed magnitudes by the maximum magnitude achievable per direction. Hence, they belong to the [0,1] range. The rescaling factor as a function of the gradient direction can be derived mathematically. We leave these details for the extended version of this research.

c) Per pixel edge magnitude and direction. The edge magnitude for each pixel corresponds to the renormalized magnitude as described previously. The edge direction is perpendicular to the gradient direction. Besides, the gradient direction differs by 180° depending whether the contrast goes from dark to light or vice versa, which is a distinction that we do not need (we only focus on edges as such). Therefore, we limit the allowed edge direction to the range -90° to 90°, where directions at these two extreme values both depict verticality.

Visualization and dynamicity-related metrics. We can now visualize and compute a few metrics about the described processing of images.

a) Edge image. The most “natural” visualization that we can produce is an image of the edge magnitudes, colored as a function of the edge direction. That is, for each pixel, the edge direction determines a color to use, which we choose ranging from green (-90°) to blue (-45°) to white (0°) to red (45°) to green again (90°), downscaled by the edge magnitude, such that surfaces with no edges are blacked out.

b) alpha-main edge image. The edge image generally suffers from having many low-magnitude edges, that appear relatively dark and are hard to distinguish with the naked eye from completely edgeless surfaces. To solve that issue, we can remove the downscaling by the edge magnitude. However, this has the effect of revealing only the edge direction for all the pixels, thus producing an extremely chaotic image as even edgeless surfaces are colorized, usually with seemingly random colors in the defined palette. Following our experimentation, we found interesting to visualize only what we call the alpha-main edges, that is, we only color the pixels with the largest magnitude, by thresholding on the proportion alpha of the total magnitude they represent. We found that setting alpha to 50% offers a good compromise.

c) Circular histogram. We define bins of 5° ranging from -92.5° to 92.5° and aggregate the magnitudes of the pixels whose direction falls within each bin, then normalize the values by scaling them down by their sum, to produce a classic density histogram. The values obtained per bin are assigned to the angle at the center of the bin, and are interpolated linearly to produce a continuous visualization between those central angles, rather than rectangular bins. We then wrap the histogram around a half-circle, and symmetrize it for the sake of completeness. By doing so, we obtain a circular version of the histogram which indicates, for a given direction, how much magnitude was computed on the image, and this is represented pointing from the origin towards the direction in question on the histogram, making it a very intuitive and straightforward tool to understand the directions in the image.

d) Histogram radius. We call the histogram radius the maximum value computed previously to build the histogram, that is, the normalized maximum total magnitude accumulated in a 5°-degree bin. A large value indicates a concentrated distribution in a particular direction, while a small value is the sign of a more uniform distribution across the directions.

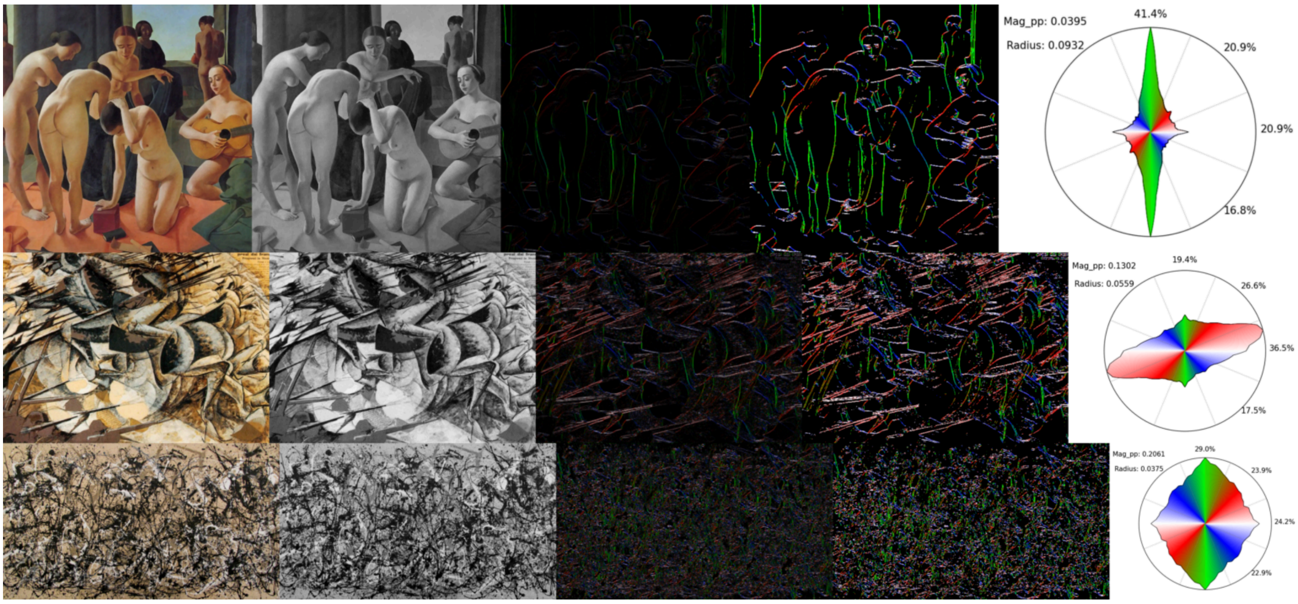
e) Proportion of verticality, horizontality, ascending and descending diagonality. For a highly compressed view of how much the main directions are present relatively to each other, we compute the proportion of the histogram contained between -67.5° and -22.5° to assess descending diagonality, -22.5° and 22.5° to assess horizontality, 22.5° and 67.5° to assess ascending diagonality, the remaining proportion assessing verticality. While the shape of the histogram carries more information, this simplified metric offers a compromise between an easy quantification and interpretation of the edges of the painting.

f) Average magnitude per pixel. As additional metric, we compute the average magnitude per pixel as the sum of all the amplitudes, divided by the number of pixels, regardless of the directions. This metric helps quantifying how many edges are present within the image, on a global level. Small values indicate that few edges are represented, while large values indicate many edges.

g) Distance between images. To compare the dynamicity induced by the edges and directions of two paintings, we compute the L1-distance (i.e. Manhattan distance) between their normalized histograms (where their total area equals 1). This corresponds to the space between the two histograms if they are represented on top of each other.

Results

a) Examples of results. Some examples of results obtained following the process described the previous section are represented hereafter. The difference between the edge image and the alpha-main edge image can be clearly seen, the latter being used only for visualization purposes. For the first painting, we can observe that most salient edges are relatively vertical given the presence of many people standing, which is correctly translated into the circular histogram, which reports over 40% of its mass in the vertical direction. For the second painting, we can see that the dominant direction is ascending diagonally, which is depicted rightfully with the red edges and the tilted shape of the circular histogram, with most of its mass pointing in the direction of roughly 20°. For the last painting, the extremely complex and dense patterns yield a much higher average magnitude per pixel (0.21 vs 0.13 and 0.04), and these patterns spread across all directions relatively uniformly, which is represented on a more roundish histogram and roughly equivalent proportions in the four main directions (23% to 29%), with a comparatively smaller radius than the other artworks (0.04 vs 0.06 and 0.09).



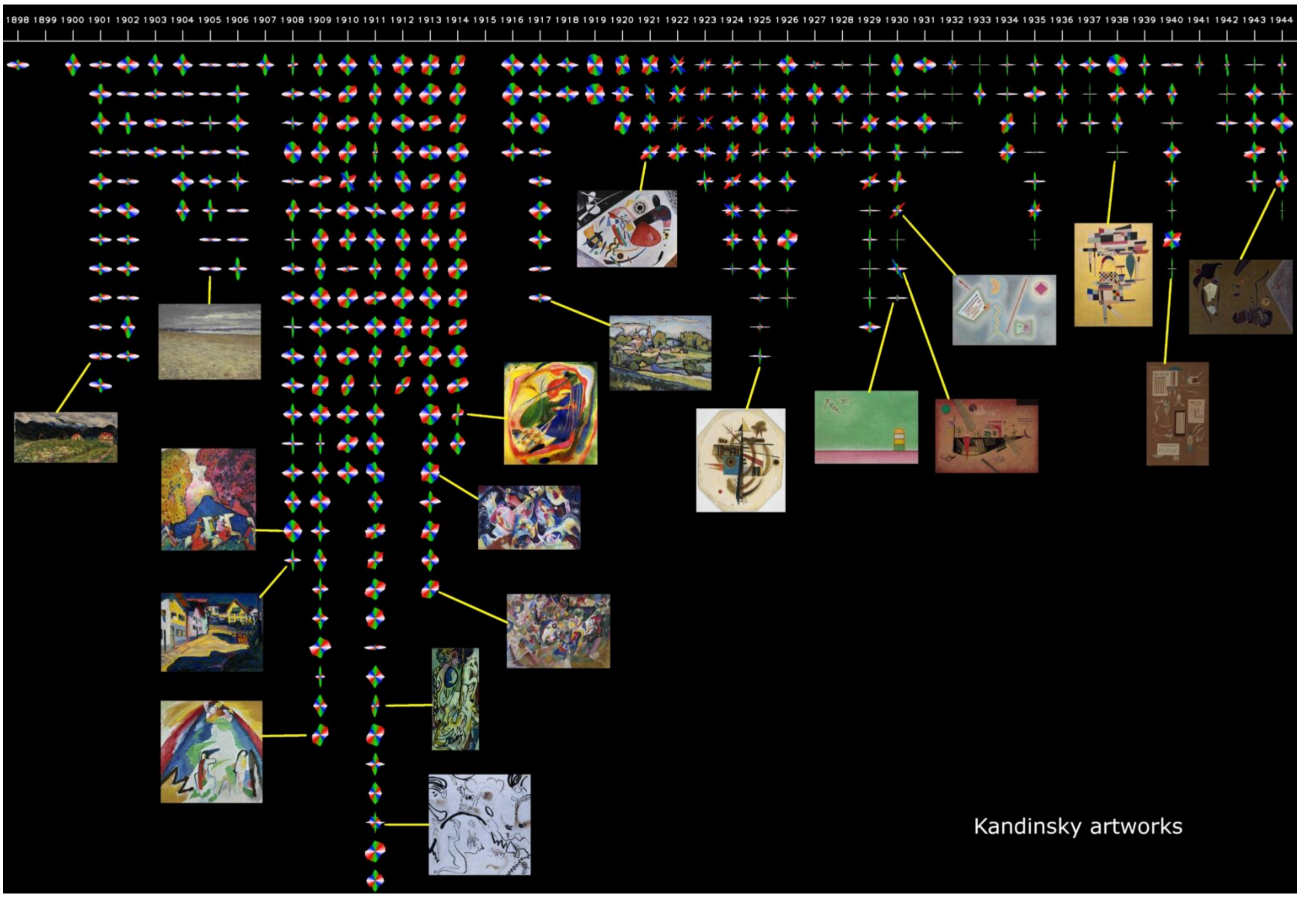
***Examples of results** of the edge detection process and visual representation with a circular histogram. First line: Felice Casorati, Concerto, 1924. Second line: Umberto Boccioni, Charge of the Lancers, 1915. Third line: Jackson Pollock, Autumn Rhythm (Number 30), 1950. First column: original image. Second column: the grayscale image used for the edge detection. Third column: the edge image, where the color represents the direction of the edge, and the brightness of the color represents its magnitude. Fourth column: alpha-main edge image (for visualizaton only), with alpha=50%, showing only the most salient edges with no downscaling on the magnitudes. Fifth column: the circular histogram aggregating the edge information and a few metrics, computed from the edge image. Mag_pp is the average magnitude per pixel.*

b) Artistic movement vs average magnitude per pixel. The three results shown above can be seen as belonging to three different artistic movements: the *calm period* for Casorati’s painting, the *futurism* for Boccioni’s, and the *abstract expressionism* through the *dripping* technique for Pollock’s work. Without being an exhaustive list of pictural movements, these intuitively drastically differ by the “amount of motion” or “dynamicity” they convey, on a global level. To further illustrate this, and show a first basic aplication of our methodology, we can represent a few paintings sorted by average magnitude per pixel, which is the metric that best captures that observation. This is represented hereafter, where a few paintings of Casorati are represented, along with a few paintings of the futurism period. Pollock’s artworks are not represented because their average magnitude per pixel are so large (above 0.2) that they would be on the far right of the Figure, forcing a heavy zoom out and making it impossible to distinguish any painting. On that matter, his artworks clearly stand apart. We can also observe that most of the paintings from the futurism (in magenta rectangles) have larger average magnitudes than most of Casorati’s artworks (in yellow rectangles), as intended by the movement itself that wants to encapsulate dynamicity in still representations. One notable exception is Casorati’s *Donna e paesaggio* (1940), which has a value within the range of most paintings of the futurism. This is because that particular painting was produced with a lot of small brushes, patterns, as they can be found in the fauvism or expressionism movements, contrary to the other, much smoother, works of Casorati represented in the Figure. Overall, this visualization accounts only for one variable, which is of course insufficient to fully characterize and distinguish major artistic movements or painters, but it shows that it can already quantify and validate some observations that could otherwise only be qualitatively assessed.

***Calm vs futurism comparison** in terms of average magnitude per pixel. Casorati's paintings (yellow frames) are considered as "calm" while images from the futurism (magenta frames) are considered "dynamic". This translates to lower (respectively larger) average magnitude per pixel. Pollock's artworks would stand far on the right, with values above 0.2. Artworks from the futurism generally display more dynamicity, represented by more edges and thus a larger average magnitude per pixel.*

c) Studying the evolution of an artist: the case of Kandinsky. Another way to use our method is to represent the evolution of the dynamicity of an artist's paintings over time. For that purpose, we choose to study the most influential paintings of Wassily Kandinsky, whose famous, prolific and diversified work constitutes an ideal testbed for us. We collect the images referenced as most emblematic paintings of Kandinsky on Wikipedia, yielding a corpus of 327 images. We notice that some of the images not only show the painting itself, but also (a portion of) the frame around it. Obviously, analyzing the images with the frame gives clear horizontal and vertical edges at the intersection of the frame and the painting, which artificially boosts these directions and tends to produce more "+-shaped" circular histograms. Therefore, we manually inspect all the images and crop them to keep only the paintings. As a result, 89 images are cropped. Furthermore, to remove similar side effects that were produced by the digitization system or leftovers of the frames that we might have missed, after computing the edge magnitude for each image, we zero out the magnitudes of the 1% leftmost, rightmost, uppermost and lowermost pixels.

For the sake of clarity, we only represent the circular histograms along with a few artworks, sorted by date of production, hereafter. In this visualization, several periods of Kandinsky's career can be distinguished through these circular histograms. We can see that, in his early career until 1906, most of the histograms have a dominant horizontal component and usually not much verticality, as depicted in *Tunis, Coastal Landscape I* (1905). This is a sign of low dynamicity and gives a sense of stability in the paintings. Then, from 1908 to 1914, we can observe that most histograms are much more roundish and uniform, such as in *Blue Mountain* (1908). As for Pollock's art, this indicates a fuzzy dynamics, where each direction is represented equally and complex patterns are interleaved. We can also see that some paintings start to have a slightly dominant direction in the ascending diagonal, like *Draft 3 for Composition VII* (1913), which drives the eye in a slighty more dynamic way. This corresponds to Kandinsky's Blue Rider period. We then observe, from 1916 to 1920, a small mix of the previous types of histograms already observed. This corresponds to a period where Kandinsky was in Russia and was teaching more than he was painting. From 1921 to 1925, we can observe many "x-shaped" histograms, where diagonal directions (both ascending and descending) are dominant and are characteristic of a somewhat more dynamic artwork. This particular pattern reflects Kandinsky's increasing inclination for lines and geometric forms, such as in *Small worlds I* (1922) or *Circles in a Circle* (1923). After 1926, we observe a switch to mostly "+-shaped" histograms, indicating primarily horizontal and vertical edges. This is often the sign of a retrieved calmness and stability, sometimes to the extremes, such as in *Green Void* (1930). During that period, we sporadically note a few x-shaped histograms, uniform, or horizontally dominant ones as observed previously.



***Temporal evolution of Kandinsky** through the circular histograms of 327 of his most prominent paintings. We can observe various shifts in the overall shape of the histograms at different times, reflecting Kandinsky’s various artistic periods.*

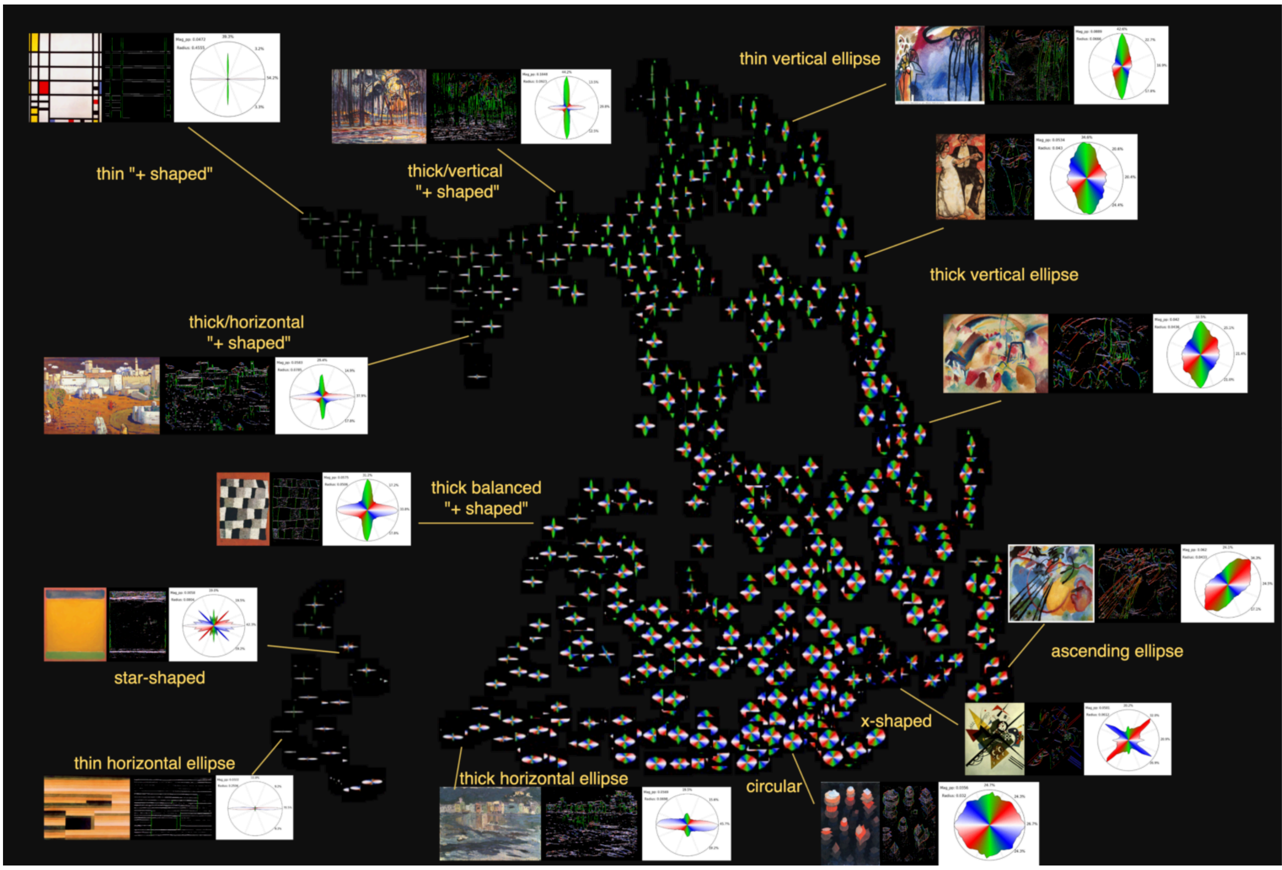
From a quantitative point of view, there is an average radius difference before 1925 and after (0.0487 vs 0.0851), consistent with the uniform vs +-shaped histograms observation. The average magnitude per pixel is larger before 1925 (0.0668 vs 0.0482). These comparisons are validated as highly statistically significant by two-sample t-tests. We did not find any strong correlation between the radius and the average magnitude per pixel, i.e. for a given radius range, the associated magnitudes per pixel may usually vary from low to large values.

While in the case of Kandinsky, we can match known periods of his life to the dynamicity (seen as the shape of the circular histograms) of his paintings, this kind of visualization might help discovering previously unknown transition phases of many other (possibly less well-studied) artists.

d) A typology of histogram shapes and visual clustering of abstract paintings. The previous section already reveals various types of histograms that can be obtained when studying a large corpus of paintings. In this section, we push the analysis further, by considering over 1000 abstract paintings and studying the distribution of the circular histograms obtained. By doing so, we establish a typology of such histograms and we gain some insight on the similarities between paintings of different artists, at different times.

We consider 6 of the most influential abstract painters of the 20th century and download their paintings from WikiArt: Mark Rothko (163), Jackson Pollock (88), Kazimir Malevich (280), Paul Klee (193), Piet Mondrian (94), and Wassily Kandinsky (228), totalling 1046 artworks. In the case of Kandinsky, we do not reuse the Wikipedia images, in order to analyze paintings from a same data source. To limit the frame effect mentioned in the previous section, and to avoid checking and cropping manually over a thousand images, we automatically zero out the magnitudes of the 5% left/right/upper/lowermost pixels. This does not prevent the frame effect when the frame is very large, but still removes frame effects in many images where the frame is relatively thin, or when the digitization system produced artefacts at the borders of the paintings.

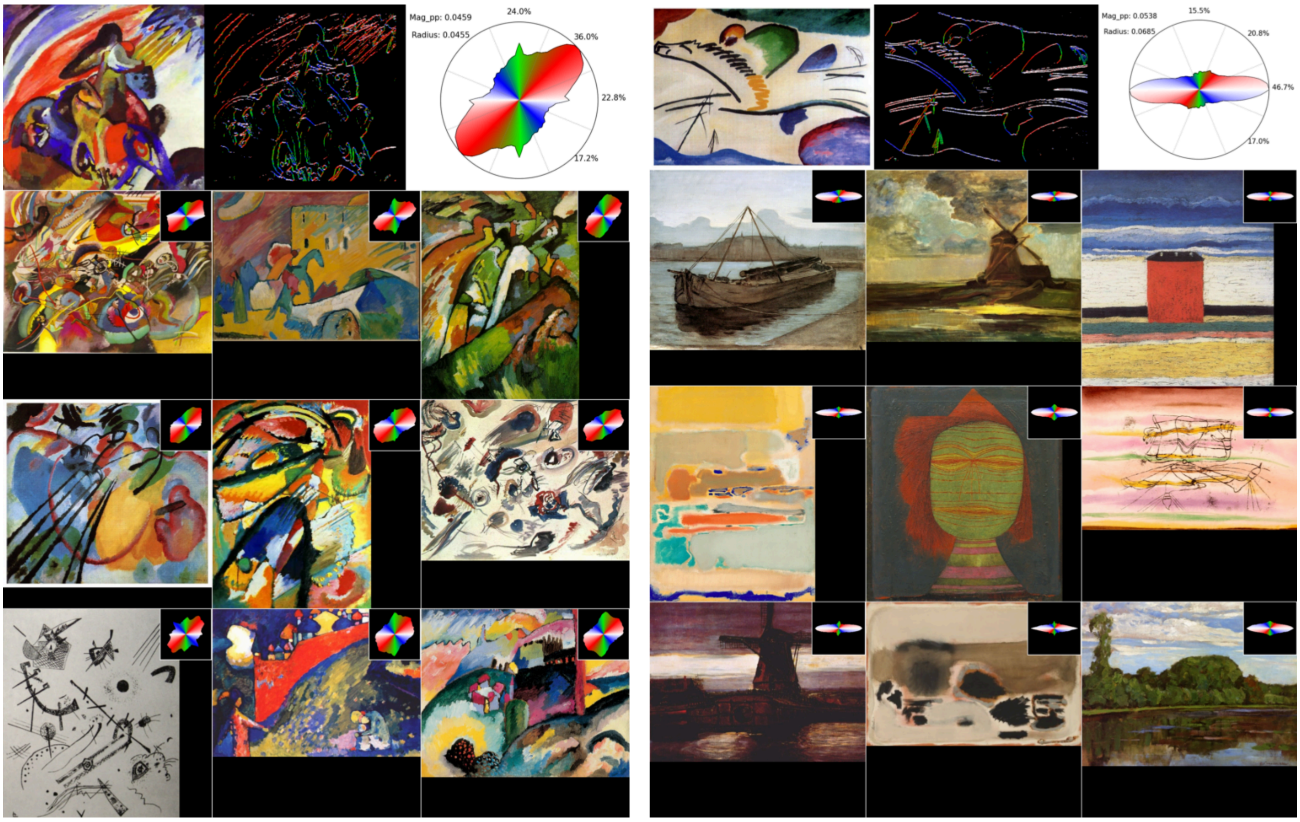
We compute the circular histograms of all the images, which can be assimilated to 36-dimensional vectors. We then project the collection of these vectors in a 2D plane with UMAP, by specifying the L1-distance, as explained in the Method section, as proximity measure to use in order to preserve as much as possible the local and global structure of the original space. Finally, the histogram images are represented where their projected vectors land in the 2D plane. We perform these operations through the PixPlot software, which then allows a dynamic WebGL visualization of the results. We provide access to this visualization at http://bit.ly/4etv4Tm (or here, or even here for the original images instead of the histograms). An overview of this visualization is shown hereafter. We observe various typical shapes in the histograms, usually characterized by different types of paintings. These shapes might have a dominant direction, e.g. vertical, horizontal, diagonal, yielding ellispoid-like histograms, or a pair of dominant direction, e.g. vertical and horizontal, or ascending and descending diagonals, yielding "+"-shaped or "x"-shaped histograms, all of which might be more or less thin or thick, depending on the prominence of the non-dominant direction. When all the directions are distributed uniformly in the painting, the associated histogram has a circular shape, as no dominant direction stretches it.



Typology and visual clustering of histograms of abstract paintings. *We project the histograms in the 2D plane and visualize them through PixPlot. We observe several types of histograms in this dataset, usually distributed across the different painters (Kandinsky, Klee, Rothko, Malevich, Pollock, Mondrian).*

Using the metadata (artist name, date) of the paintings in the WebGL application, we can further analyze the association between these shapes, the artists, and the period of time considered. For instance, Mondrian produced paintings with sharp “+”-shaped histograms, especially after 1918, while he produced more thin or thick horizontal ellipses beforehand. Many star-shaped histograms belong to Rothko from artworks after 1955, presumably due to subtle variations in the painting that we cannot distinguish with the naked eye on the digitized images, produced e.g. by the technique used or the type of stroke performed, that are captured by our process and that are reflected in the analysis. He transited from a more horizontal ellipse period and before that from a more vertical or balanced “+”-shaped period to gain his distinctive style. Kandinsky produced almost all types of histograms, some more prominently in some time periods than others. Pollock’s art is generally composed of many homogeneous directions, yielding mostly circular-shaped histograms. Malevich and Klee have a more uniform production, in the sense that they produced also most of the histogram types, but not necessarily at different periods of time, which can be interpreted as a sign of highly polyvalent artists, already starting from their early days.

e) Retrieval of most similar images in terms of histogram shape. Let us note that the PixPlot visualization is just that: a visualization. This means that we have to keep in mind that this represents a 2D projection of a 36-dimension set of vectors. Therefore, there is necessarily a loss of information due to the information compaction, which implies that some images that appear close to each other in the 2D visualization might not be that close in the original space. Therefore, if one wants, for instance, to find the top *K* closests image to a query image in terms of circular histogram and L1-distance, then the search should be conducted with the vectors rather than with their projection. This is what we have done, and we show some results hereafter. In that case, we show, on one hand "dynamic" images, with a main diagonally ascending direction, and on the other hand, more "calm" images, with horizontal ellipsoids as circular histograms. This method thus allows to find similarities in paintings that are relatively difficult to observe with the naked eye or to quantify precisely, and opens the door to further analyses across styles, painters, and time periods.



Top 9 most similar images to a query image. *Our process allows to find similarities in paintings in terms of "dynamicity", understood as the main directions within the paintings. Left: the query image is Kandinsky's Improvisation 12 (Rider) (1910). Right: the query image is Kandinsky's Lyrical (1911). The retrieval corpus is the same as in the previous PixPlot visualization (6 main abstract painters).*

Limitations

We briefly discuss two limitations of the present study, that would be suitable for follow-up works and improvements of our methods.

a) Technical limitation: Frames around the paintings. In the results section, we either manually cropped the images to remove the frames around the paintings, or we automatically removed a fraction of the computed quantities at the border of the image. However, the first case cannot scale to large databases, and the latter does not ensure that the frame is fully removed, or may remove too large parts of the painting. When the frame persists, it mainly influences the vertical and horizontal amount of edges within the image, which artificially pushes it towards the category of “+”-shaped histograms, which flaws the downstream analysis. Besides, in some cases, the artist paints on the frame, which thus becomes parts of the artwork itself. Should it still be removed then? It may also happen that the frame produces some shadow or border effect on the digitized image, should it be removed too? Finally, when the painter does not fill in all the space between the edges of the painting and the edge of the canvas, should we consider that the edges created at the intersection are relevant?

b) Interpretative limitation: Dynamicity reduced to histograms. While diagonal lines are often associated with a dynamic content, quantifying the dynamicity of a painting through the sole examination of the histogram and/or the average magnitude per pixel might be an oversimplification of what dynamicity is. Indeed, Pollock’s artworks contain a lot of intrinsic dynamicity, while having circular histograms. A slightly tilted version of Boccioni’s *Charge of the Lancers* would have a horizontal ellipsoid as histogram, which might falsely be associated with a static content. Conversely, some of Kandinsky’s x-shaped paintings with several dominant diagonals might be considered rather static. Let us note that our analysis does not take into account the spatial distribution of the edges within the image, which might also induce different perceptions of dynamicity. Therefore, while we believe our tools provide interesting insights on the potential dynamicity of paintings, it is important to keep in mind that edges, aggregated globally as done here, constitute only one of the modalities that can be used to characterize this complex notion.

Conclusion

Quantifying the dynamicity that can be sensed in paintings is not an easy task. In this paper, we provide one way to gain insights on this problem, by computing the edges within the images and studying their distribution. By doing so, and with the help of a few other metrics, we can show some differences between artistic styles, we can study the temporal evolution of an artist across different periods of his artwork production, and we can define a typology of histogram shapes that traverse various artists, allowing the retrieval of similarly-composed paintings, in terms of distribution of edges. We further discuss two limitations of the present work. Our method and set of tools could certainly be improved, extended or reused in many ways, which is why we are committed to release all code material, which will hopefully serve the digital humanities community.

OpenEdition vous propose de citer ce billet de la manière suivante :
AdrienDeliege (5 juillet 2024). Advances on the F.R.S.-FNRS research project “Towards a Genealogy of Visual Forms” : On edge directions and dynamicity in paintings. *Centre de Sémiotique et Rhétorique*. Consulté le 22 août 2024 à l’adresse https://doi.org/10.58079/11yav

AdrienDeliege / 05/07/2024