On the dynamicity in images through character poses and edges detection

Adrien Deliege

Colloque « Interroger le visible, images qui se répondent. Analyse outillée, IA assistée »
ENS Lyon, 21 juin 2024

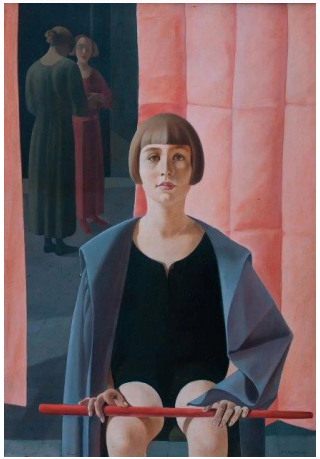
On the dynamicity in images through character poses and edges detection

Adrien Deliege

Colloque « Interroger le visible, images qui se répondent. Analyse outillée, IA assistée »
ENS Lyon, 21 juin 2024

Hypothesis

The directionality of the main "edges" ("traits", "strokes", "lines") in an image conveys a sense of dynamicity, "motion".



Felice Casorati,
Renato Gualino

Mainly horizontal and vertical edges

Static image

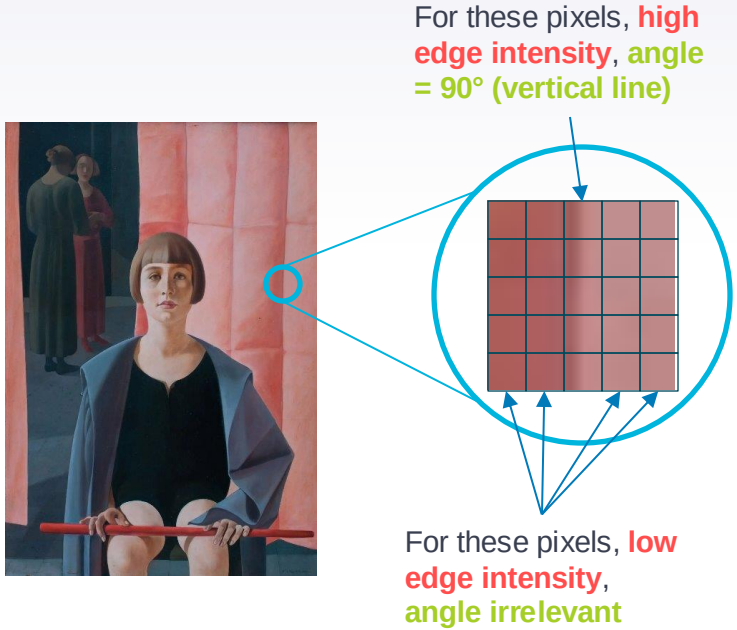
Alessandro Bruschetti,
A fondo

Mainly diagonal edges

Dynamic image

How can we measure, quantify, these observations ?

Compute edges



Note : this is just for explanation purposes. One pixel contains only one color of course, but you get the spirit of the idea.

For each pixel of the image, compute 2 things :

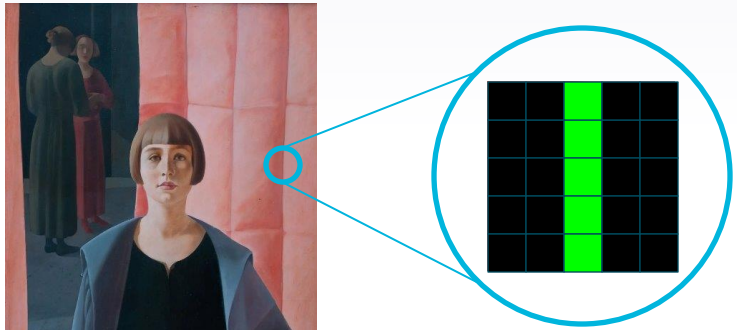
- the **angle** of the « main edge » that passes through the pixel
- the **intensity** of that « main edge »

This is done for a pixel by :

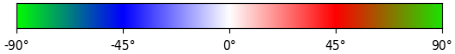
- Converting the image to grayscale
- Looking at a 3x3 tile around the pixel of interest
- Applying horizontal and vertical **Sobel filters** on the tile. This gives two values, Sx, Sy.
- Computing the angle and the intensity of the gradient vector S = (Sx, Sy).

(I do a couple of other normalization/processing steps, but these are irrelevant details for this slide)

Compute edges



If edge intensity is too low, I black out the pixels.
Else, I color the pixels depending on the angle computed. Legend :



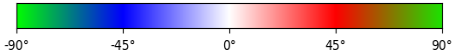
Measured with respect to a horizontal arrow pointing to the right

For each pixel of the image, compute 2 things :
- the **angle** of the « main edge » that passes through the pixel
- the **intensity** of that « main edge »

This is done for a pixel by :
- Converting the image to grayscale
- Looking at a 3x3 tile around the pixel of interest
- Applying horizontal and vertical **Sobel filters** on the tile. This gives two values, Sx, Sy.
- Computing the angle and the intensity of the gradient vector S = (Sx, Sy).

(I do a couple of other normalization/processing steps, but these are irrelevant details for this slide)

Compute edges



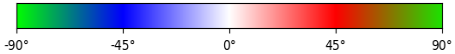
What does « too low » mean, for an edge intensity?

In this case, I proceeded as follows:
- Each pixel of the image has an edge intensity
- Let's say that the sum of all these intensities is *T*
- I order the list of computed edge intensities
- I black out the pixels with the smallest intensities, until the sum of their intensities reaches 50% of *T*.

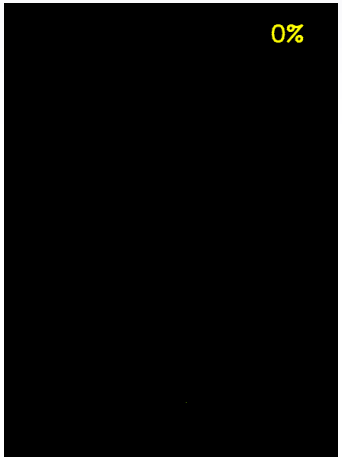
Thus, I keep only the pixels with the largest intensities such that their combined edge intensities amount for 50% of the sum of all the edge intensities of the image.

This process is a **visualization** process only. For further computations, I still use the full information, I do not delete anything. This is just for a better visualization.

Compute edges

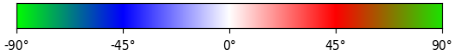


What does « too low » mean, for an edge intensity?



Keeping various fractions of total edge intensities

Compute edges

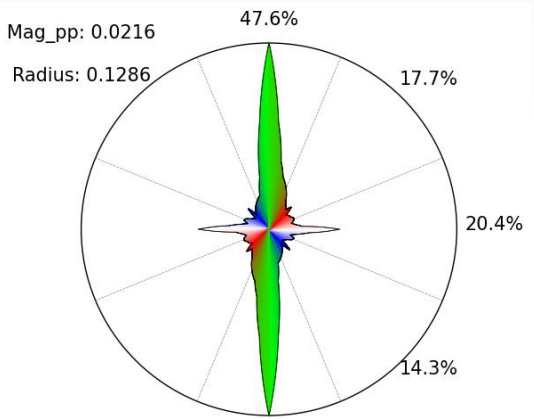


If we just light up pixels proportionnally to the edge intensity

Very dark.
Not useful for visualization.
But useful for computations.

Compute edges and histogram

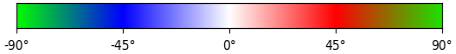
Let’s aggregate the intensities, binned by directions.



Circular normalized histogram

In this case, mostly vertical edges, and a bit of horizontal edges.

« Mag_pp » : magnitude/intensity per pixel. Large values indicate that many pixels have a large edge intensity. Comparing this value for different images indicate, at a general level, which image has more edge intensity (regardless of image size).



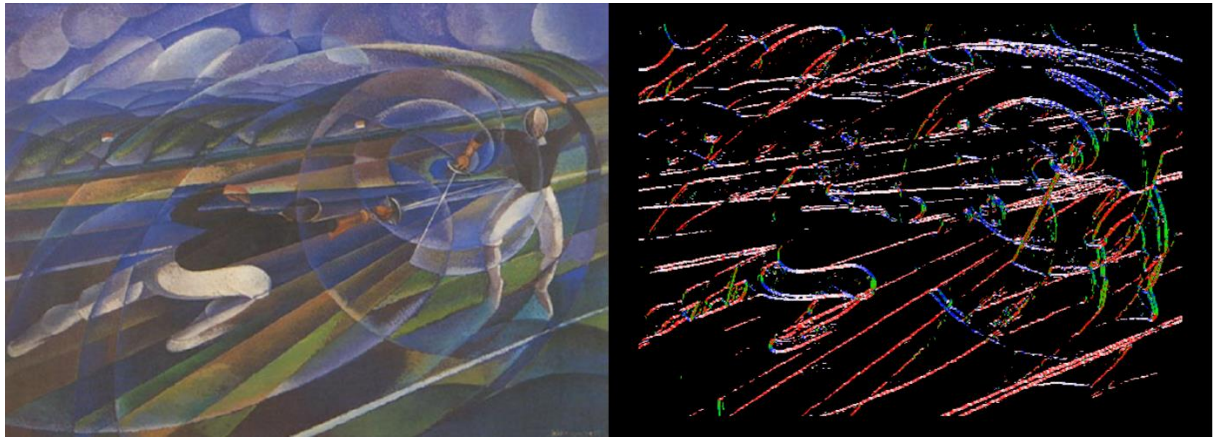
Normalized : total area (taking radius into account for scale) equals 1.

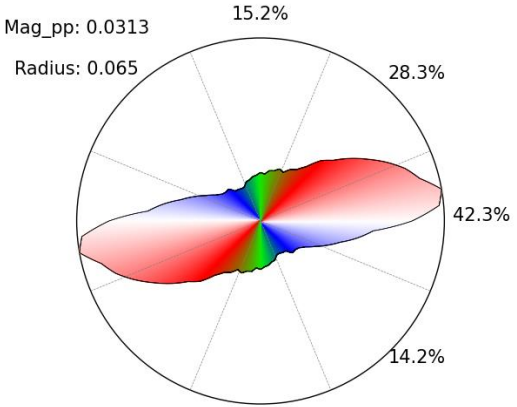
Compute edges and histogram

Another example.

In this case, the directionality is much more diagonal (ascending).

Mag_pp is also larger, indicating on general more edge intensities per pixel.



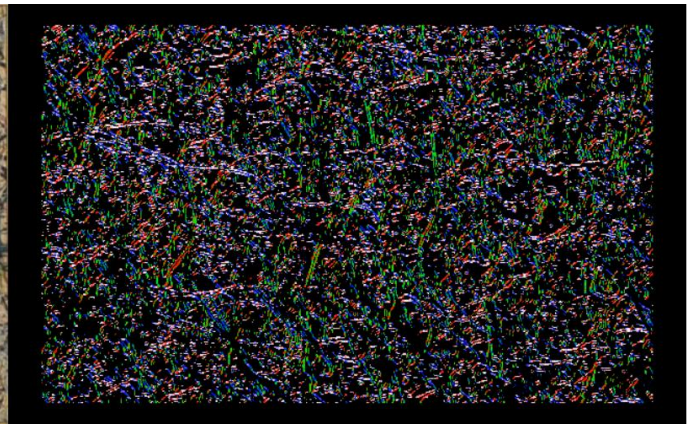


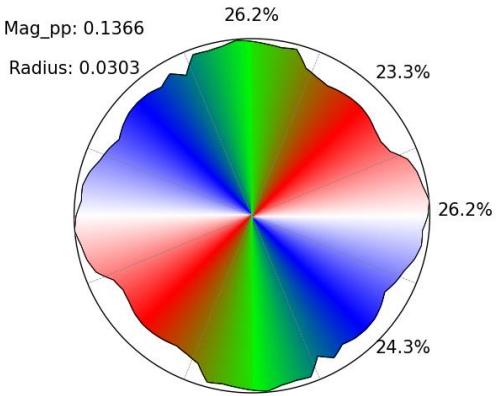
Compare images from histograms

Another example (Jackson Pollock).

In this case, the directionality is much more chaotic, there is no dominant direction, the distribution is almost uniform (as also indicated by the percentages in the four quadrants, all close to 25%).

Mag_pp is also much much larger, indicating on general a lot of edge intensities per pixel.





Compare images from Mag_pp

Futurist images generally have a higher average intensity per pixel than calm images. Which makes sense, but it's good to see it confirmed here.

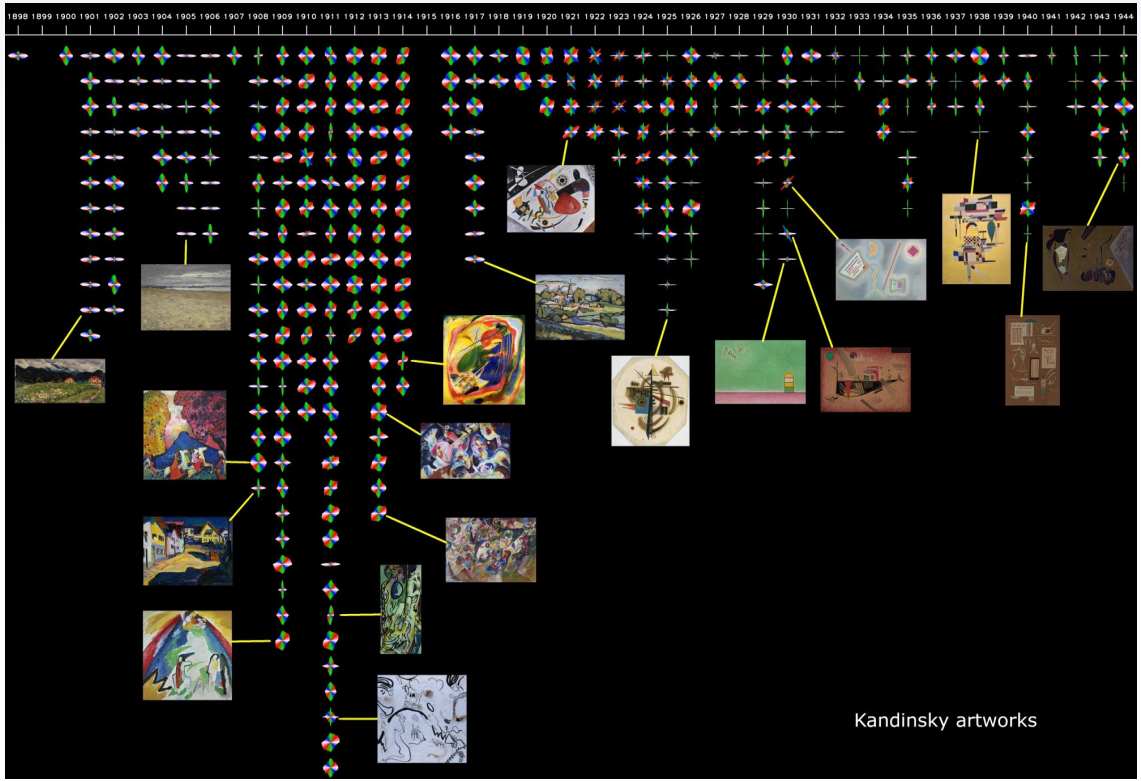
Temporal evolution of an artist

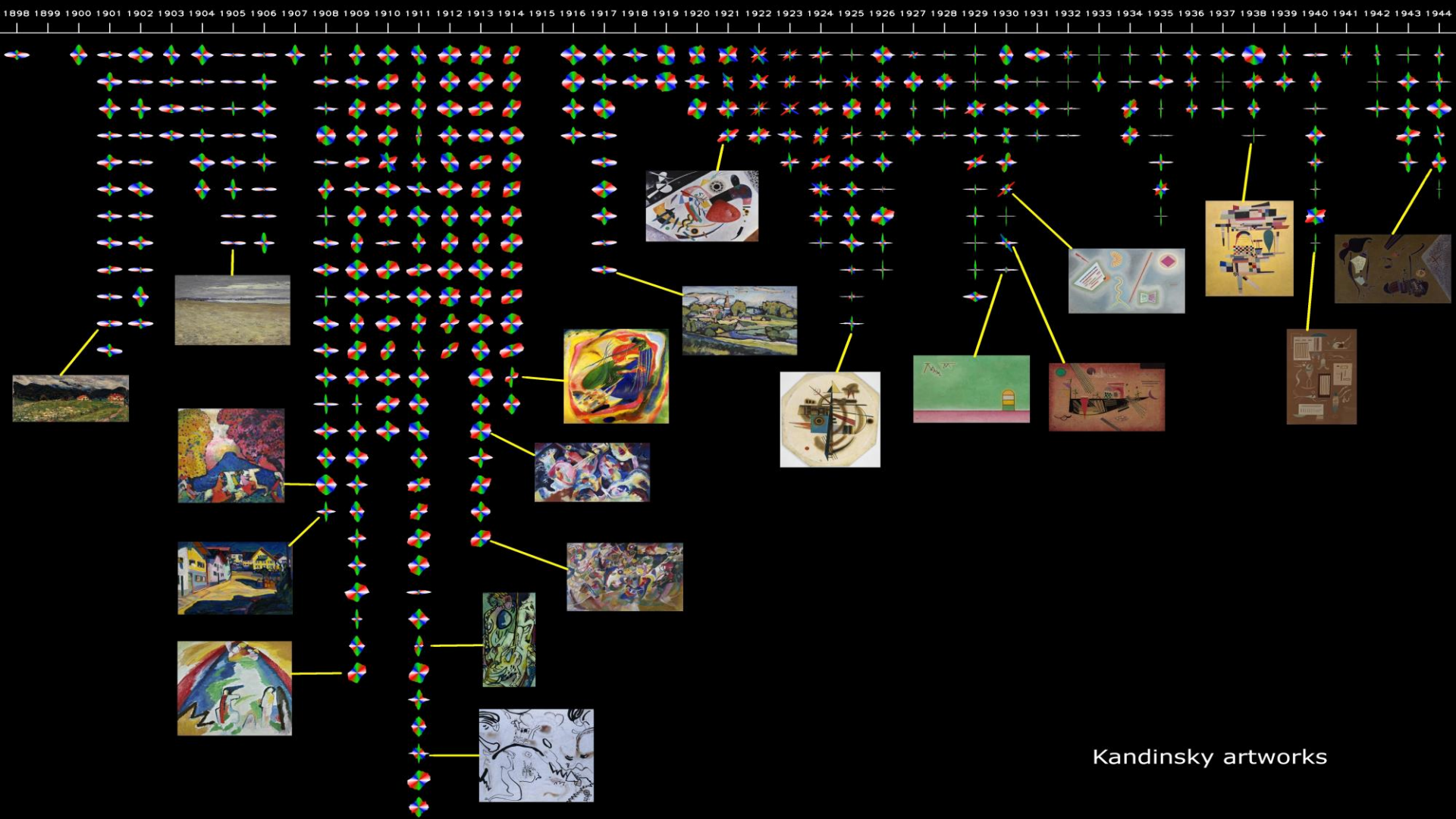
Kandinsky artworks

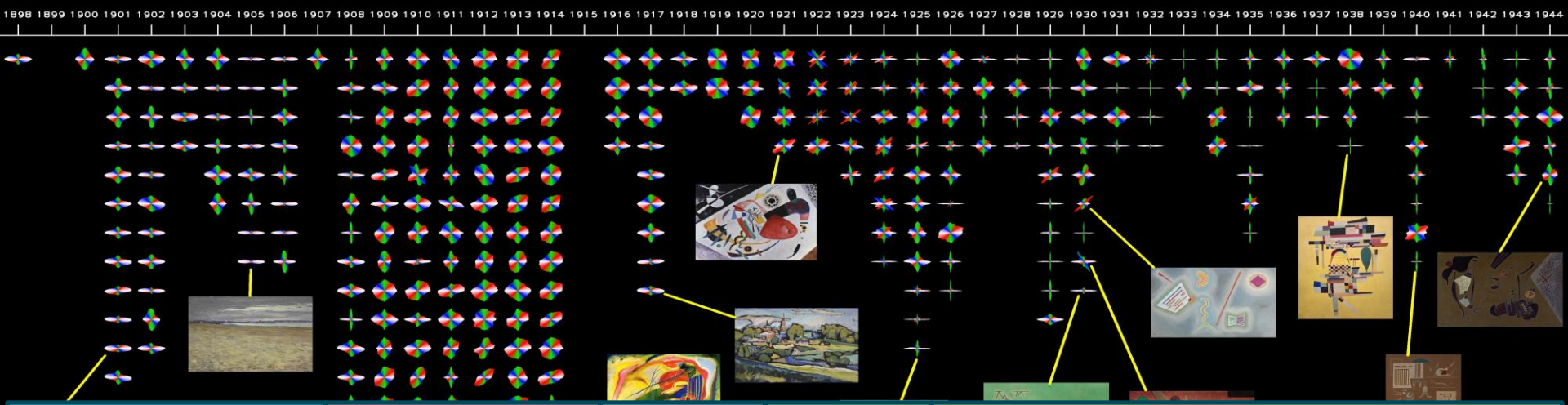
327 images (Wikipedia)

Sorted temporally

Various shifts ocurred

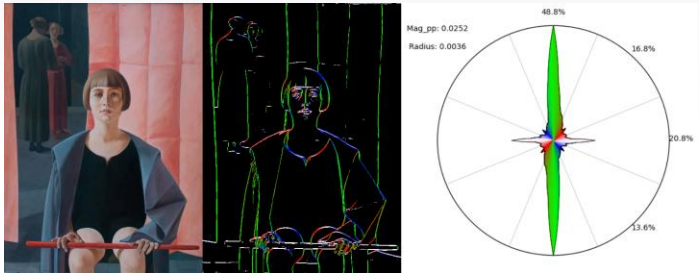


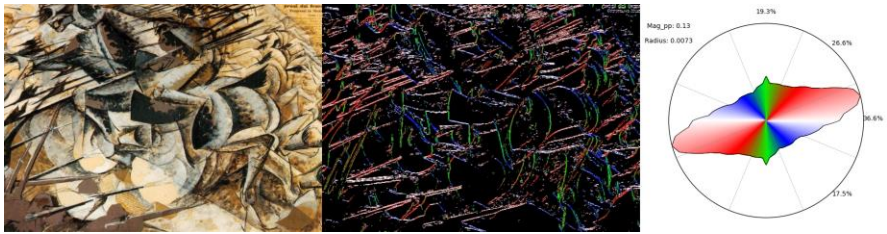

1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944
Kandinsky artworks

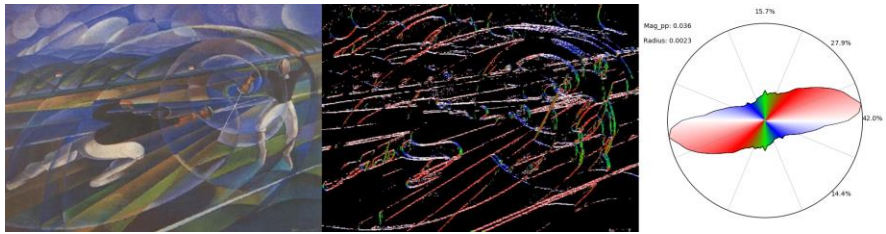


Mostly horizontal histograms	Mostly circular or diagonal ascending histograms	Mix of horizontal and circular	X-shaped	Mostly +shaped

Compare images from histograms

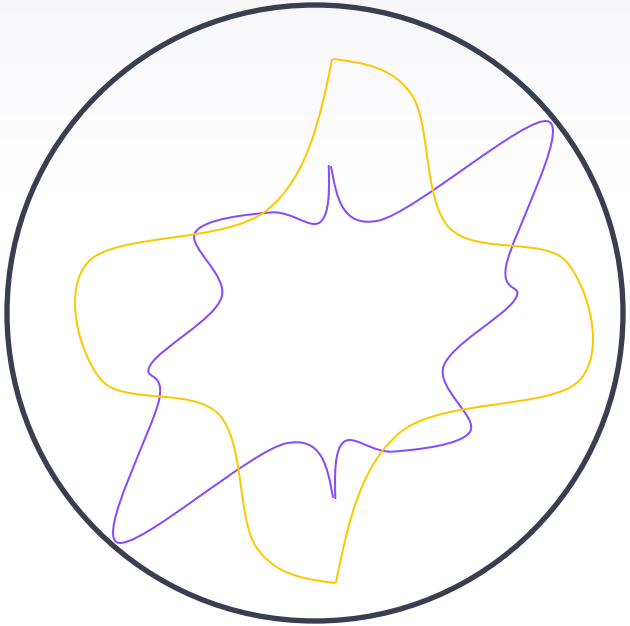




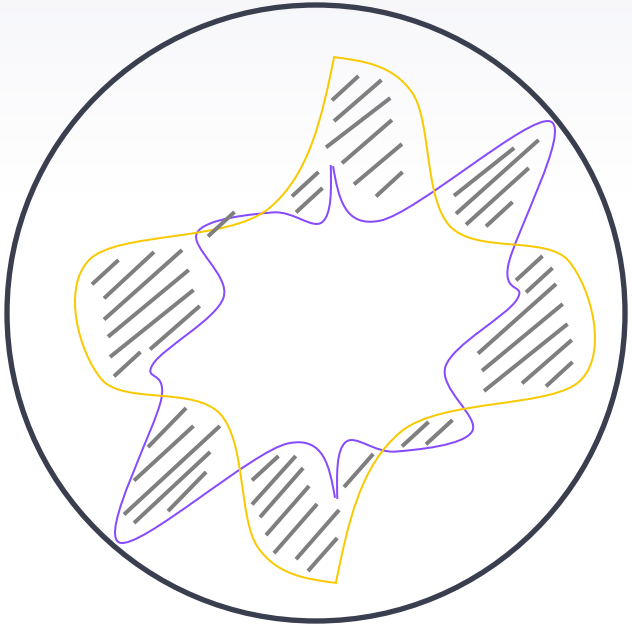


Compare images from histograms

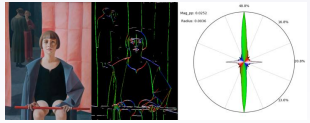
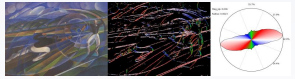
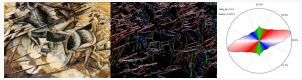
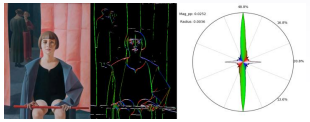
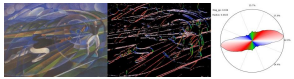
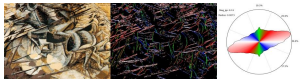
Let’s imagine the yellow and purple lines represent the circular histogram of two different images



The distance between the two images will be the distance between their histograms, which corresponds to the shaded area.



Compare images from histograms

Distance				
	0	0.14	0.71	0.62
	0.14	0	0.78	0.69
	0.71	0.78	0	0.15
	0.62	0.69	0.15	0

Typology of histogram shapes ?

Take a database (set of images) : 1046 images (WikiArt)

« Abstractists » : **Klee, Kandinsky, Rothko, Malevich, Mondrian, Pollock**

Compute all histograms

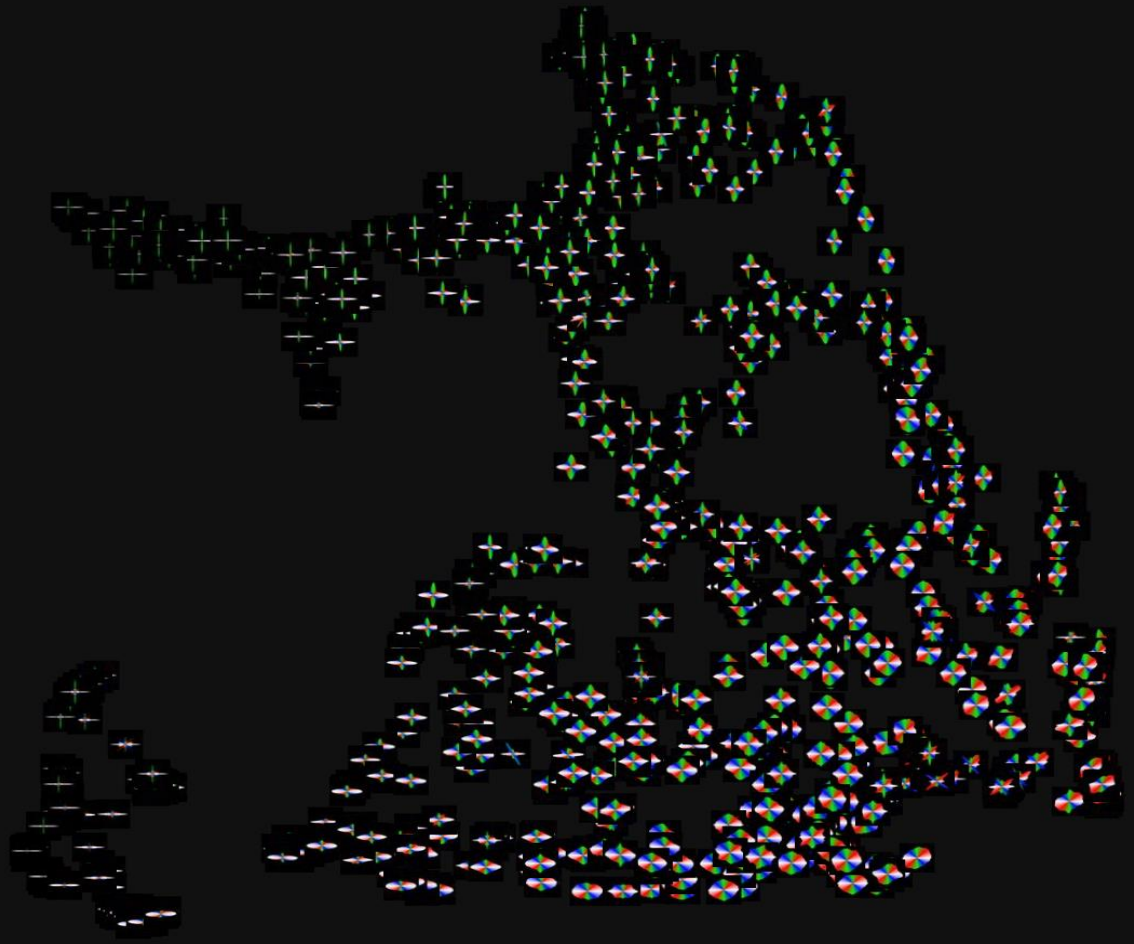
Compute distances between them

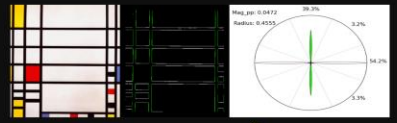
Cluster those that are « close » to each other, projected in a 2D plane. (UMAP)

Analyze the visualization → Use **PixPlot** for an interactive visualization in a web browser

https://adriendeliege.z6.web.core.windows.net/outputs/abstractists1/index.html

(or http://bit.ly/4etv4Tm)

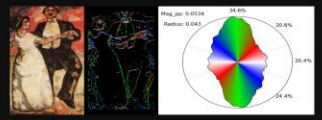




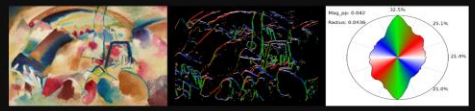
thin "+ shaped"

thick/vertical
"+ shaped"

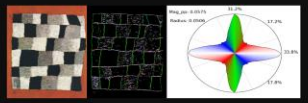
thin vertical ellipse



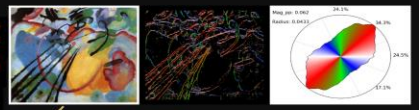
thick vertical ellipse



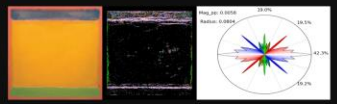
thick/horizontal
"+ shaped"



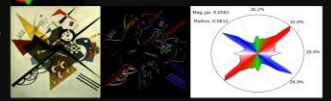
thick balanced
"+ shaped"



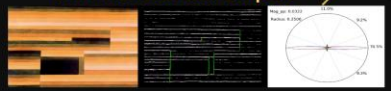
ascending ellipse



star-shaped

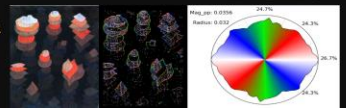


x-shaped



thin horizontal ellipse

thick horizontal ellipse



circular

Try it for yourself

https://adriendeliege.z6.web.core.windows.net/outputs/abstractists1/index.html

Or http://bit.ly/4etv4Tm

Retrieving closest images

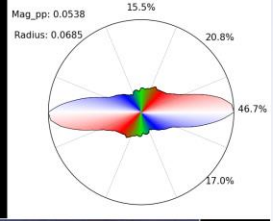
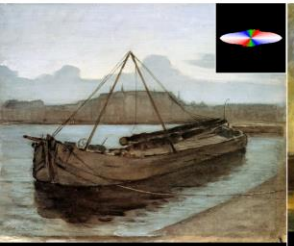
PixPlot is a *visualization* tool.

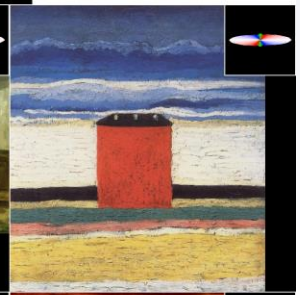
There is a dimensionality reduction.

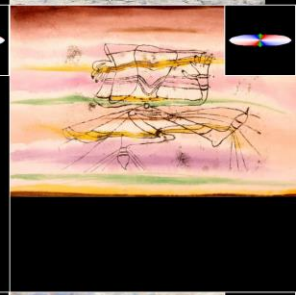
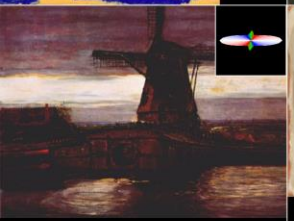
This means there is a loss of information.

To compute « closest images », proceed in the original histogram space.

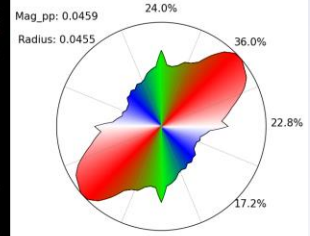



Mag_pp: 0.0538
Radius: 0.0685
15.5%
20.8%
46.7%
17.0%




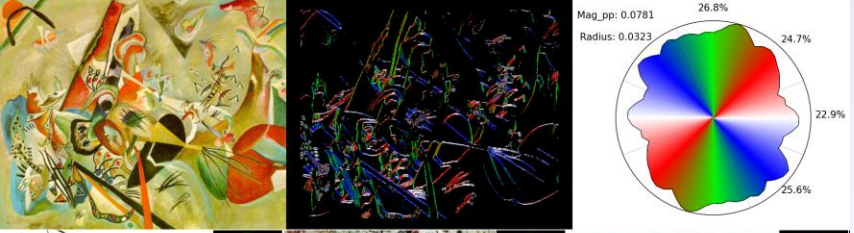




Mag_pp: 0.0459
Radius: 0.0455
24.0%
36.0%
22.8%
17.2%

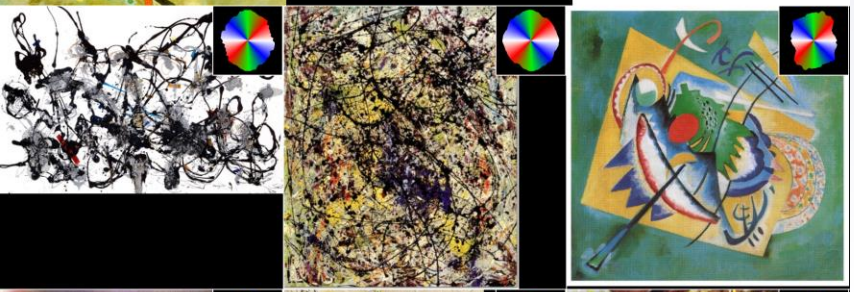


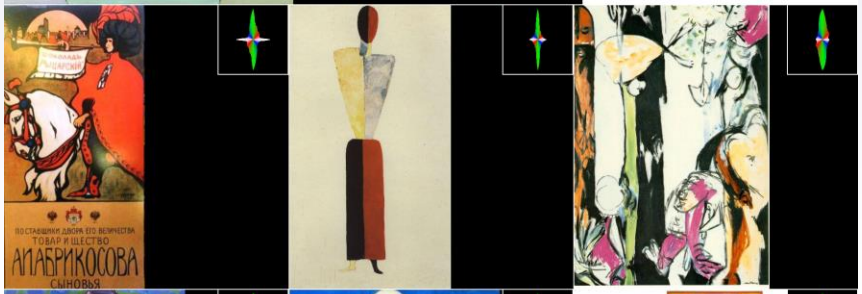


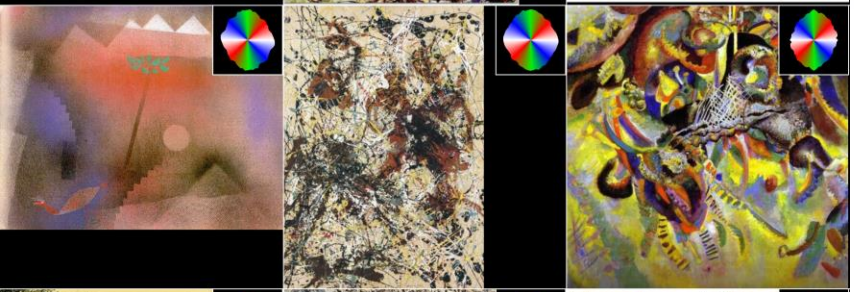


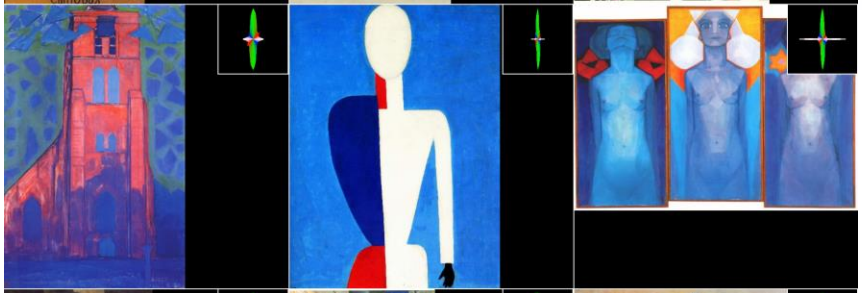


Mag_pp: 0.0781
Radius: 0.0323
26.8%
24.7%
22.9%
25.6%

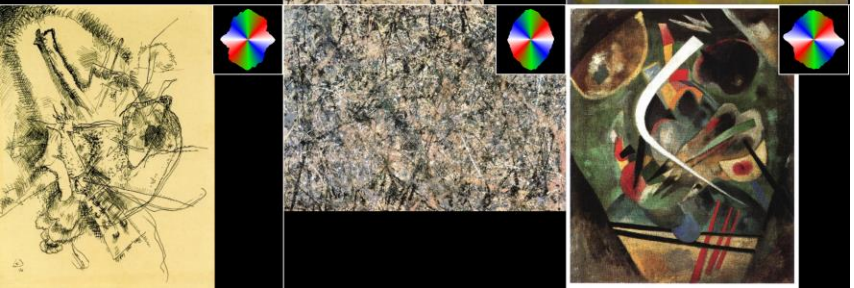

Mag_pp: 0.0503
Radius: 0.119
46.4%
19.5%
17.3%
16.8%

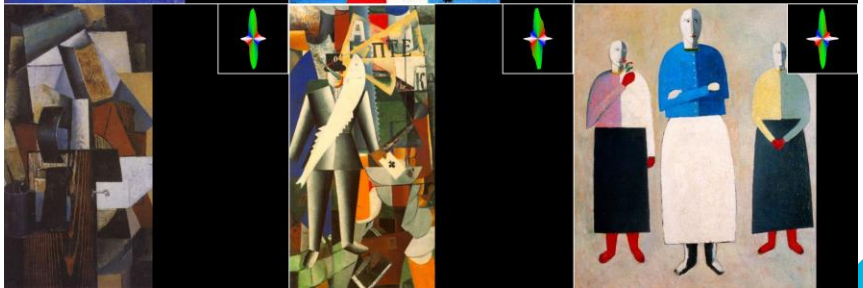



ПОСТАВЩИКИ ДВОРА ЕГО ВЕЛИЧЕСТВА
ТОВАРИЩЕСТВО
АЛАБРИКОСОВА
СЫНОВЬЯ

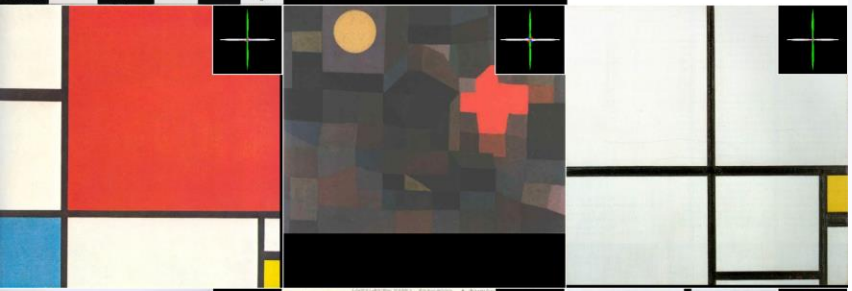


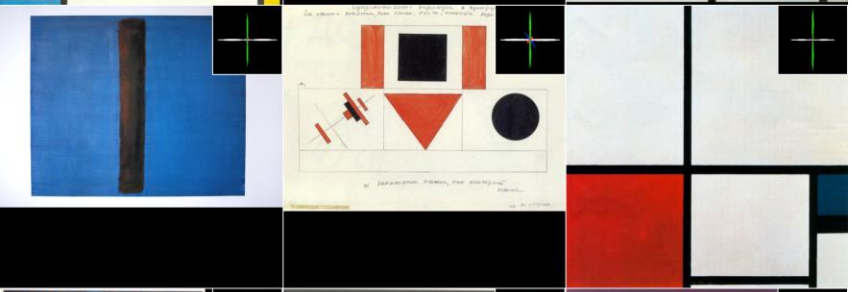


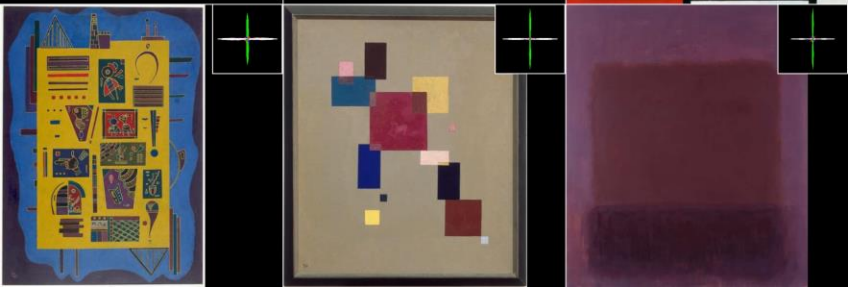


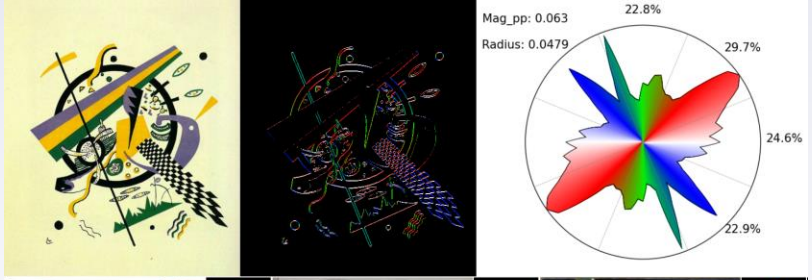


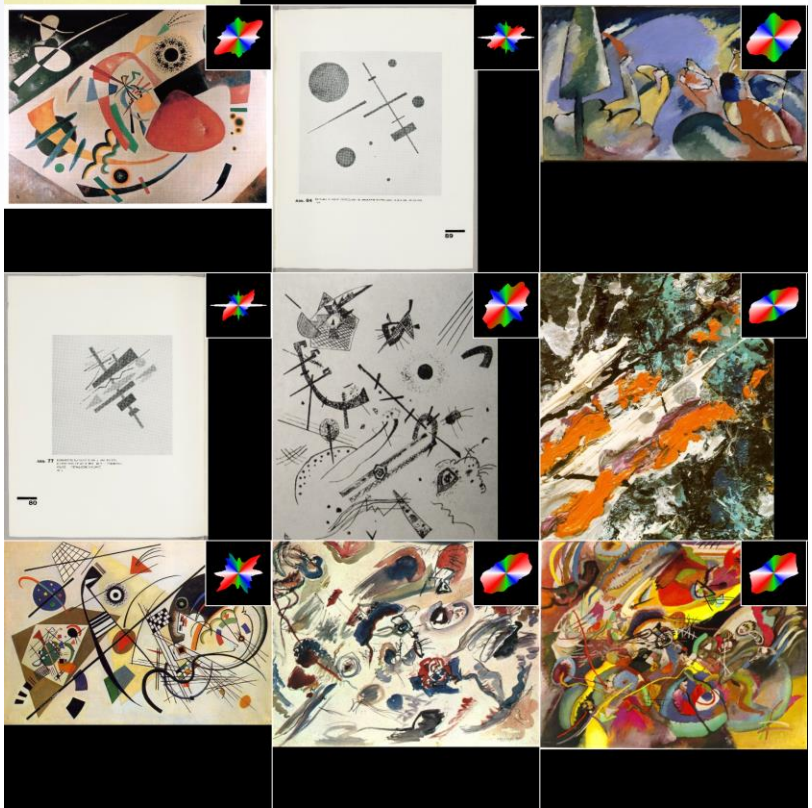

Mag_pp: 0.1381
Radius: 0.2005
40.8%
14.6%
29.2%
15.4%







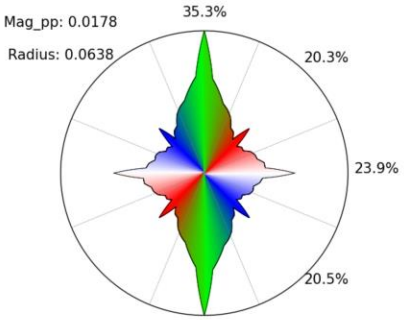

Mag_pp: 0.063
Radius: 0.0479
22.8%
29.7%
24.6%
22.9%

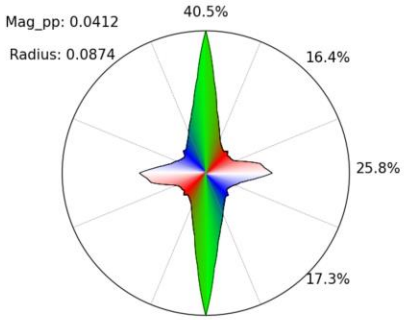


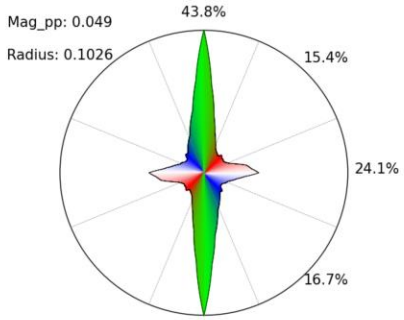
Limitations -- Technical

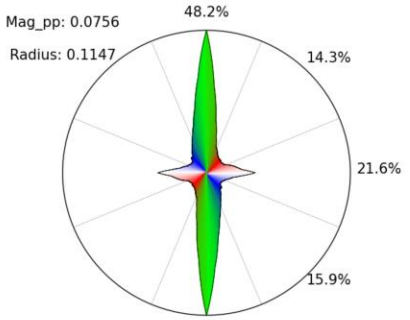
Different histogram shapes for different image resolutions

Effect of image resolution







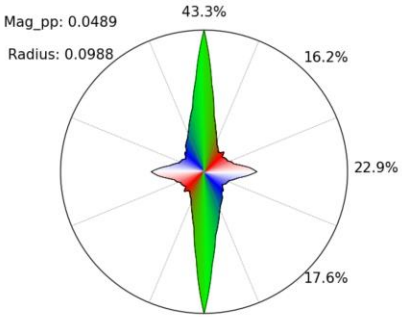


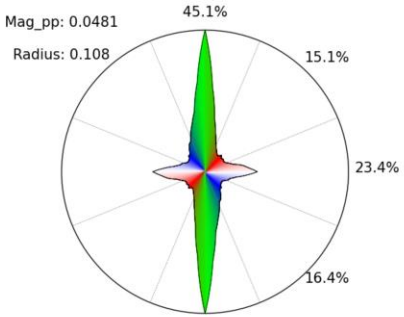
Limitations -- Technical

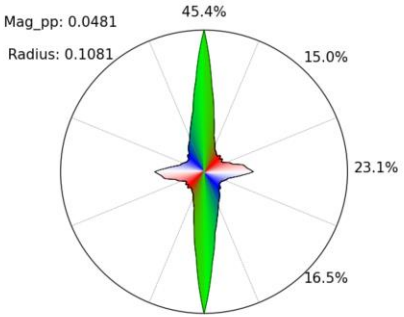
Effect of image resolution

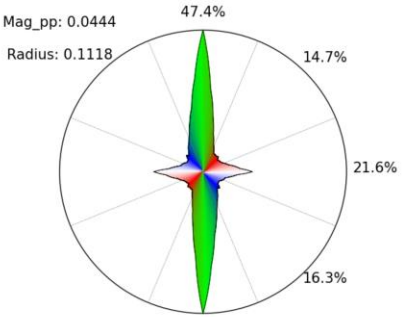
Different histogram shapes for different image resolutions

Problem vanishes if images first rescaled to common dimensions before the analysis
→ But do we lose some info?



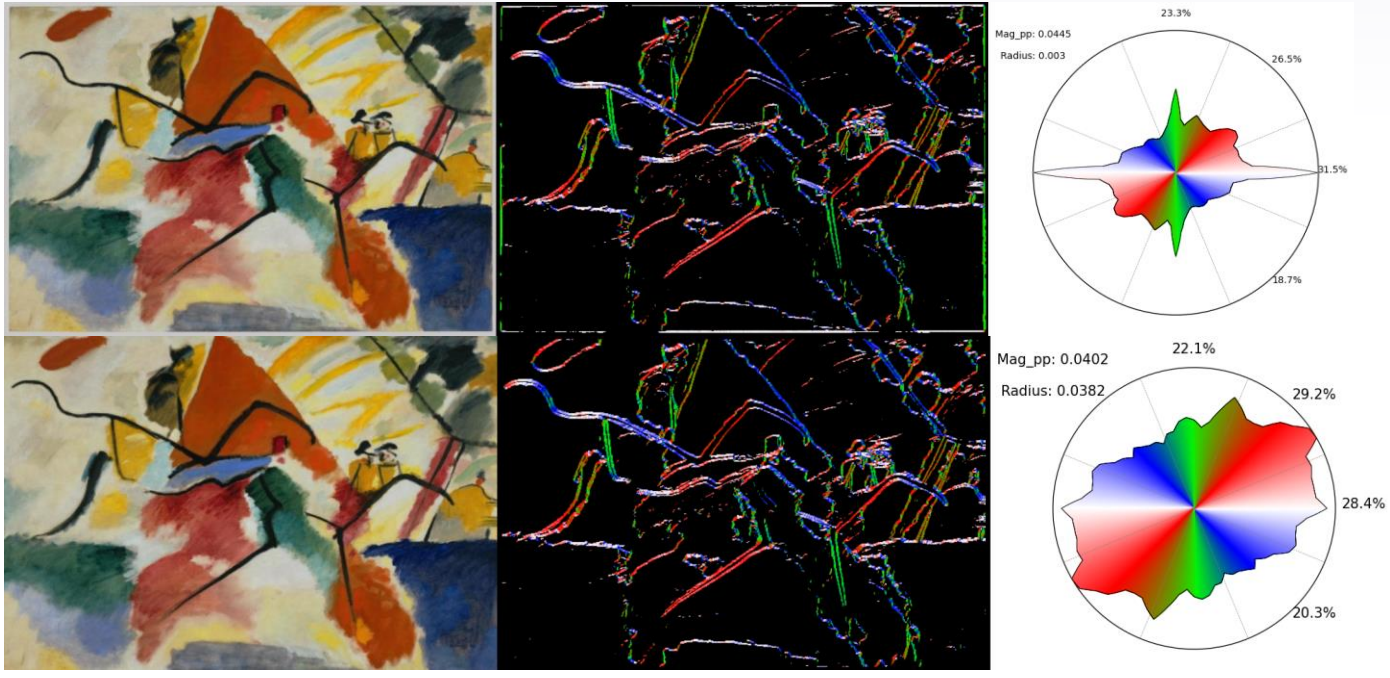






Limitations -- Technical

Sensitivity to frames



Limitations -- Technical

Sensitivity to frames

Idea : Zero-out computations of x% bordermost pixels

Limitations -- Technical

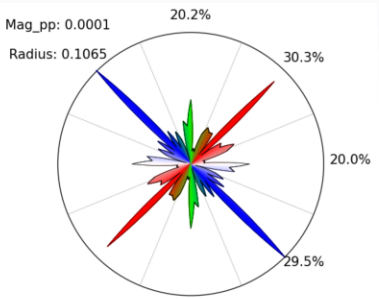
Sensitivity to frames

Idea : Zero-out computations of x% bordermost pixels

But… is the frame part of the artwork?
Is this x% constant?
Is there a shadow/artefact due to the frame?
Is the border always rectangular?
Is it a frame at all?

Limitations -- Technical

Conversion to grayscale : information loss
Compression artefact : patterns appear

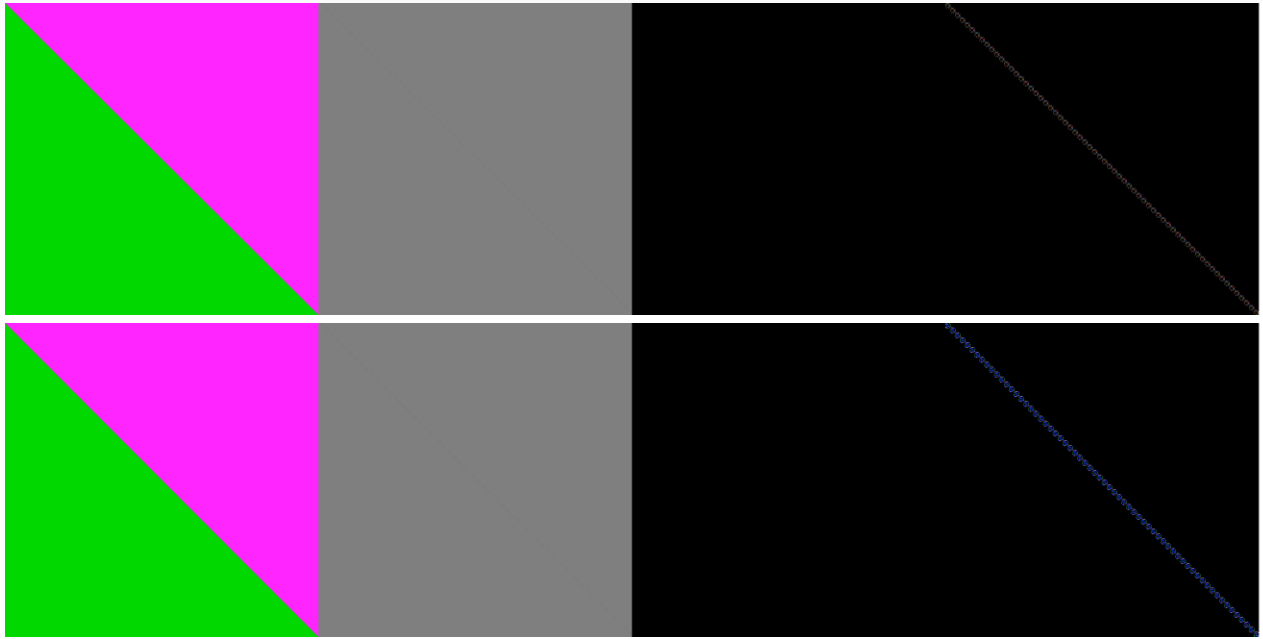


Limitations -- Technical

Or use a better technique? Anyway, generally, what is the « correct » result?

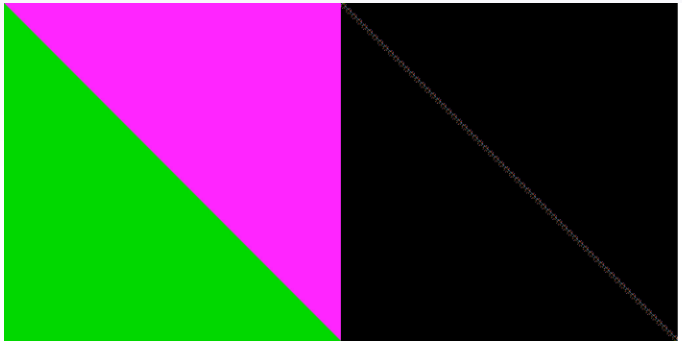
Conversion to grayscale : information loss
Compression artefact : patterns appear

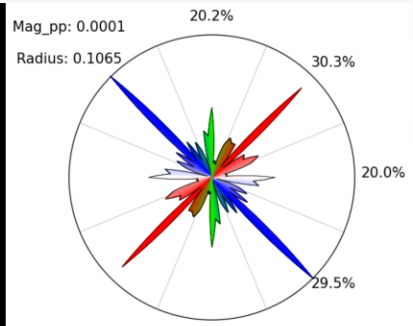
Use some Gaussian Blur? But might blur real strokes and subtelties of the artists barely visible in digitized images?

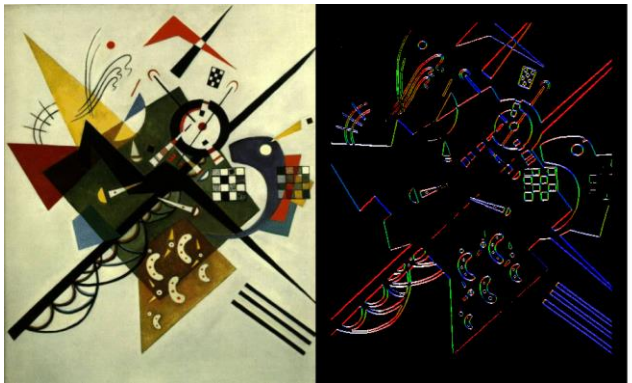


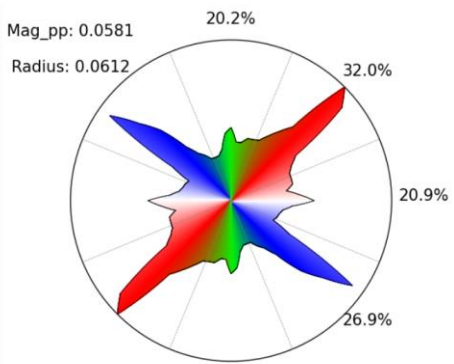
Limitations -- Technical

Combine histogram shape + mag_pp in retrieval?





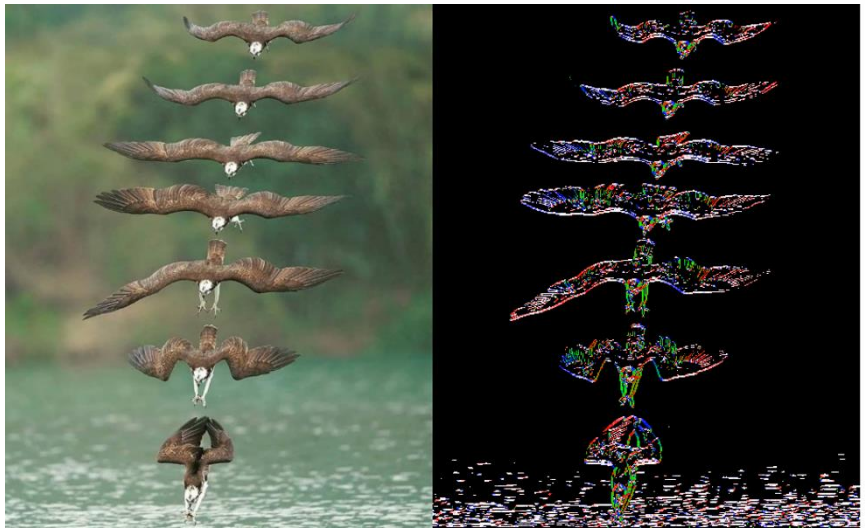


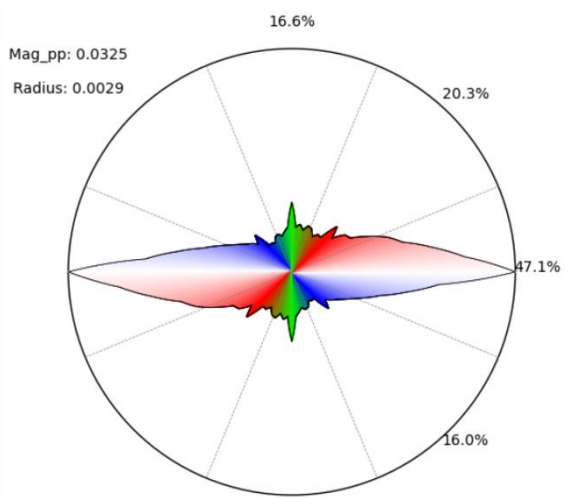


Similar shape
Highly different mag_pp

Limitations -- Interpretative

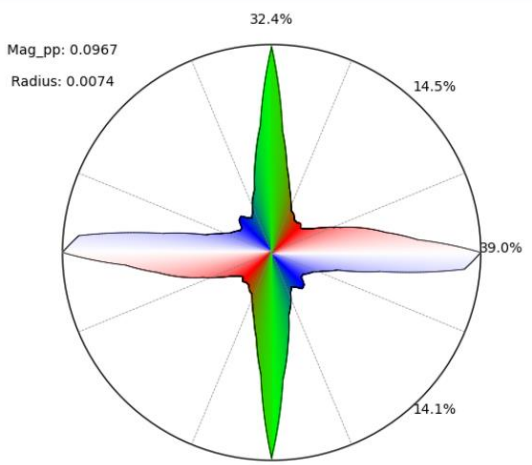
The direction of edges may not reflect the direction of a movement. For example, on the image below, the main direction (from the edges) is horizontal, while the movement of the eagle is vertical. Plus, it's funny because the best AIs just tell me that they see « a group of eagles flying » (but that's another topic).





Limitations -- Interpretative

In this image, we can « feel » the ascending motion of the woman in her bubble, but there is no way we can compute it from the edges, which shows mostly horizontal and vertical directions, as for most everyday-life images. These edges are not triggered by the movement itself.



Conclusion

Beyond the tool and the qualitative validation… the discovery of **new results ?**

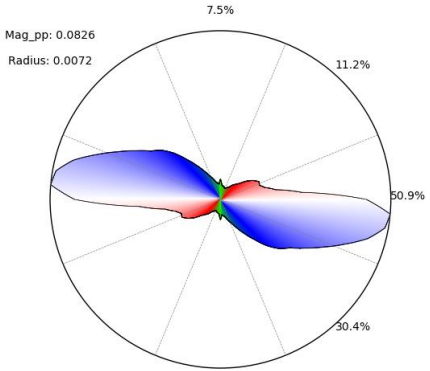
- Shifts for many artists?
- Artists comparisons?
- Style differences?
- Who started a new trend and when?

Other needs from the Digital Humanities community ?

Conclusion

This is an interesting research direction, let’s see where it goes, but I can imagine that it could become a useful tool for the Digital Humanities community.

Let’s wrap up with this beautiful DALL•E 3 image, simply prompted « Motion ».



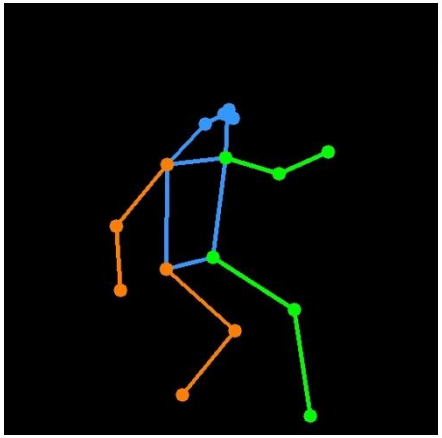
On the dynamicity in images through character poses and edges detection

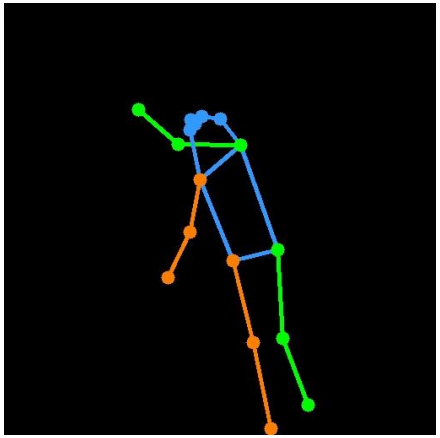
Adrien Deliege

Colloque « Interroger le visible, images qui se répondent. Analyse outillée, IA assistée »
ENS Lyon, 21 juin 2024

This is a brief summary

More technical details are given in our blog post here *https://ceserh.hypotheses.org/3929*

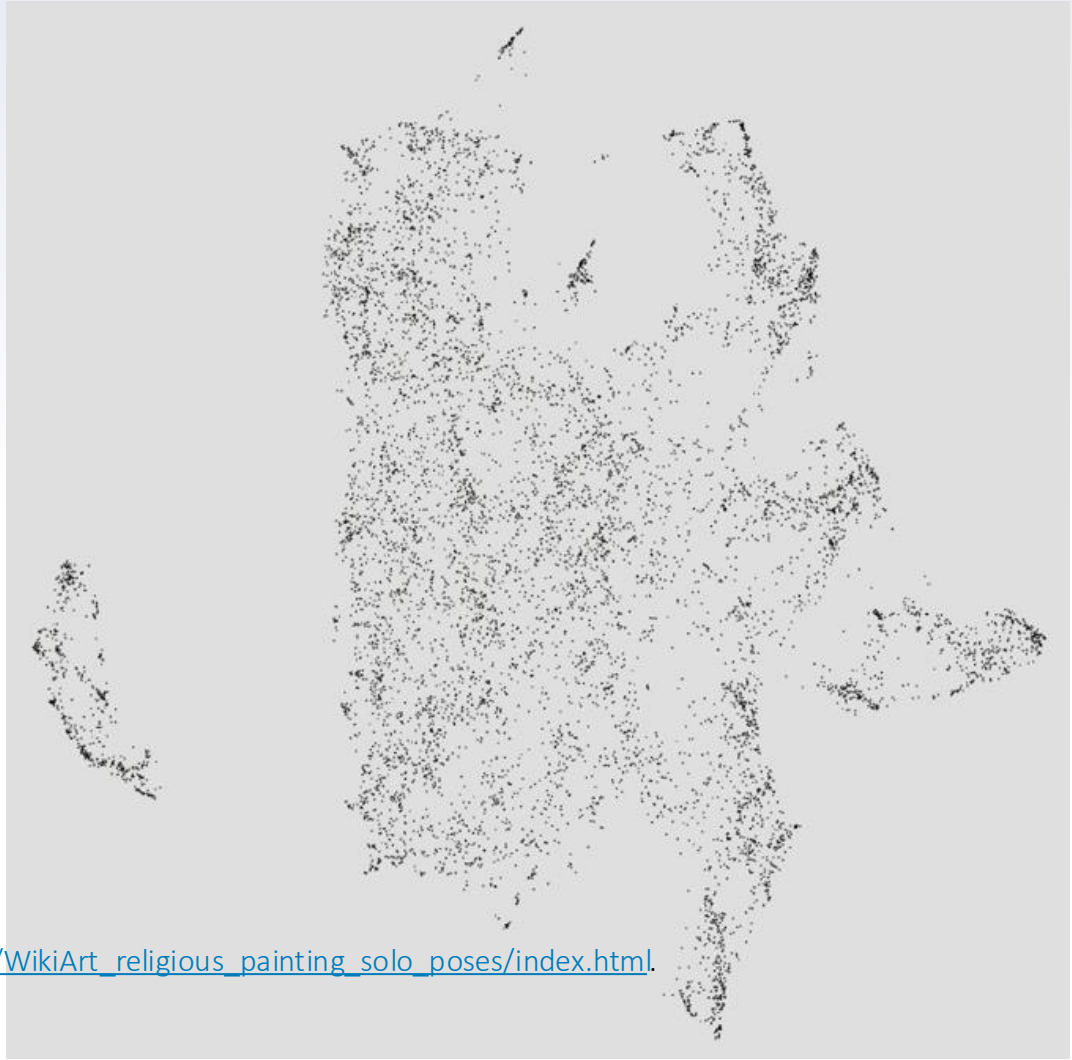




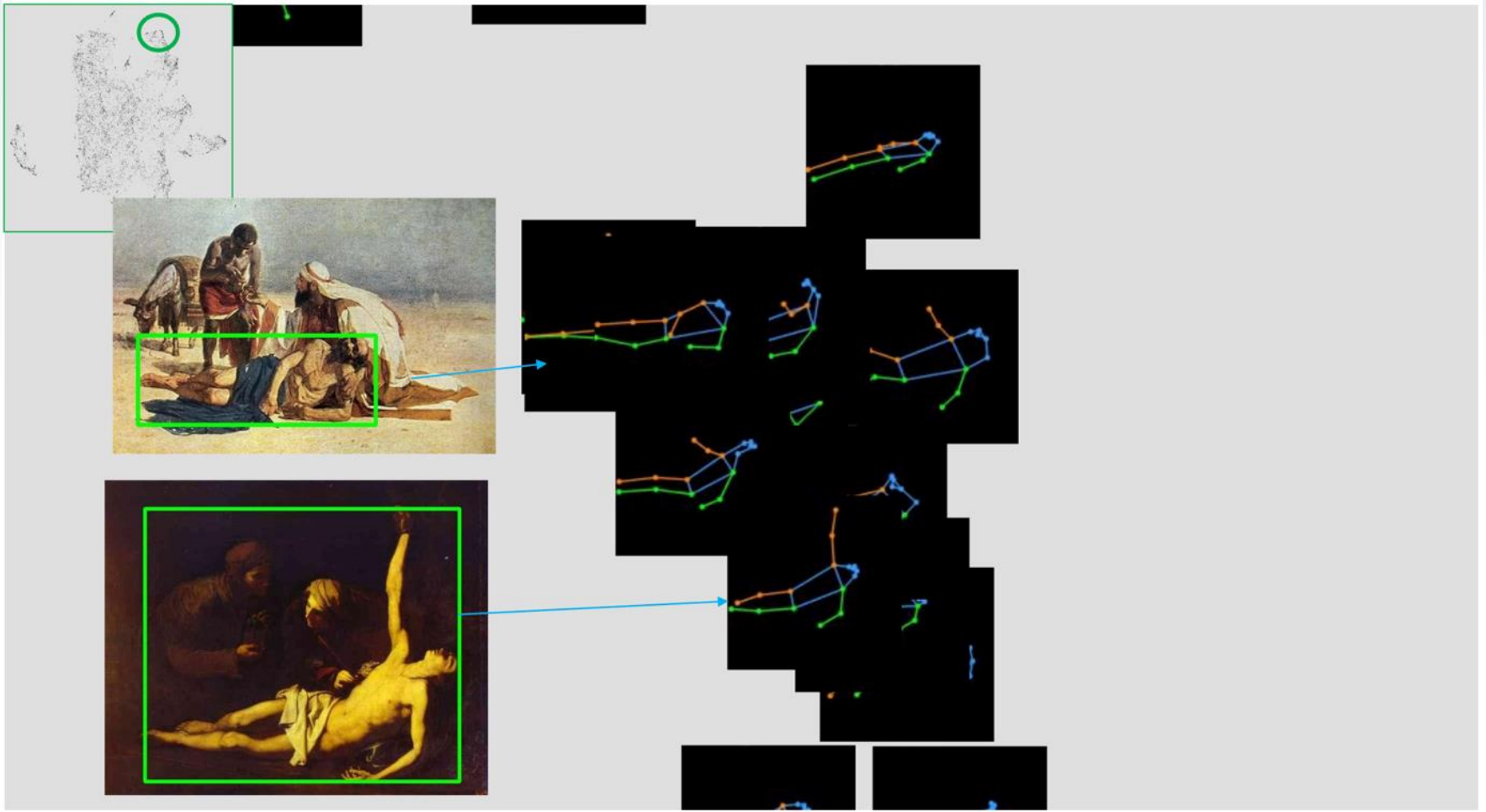
Pose extraction with MMPose.

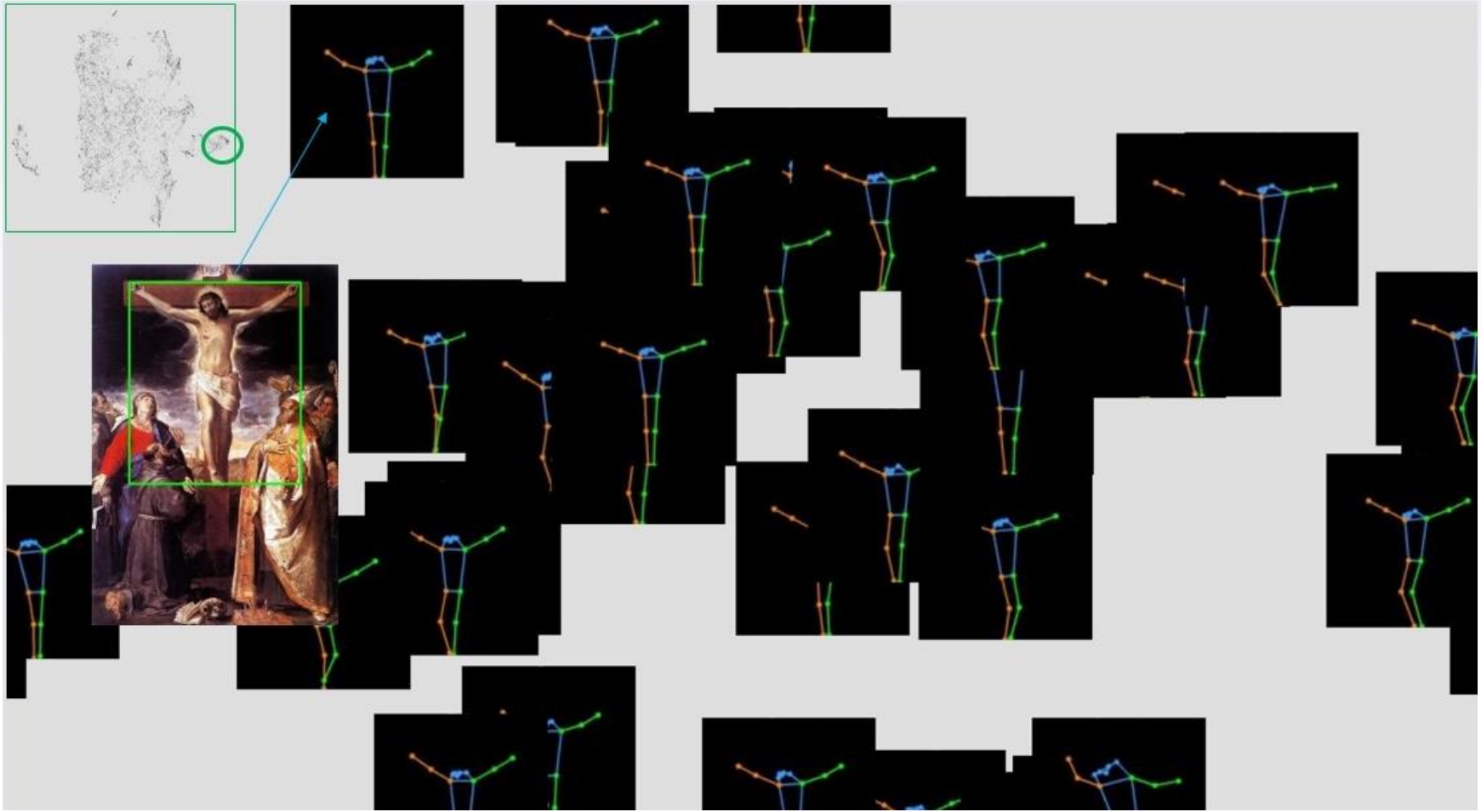
PixPlot visualization

8,599 individual poses from
5,269 religious paintings

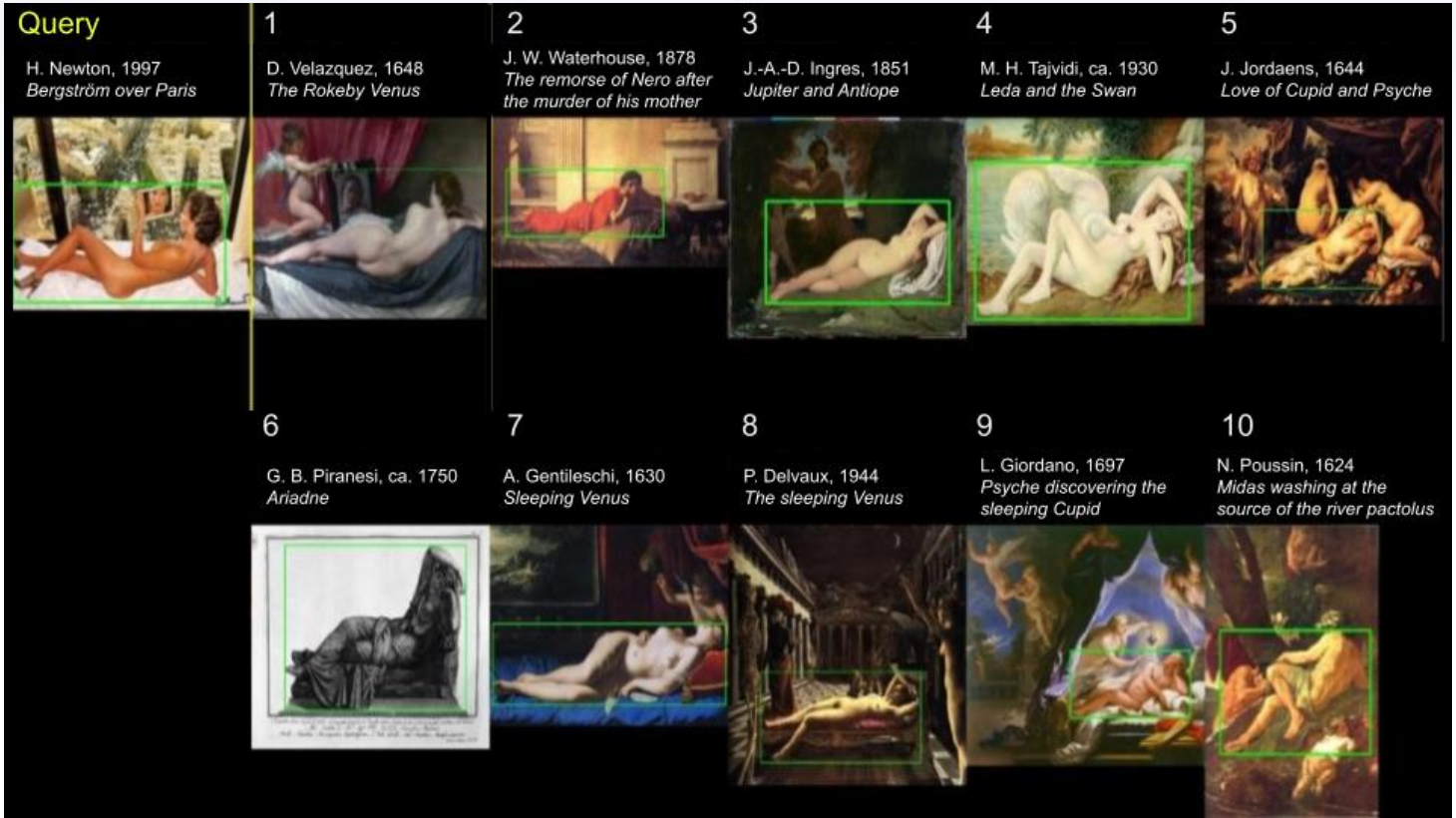


https://adriendeliege.z6.web.core.windows.net/outputs/WikiArt_religious_painting_solo_poses/index.html.

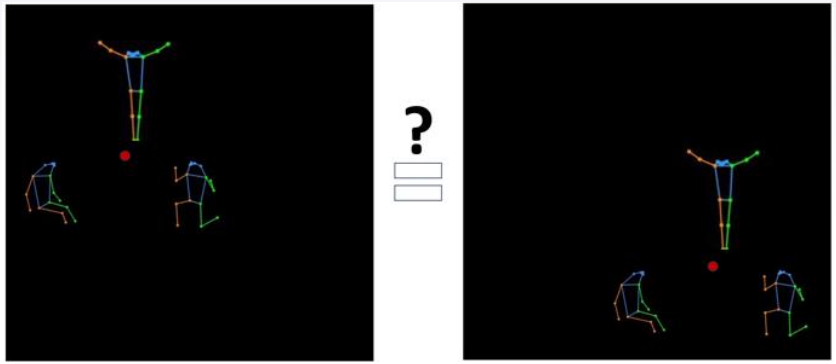


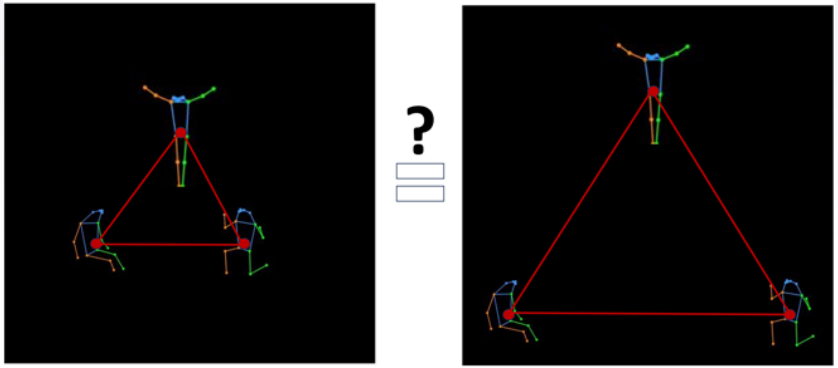


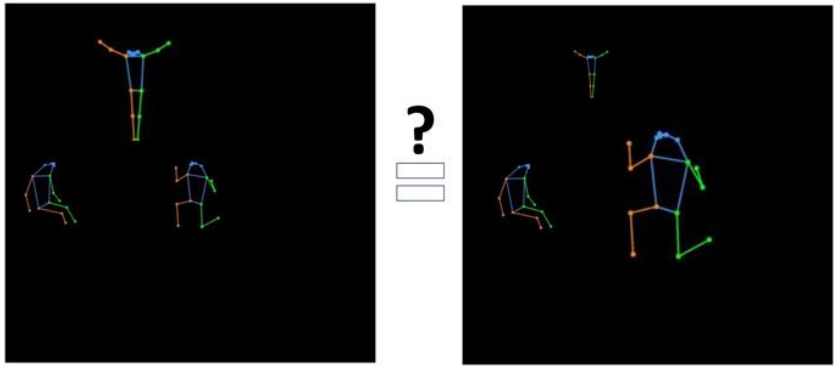
Retrieving similar poses

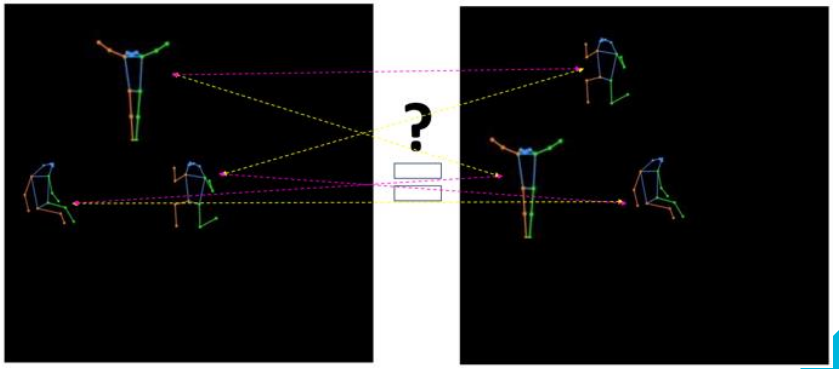


How to deal with multiple poses?

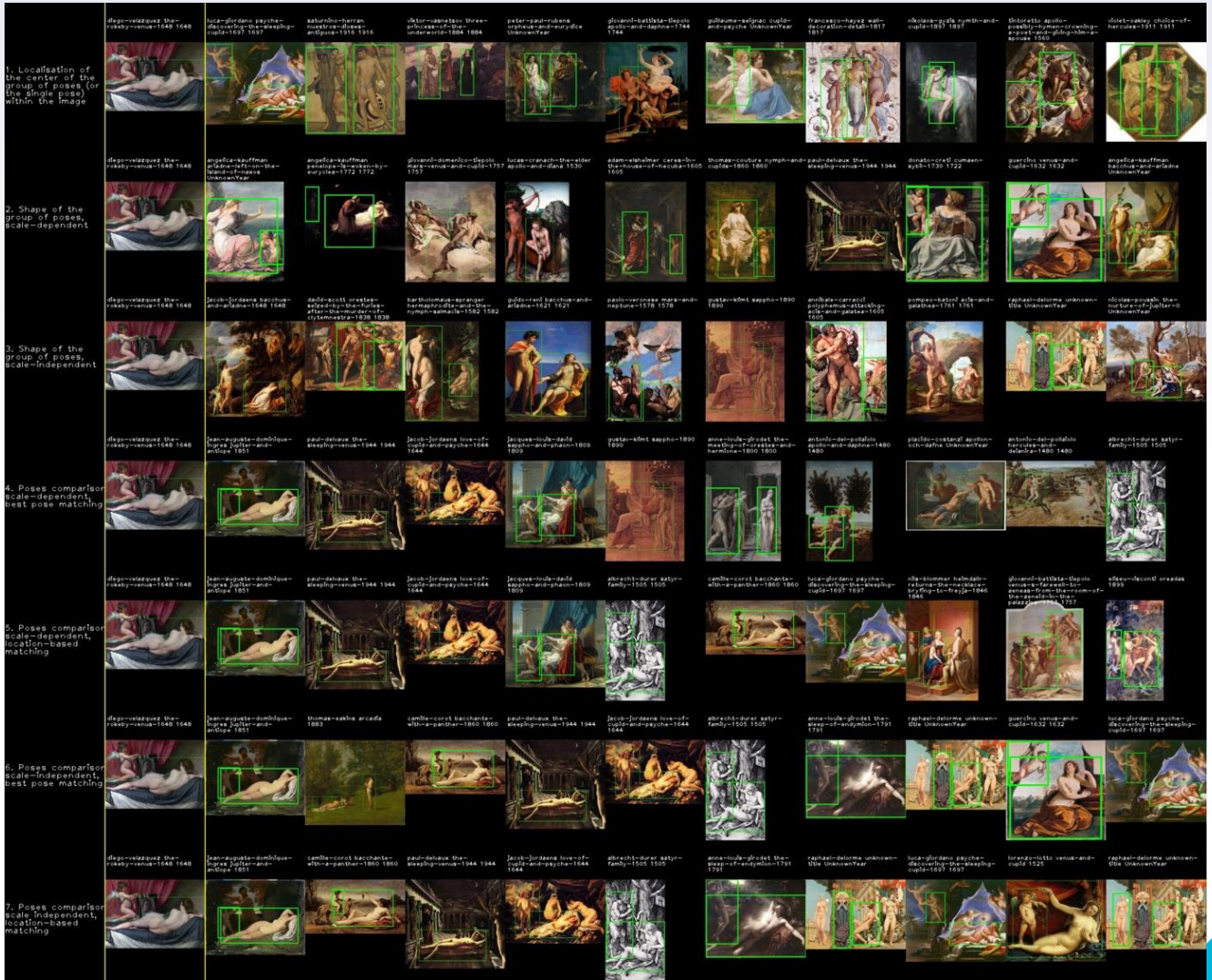




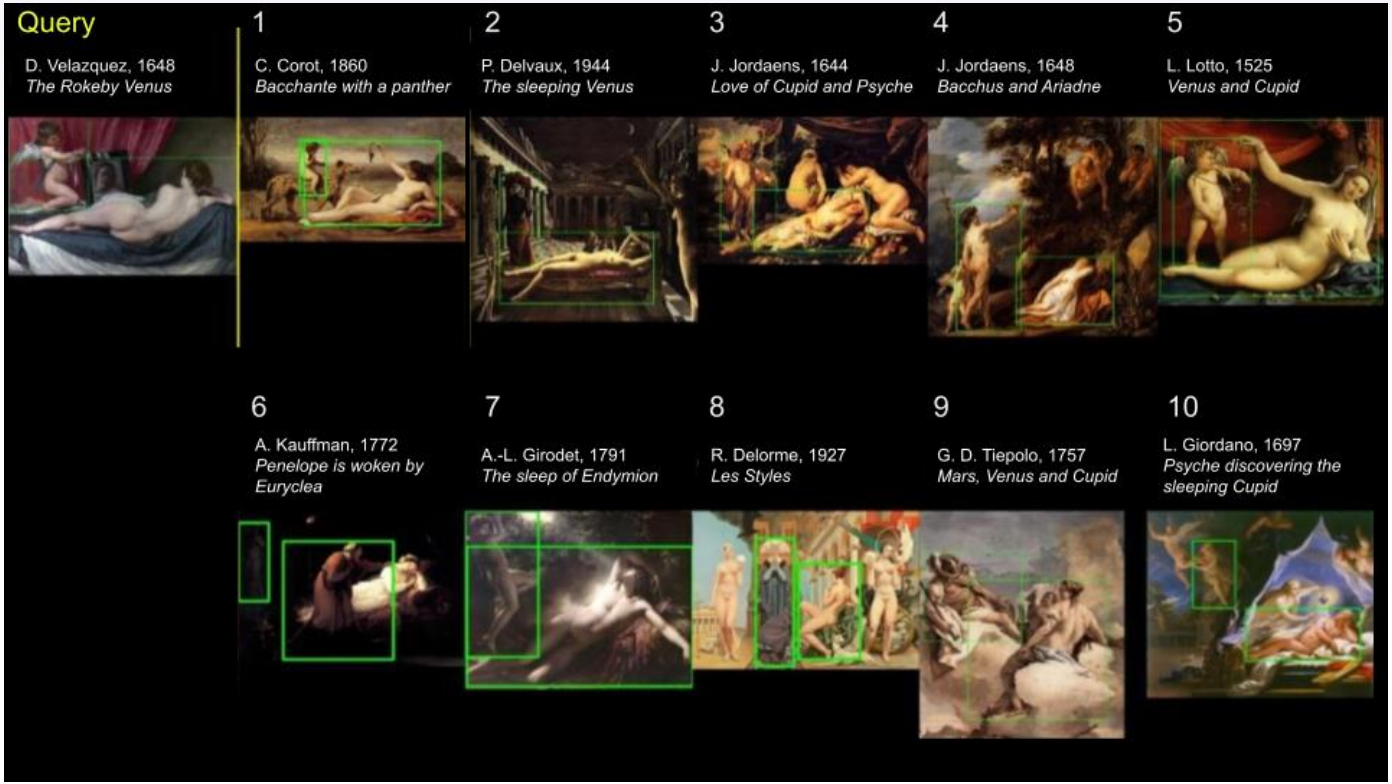




Each choice yields
a different retrieval



Combining criteria



Info beyond motion?