

Image statistics for clustering paintings according to their visual appearance

M. Spehr¹ and C. Wallraven¹ and R. W. Fleming¹

¹Max-Planck-Institute for Biological Cybernetics, Tübingen, Germany

Abstract

Untrained observers readily cluster paintings from different art periods into distinct groups according to their overall visual appearance or “look” [WCF08]. These clusters are typically influenced by both the content of the paintings (e.g. portrait, landscape, still-life, etc.), and stylistic considerations (e.g. the “flat” appearance of Gothic paintings, or the distinctive use of colour in Fauve works). Here we aim to identify a set of image measurements that can capture this “naïve visual impression of art”, and use these features to automatically cluster a database of images of paintings into appearance-based groups, much like an untrained observer. We combine a wide range of features from simple colour statistics, through mid-level spatial features to high-level properties, such as the output of face-detection algorithms, which are intended to correlate with semantic content. Together these features yield clusters of images that look similar to one another despite differences in historical period and content. In addition, we tested the performance of the feature library in several classification tasks yielding good results. Our work could be applied as a curatorial or research aid, and also provides insight into the image attributes that untrained subjects may attend to when judging works of art.

1. Introduction

Judging and understanding a painting is a complex process [GH89]. A trained expert is able to draw on knowledge of related works of art, consideration of technique, historical context and an understanding of the allegorical meanings of specific objects, people and their arrangements to come to a deep understanding of an artwork. However, there are some visual attributes of paintings that are equally accessible to all observers with normal vision. The distributions of lights and darks, the spatial complexity or sparsity of the painting style, and the overall composition of the work are examples of attributes that form part of the artist’s aesthetic vision of the work of art, but which do not necessarily need specialist knowledge to appreciate. Together these attributes are responsible for the overall visual appearance or “look” of a painting, the initial perceptual ‘gist’ [OT06] or impression that even untrained observers immediately apprehend. It is interesting to ask which image features are responsible for this “naïve” visual impression of art.

In previous work [WCF08], we asked naïve subjects to cluster works of art according to their historical period or style. While subjects performed well with some style periods

(e.g. successfully separating abstract paintings from other styles), many of the clusters produced by the subjects appeared to be based on a grouping by visual similarity. For example, it was common for subjects to group dark portrait paintings together irrespective of whether they were painted in renaissance, baroque or romantic periods. We also found little correlation of several, low-level computational features with the perceptual data. In this paper we therefore try to identify image measurements that can be used for clustering artworks into distinct visual classes much like the untrained subjects in our previous experiments.

The specific goals are twofold. First, to create a system that takes as input a database of images of paintings, and returns as output a clustering of those images according to their visual similarity. This could have applications in ‘appearance-based’ image retrieval in general, or as an aid to curators seeking to create a comparative exhibition in which paintings are organized according to their visual likenesses as opposed to historical or thematic considerations. Second, by emulating human performance, we hope to learn something about the types of image measurements that observers might use when making judgements about works of art.

The general approach we take is to start with a database of images that span a range of artistic styles. We take each image in the database and apply a barrage of image analysis methods intended to capture aspects of the work's colour properties, spatial scale properties, composition and content. The outputs of these analyses are represented in a compact feature vector or 'signature' for each image in the database. We then take these signatures as input to unsupervised clustering methods, to return a set of clusters derived from the similarity between image signatures.

The rest of the paper is organized as follows. In Section 2, we summarize the various image features used to create the image signatures. In Section 3, we describe how these signatures can be used to form groups of similar images using standard unsupervised clustering methods. In Section 4, we evaluate the system's classification performance both in unsupervised and supervised clustering tasks. In Section 5 we discuss the limitations of the method and speculate about potential improvements and future applications.

2. Feature Extraction

The visual appearance of a painting depends on many factors including the pigments available to the artist, brush technique, compositional considerations, and of course the *subject* of the work. To capture such a wide range of attributes necessarily involves applying a wide range of features, each designed to measure a different aspect of the image. In this section, we organize the features we use by the kinds of attributes that they are meant to measure, with a general progression from simple, low-level image features to more complex features that are intended to capture aspects of composition and semantic content. However, as our implementation uses more than 200 features, not all are described in complete detail. Many of the features represent the overall distribution of a certain characteristic using summary statistics (mean, variance, skewness, kurtosis). Other features capture the spatial layout of a given attribute in a low-resolution map. These maps were reshaped to a square aspect ratio to ensure that all image signatures were the same length. Examples of some of the features derived from one image are shown in Figure 1.

2.1. Image Database and Pre-Processing

We developed and applied our method using the image database *10.000 Meisterwerke der Malerei* ("10,000 Masterpieces of Painting"), which is available as part of the Yorck project from: http://wiki.directmedia.de/index.de/index.php/The_Yorck_Project.[†] This contains JPEG images of paintings from a wide range of historical periods, along with style labels that

[†] Note that this is a different database from the one used in our previous experiments. This was due to requiring higher resolution

we used to evaluate the unsupervised clustering. From this database we selected 772 images from the following eight periods ranging from C11th to early C20th: *Gothic, Renaissance, Baroque, Classicism, Romanticism, Impressionism, Post-impressionism* and *Art Nouveau*. Each set contained 100 images, except for the *Art Nouveau* class, which contained 72. Prior to the extraction of the image features, the images were down-sampled so that the shorter dimension was 512 pixels.

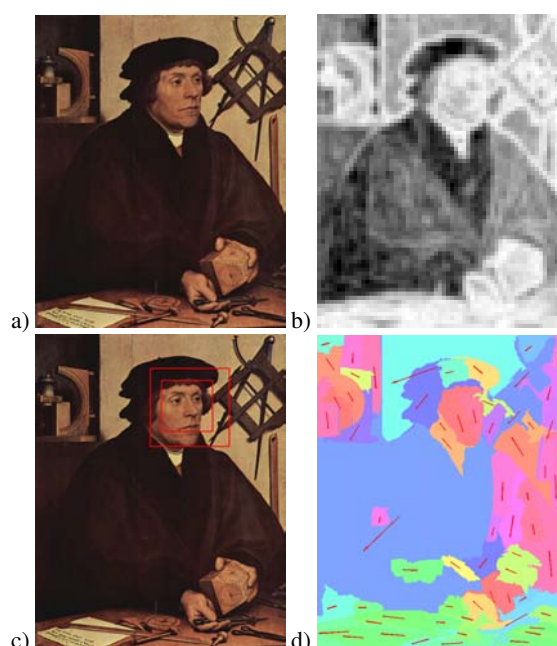


Figure 1: Example features from one image. a) low-resolution version of the image; b) entropy map; c) output of the Viola-Jones face detector; d) orientations of the segments in the image.

2.2. Colour Distributions

Colour usage is one of the most important characteristics of a painting. Artists select and combine colours to achieve particular effects: for example, using a narrow palette to achieve a hazy appearance in a landscape, or applying bold hues to drapery to draw attention to an important figure. We capture the distribution of colours using summary statistics (mean, variance, skewness, kurtosis and entropy) applied to each colour channel in RGB, HSV, Luv and Lab colour spaces. Although simple to compute, when applied to typical photographs, colour histogram features such as these are known

images, and the desire to exclude abstract paintings, which are often too trivial to distinguish from other works

to be quite powerful predictors of perceived image similarity [WCF08] and are commonly used in content-based image retrieval [BPvdH07, IGSb, IGSa]. We therefore believed these would be useful features for clustering artworks too.

Additionally, to represent the dominant colours in the painting, we cluster the pixel values in Luv colourspace and retain the coordinates of the first 4 centres of mass. To capture the primary sources of variation in colour space for each image, we apply PCA to the RGB pixel values, and use the orientations of the 2nd and 3rd eigenvectors and the ratios of the first 3 eigenvalues. Together these features are intended to summarize the ‘palette’ of the work.

2.3. Textural Properties, Spatial Scale and Orientation Structure

Depending on the brush technique and the content of the image, different paintings tend to be dominated by features at different spatial scales. For example, a pointillist landscape is densely populated with small features, while a gothic painting often contains large near-homogeneous regions. To distinguish these types of paintings we need features that measure the ‘textural’ qualities of the image. Similarly, certain structures within the depicted scene tend to produce characteristic orientations in the painting. For example, a landscape with a pronounced horizon will contain more horizontal energy, while a painting of a church facade will tend to contain more vertical energy. To capture these properties of images, we used the following measures.

Entropy maps Entropy is a measure of how much unpredictable variation is present in a signal and has proven helpful in describing variations in artistic style in Van Gogh’s paintings, for example [RFS08]. We computed the local pixel entropy across the image with a sliding window. The resulting entropy maps for each colour channel were down-sampled to 25 x 25 pixels. We also computed summary statistics for each entropy map.

Structure tensor features Another useful measure of local image structure is the structure tensor [HBS92]. This operator effectively performs PCA on the local image gradients to yield an estimate of the dominant local orientation and contrast at each point in the image. We derived 3 quantities from the output of structure tensor:

1. *local orientation*: the orientation of the dominant eigenvector,
2. *local isotropy*: the ratio of eigenvalues and
3. *local gradient magnitude*: the mean of the eigenvalues.

For each of these quantities we included in the image signature a low-resolution spatial map of the feature and summary statistics for each colour channel.

Amplitude spectra We represented the global spatial frequency and orientation content of the images using features derived from the Fourier spectrum. The images were Hanning-windowed prior to application of the FFT. It is well known that for natural images, energy falls off approximately proportionally to $1/f$ (see [GF07] for an interesting application of this idea to art as well as [TMJ99] for an analysis of Pollock’s paintings). [vdSvH96] have shown that differences in the fall-off between images can be captured by fitting a suitable function to orientation slices through the amplitude spectrum. In our case, this fit yields two parameters for each of 226 slices.

We also fit a second 7 parameter model that we have developed to represent the orientation content of natural image spectra. The model is based on Lamé curves (superellipses), and is designed to capture the fact that anisotropic natural spectra typically have a small number of clearly dominant orientations. The parameters of the model allow us to represent the dominant orientations, how dominant they are and the fall-off as a function of spatial frequency.

We employed one other measure of the variations in amplitude spectra between images, which was based on a PCA decomposition of the spectra of all images in the database. For each image we stored the first ten coefficients of the projection of the spectrum onto the decorrelated basis set. This is a particularly compact representation of how a given image deviates from the mean spectrum.

2.4. Segmentation and Composition

Textural statistics of the variety just described are good for capturing the ‘granularity’ of images, but they are blind to the larger scale organization of images. Both the natural world and paintings are typically arranged into discrete units that are organized by natural and aesthetic forces into meaningful shapes and configurations. In order to summarize the mid-level structural and compositional statistics of the paintings in our dataset we applied a number of features derived from segmentation algorithms. We obtained similar results with two standard segmentation methods: *k-means clustering* [Bis06], and *normalized cuts* [SM97].

Number and sizes of segments One simple summary of the structure of the images can be derived from the number of segments and the distribution of their sizes. For example, a landscape scene with a large open sky will produce fewer and larger segments on average than a crowded battle-scene. We therefore measure the following properties related to the size and number of segments returned by the segmentation algorithms:

1. Summary statistics of the distribution of segment sizes
2. low-resolution map of spatial distribution of segment sizes
3. low-resolution map of spatial distribution of number of segments per area

Segment shapes In addition to the number of segments and their spatial arrangement, we also require a method for representing their characteristic shapes. We applied several different methods for representing the shape properties of the extracted image segments:

1. *Fourier shape descriptors* For each segment, we compute the Fourier spectrum of the boundary coordinates, to capture the scales present in the segment's outline. We summarize this with the mean amplitude and phase spectrum across all segment shapes in the image
2. the *compactness* of the segments is defined by the difference in the length of its perimeter from a circle that encloses the same area. We stored summary statistics computed from the compactness of all segments in the image
3. the *principal axis* of the segment is a representation of the orientation and elongation of each segment, which is computed by applying PCA to the coordinates of the perimeter. We stored summary statistics of the principal orientation and elongation across all segments in the image, as well as a low-resolution map of the spatial distribution segment orientations

Extended edges The presence of extended edges in images can also be highly diagnostic. We therefore applied a standard connected components algorithm to the gradient fields that are produced in the pre-processing stages of the segmentation algorithms. The lengths and orientations of the edges that survived the connected-components pruning were stored in a low-resolution map.

2.5. Semantic Content

Human vision is exquisitely adapted to recognizing objects and scenes. It is well known that with even extremely brief presentation, we are able to recall subsequently whether we have seen a particular scene before, or detect the presence of a particular object class (e.g. animals) [TFM96]. We therefore need some measures that correlate with aspects of the semantic content of the image. Note, however, that we do not need to be able to label the precise contents of the image, but rather to measure *similarity* in content between pairs of images. We used the following set of measures.

Template Matching and Aspect Ratio Due to compositional constraints, in many cases, images of similar scenes correlate on a pixelwise basis. For example, portraits tend to have a bright central face with a dark surround, while landscapes tend to have a bright sky above a darker ground. This means that the raw image data itself can be a useful predictor of similarities in semantic content. Accordingly, we included a reshaped, low-resolution (25 x 25 pixel) version of the image as one of our features.

Another very simple feature that correlates to some degree with content is the aspect ratio of the image. Indeed, in

common parlance, vertically oriented images are called 'portrait' format, while horizontal images are called 'landscape' format. This is only a very coarse classification, of course, as still-lives and group scenes are also usually painted in landscape format, and some styles (particularly Art Nouveau) favoured extreme aspect ratios. Nevertheless, we found the ratio of the horizontal to vertical edge length to be a useful feature for grouping images by visual similarity.

Scene Gist A more detailed scene classification scheme is the 'gist' operator [TMFR03], which represents the spatial variations in the spatial frequency and orientation content. Torralba and colleagues have shown that this is remarkably effective at distinguishing between different common classes of scene (such as beach scenes versus street scenes). We applied this operator to our dataset.

Face Detection Given the importance of human figures in Western art, another useful source of information about the content of an image is the number, sizes and locations of faces within the painting. Still-lives and open landscapes contain no faces, standard portraits contain usually exactly one face that fills the middle third of the image, while group portraits usually contain several smaller faces. To capture this, we applied the Viola-Jones face detector [VJ04] to the image, and stored the number of faces found along with a low-resolution map to capture their location and sizes. When applied to paintings, the face detector does not work as well as when applied to photographs. Nevertheless, the feature values were a useful additional indicator of probable scene content.

3. Clustering

Taken together, the features described in the previous section form a 'signature' (feature vector) for describing the appearance of each image. This effectively embeds the dataset in a high-dimensional space (although lower than the original image space). The goal is to take this abstract summary of the visual appearance of the images, and cluster the images based on their similarities, i.e., their proximity in this feature space. To do this, we measure the differences between images in this feature space using the Euclidean norm. From these distances, we can define a similarity matrix by transforming the Euclidean distances with a Gaussian kernel function. To achieve equal weighting of all features, distances were normalized by the maximum distance for each single feature. We use the resulting similarity matrix as input to standard unsupervised clustering methods. An example similarity matrix for one feature (*entropy map*) is shown in figure 2.

Spectral Clustering Having defined the image features and a way to measure similarities one can now look into the discrimination powers that lie within each candidate feature. To get a first impression of what a particular feature is good at

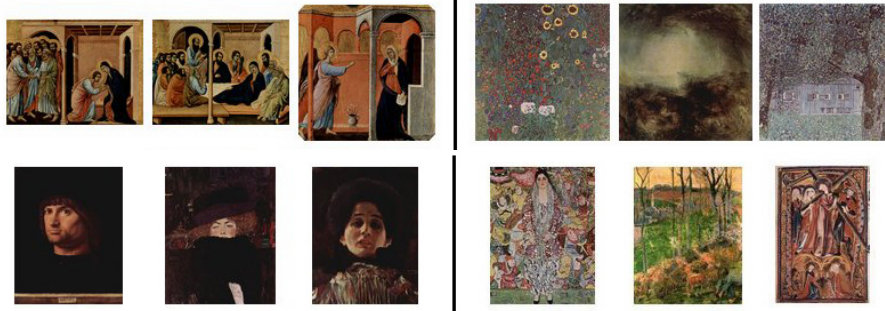


Figure 3: Example images from the extremes of two features, as revealed by spectral clustering. Top: distribution of straight lines. Note the gothic images with prominent extended edges. Bottom: spatial entropy map. The portraits all contain a large central entropy maximum.

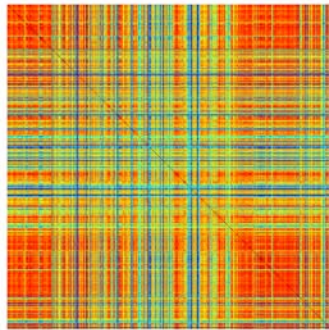


Figure 2: Example similarity matrix for the entropy map feature. The matrix contains one row and one column for each image in the dataset, and they are ordered by the style labels arranged chronologically. Blue values have low similarity, red have high similarity. Early and late paintings are most similar in terms of their spatial entropy distributions.

measuring, one typically wants a bipartitioning of the whole dataset that places the extremes—defined by each feature—at opposite ends of each partition. Spectral clustering [Bis06] provides a straightforward way for achieving this based on a similarity matrix. Examples images from the extremes of two features are shown in figure 3.

Kernel PCA and K-means Although the separation of the data-set along single features is promising, no single feature from our feature-set can capture all of the ‘visual similarity’ of two images. Therefore we need to combine features to achieve a global clustering that takes into account all the measured attributes of an image. The dimensionality of the feature-space is high and many of the features correlate to some extent with one another. Taking the similarity matrix

as input, the kernel PCA algorithm [Bis06] produces new representations for the images in a decorrelated feature space. To reduce the dimensionality, the first 8 novel dimensions were taken into account (8 being the number of *style* labels in the dataset). The result of this procedure is a point cloud in 8 dimensional space that represents the whole image dataset. We then use *K-means* clustering to infer structure from it and obtain the output groups according to the new properties.

Example results are shown in the collages in figure 4 for three out of the 8 clusters. Despite differences in both the style and content between images within a cluster, the images share a certain visual ‘look’: pastel-shaded, pastural images; dark portraits; and flat, geometrical church paintings can be grouped together automatically with the features we have presented here. The remaining five clusters ranged in quality: some overlapped clearly with these clusters (e.g. another cluster also contained many dark portraits), although others were more heterogeneous such that it is difficult to verbalize what the images have in common. We observed that this grouping is far better in coherence than the ones created by the simple, low-level measures in [WCF08] which were much more heterogeneous.

4. Relationship to Style labels

Having clustered the images using features intended to capture the overall ‘look’ of the painting, it is interesting to ask how this relates to the stylistic period in which the work was painted. On the one hand, conventions related to the intended appearance of the work are one of the defining characteristics of a given period (e.g. visible brushstrokes in impressionist paintings), so we might expect a strong correlation. On the other hand, certain qualities are universal (e.g. portraits against a dark background), so we might expect the clustering to cut across style boundaries. Using the clusters obtained in the previous section together with the ‘ground truth’ labels of the art periods in the database, we can there-



Figure 4: Collages showing the images in three of the eight clusters.

fore evaluate how well *unsupervised* methods correspond to historical art periods. Complementing this, we can also ask how *supervised* methods would perform: since we have the

class labels and feature vectors for each painting, we can directly train classifiers on a subset of the database to discriminate between different art periods.

4.1. K-means cluster evaluation

In table 1, we show the discrimination ability of the clusters regarding the original *style* labels. Specifically, we report the frequency of images of a specific *style* within each cluster.

In general the clusters cut across styles rather than rigidly adhering to them with the exception of Cluster 8 which almost exclusively contains Gothic paintings. Reading the table column-wise we observe that a large percentage of Art Nouveau is concentrated in Cluster 6 (which also contains Gothic) images. These results make sense as the *content* varies greatly within a style, and we know from our previous experiments that this also plays a major role in the ‘appearance’ of a work [WCF08]. In future work, it would be interesting to use a dataset that also has ground-truth labels for content classes to look into content-based clustering for which our features might be better suited.

4.2. SVM classification

After having performed *unsupervised* classification, we also trained a support vector machine based on the “style” ground truth labels.

Feature selection For the classification, we first needed to find good feature vectors in our collection of features which were able to *discriminate* between classes. In order to do this, a class contrast function CC was introduced:

K_i	Similarity/Kernel matrix for feature(-combination) i
S	Samples
S_A	Samples class A
S_B	Samples class B
$K_i(S_A, S_B)$	$\sum_{s_a \in S_A} \sum_{s_b \in S_B} sim_i(s_a, s_b)$
$\overline{K_i(S_A, S_B)}$	$\frac{K_i(S_A, S_B)}{ S_A \cdot S_B }$

$$CC(A, B) = \frac{0.5 \cdot (\overline{K_i(S_A, S_A)} + \overline{K_i(S_B, S_B)}) - \overline{K_i(S_A, S_B)}}{\overline{K_i}} \quad (1)$$

CC determines how well feature i separates classes A and B from each other. The best results of CC across all features are given in table 2. The features that provide the highlighted contrast values are almost exclusively provided by the statistical colour descriptor testifying to the importance of colour in separating different categories of art. We can also observe that the Renaissance, Baroque, and Impressionist periods all have rather low values across the board indicating that the discriminative power of our features is not good enough to really separate them.

	Gothic	Renaissance	Baroque	Romanticism	Classicism	Impressionism	Postimpressionism	Art Nouveau
Cluster 1	0	2	13	32	6	25	11	2
Cluster 2	0	6	17	26	25	4	1	8
Cluster 3	14	8	8	8	9	23	39	3
Cluster 4	20	20	17	7	28	4	3	6
Cluster 5	11	21	8	8	17	30	16	13
Cluster 6	16	4	4	4	1	8	12	34
Cluster 7	16	31	31	12	14	5	15	4
Cluster 8	11	2	2	3	0	1	3	2

Table 1: Table showing how many images in each of the 8 clusters fall into one of the 8 style labels. Numbers in bold indicate strong contributions of the particular cluster.

	Gothic	Renais.	Baroque	Classicism	Romanticism	Impr.	Postimpr.	Art N.
Gothic	0	0.2983	0.4322	0.6347	0.5161	0.2916	0.4622	0.5577
Renais.	0.2983	0	0.1128	0.1996	0.2196	0.2910	0.4303	0.2857
Baroque	0.4322	0.1128	0	0.1191	0.1529	0.3557	0.4559	0.2597
Classicism	0.6347	0.1996	0.1191	0	0.1137	0.3233	0.4948	0.2749
Romanticism	0.5161	0.2196	0.1529	0.1137	0	0.1793	0.5308	0.2404
Impr.	0.2916	0.2910	0.3557	0.3233	0.1793	0	0.2436	0.2870
Postimpr.	0.4622	0.4303	0.4559	0.4948	0.5308	0.2436	0	0.5936
Art N.	0.5577	0.2857	0.2597	0.2749	0.2404	0.2870	0.5936	0

Table 2: Class-contrast function CC showing the best values across all features. Numbers in bold indicate strong contrasts which should lead to good discrimination performance.

Classification performance On the basis of this function for each class combination, the single best feature was determined and a new similarity (kernel) matrix K was created using a linear combination of all similarity matrices. In this study, equal weights were used so that an average kernel matrix K of the best individual features was created. For future work, variable weights could be introduced to enhance classification performance. We used 550 of our images as training examples and tested on 222. As SVMs are inherently designed for a two-class classification problem, we trained several SVMs on different bipartitions of the dataset in order to perform a multi-class classification. The winning class was then selected by a group vote of all SVMs.

Chance performance of the combined SVMs for 8 classes is $\frac{1}{8}$. The best performance that was achieved after training was 0.45 meaning almost half of the test images were well classified and the other part misclassified. This pattern, however, is critically dependent on the art period in question: as table 3 shows, the Gothic period, for example, is classified very well, followed by Classicism and Postimpressionism, whereas performance is rather weak (although above chance) for the Baroque, and Romanticism periods.

Figure 5 shows the confusion matrix for the classification results as a color-coded representation. Baroque images are often identified as Classicism and Romanticism, Gothic as Renaissance, Romanticism as Classicism and Impressionism as Romanticism. It is interesting to note that the performance with which different periods are discriminated using our features is related to their historical proximity as well as their intuitive visual similarities. Note also, that the SVM classification cannot be fully predicted from table 2 alone: we found high values of CC within the Art Nouveau period, for ex-

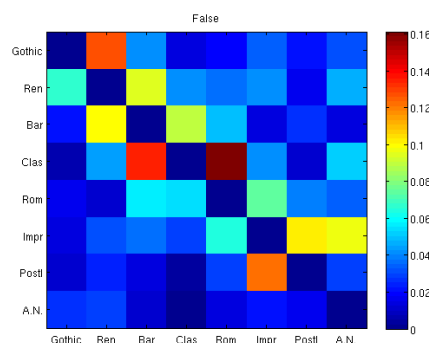


Figure 5: Ratio of false positives - (Row-)Class A is falsely classified as (Column-)Class B. The axes follow the historical progression from Gothic to Art Nouveau from top to bottom, and left to right, respectively.

ample, even though the final classification performance was rather weak. Overall, however, our results are very encouraging given the difficult nature of the classification task.

5. Conclusion

We have presented a method for automatically clustering images according to the overall visual appearance or “look”, much as untrained observers do. Because the appearance of paintings is complex and spans many aspects ranging from colour content to semantics, we argued for using a large number of features, each of which is insufficient to capture appearance on its own, but which when taken together can

Gothic	Renaissance	Baroque	Classicism	Romanticism	Impressionism	Postimpr.	Art Nouveau
0.7144	0.3079	0.2116	0.5865	0.2415	0.3582	0.5636	0.2642

Table 3: Classification performance per art period

parse a database of images into visually meaningful groups. We have shown that such an “appearance-based” clustering is affected by, but is not the same thing as a clustering based on distinct artistic periods or styles. Much like human observers, the system confounds style and content when assessing the similarity of two images.

The method is far from perfect. Some of the clusters have a very intuitive visual quality (e.g. dark portraits or hazy landscapes, see figure 4). However, other clusters are more heterogeneous, containing images whose principal shared attribute is not belonging to one of the other well-defined classes. Interestingly, in our previous experiments [WCF08], some subjects reported forming ‘miscellaneous’ groups to classify images that did not belong with the others. Thus these failures may reflect a key aspect of the data, and in the near future we will correlate the perceptual data with our collection of features. Nevertheless, our work here represents a significant step forward in terms of capturing higher-level features and concepts along which participants might group works of art compared to the clustering obtained, for example, in [WCF08].

We also tested the ability of the features for classifying paintings into the historical art periods both in an unsupervised and a supervised fashion. Performance in both cases was well above chance with clear variations in discrimination across art periods: paintings from the Gothic period, for example, were easy to separate (a result which agrees well with the perceptual data found in [WCF08]). As we did not specifically work on efficient methods for integrating the features that we implemented, this represents a large area for improvement of classification performance (see, for example, [DvdLB06])—we expect different features to be important for separating different art periods.

One of the more obvious features that is missing from the current implementation is a representation of the relative ‘importance’ or ‘saliency’ of different regions in the image. Another obvious respect in which the method could be improved would be to define better representations of the composition and content of the image, particularly by improved segmentation or object detection schemes.

We believe that by refining the features used to measure intermediate and high-level image properties, a system in the spirit of the approach we have presented here may be able to perform complementary clusterings according to style and content. This would be a particularly useful tool for archiving and image retrieval.

References

- [Bis06] BISHOP C.: *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc. Secaucus, NJ, USA, 2006.
- [BPvdH07] BEREZHNOY I., POSTMA E., VAN DEN HERIK J.: Computer analysis of Van Gogh’s complementary colours. *Pattern Recognition Letters* 28, 6 (2007), 703–709.
- [DvdLB06] DEAC A., VAN DER LUBBE J., BACKER E.: Feature Selection for Paintings Classification by Optimal Tree Pruning. *Lecture Notes in Computer Science 4105* (2006), 354.
- [GF07] GRAHAM D., FIELD D.: Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities. *Spatial Vision* 21, 1 (2007), 149–164.
- [GH89] GOMBRICH E., HANS E.: *The Story of Art*. Phaidon Oxford, 1989.
- [HBS92] HANCOCK P., BADDELEY R., SMITH L.: The principal components of natural images. *Network: Computation in Neural Systems* 3, 1 (1992), 61–70.
- [IGSa] ICOGLU O., GUNSEL B., SARIEL S.: Arthistorian: An integrated indexing and personalized browsing system for art paintings.
- [IGSb] ICOGLU O., GUNSEL B., SARIEL S.: Classification and indexing of paintings based on art movements. In *Proc. of EU-SIPCO*, pp. 749–752.
- [OT06] OLIVA A., TORRALBA A.: Building the gist of a scene: the role of global image features in recognition. *Visual Perception: Fundamentals of Awareness, Multi-Sensory Integration and High-Order Perception* (2006).
- [RFS08] RIGAU J., FEIXAS M., SBERT M.: Informational dialogue with van gogh’s paintings. In *Computational Aesthetics 2008* (2008), pp. 115–122.
- [SM97] SHI J., MALIK J.: Normalised cuts and image segmentation. In *Proceedings of CVPR’97* (Puerto Rico, June 1997).
- [TFM96] THORPE S., FIZE D., MARLOT C.: Speed of processing in the human visual system. *Nature* 381 (1996), 520D–522.
- [TMFR03] TORRALBA A., MURPHY K., FREEMAN W., RUBIN M.: Context-based vision system for place and object recognition. *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on* (2003), 273–280.
- [TMJ99] TAYLOR R., MICOLICH A., JONAS D.: Fractal analysis of pollock’s drip paintings. *Nature* (1999), 422–422.
- [vdSvH96] VAN DER SCHAAF A., VAN HATEREN J.: Modelling the power spectra of natural images: Statistics and information. *Vision Research* 36, 17 (1996), 2759–2770.
- [VJ04] VIOLA P., JONES M.: Robust Real-Time Face Detection. *International Journal of Computer Vision* 57, 2 (2004), 137–154.
- [WCF08] WALLRAVEN C., CUNNINGHAM D. W., FLEMING R.: Perceptual and computational categories in art. In *International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging* (Aire-la-Ville, Switzerland, 06 2008), Brown P., (Ed.), Eurographics Association, p. 1.