

Perceptually-Inspired Artistic Genre Identification System in Digitized Painting Collections

Razvan George Condorovici*, Corneliu Florea, Ruxandra Vrânceanu,
and Constantin Vertan

The Image Processing and Analysis Laboratory,
University "Politehnica" of Bucharest,
Bucharest, Romania

{rcondorovici, cflorea, rvranceanu, cvertan}@alpha.imag.pub.ro

Abstract. This paper presents an automatic system for the recognition of artistic genre in digital representations of paintings. This solution comes as part of the recent extensive effort of developing image processing solutions that facilitate a better understanding of art. As art addresses human perception, the current extracted features are perceptually inspired. While 3D Color Histogram and Gabor Filter Energy have been used for art description, frameworks extracted using anchoring theory are novel in this field. The paper investigates the possible use of 7 classifiers and the resulting performance, as evaluated on a database containing more than 3400 paintings from 6 different genres, outperforms the reported state of the art.

Keywords: Paintings, Image Classification, Artistic Genre, Anchoring Theory, 3D Color Histogram, Gabor Filters.

1 Introduction

Art has accompanied the human being while attempting to recreate “the beauty” through his entire history. With the late growth of computers usage in daily life, the art world began to be dissected by intelligent systems. Tremendous efforts were put lately into creating automatic image processing solutions in order to facilitate a better understanding of art [1], either by obtaining high quality and fidelity digital versions of art paintings [2], either by targeting subjects like image analysis and diagnostics or virtual restoration, color rejuvenation, pigment analysis, brush stroke analysis, lightning incidence, perspective anomalies detection, three dimensional space recovery, craquelure analysis or painting authentication, etc. as discussed in the review of Stork et al. [3].

* This work was supported by the Romanian Sectoral Operational Programme Human Resources Development 2007-2013 through the European Social Fund Financial Agreements POSDRU/107/1.5/S/76903 and POSDRU/89/1.5/S/64109.

A crucial aspects for artwork understanding is to successfully place it in a context, and the broadest possible is the corresponding artistic genre. The state of the work in automatic identification of the artistic genre of a painting, while witnessing noticeable results, still offers space for improvements. Our current proposal lies into this category, namely automatic artistic genre identification. Solutions to this problem can be divided in two classes: based on very simple, primitive, features and respectively using high level, complex features.

Systems with low-level features, were proposed by Günsel et al. [4] that discriminates among three genres using six basic features extracted only from the luminance image and by Zujovic et al. [5] relied on a set of gray-level features for a five genre classification. The downside of these methods is the low number of paintings used to test the systems (107 for [4] and 353 for [5]).

Acknowledging the task difficulty, the solutions from the second class introduce larger sets and higher complexity of the features. For instance, Li and Wang [6] relied on Multiresolution Hidden Markov Models to classify Chinese ink painters. Widjaja et al. [7] identified four painters based on selected skin samples with a reported accuracy of 85%. More recently, Shamir et al. [8] using an extensive set of features, discriminate among 9 schools of art within 3 artistic currents with a reported accuracy of 77% on a database of 517 images. Yet, the use of complex features opened the way for high accuracy only in the narrow cases (e.g. specific artistic identification) and within limited variation.

Motivated by the findings of Zeki et al. [9], we address the problem from a perceptual point of view and construct our features to be highly correlated with human perception. To prove the efficiency of the introduced system we test it on a un-restrictive database of ≈ 3400 paintings from 6 artistic currents yielding high within-current and cross-current variation.

The motivational overview of the proposed system and the descriptive features are presented in section 2; the data set and the classifier choice are presented in section 3. Finally, the results obtained with the proposed system are discussed in section 4, while the last section is dedicated to some conclusions.

2 Feature Extraction

The starting point in our construction is the work of Zeki [9], who showed that different elements of visual art such as shapes, colors, and boundaries are processed by different pathways and systems in the brain, designed to interpret each aspect of the art and there is no single central mechanism that receives and interpret visual art, but instead, pieces of information received from a painting are selectively redistributed to more specialized centers for processing. Next, Ramachandran and Herstein [10] identified as the key for understanding of the art perception, the identification of the perceptual processes, rather than analysis of the aesthetic properties, augmenting Zeki's statement ("the painter does not paint with his eyes, but with his brain"). Thus we divided our set of features into three categories, each closely connected with one of the mentioned perceptual elements: lightness perception and shape extraction, color distribution and

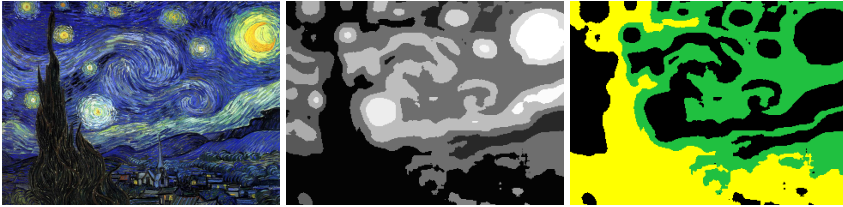


Fig. 1. Original Painting (left), the extracted frameworks (center) and the two largest frameworks marked with green/yellow (shades of gray) used in classification (right)

texture and edge analysis. For the image shapes and lightness description, we rely on the anchoring complex image decomposition derived from the gestalt (shape) theory, for the color we use the 3D *Lab* Color Histogram, while for texture and edge, the Gabor energy is employed. The features vector is obtained by concatenating the features presented in Table 1, extracted for each painting.

2.1 Anchoring Theory and Frameworks

While many studies attempted to explain and to mimic the human perception of lightness and scene decomposition, no definite model exists. Yet the reformulation by Gilchrist et al. [11] of the anchoring theory for complex scene proved to pass many perceptual tests and explained many phenomena. The anchoring theory asserts that complex scenes are perceived by the human beings in terms of consistent areas, called *frameworks*. Such a framework is defined as a region of common illumination. When it comes to image perception, the human brain estimates the lightness within each framework through the anchoring to the luminance perceived as white, followed by the computation of the global lightness. We claim that scene decomposition in frameworks is crucial for an inside view into painter scene composition procedure. Artistic genres do differentiate themselves by the scene composition technique.

In terms of images, the first computational model for the anchoring theory was provided by Krawczyk et al. in [12] for rendering high dynamic images. We follow this procedure, which relies on segmentation with mean-shift for initial decomposition, followed by bilateral filtering for removing very small anchors. A painting and the extracted frameworks are shown in Fig. 1.

While the frameworks are the shapes that holistically describe a painting, for matters of classification we extract specific parameters. Thus for each of the detected frameworks we have computed the region cardinality, articulation as defined in [12], mean framework value and region's center position (features detailed in Table 1). Our study shows that these 4 parameters for the most significant two frameworks (showing the relationship between the two most important areas of the painting) leads to optimal classification results.

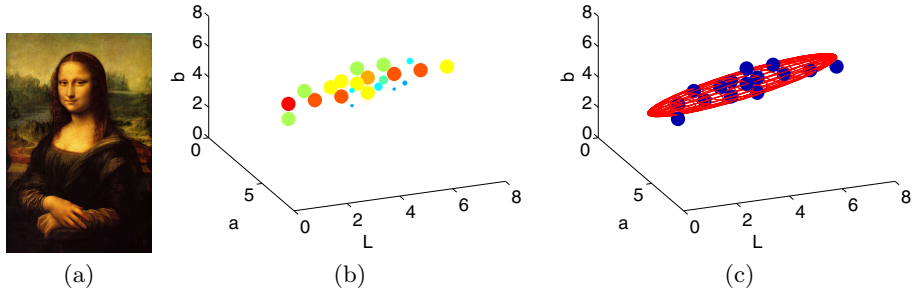


Fig. 2. Example of 3D Histogram Modelling. (a) Original input image. (b) Initial $8 \times 8 \times 8$ 3D *RGB* Histogram (redness of dots means larger bins). (c) Minimum Volume Enclosing Ellipsoid for the remaining histogram modes.

2.2 3D Color Histograms

Rappaport [13] noted that different artistic currents approximately match different historic periods and techniques known at the time and each of the currents is formally described as inspiring specific sentiments that are subsequently associated with colors. An accurate description of the color palette of a painting may be found in the 3D Color Histogram [14], which was shown to be precisely connected to a scene (thus painting) structure. The natural choice for computation is the *Lab* color space, as it exhibits the most perceptual-like inter-color distances.

For the 3D Color Histogram and the color palette description, we use a reduced set of parameters, extracted as (also exemplified in Fig. 2):

1. Compute the 3D Histogram into N^3 bins. Ignore bins smaller than certain threshold.
2. Label the histogram modes and disregard modes with fewer than a certain threshold bins. A histogram mode consists in adjacent histogram bins.
3. Keep only the biggest histogram mode and compute the Minimum Volume Enclosing Ellipsoid for the cloud of 3D points as described in [15].
4. For the obtained ellipsoid compute the descriptive parameters.

Our tests showed that keeping the largest mode is sufficient for achieving optimal classification performance. This can be explained by the fact that the main discrimination between arts movement is not necessarily offered by the utilized colors, but by the colors distribution. For example, a scarce mode containing few pixels reflects the usage of a wide range of colors, while a large and dense mode corresponds to the usage of a single main color. The ellipsoid containing the greatest number of image pixels was described through a set of parameters presented in Table 1. Various values for N (the number of histogram bins) were tested and it showed that for values higher than $N = 6$ the computation times drastically increases, while the performance remains virtually constant (less than 1 % detection rate improvement).

2.3 Gabor Filter Based Features

The complementary aspect in terms of both human perception and in art theory [16] found to be relevant for detecting a genre is edginess. Here we envisage both magnitude (as for example, in Impressionism objects tend to have softer edges, while in Cubism edges are more sharp) and orientation.

To assess edginess, we use Gabor filters, as they are known to mimic human perception in terms of contour analysis [17]. We have used two banks of Gabor filters, one symmetric and one with antisymmetric Gabor kernels, each containing 24 Gabor filters obtained with 3 different spatial frequencies and 8 orientations, as it arise as the optimal choice from several classification experiments. Combining the feature images obtained using the two filters banks, the Gabor Energy Image is obtained and the vector of Total Gabor Energy is then computed for each of the 24 vector filters.

3 Database, Features and Classifier Design

One of the most important challenges when it comes to paintings genre analysis is the lack of a standard public database [1]. We have evaluated the performance of our system on a database containing 3410 paintings belonging to 6 different art movements (Renaissance, Baroque, Rococo, Romanticism, Impressionism and Cubism), from more than 600 authors. These genres were chosen such so to span both separable and easy to discern classes (like Cubism and Renaissance) and mixed and hard to separate ones (Baroque and Renaissance). The images were acquired from various sources (e.g. scanning art albums) hence lack cohesion in acquisition conditions¹.

The classification of paintings into genres was tested for seven classifiers implemented in the open-source machine learning library Weka [18] and presented in Table 2. For testing, a 10-fold cross validation technique was used; the 10-fold rule was applied on each current. The parameters involved in the feature definition were tuned on a separate database with 600 images, uniformly distributed across all artistic genres. The detection rate for a genre is defined as the number of correctly identified images from that genre normalized to the total number of paintings of the genre. When the overall results are presented, the detection rate is taken as the averaged of the all envisaged genres.

In all the subsequent experiments, we have used the Bagging classifier which offered the best results and is also a reasonable-fast implementation.

The performance of the 3D color histogram features was evaluated for four color spaces, using these features alone. While the *HSV* and *HMMD* (which are also known to be perceptual) color spaces offered a detection rate of 49.82% respectively 49.50%, the *RGB* provided 54.45% and *Lab* histograms offered the best discrimination leading to a 55.60% detection rate, hence proving the perceptual claim.

¹ The list of paintings, as well as the code used for classification will be made available at http://alpha.imag.pub.ro/common/staff/r_condorovici/rc_paint.htm

Table 1. Features extracted from painting and used for classification

Feat. No.	Type Feature	Description	Interpretation	No. Vals.
1-2	<i>Anchor:</i> Region Cardinality	No. of pixels belonging to the current framework	Indication about the spatial relationship in the painting	2 Fmwks $\times 1=2$
3-4	<i>Anchor:</i> Articulation	Measure of the dynamic range in the framework	The luminance variation inside the current region	$2 \times 1 = 2$
5-6	<i>Anchor:</i> Mean Value	The average luminance of the current framework	Indicates the luminance value of the current region	$2 \times 1 = 2$
7-10	<i>Anchor:</i> Center pos.	Position of the region center inside the image	Indicates the position of the framework in the painting	$2 \times 2 = 4$
11-19	<i>3D Hist:</i> Directions	Principal directions of the ellipsoid, computed as the eigenvectors of the matrix representing the ellipsoid	Indication about the main colors used in the painting	9
20-22	<i>3D Hist:</i> Geometric Center	Coordinates of the ellipsoid center	Indication about the painting dominant color	3
23-25	<i>3D Hist:</i> Axis length	Length of the ellipsoid axis, computed as the eigenvalues of the matrix representation	Indication about the color gamut size	3
26-28	<i>3D Hist:</i> Center of mass	Center of mass coordinates, computed as the weighted average of the current mode bins	The painting dominant color; together with geometric center, offers information about the mode's shape	3
29	<i>3D Hist:</i> Geometric Volume	Ellipsoid's volume	Information about the gamut's size	1
30	<i>3D Hist:</i> Mass	Number of pixels belonging to the current mode	Information about the colors variation in the painting	1
31	<i>3D Hist:</i> Punctual Density	Average bin load for the current mode, computed as the ratio between the number of bins and the number of pixels belonging to the current mode	Information about the colors variation (e.g. a high percentage of pixels located in a small volume mode indicates a small number of colors used in the painting)	1
32	<i>3D Hist:</i> Volumetric Density	The mass to the geometric volume ratio	Indication about the shape of the mode	1
33	<i>3D Hist:</i> Number of modes	Number of significant modes in the histogram	Information about the color variation in the painting	1
34-57	<i>Gabor</i> Energy	The normalized bins of the Gabor energies	Magnitude of the specific orientation and band	$3 \text{ scales} \times 8 \text{ orient} = 24$

Table 2. Average Detection Rate (ADR) for tested classifiers (Logistic Regression – LR, Multilayer Perceptron – MLP, Sequential minimal optimization – SMO, Bagging – Ba, LogitBoost – LB, Decision Table – DT, Random Forest – RF). Details regarding the implementation of the classifiers are to be found in [18] and references therein.

Classifiers	LR	MLP	SMO	Ba	LB	DT	RF
ADR [%]	64.63	57.71	57.68	65.28	57.33	45.45	61.56

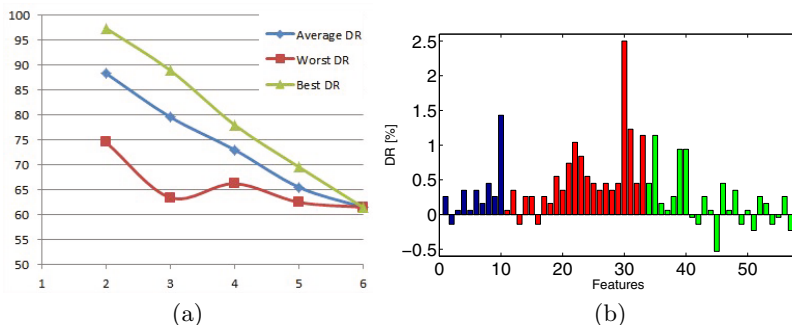


Fig. 3. (a) Average, best and worst detection rate (DR) with respect to the number of classes. (b) Individual features (as named in table 1) contribution: the amount of decrease in DR if the specified feature is removed form the system (higher values show more significant contribution).

As a replacement for Gabor energy we considered the Histogram of Oriented Gradient (HOG) as it was basis for reliable aesthetic measures [19]. Yet the Gabor Energy features offered a marginally better (+2%) detection rate.

To see the contribution of each feature, we removed each one of them and re-classified the database; results may be seen in figure 3 (b). Noting that the worst feature is one of the Gabor energies, we may conclude that redundancy is low in the proposed solution.

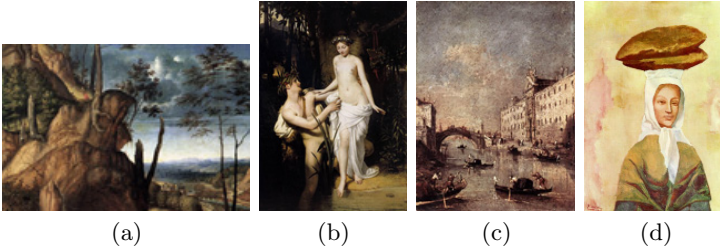
4 Results and Discussions

In the first performed experiment we assessed the behavior of our system with respect to the number of classes. As expected, the overall detection rate decreases with the increase of possible artistic genres, but the results remain acceptable for all tested genres. All possible combinations of the following six genres were tested: Renaissance, Baroque, Rococo, Romanticism, Impressionism and Cubism. In Fig. 3 the average, the lowest and the highest detection rates are presented for each possible number of classes.

It can be seen that for some numbers of classes (genres) the difference between the best and the worst detection rate can be significant. This can be explained through the genres that were chosen; if the paintings belonged to similar genres, (e.g. Baroque, Rococo and Romantism - 63.87 %) the detection rate can

Table 3. Database content and Confusion Matrix computed for a 10-fold run for the entire database

	Ren	Bar	Roc	Rom	Imp	Cub	Total	Error	DR
Renaissance	312	98	14	18	19	37	498	186	62.65
Baroque	106	428	59	65	21	13	692	264	61.84
Rococo	43	111	162	48	21	50	435	273	37.24
Romanticism	29	86	40	221	18	8	402	181	54.97
Impressionism	13	7	6	22	511	143	702	191	72.79
Cubism	18	14	7	2	102	538	681	143	79.00

**Fig. 4.** Examples of incorrectly classified paintings. (a) Renaissance as Rococo: L. Lotto "St. Jerome in the Desert" (b) Romanticism as Baroque: J.D. Court "Young Girl at the Scamander River" (c) Rococo as Romanticism: F. Guardi "Rio dei Mendicanti" (d) Cubism as Impressionism: P. Picasso "Woman with Loaves".

be smaller, while if the paintings belonged to more distinct genres (Cubism, Renaissance and Romantism - 88.97 %), a higher detection rate is obtained.

Figure 4 shows examples of incorrectly classified paintings. In general, the miss-classification occurs between more similar genres, harder to discriminate even for a human user. The similarities between different artistic genres can also be observed from the confusion matrix presented in Table 3.

4.1 Comparison with State of the Art

The main difficulty in comparing the proposed approach with state of the art solutions [5], [4], [8] is the use of different databases. We stress that our database is ≈ 10 times larger than any reported solution. Each author has used, apart the different number of image examples per class, different classes. The exact number of genres, as well as the number of paintings in the database used and the average detection rate may be seen in Table 4. To directly compare the performance we computed how much better is one than random guessing, so we normalized the reported DR by chance (i.e. $1/\text{no of genres}$). One may easily see that we outperform state of the art solutions in terms of accuracy.

To see how other solutions performed on our database, we implemented and tested the solution from [4]. Upon testing the solution on our extensive database we obtained rather different results from the ones reported by the authors on

Table 4. Comparison with state of the art: number of art movements, number of paintings, detection rate and detection rate normalized by chance. Higher values points to better results.

Author	Genres	Database	DR	DR/random chance
<i>Proposed</i>	6	3419	65.28%	3.9
Gunsel et al. [4]	3	107	91.66%	2.75
Zujovic et al. [5]	5	353	68.3%	3.41
Shamir et al. [8]	3	517	77%	2.31
Non-experts humans [20]	11	275	68%	7.48

Table 5. Detection rate comparison between the proposed solution and other similar solutions from the recent literature; all tests are performed on the current 3410 image database.

Genre	Ren	Bar	Roc	Rom	Imp	Cub	Overall
<i>Proposed</i>	62.65	61.84	37.24	54.97	72.79	79.00	65.28
Gunsel et al. [4]	47.2	34.3	13.8	20.3	49.2	48.5	35.5

their database. On our database, Gunsel’s solution offered an overall DR of 35%. The explanation lies in the larger database used here with more paintings, acquired with more variability. The results of the comparison are shown in Table 5, proving that, indeed, our solution outperforms the solution in [4].

Wallraven et al. showed in [20] that non-expert human users can perform a better classification than automatic systems, but the performance is still far from perfect, as can be seen in 5.

While computation time is, at current stage of the project, not of interest, we note that a painting classification query takes in average 5400 msec on Intel 2.7 GHz in Matlab code, out of which ≈ 5000 are required by the bilateral filtering.

5 Conclusion

In this paper we proposed a perceptually inspired system for automated analysis of paintings and applied for discrimination among six artistic currents. While the feature extracted were specifically selected as being relevant for human perception, the classifier was selected by comparison. The system proved to outperform state of the art methods in terms of classification accuracy. Although the results are satisfactory, a 100% DR is practically impossible to achieve yet, as long as the separation between some genres is not always very clear even for art historians.

As continuation paths, we envisage two directions. First the system has to be further refined for even more improve accuracy and tested for more currents and paintings. Also in the same direction the procedure is useful to tackle problems like specific painter description or discrimination among different painters. The second direction is identifiable taking into account that our system is perceptually oriented and thus it may be used for other kind of images (e.g. natural, comics streams) analysis.

References

1. Cornelis, B., Dooms, A., Cornelis, J., Leen, F., Schelkens, P.: Digital painting analysis, at the cross section of engineering, mathematics and culture. In: Proc. of EUSIPCO, pp. 1254–1259 (2011)
2. Martinez, K., Cupitt, J., Saunders, D., Pillay, R.: Ten years of art imaging research. Proceedings of the IEEE 90(1), 28–41 (2002)
3. Stork, D.G.: Computer vision and computer graphics analysis of paintings and drawings: An introduction to the literature. In: Jiang, X., Petkov, N. (eds.) CAIP 2009. LNCS, vol. 5702, pp. 9–24. Springer, Heidelberg (2009)
4. Günsel, B., Sariel, S., Icoğlu, O.: Content-based access to art paintings. In: Proc. of ICIP, pp. 558–561 (2005)
5. Zujovic, J., Gandy, L., Friedman, S., Pardo, B., Pappas, T.: Classifying paintings by artistic genre: An analysis of features & classifiers. In: Proc. of IEEE MMSP, pp. 1–5 (2009)
6. Li, J., Wang, J.: Studying digital imagery of ancient paintings by mixtures of stochastic models. IEEE Trans. on Image Processing 13(3), 340–353 (2004)
7. Widjaja, I., Leow, W.K., Wu, F.: Identifying painters from color profiles of skin patches in painting images. In: Proc. of ICIP, pp. 845–848 (2003)
8. Shamir, L., Macura, T., Orlov, N., Eckley, D.M., Goldberg, I.G.: Impressionism, expressionism, surrealism: Automated recognition of painters and schools of art. ACM Transactions on Applied Perception 7(2), 1–17 (2010)
9. Zeki, S.: Inner Vision. Oxford University Press (1999)
10. Ramachandran, V., Herstein, W.: The science of art: A neurological theory of aesthetic experience. Journal of Consciousness Studies 6, 15–51 (1999)
11. Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., Spehar, B., Annan, V., Economou, E.: An anchoring theory of lightness perception. Psychological Review 106(4), 795–834 (1999)
12. Krawczyk, G., Myszkowski, K., Seidel, H.P.: Lightness perception in tone reproduction for high dynamic range images. In: Proc. of EUROGRAPHICS, Computer Graphics Forum., vol. 24 (2005)
13. Rappaport, A., Rapaport, A.: Color preferences, color harmony, and the quantitative use of colors. Empirical Studies of the Arts 2(2), 95–111 (1984)
14. Novak, C.L., Shafer, S.: Anatomy of a color histogram. In: Proc. of CVPR, pp. 599–605 (1992)
15. Moshtagh, N.: Minimum volume enclosing ellipsoid. Convex Optimization (2005)
16. Melcher, D., Cavanagh, P.: Pictorial cues in art and in visual perception. In: Art and the Senses, pp. 359–394. Oxford University Press (2011)
17. Daugman, J.: Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. Journal of the Optical Society of America A 2(7), 1160–1169 (1985)
18. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: an update. ACM SIGKDD Explorations Newsletter 11(1), 10–18 (2009)
19. Redies, C., Amirshahi, S.A., Koch, M., Denzler, J.: PHOG-derived aesthetic measures applied to color photographs of artworks, natural scenes and objects. In: Fusiello, A., Murino, V., Cucchiara, R. (eds.) ECCV 2012 Ws/Demos, Part I. LNCS, vol. 7583, pp. 522–531. Springer, Heidelberg (2012)
20. Wallraven, C., Fleming, R.W., Cunningham, D.W., Rigau, J., Feixas, M., Sbert, M.: Categorizing art: Comparing humans and computers. Computers & Graphics 33(4), 484–495 (2009)