Recognizing Surreal Compositions in Digitized Paintings

Laura Florea(Corresponding author)[1,*], Corneliu Florea[1] and Mihai Badea[1]
[1]*University Politehnica of Bucharest, Image Processing and Analysis Laboratory, Bucharest, Romania*
[*]Corresponding author (E-mail : laura.florea@upb.ro)

***Abstract*—This paper addresses the problem of recognizing illogical object juxtaposing in the specific form of classifying digitized paintings in art movements. More precisely we distinguish between realism and surrealism movements. We propose a system based on feature extraction and machine learning that is able to understand the scene in the digitized paintings and to classify the art works from the two movements. We will show that global GIST features being able to access the image content, complemented by random forest classifier, give good accuracy for this task.**

***Key words* - Painting classification; Realism; Surrealism; GIST; Random Forest.**

I. Introduction

Due to the remarkable expansion of the Internet and of digital cameras during the recent years, almost anyone can have much easier access to digitized versions of different works of art. Many efforts have been done to digitize fine art with high fidelity ([14], [18], [24]) and now art experts have an unique opportunity to analyze art works at a larger scale. The scientific community realized that a computer based solutions for analyzing large collections of paintings from various artists or different painting styles is needed. These solutions have not only a theoretical importance, but also a great practical significance, especially for the art community. The usage of advanced algorithms and image analysis techniques would allow the study of paintings in more details than a typical art historian could do. By exploiting even finer features and structures that normally would not be accesible for a human expert, the usage of computer based solutions make possible a new level of analysis. The numerical features give by those solutions could be used for comparing paintings, painters or even painting styles or schools from different countries and time periods. Also automatic classification and annotation of large collections of images for retrieval purpose can be achieved. Those are the reasons why many research has been done in this area in the recent years.

In order to understand a work of art, one should place it into the appropriate context, which, in the case of paintings, means the artistic genre or art movement. This task is easy to be done in some cases but sometimes there are no clear borders between different genres. In this case art historians group the visual art creations by common elements used by artists. Usually painters belonging to the same art movement rely on the same techniques and approach the same themes [1].

A. Related work

Artistic genre recognition is sometimes quite difficult even for experts due to variations within a current. To address this task, automatically two types of systems were previously used: relying on low level features (such as pixel luminance or color, number and sharpness of edges, etc) and relying on high level features.

Gunsel et al. [8] propose a system with low level features that discriminate among three genres (Classicism, Cubism and Impressionism) using six basic features extracted only from the luminance image. Zujovic et al. [25] use a set of gray-level features for a five genre classification (Abstract Impressionism, Cubism, Impressionism, Pop Art and Realism). These methods used a low number of paintings to test the systems (107 paintings for [8] and 353 paintings for [25]). Condorovici et al. [5] extract three categories of features (composition, color and edginess) described paintings from a database with more than 5000 digital painting images from six different genres (Baroque, Renaissance, Rococo, Romanticism, Impressionism and Cubism).

The second type of system acknowledge the difficulty of the task. In this case larger databases and higher complexity features are introduced. Shamir et al. [19] rely on an extensive set composed of 548 features. Using a Fisher criterion filtering the most discriminative 83 ones were selected and coupled with a weighted nearest neighbor classifier in order to discriminate among 9 schools of art within 3 artistic currents (Impressionism, Expressionism and Surrealism). The reported accuracy was of 77% on a database of 517 images. Classemes [21] framework are used by Arora and Elgammal [2] to describe paintings and classify them in seven currents (Renaissance, Baroque, Impressionism, Cubism, Abstract, Expressionism and Popart) using a Bag of Words schema with Support Vector Machine classifier. A more complex set of features that mimics the human perception was used by Condorovici et al. [6]. Six art movements (Renaissance, Baroque, Rococo, Impressionism, Cubism, Romanticism) are analyzed.

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. One can notice that in all those cases, the genres are quite different. In the current paper we intend to distinguish between two art movements that are quite similar: Realism and Surrealism. The movements are often similar in terms of colour palette and texture yet they do differ in terms of composition, as surreal paintings depict non-natural facts and objects. Surrealist artists painted illogical scenes, sometimes with photographic precision, created strange creatures from everyday objects [3]. By contrast, realism artists painted scenes as close as possible to the reality. Examples of paintings from the two currents can be seen in Fig. 1. Consequently, the automatic system that has to make the difference between these two art movements should be able to detect the illogical juxtapositions from the surrealist paintings, thus it has to understand the scene.

Thus the remainder of the painting is organized as follows: in section 2 the used features are discussed. The database and the classification methods are presented in section 3, while experiments and results are presented in section 4. Finally, the last section is dedicated to conclusions.

II. Features Extraction

The solution proposed in this paper is based on the classical approach: first relevant features are extracted from digitized paintings, then a classifier is chosen to discriminate between categories. In this section the set of extracted features will be presented. Because the art movements that we want to separate are quite similar, one needs features that are able to understand the global scene, not features based only on luminance, colors, number of edges or their orientation.

A. Histogram of Oriented Gradients(HOG) and Pyramid Histogram of Oriented Gradients (PHOG)

The Histogram of Oriented Gradients (HOG) has been introduced by Dalal and Triggs [7] for pedestrian detection and has been used in Computer Vision and Image Processing for a lot of other tasks proving to be a very useful feature. The HOG descriptor technique counts occurrences of gradient orientations in localized portions of an image. It decomposes the image into small cells, computes the histogram of oriented gradients in each cell, normalizes the result using a block-wise pattern, and returns the descriptor for each cell.

Relying on HOG and the image pyramid representation, Bosch et al. introduced the Pyramid Histogram of Oriented Gradients (PHOG) [4]. The method divides the image into a sequence of increasingly finer spatial grids. A HOG vector is computed for each grid cell at each pyramid resolution level. The final PHOG descriptor for the image is a concatenation of all the HOG vectors.

B. GIST descriptor

Another global descriptor that received increasing attention in the context of scene recognition is the GIST descriptor.

The GIST descriptor was first proposed by Oliva and Torralba [16], [17] and was named Spatial Envelope. The idea was to develop a model based on a very low representation of the image for the recognition of real scenes without using any form of segmentation. In order to do this, a set of perceptual dimensions are defined: naturalness, openness, roughness, expansion, ruggedness. These dimensions represent the dominant spatial structure of the scene. The authors show that these dimensions may be reliably estimated using spectral and coarsely localized information. The image is divided in a 4 by 4 grid. For each of the regions the orientation histograms are extracted. These low-level features quantify higher-level semantic properties of the scene and ignore the local objects in the scene.

The GIST descriptor was successfully used for retrieving images of the same landmarks by Li et al. [12] or for image completion by Hayes and Efros [9]. Torralba, Weiss and Fergus [20], [23] developed different strategies to compress the GIST descriptor.

C. Local Binary Patterns (LBP) and Pyramid Local Binary Patterns (PLBP)

Local Binary Pattern (LBP) was introduced by Ojala et al. [15] is a simple yet very efficient texture operator. In its simpler form, it assigns a label to every pixel of an image by thresholding the 3×3-neighborhood of each pixel with the center pixel value and considering the result as a binary number. Next, the possible values are quantized to a set of 58 values by clustering together the values corresponding to more uniform patches.

By cascading the LBP information of hierarchical spatial pyramids, PLBP descriptors take texture resolution variations into account [10],[11].

III. Database and classifiers

We have evaluated our system on a database containing a total of 549 digitized paintings. There are 307 paintings from 8 artists belonging to the realism movement and 242 paintings from 5 artists belonging to surrealism movement. The images were acquired from various sources (e.g. scanning art albums, Internet) hence deliberately lacking cohesion in the acquisition conditions.

A. Classifier design

The classification of paintings into the two art movements was tested for two popular types of classifiers: Nearest Neighbor (NN) based on standard Euclidean distance and random forest (RF).

For testing, a 4-fold cross validation technique was assumed; the 4-fold rule was applied for each artistic movement.

B. Performance measure

The detection rate (DR) for each movement is defined as the number of correctly identified images from the given movement normalized to the total number of paintings of the movement.

IV. Results

In order to asses the optimal features and classifiers we performed a series of experiments. The overall detection rates in each case can be seen in Table I. As expected the simple, local descriptors give poor results, not being able to make the difference between similar scenes that differ only by meaning. The LBP descriptor gives marginally better results than HOG. However by complementing the local descriptors with

Figure 1. Example of digital paintings from the database, belonging to realism (upper row) and surrealism (lower row).

Table I
AVERAGE DETECTION RATE FOR THE TESTED CLASSIFIERS: NEAREST NEIGHBOR WITH 1,3 AND 7 NEIGHBORS (1–NN, 3–NN AND 7–NN RESPECTIVELY), AND RANDOM FOREST (RF) AND THE FEATURES: HISTOGRAM OF ORIENTED GRADIENTS(HOG), PYRAMID HISTOGRAM OF ORIENTED GRADIENTS (PHOG), SPATIAL ENVELOPE (GIST), LOCAL BINARY PATTERNS (LBP) AND PYRAMID LOCAL BINARY PATTERNS (PLBP)

Classifier	**Features**				
	HOG	***PHOG***	***GIST***	***LBP***	***PLBP***
1–NN	58.11	59.95	68.91	59.73	67.82
3–NN	59.01	60.65	70.61	59.92	67.58
7–NN	57.21	62.29	72.21	59.89	67.86
RF	58.28	64.85	73.73	62.46	71.28

hierarchical spatial pyramids, the detection rate is increased since the spatial structure and thus content is accessed too. The global descriptor GIST gives the best result no matter the used classifier.

As for the classifiers, the random forest gives the top accuracy, but the nearest neighbor, based on simple Euclidean distance, is not far behind. We must stress that for this test we used the implicit random forest parameters without any tuning (100 trees and the number of feature variables - dimensions - to select at random for each decision split is set to the square root of the total number of variables).

Given the results in Table I one may note that the GIST descriptor and the PLBP features give comparable good results.

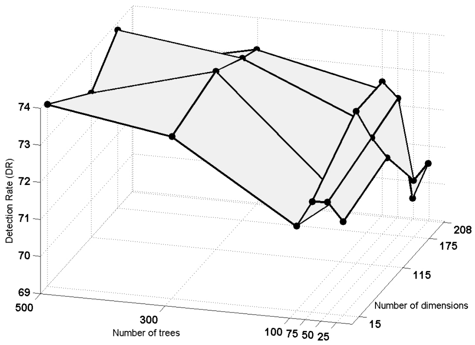
Table II
CONFUSION MATRIXES FOR EACH TYPE OF FEATURE USED.

Features	**PLBP**		**GIST**	
Realism	220	87	235	72
Surrealism	57	185	50	192

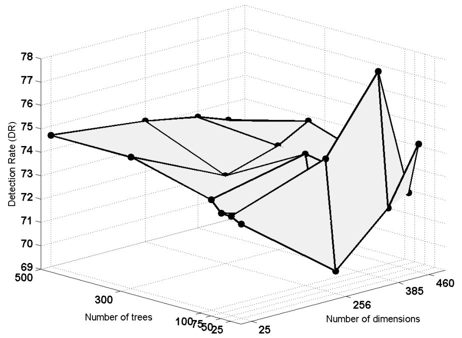
So we further test these two types of features, along with the classifier that gave the best results: random forest.

We search for the best parameters for the random forest classifier in the case of the GIST and PLBP features separately. The results can be seen in Fig. 2 a) and b). If the GIST descriptor is used, the best obtained detection rate was 77.79% and is obtained for 50 trees and 385 randomly selected dimensions. However results over 74.5% are obtained also for smaller number of trees and randomly selected dimensions (e.g. a detection rate of 74.65% is obtained for 50 trees and 256 dimensions). In the case of PLBP features the best obtained detection rate is 73.90% and is obtained for 500 trees and 15 randomly selected dimensions.

In order to asses how the proposed solution behaves for each tested art movement, we computed the confusion matrix which is detailed in Table II for each type of feature. One can see that the errors are almost equally distributed between the two art movements. Also the behavior of the two features is similar.



(a)



(b)

Figure 2. Achieved detection rates when the parameters of random forest classifier were searched for in the case of PLBP features (a) and GIST descriptor (b) respectively.

V. Conclusions

In this paper we proposed a method to automatically classify digitized paintings belonging to realism and surrealism movements. The extracted features were specifically selected to describe the meaning of the painting, not the local context of objects inside it. We tested our system on a database containing similar number of images for each movement as the ones used in other state of the art papers. The achieved results are satisfactory for now, considering that a detection rate of 100% is impossible to achieve even for art historians as long as the separation between the movements is not always very clear.

As continuation paths we envisage two directions. First, the system has to be further refined for even more improved accuracy and tested for more paintings in each movement. Since the surrealism movement has images that resemble also the abstract one, this third genre may be added to the system.

Acknowledgements

This work was supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNCS - UEFISCDI, project number PN-II-RU-TE-2014-4-0733.

References

[1] L. S. Adams, The Methodologies Of Art: An Introduction, Westview Press, 1996.

[2] R.S. Arora, A. Elgammal, "Towards automated classification of fine-art painting style: a comparative study," Proceedings of the International Conference on Pattern Recognition (ICPR), 2012, pp. 3541–3544.

[3] R. Barnes et al, Twetieth-century art book, Phaidon Press, ISBN-10: 0714838500, 1999.

[4] A. Bosch, A. Zisserman, X. Munoz, " Representing Shape with a Spatial Pyramid Kernel, " Proceedings of the ACM International Conference on Image and Video Retrieval, 2007.

[5] R. Condorovici, C. Vertan, L. Florea, "Artistic genre classification for digitized painting collections," UPB Scientific Bulletin, Vol. 75, Issue 2, 2013, pp. 75–86.

[6] R. Condorovici, C. Florea, C. Vertan, " Automatically classifying paintings with perceptual inspired descriptors,"Journal of Visual Communication and Image Representation, 2015, vol.26 pp. 222–230.

[7] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection," Proceedings of the Computer Vision and Pattern Recognition (CVPR), 2005, pp. 886–893.

[8] B. Gunsel, S. Sariel, O. Icoglu, "Content-based access to art paintings," Proceedings of International Conference on Image Processing (ICIP), 2005, pp. 558–561.

[9] J. Hayes and A. Efros, "Scene completion using millions of photographs,"ACM Transactions on Graphics, Vol. 26, Iss. 3, 2007, Art no. 4.

[10] Y. He, N. Sang, C. Gao, "Pyramid-Based Multi-structure Local Binary Pattern for Texture Classification," Proceedings of the Asian Conference on Computer Vision (ACCV), Vol. 6494 LNCS, 2010, pp. 133–144.

[11] X. Qian, X.S. Hua, P. Chen, L. Ke, "PLBP: An effective local binary patterns texture descriptor with pyramid representation," Pattern Recognition, Volume 44, Issues 10–11, 2011, pp. 2502–2515.

[12] X. Li, C. Wu, C. Zach, S. Lazebnik, and J.M. Frahm, "Modeling and recognition of landmark image collections using iconic scene graphs," Proceedings of the European Conference on Computer Vision (ECCV), 2008, pp. 427–440.

[13] C.C Chang, C.J. Lin, "LIBSVM : a library for support vector machines," ACM Transactions on Intelligent Systems and Technology, Vol. 2, Issue 3, pp.27:1–27:27, 2011, Software available at http://www.csie.ntu.edu.tw/˜cjlin/libsvm.

[14] K. Martinez, "High resolution digital imaging of paintings: the vasari project," Microcomputers for Information Management, Vol. 8, Issue 4, 1991, pp. 277–283.

[15] T. Ojala, M. Pietikainen, D. Harwood, "A comparative study of texture measures with classification based on feature distributions," Pattern Recognition, vol. 29, no. 1, pp. 51–59, 1996.

[16] A. Oliva, A. Torralba, "Modeling the shape of the scene: a holistic representation of the spatial envelope," International Journal of Computer Vision (IJCV), Vol. 42, Issue 3, 2001, pp.145–175.

[17] A. Oliva, A. Torralba, "Scene-centered description from spatial envelope properties," Proceedings of the Second International Workshop on Biologically Motivated Computer Vision, BMCV, Springer-Verlag, 2002, pp. 263–272.

[18] A. Ribes, F. Schmitt, R. Pillay, C. Lahanier, "Calibration and spectral reconstruction for crisatel: an art painting multispectral acquisition system," Journal of Imaging Science and Technology Vol. 49, Issue 6 , 2005, pp. 563–573.

[19] L .Shamir, T. Macura, N. Orlov, D.M. Eckley, I.G. Goldberg, "Impressionism, expressionism, surrealism: Automated recognition of painters and schools of art," ACM Transactions on Applied Perception, Vol. 7, Issue 2, 2010, pp. 1–17.

[20] A. Torralba, R. Fergus, Y. Weiss, "Small codes and large databases for recognition," Proceedings of the Computer Vision and Pattern Recognition (CVPR), 2008.

[21] L. Torresani, M. Szummer, A. Fitzgibbon, "Efficient object category recognition using classemes, " Proceedings of the European Conference on Computer Vision (ECCV), 2010, pp. 776–789.

[22] Z. Wang, B. Fan, F. Wu, "Local intensity order pattern for feature description," Proceedings of the International Conference on Computer Vision (ICCV), pp. 603–610 2011.

[23] Y. Weiss, A. Torralba, R. Fergus, "Spectral hashing," Advances in Neural Information Processing Systems, pp.1753–1760, 2009

[24] A. Woll, J. Mass, C. Bisulca, R. Huang, D. Bilderback, S. Gruner, N. Gao, "Development of confocal x-ray fluorescence (xrf) microscopy at the cornell high energy synchrotron source," Applied Physics A Vol. 83, Issue 2, 2006, pp. 235–238

[25] J. Zujovic, L. Gandy, S. Friedman, B. Pardo, T. Pappas, "Classifying paintings by artistic genre: An analysis of features and classifiers," IEEE Proceedings of Multimedia Signal Processing, 2009, pp. 1–5.