J. Vis. Commun. Image R. 26 (2015) 222–230

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Automatically classifying paintings with perceptual inspired descriptors ☆

Răzvan George Condorovici, Corneliu Florea *, Constantin Vertan

The Image Processing and Analysis Laboratory (LAPI), University "Politehnica" of Bucharest, Splaiul Independeţei, 313, Bucharest, Romania

ARTICLE INFO

Article history:
Received 21 January 2014
Accepted 24 November 2014
Available online 2 December 2014

Keywords:
Painting
Perceptual image features
Genre recognition
Anchoring theory
Dominant Color Volume
Gabor filters
Image classification
Late fusion

ABSTRACT

We propose a framework for the automatic recognition of artistic genre in digital representations of paintings. As we aim to contribute to a better understanding of art by humans, we extensively mimic low-level and medium-level human perception by relying on perceptually inspired features. While Gabor filter energy has been used for art description, Dominant Color Volume (DCV) and frameworks extracted using anchoring theory are novel in this field. To perform the actual genre recognition, we rely on a late fusion scheme based on combining Multi-Layer Perceptron (MLP) classified data with Support Vector Machines (SVM). The performance is evaluated on an extended database containing more than 4000 paintings from 8 different genres, outperforming the reported state of the art.

1. Introduction

George Bernard Shaw said that "without art, the crudeness of reality would make the world unbearable" acknowledging that art has accompanied the human evolution through his entire history. With the late growth of computers usage in daily life, the art world began to be dissected by artificial, intelligent systems. Tremendous efforts were put lately into creating automatic image processing solutions that facilitate a better understanding of art [1], either by obtaining high-quality and high-fidelity digital versions of paintings [2], either by targeting subjects like image analysis and diagnostics, 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 [3]. While gathering more than 20 years of intensive research, digital investigation of visual art has not yet answered all questions.

A crucial aspects for artwork understanding is to successfully place it into a context. Typically, two cases are envisaged: a narrower one which is to nominate the painter and a broader one, namely recognizing the artistic genre. The state of the art in automatic identification of the context of a painting, although witnessed noticeable results, still offers space for improvements. The current proposal lies into the second category, as we describe a system for automatic identification of artistic genres. We consider that previous attempts tackled the problem from a computer vision point only, ignoring a perceptual point of view, in which features and machine learning systems are build to match the human perception.

1.1. Related work

In reviewing solutions to the artistic context for recognition problem (both painter and genre), we identify the very typical pattern recognition approach: first, using features, digitized paintings are described, than a learning scheme is employed to extract common and respectively discriminative traits among envisaged classes. A condensate overview of the state of the art methods may be followed in Table 1.

Automatic identification of the painter proved to be more popular in the early stages. Thus, adopting the cosine transform for extract repetitive texture features linked to a Naive Bayes Classifier (NBC), Keren [4] identified several painters. In the same line, Li and Wang [5] used 2-dimensional Multiresolution Hidden Markov Models (MHMM) over wavelet extracted features to classify five Chinese ink painters. Widjaja et al. [6] identified four painters based on selected skin samples (described from both color and

☆ This paper has been recommended for acceptance by M.T. Sun.

* Corresponding author.
E-mail addresses: rcondorovici@imag.pub.ro (R.G. Condorovici), corneliu.florea@upb.ro (C. Florea), constantin.vertan@upb.ro (C. Vertan).

http://dx.doi.org/10.1016/j.jvcir.2014.11.016
1047-3203/

texture point of view) with a reported accuracy of 85%. More recently, Khan et al. [7] combined color and shape information in a Bag of Word (BoW) approach to recognize among 10 painters out of 400 images. Yet, taking into account that a painter is consistently more conservative in the approached themes and in the techniques than his peers, even from the same current, the painter recognition is rather less intricate when compared with the genre recognition, which requires an extended level of abstraction. Furthermore, due to the finite work capacity of any human, the amount of paintings authored by a single artist is also limited, thus in painter recognition cases, the database are confined to less than 500 examples (with a maximum of 50 examples per class/painter).

The other direction of the context recognition, namely the artistic genre recognition, is more difficult, sometimes even for the specialists, due to the natural variation within the artistic genres. In this direction, we identify two types of systems: relying on low level features (such as pixel luminances and color means or as total edginess) and relying on high level features. Systems with low-level features were proposed by Gunsel et al. [8], which dissociate three genres based on six basic features extracted only from the luminance image and by Zujovic et al. [9] who relies on a set of gray-level features for a five-genre classification. The downside of these methods is the reduced number of paintings used to test the systems (107 paintings for [8] and 353 paintings for [9]).

Acknowledging the task difficulty, the solutions from the second class introduce larger sets and higher complexity of the features. More recently, Shamir et al. [10] adopted an extensive set of 548 features, out of which, by means of the Fisher criterion filtering, selected the most discriminative 83 ones, coupled with a weighted nearest neighbor (WNN) classifier; as a result they discriminated among 9 schools of art within 3 artistic currents for a reported accuracy of 77% within a database of 517 images. In the same line, Arora and Elgammal [11] described paintings with Classemes introduced by Torresani et al. [12] framework and distributed them in 7 currents by means of a Bag of Words (BoW) schema with a Support Vector Machine (SVM) as classifier. Yet, the use of complex features opens the way for high accuracy only in narrow cases (e.g. specific artistic identification) and within confined variation.

In all the mentioned methods, the results are somehow restricted in generalization due to the limited size of the database (i.e. less than 1000 examples).

1.2. Paper structure

To motivate our construction we recall that Michelangelo wrote down in Middle Ages that "a man paints with his brains and not with his hands". Furthermore, although computer based discrimination among artistic genre is difficult, Wallraven et al. [13] noted that non expert humans still achieve considerably larger scores than computers. Thus, we claim that the key to better accuracy is to rely on features compatible with human perception.

We addressed the problem from a perceptual point of view and we constructed the descriptors to be highly correlated with human perception, thus encoding the major classes of perceptual features: luminance and shapes, color and, respectively, texture and edge.

To ensure proper coverage of this problem, we propose a new color descriptor named Dominant Color Volume and for the dominant luminance levels, we introduce the anchoring theory into the art digital analysis. For recognition, we employed a late fusion scheme, as the human process first each category of data and then aggregate the results. The efficiency of the proposed system is tested on a un-restrictive database of some 4200 paintings from 8 artistic genres yielding high within-current and cross-current variation.

In the continuation of this paper, the motivational overview of the proposed system and the descriptive features are presented in Section 2; the data set and the classifier details are given in Section 3. Finally, the results obtained with the proposed system are discussed in Section 4, while the last section is dedicated to conclusions and perspectives.

2. Feature extraction

There were many attempts to unravel the human understanding of art from a neuro-scientific point of view. The first significant results were disclosed by Zeki [14], 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.

Ramachandran and Herstein [15] identified as the key for understanding the art perception to be the identification of the perceptual processes, rather than the analysis of the aesthetic properties, augmenting Zeki's tweak on Michelangelo 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 important perceptual elements: lightness perception and shape extraction, color distribution and, respectively, texture and edge analysis.

For the image shapes and lightness description we relied on the anchoring complex image decomposition, derived from the gestalt (shape) theory; for the color, we computed the Minimum Volume Enclosing Ellipsoid over the 3D *Lab* color histogram to get the Dominant Color Volume (DCV), while for textures and edges we employed the Gabor energy. These features are presented in Table 2 and are extracted for each painting.

2.1. Anchoring theory and frameworks

Although many studies attempted to explain and to mimic the human perception of lightness and scene decomposition, no

Table 1
Artistic genre/painter recognition methods: main differences.

Method	Recognizes	No. of classes	Desc. level	Features	Learning scheme
Keren [4]	Painter	5	High	Spectral (Cosine)	Naive Bayes
Li and Wang [5]	Painter	5	High	Wavelet	2D-MHMM
Widjaja et al. [6]	Painter	4	High	Color, Skin texture	SVM
Khan et al. [7]	Painter	10	Low	Color, Shape	BoW
Gunsel et al. [8]	Genre	3	Low	Luminance, Color	PCA-SVM
Zujovic et al. [9]	Genre	5	Low	Texture, Edge, Color	AdaBoost
Shamir et al. [10]	Genre/painter	3/9	High	Edge, Texture	WNN
Arora [11]	Genre	7	High	Classemes	BoW/SVM
Condorovici et al. [34]	Genre	6	High	Dominant Color, Anchors (Shape), Gabor	Bagged ensem-ble of trees
Proposed	Genre	8	High	Dominant Color, Anchors (Shape), Gabor	Late fusion