Computer Vision Models to Categorize Art Collections According to Visual Content

A New Approach to the Abstract Art of Antoni Tàpies

PILAR ROSADO

This study uses computer vision models, which to some extent simulate the initial stages of human visual perception, to help categorize data in large sets of images of artworks by the artist Antoni Tàpies. The images have been analyzed on the basis of their compositional, chromatic and organizational characteristics, without textual notes, so that the analogies found may take us closer to, and help us to understand, the creator's original values. The system as programmed can assist the specialist by establishing analogies between different artists or periods using the same criteria.

STATE OF THE ART

The need of semantic image classification is becoming increasingly important to support effective image database indexing and retrieval. The goal of the project described here will be to develop an image classification system that reduces the amount of manual supervision required as well as the computational cost of training the classifier.

This application is novel and the results extend previous findings on the use of these models to facilitate pattern recognition in image collections of natural scenes (with photographs of landscapes, interiors, cityscapes) [1], object detection [2] and abstract images of natural items (with photographs of water, sky and stones) [3,4]. Computer vision has previously been employed to classify artists' paintings, in all cases applying machine learning techniques to prior classifications performed by art experts [5,6].

The method developed to establish a visual taxonomy is totally automated and requires no previous intervention. However, at the time of the project's execution, no references were found in the literature regarding the use of these models to study patterns in abstract painting, and this positive evaluation should contribute to confirming their value in this field.

We suggest the need to develop a new way of studying the collections of art institutions based on the possibilities that computer vision techniques place at our disposal, and that this alternative point of view will increase the interest of the general public and experts alike.

METHODOLOGY; CONSTRUCTION OF A VISUAL VOCABULARY

Here we explain how information regarding the global characteristics of an image can be extrapolated purely on the basis of a specific group of pixels.

In 1997, the neurobiologist Tanaka [7,8] provided new information about the neural mechanisms of visual object recognition in primates. Considerable evidence shows that object recognition makes use of neurons in the lower temporal cortex that respond to features of intermediate complexity. These features are invariant with a wide range of changes in scale, location and illumination, while being very sensitive to particular combinations of color, texture and local shape properties.

In 2000, the computer vision expert Lowe [9,10] described a computer vision system for performing object recognition that achieved feature integration in a manner similar to the process of serial visual attention shown to play an important role in object recognition in human vision. The approach is called a scale invariant feature transform (SIFT) (see supplemental Appendix 1 online). The features describe image regions around a node using a descriptor computed from a 16×16-pixel array. The computation of SIFT local descriptors is performed on the nodes of a regular grid superimposed on the image [11,12]. Considering a generic collection of images (Fig. 1a), a large collection of SIFT descriptors is available (Fig. 1b).

In order to build up a "visual vocabulary" on which to base image description, we followed a procedure similar to that used in automatic text analysis, known as the "Bag-of-Words" (BOW) model, because every document is represented as a distribution of frequencies of the words in the text, without considering the syntactic relationships among them. In the field of images, we will refer to such representations as a "Bag-of-Visual-Terms" (BOV) model [13–16].
(see Appendix 2). Construction of the "visual vocabulary" is performed by means of clustering. More specifically, the K-means algorithm [17] (see Appendix 3) is applied to a representative set of SIFT descriptors obtained from the image collection, and the vectors of means of each cluster are taken as "visual words" (Fig. 1c). We use the Euclidean distance in the grouping and quantization processes, and the number of clusters is chosen depending on the desired size of the vocabulary.

Thus, every image is represented as a distribution of frequencies of the "visual words" in the collection, without considering the spatial relationships among them. In order to complete this approach, we have implemented probabilistic latent semantic analysis (PLSA), a methodology originated through text mining [18]. Extending PLSA to image analysis involves considering images as documents in a visual vocabulary established by means of a quantization process, as previously mentioned [19]. PLSA is a generative statistical model that associates a latent aspect according to the occurrence of a visual word in a given image. The PLSA will detect categories of objects or formal patterns within the images based on the co-occurrence of certain "visual words." For example, there is, in every frontal portrait painting, the concurrence of two eyes, a nose, a mouth and a forehead. The system will make a group of all the images containing a frontal portrait under the heading "portrait" (Fig. 1d).

An added benefit of the categorization of the actual collection itself is that, once the classes have been established that are contained within the whole, the system is capable of learning and, in the case of a problem image, the system will be able to determine the latent aspect to which it is associated with the greatest probability [20,21]. Figure 2 summarizes the process to be followed in order to obtain the latent aspects of a collection of images.
Antoni Tàpies (Barcelona, 1923–2012) is one of the main representatives of the trend known as Informalism. We have analyzed a set of 434 digitized images of paintings, drawings, books and engravings by Tàpies [22] and have been able to determine underlying aspects by means of which the whole collection can be catalogued. Figure 3 shows a sample of the collection.

The Antoni Tàpies database that we used for our analysis contains collections of shapes that the artist linked together because he considered that analogies of meaning existed...
between them. The detection of patterns between the items that the works are composed of helps to unveil certain relationships that the artist established in their configuration to constitute his personal visual language. The fact that works that belong to different periods are included helps to determine the conformations that the artist maintained over time and those that he discarded. In brief, the results can bring us closer to the artist's creative process.

Throughout the Tàpies collection, we identified 13 latent aspects. Figure 4 shows 6 of these represented by a selection of meaningful images. Latent aspect 1 shows images that contain subtle details upon a flat background; they present an important amount of light-colored space, with elements placed in a very purposeful and measured way in the center or in a compositional balance. Latent aspect 2 contains images of works created from edge to edge of the support, with vigorous strokes of dark colors on a slightly lighter ground. Latent aspect 3 shows more complex works that present highlighted motifs by means of textured brushstrokes of medium density combined with a variety of lines and stains deployed over a lighter ground. Latent aspect 4 associates works that present a dotted surface or in which the paint seems to have been sprayed on, sharing a finish that we might call a granular texture. Latent aspect 5 presents images with few strokes, wide and gestural, preferentially occupying the center of the space; these are dark and thick strokes that show up clearly against the paler ground. Latent aspect 6 brings together works in which the shapes are deployed upon indistinct atmospheres characterized by a vaporous and hazy treatment.

In this way we could automatically classify the images in other collections, following the aesthetic criteria of the categories found in the Tàpies collection. We wish to underline the importance of this fact, which would make it possible to establish comparisons between different artists maintaining the same criteria (Fig. 2i).
VISION OF THE WHOLE

Artists are image generators, constantly producing images in their creative process. The visual artist makes use of the formal categories of images to identify what is universally significant in a particular image and does this in a way that is unquestionably personal. Artists’ image collections show connecting links, formal relationships constituting families that share common meanings.

The organization and categorization of a large collection of images gives us an inestimable global vision of the whole. Within the field of artistic creation and its instruction thereof, it is common practice to place groups of photographic images, sculptures or paintings next to each other, and it is surprising to see how, among the visual objects, a dialogue will arise that may be mutually supportive or contradictory. It is a type of knowledge that is hard to explain in any other way and especially difficult with words but when experienced visually, as a whole, produces an immediate understanding of its meaning. Thus, the artist, teacher or student of art can have information at their disposal that would not be accessible if those objects were treated in isolation. This aspect is particularly valuable when the content of the images to be studied is abstract, because here the theme, the subject matter, the meaning or the purpose is not the result of social accord but of visual resonances and synchronicities that the artist relates and links together.

CONCLUSIONS

Computer vision applications are producing valuable returns in many fields of knowledge and, as this study shows, they can also be used with optimal results in the discovery of visual analogies in images of abstract works of art. The difficulty of interpreting the results of searches on "nonfigurative art" is compounded by our need to identify formal features and compositional and structural values rather than known objects or scenes. In view of the results, however, it can be concluded that, with a large collection of images, the system tested would allow us to make

1. a formal preclassification providing a global vision of the whole; and
2. a novel and alternative point of view unconstrained by any aesthetic preconceptions conditioning the user’s prior knowledge of the study sample.

Although the system as programmed is naturally no substitute for expert criteria, it can support the specialist by establishing analogies between works of art and identifying latent aspects in large collections. It allows researchers to repeat searches on collections of the same artist in different periods or collections of different artists or periods using the same criteria.

In this respect, the BOW and PLSA models can contribute to tapping the knowledge contained in art collections according to visual content. Because our human capacity to analyze large numbers of images is limited and because many artworks are simply not available to us, it might be argued that there is no other way to conduct an art study of this kind. And although classification is an implicitly subjective activity that proceeds according to one criterion at the exclusion of others, digital technology still offers a unique means of storing and approaching visual content whose sheer volume would otherwise be impossible to process.

Note that the tools presented here may interest two kinds of user: artists reviewing their own creative process from one period of time to another and art scholars wishing to study a series of artists within a given period of history or the art of different periods without having to move a single work from its original location.

Acknowledgments

This research project was financed by AGAUR (Management Board for University Scholarship and Research Funds of the Generalitat de Catalunya) 2012–2013 and by the research group BRAC (Barcelona Research Art Creation, University of Barcelona).

We gratefully acknowledge the amiability of the archives of the Tàpies Foundation in Barcelona and give our special thanks to the foundation’s directors, Laurence Rassel and Carles Guerra.

References and Notes


15. Fei-Fei and Perona [12].


20. Bosch [1].

21. Quelhas et al. [19].


**Image Credits**

**Fig. 3.**


Image 2: Antoni Tàpies, *Blanc, negre i grafismes*, 1988 (*White, Black and Graphics*). Paint and grattage on glued cardboard, 206 × 150.5 cm. (Credits A).


**Fig. 4.**

**Latent aspect 1:**


Image 2: Antoni Tàpies, *Cartes per a la Teresa*, 1974 (*Letters to Teresa*). Lithograph by Antoni Tàpies. (Credits B).


**Latent aspect 2:**


**Latent aspect 3:**


**Latent aspect 4:**

Images 1, 3 and 4: Antoni Tàpies, *Tres aiguaforats*, 1943. Text by Joan Brossa; 3 etchings by Antoni Tàpies. (Credits B).


**Latent aspect 5:**

Image 1: Antoni Tàpies, *Série negre Núm. XIV*, 1967 (*Black Series No. XIV*). Ink and pencil on wrapping paper, 36.5 × 24.5 cm. (Credits A).


**Latent aspect 6:**

Image 1: Antoni Tàpies, *Estós*, 1958 (*Sketch*). Pencil on paper on lithographic background, 70 × 100 cm. (Credits A).


Reproduction, distribution, public communication or transformation of the images, fully or in part, is forbidden without the written permission of the copyright holders

**Credits A:** © Fundació Tàpies, Barcelona / VEGAP. Photo: © Gasull Fotografia, 2016.

**Credits B:** © Comissió Tàpies, VEGAP. Photo: © Lluís Bover, 2016.

PIÑAR ROSADO RODRIGO has a PhD in fine arts and a master’s degree in biology. Her interest is focused on how new technologies can modify the way we look at the world and on the creative possibilities that are placed within our reach.