# Annotation and search in a collection of art images

Concepción Pérez de Celis Herrero[1], Jaime Lara Alvarez[2], Gustavo Cossio Aguilar[1]
[1]Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla
[2]Área de Educación a Distancia, Universidad Tecnológica de Puebla
{cperezdecelis, gcossio}@cs.buap.mx, jimmy.tec@gmail.com

**Abstract**
In this paper we discuss a tool for annotation and search in a collection of art images. Multiple existing Thesaurus may be used to support this process, including the Art and Architecture Thesaurus, WordNet, ULAN and Iconclass. We discuss knowledge-engineering aspects such as the annotation methodology, and the use of taxonomic dynamic search. The annotation and search process are illustrated with an application scenario.

**Keywords:** Images recovery by content, dynamic taxonomies, CBIR

## 1. Introduction

Historically, curators have retrieved images by first manually annotating them with keywords. Given a query, these annotations are used to retrieve appropriate pictures. Underlying this approach is the belief that the words associated (manually) with a picture essentially capture the semantics of the picture and any retrieval based on these keywords will, therefore, retrieve relevant pictures. Since manual image annotation is expensive, there has been great interest in producing automatic ways to retrieve images based on content.

Queries based on image concepts like color or texture, have been proposed for retrieving images by content, but most users find it difficult to query using such visual attributes. Most people would prefer to pose text queries and find images relevant to those queries. For example, one should be able to pose a query like *find me sad people*. This is difficult, if not impossible, with many of the current image retrieval systems. This paper aims to address the problem of art image retrieval (AIR) and introduce a methodology to intent to reduce the semantic gap in image annotation of art images from a large art image collection.

## 2. Content based image retrieval (CBIR)

The explosive growth of digital media and its usage on the web has opened both great challenges and opportunities to computer science researchers. One such example is the lively research area content-based image retrieval (CBIR), which aims to effectively and efficiently locate a few relevant images in a large database according to a user's query concept.

Luis Enrique Sucar and Hugo Jair Escalante (eds.): AIAR2010: Proceedings of the 1st Automatic Image Annotation and Retrieval Workshop 2010. , volume 1, issue: 1, pp. 74-82

This paper will focus on another related, but different topic: Art image retrieval (AIR), which is of great interests to us because of its application potentials and technical challenges. In recent years, many museums, galleries and commercial companies have put their art images on the web for online exhibitions or selling. The WWW technology has created dramatic flexibilities in helping people freely share, exchange, enjoy and purchase paintings. To facilitate the search of paintings, one of the main standards for the cataloging of cultural objects is the one proposed by the Getty Research Institute, namely, Categories for the Description of Works of Art (CDWA) [1]. Besides being a very complete and accepted model, from our very own point of view, it does not comply with all of our possible descriptions of complex objects that are required in the particular case of Industrial heritage objects and New Media Art.

A second model considered was the Cataloging Cultural Objects model (CCO). A guide to describe cultural works and their images was published as a draft in February 2005, and its final version was published in 2006 [2, 3]. These two models categorized all the art images by styles (e.g. classical art, impressionism) or other information such as names of artists. However, cataloging-based search might not be sufficient in capturing users' personal interests.

## 2.1 Properties of Art Image Retrieval

From a technical point of view, AIR encounters many challenges that are also common in CBIR, like high-dimension feature space, nonlinear distributions, insufficient training examples and the gap between low-level features and high-level contents.

Thus previous research on CBIR has provided a good basis for our work on AIR. Moreover, there are some other unique characteristics coming with AIR. In the following work we will discuss some major issues.

# 3. Information Retrieval of Artistic Objects

One of the objectives that our approach tries to achieve is the search and information retrieval through an analysis of a textual description carried out for different pieces from the database of the artistic objects.

In art works databases, information retrieval could be done by the cataloging descriptors: author, title, or pictorial tendency. However, in many cases, users would like to retrieve some or a group of works through the objects painted, photographed or drawn. In other words, if a user were interested in retrieving pieces of work that include women, as response, he would expect for pieces in which feminine figures appear: Girls, young women or adult women. The type of search mentioned implies the use of facets that allows proper retrieval of items searched. The types of query required should then allow retrieval of information, either through the already-existing facets in the system or defined at the time of execution through search equations in which the user is able to include a series of words that may indicate the elements that define a concept, which must be found in the description of the piece, so that it can be retrieved (simple concept). For example,

**Urban images = vehicles + buildings + people**

In both cases, the defined concepts may become permanent upon the user's request. By doing so, some feedback can be obtained from the user. On the other hand, as an object can be a part belonging to more than one concept, it is the intention to be able to categorize the order of importance of the object regarding the concepts related to it.

To support the structure and efficient retrieval of the artistic objects, and given the complexity in the multiplicity of the possible interpretations and contents that can be linked to the objects, we recommend the use of a faceted classification process.

Under the previous premise, faceted taxonomies can take advantage of the ways in which the metadata act. Metadata are the collection of structured information about any type of object or any part of that object, e.g. the name (or title) of a work art, the author, age, image, textual description (denotation), its interpretation (connotation), or its gender. All these are some kind of metadata associated with a cultural object.

In our approach, we have implemented the faced classification, and thus, the metadata model proposed in [4,5] was extended with the metadata of its gender, connotation, and denotation, so as to include its textual description as a part of the requests of the objects. This fact determines the need of a methodology that allows us to generate the taxonomic structure of the facets on which the different terms included in the textual description of the objects are going to be classified.

In this natural way, every element of the metadata structures can be incorporated as a concept of the faceted taxonomy and can be retrievable through a search engine. The importance of taxonomies is based on the possibility of their use as the triplet {classification diagram, semantic interpretation, knowledge map} [6] in the organization and later retrieval of possibly stored objects in a database. As a consequence, we have the opportunity to access to a request of an object under any of the dimensions in which they were classified. If we consider the example previously given, we could retrieve a cultural object by its gender, connotation, denotation, or simply surf the different facets, because they are orthogonal among themselves.

Sacco [7] introduced the concept of dynamic taxonomy, along with the notion to bear the incorporation of facets that require themselves an independent taxonomy for their description. *Dynamic Taxonomies* [8] (*DT*, also recently known as *Faceted Search Systems*) are a general knowledge management model based on a multidimensional classification of heterogeneous data *objects*, and are used to explore/browse complex information bases on a guided, yet unconstrained, way through a visual interface. The model is primarily concerned with user-centered access, and object classification is not addressed in the base model. As our interest is the retrieval of the works of art based on the textual description of their images, we will use the fundamentals of the dynamic taxonomies, but will extend the domain of data modeling established by Sacco, including textual objects within the metadata.

In the literature, we can find two interesting approaches to the treatment of textual objects. The first one presupposes the existence of several topics on which a taxonomy is to be developed, and in this way, the job of the developed algorithms is to extract the terms and concepts linked to the topics in question from the documents. In this approach, each of the several derived taxonomies requires a group of key terms, linked to the concepts based on which the classification of the documents of the target objective will be carried out.

The second approach, in contrast, focuses on determining the facets and defining their taxonomy, departing from the textual analysis of the documents of the collection. In this case, some techniques for text analysis are used. In particular, in [9], an algorithm is described, used by the Flemish project of the University of Berkeley for the automatic generation of facets about the corpus in the English language, using WordNet.

## 4. Faceted Indexing and Classification

The purpose of indexing is to identify the concepts that represent the content and translation of documents into a way that can be computably managed [10]. The use of a faceted classification will assist us in analyzing the text, and will allow us to ensure that we are entering all relevant items that should be indexed.

We would now present a brief description of the three facets, genre, denotation, and connotation, which we had already mentioned in the previous paragraphs. Using these three facets, we would extend the classification of the artistic objects.

- Genre (Artistic subject): Each of the different categories or kinds that can be used to order the works depending on common features, such as shape and content. For example, the portraits and self-portraits, landscapes, religion, mythology, etc.
- Denotation (Content): Concepts that are shown in an artistic work.
- Connotation (Interpretation): Concepts reflected or transmitted by the artistic work. For example, the word *rain* has the connotative meaning of *sadness* or *melancholy*.

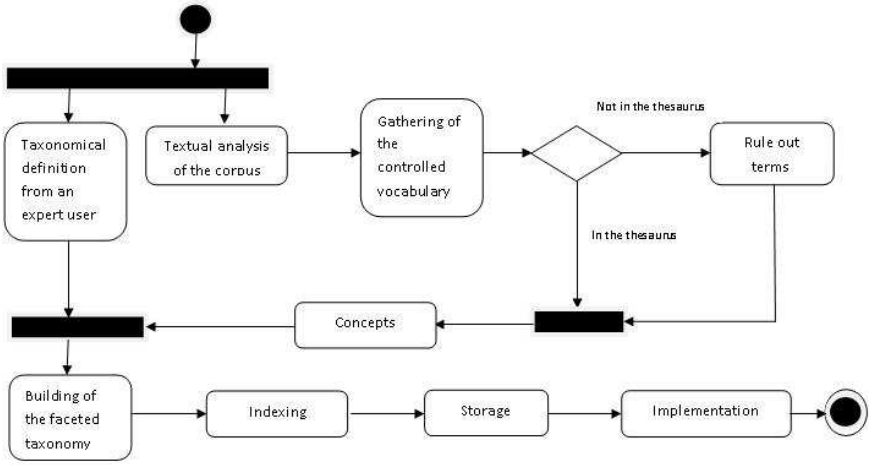
From these three facets, only the facet Genre was defined by an expert user. The other two facets were obtained automatically.

To develop an information system and a motor search for the retrieved objects in a collection of 500 works of art using their textual description in Spanish [11], our methodology uses some techniques and algorithms used in Information Retrieval (IR), to generate, in a semi automatic way, a faceted classification that allows us to link a resort of a document to the different facets. By doing so, we could obtain a controlled vocabulary extracted from the processing of different texts of the target collection. Subsequently, with the help of the controlled vocabulary and a thesaurus, the concepts were determined. These concepts were structured hierarchically with the aid of the thesaurus; in particular, we used the following: *Human Emotions Thesaurus* (HET)[1], online thesaurus of the human emotions developed by clinical researchers of the University of British Columbia and the *Tesauro Europeo de la Educación*[2], a thesaurus structured in Spanish created in 2003 by the Commission of the European Communities, which contains 3516 terms and 2100 relations.

Finally, the objects are indexed [12] with the faceted taxonomy. This allowed us to generate a faceted classification system. The objects, the faceted taxonomy, and

[1]Human Emotions Thesaurus. http://www.slais.ubc.ca/COURSES/libr512/04-05-wt2/thesauri/het/

[2] Tesauro Europeo de la Educación. http://redined.r020.com.ar/es/

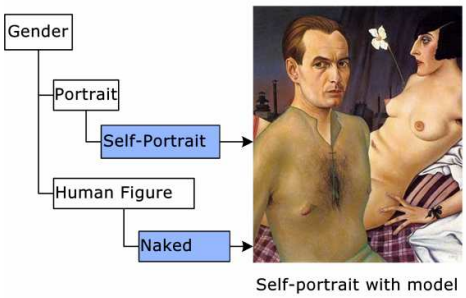



indexes are stored in a database and integrated to the system that enables the user to explore and, at the same time, refine the search through a navigation tree. In Figure 1, the process is shown through an activity diagram.

*Figure 1. Activity diagram of the dynamic taxonomy generation process*

## 5. User Interface Design

The exploration by the user will be realized by means of a navigation based search [13] (taxonomic), which allows to guide the user so that he can interact and tune his query. In addition, it allows to structure, present, and access to the information in more than one dimension, which benefits the user, because the user can easily and intuitively locate and explore the information by means of the different approaches provided by the facets. The navigation based search allows representing the characteristic of dynamics taxonomy visually. In Figure 2, we have shown an example of an object classified under two facets.




Index table

| Concept | Artwork |
|---|---|
| Gender | Self-portrait with model |
| Portrait | Self-portrait with model |
| Self-portrait | Self-portrait with model |
| Human Figure | Self-portrait with model |
| Naked | Self-portrait with model |

***Figure 2.*** *An object with multifaceted classification and their index table*

One of the core principles of dynamic taxonomies is to restrict the available filtering options in the given focus to only those that will lead to a nonempty result set. Hence, the user can never run into a situation with zero results. This is opposed to the process in a typical advanced search situation, where first, a complex Boolean query is constructed, which is then evaluated on demand. This, however, can result in empty result sets, often without further indication on which part of the query could be relaxed to retrieve some results. The exclusion of potentially frustrating situations by design is often referred to as poka-yoke principle (see Figure 3).

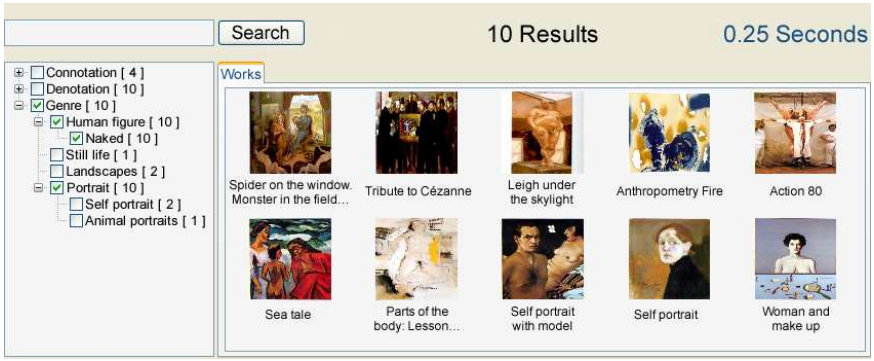



*Figure 3. Interface developed for the search and exploration of works of art*

Figure 4 shows a screenshot of TESEO. The current query is at the top. The extension of this query, that is, the current focus, is at the right, where each object is represented by a thumbnail or a text snippet depending on its type. The dynamic taxonomy is at the left, in the form of concept trees whose nodes are expanded on demand.

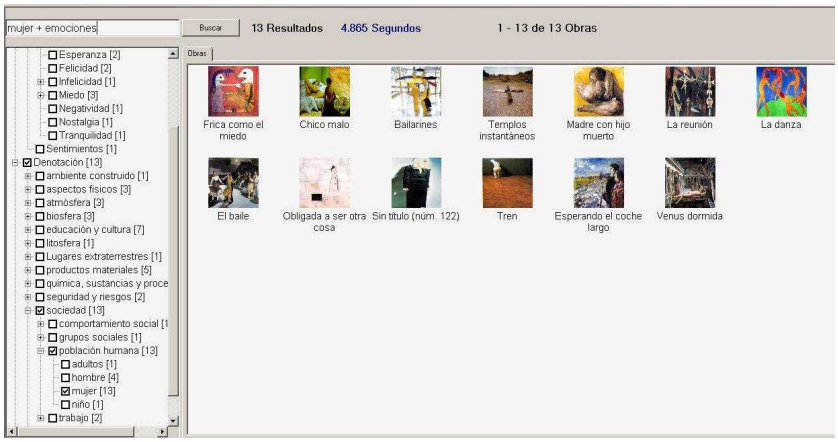



*Figure 4. TESEO navigated search*

In Figure 5, we have summarized our approach for automatic annotation of the images of the pieces of an Art Collection based on their textual description, and in

this way, can described an object in more detail. With this, every artwork object should index not only with the information contained in textual descriptions, but should also have the hierarchical information of the concepts that lie in the dynamic taxonomy.

## 6. Results and Future Work

We proposed a methodology, using IR techniques, to extract concepts and structures and facets to manage information and enable further research and exploration.

Taxonomies were constructed in the following aspects: gender, connotation and denotation, which can serve as a standard for other cases in which the use of this type of hierarchy is required.

We made and implemented algorithms that allow the retrieval of images by relying on their contents and thereby contributing to solving the semantic labeling of images, which is one of the goals of the CBIR methodology. The concept extraction process and indexing between the facets and the works was designed and subsequently implemented in AWK. This module can be used on any set of textual descriptors. With regard to the construction of dynamic taxonomies and navigation, the following operations were implemented: Zoom-In, Zoom-Out, Pivot, Shift, and Query by Example (QBE).

Dynamic taxonomies generated, unlike those reported in the literature, are not on a closed domain. An example is the taxonomy *denotation*, in which there is an open domain, associated with a thesaurus for expanding the domain of the facets obtained from the analysis of unstructured text.

Among the extensions and improvements that we have referred to are that of incorporating: The Art & Architecture Thesaurus ® developed by the Getty Research Institute (GRI), and translated into Spanish by the Documentation Center of Patrimonial Assets (under the Directorate of Libraries, Archives and Museums - Dibam), the Thesaurus Spanish Historical Heritage, the Multilingual Thesaurus of Nations, and the Extended Multilingual Wordnet (University Alicante).

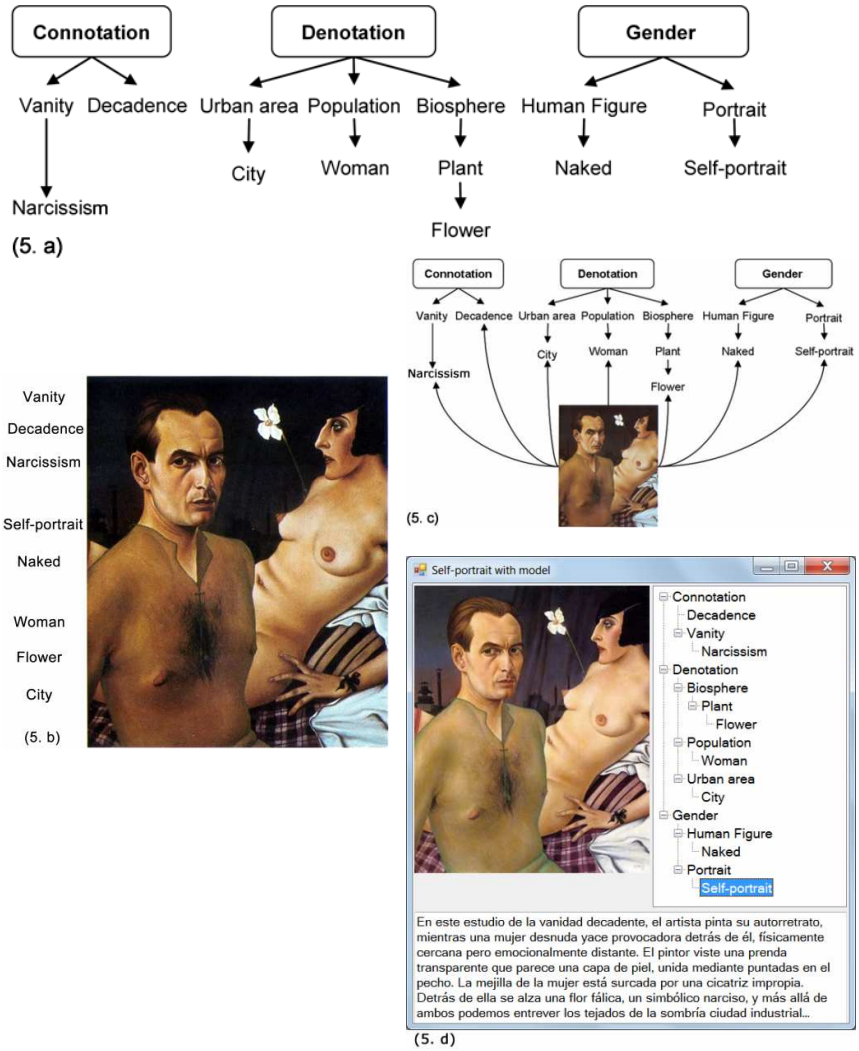



*Figure 5. Summary of the approach. a) An example of the facets whit its taxonomies b) Terms included in the controlled vocabulary that will be indexed to the object c) Concepts related whit the taxonomy d )Art work indexed and it's view in the user interface.*

## References


1. Categories for the Description of Works of Art (CDWA), J. Paul Getty Trust, College Art Association, http://www.getty.edu/research/conducting_research/standards/cdwa/index.html (accessed 1 august 2010).
2. Baca, M.; Harpring, P.; Lanzi, E.; McRae, L.; Whiteside, A., Eds. Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images; American Library Association: Chicago, 2006
3. Gilliland-Swetland, Anne J. "Setting the Stage: Defining Metadata" in Introduction to Metadata: Pathways to Digital Information, Murtha Baca, ed. (Los Angeles: Getty Information Institute, 2000), second edition. 12 pp. Available at http://www.getty.edu/research/conducting_research/standards/intrometadata/ (Third edition Summer 2006, accessed 1 august 2010).
4. Concepción Pérez de Celis Herrero, Gustavo Cossio Aguilar, Sistema de Gestión de Exposiciones: TESEO , VII Reunión del Arte Contemporáneo Grupo Español del International Institute of Conservation , Museo Nacional Centro de Arte Reina Sofia - VEGAP ,2006, ISBN 84-8026-315-6.
5. C. Pérez de Celis Herrero, G. Cossio Aguilar, R. Eguézabal Blanco ; Sistema di Catalogazione per il Museo Nacional de los Ferrocarrriles Mexicanos
6. 12th. International Congress of the International Committee for the Conservation of the Industrial Heritage, Roma, Italia 2006
7. Lambe Patrick, Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness. Chandos Publishing, 2007.
8. Sacco, G. M., Dynamic Taxonomies: A Model for Large Information Bases. IEEE Transactions on Knowledge and Data Engineering 12, 2 , p. 468-479, 2000
9. Dynamic Taxonomies and Faceted Search Theory, Practice, and Experience; The Information Retrieval Series, Vol. 25 Sacco, Giovanni Maria; Tzitzikas, Yannis (Eds.) 2009, XVII, 340 p., Hardcover ISBN: 978-3-642-02358-3
10. Stoica Emilia, Hearst Marti y Richardson Megan. Automating Creation of Hierarchical Faceted Metadata Structures. In the proceedings of NAACL-HLT, Rochester NY, Abril 2007. Available at: http://flamenco.berkeley.edu/papers/castanet.pdf. (accessed 1 august 2010).
11. Tolosa Gabriel y Bordignon Fernando. Introducción a la Recuperación de Información - Conceptos, modelos y algoritmos básicos. Universidad Nacional de Luján, Argentina.Available at: http://www.tyr.unlu.edu.ar/tallerIR/2008/docs/Introduccion-RI-v9f.pdf (accessed 1 august 2010).
12. El ABC del Arte del siglo XX, Primera edición en español 1999, Editorial Phaidon Press Limited.
13. Leise Fred. Using Faceted Classification to Assist Indexing. Contextual Analysis, 2003. Disponible en: http://www.contextualanalysis.com/pub_usingfacets.php (accessed 1 august 2010)