# Image Indexing and Retrieval for a Vietnamese Folk Paintings Gallery

Binh Pham
*Faculty of IT, Queensland University of Technology*
*2 George St, Brisbane, AUSTRALIA*
*b.pham@qut.edu.au*

## Abstract

*Cultural aspects such as history, tradition, beliefs and social customs are often depicted in paintings, both explicitly and implicitly. This paper discusses our experience and difficulties encountered while creating a multimedia Vietnamese Traditional Paintings Gallery for the purpose of dissemination and analysis of information regarding the cultural and artistic aspects of these painting . Issues regarding indexing and retrieval are analysed, which provide the motivations for our own approach which combines keyword and visual features-based methods with iconclass-based method used in arts and humanities. The use of XML-based language to express semantic content and Bayesian inference network to cater for vague queries are also discussed.*

## 1. Introduction

Rapid advance in multimedia technology which allows information in a variety of forms (text, images, video, audio) to be organised and presented together has created an impetus for technology transfer to arts and humanities. In addition, there is an increasing demand for intelligent software systems which can reason and perform in a similar manner to humans. This necessitates the development of techniques that facilitate the incorporation of human expertise, knowledge and requirements. One problem which responds to both of these demands is the study of the behaviour and effects of social culture on the development of societal achievements such as art, music, dance and architecture. For example, while paintings reflect the feelings, thoughts and aspirations of an artist, they also reflect the political, historical, economic or moralistic climate of the time (see, for example, the analysis of the evolution of Chinese paintings through various dynasties [3,6]). We have also explored models for traditional Chinese paintings [12] with the view to facilitate retrieval. Vietnamese paintings also reflect significant changes in political and economic conditions as well as strong impacts of a number of foreign cultures over the Vietnamese culture occurred during different historical periods. Although the general solution to such a study is complex and impossible to obtain, it is nevertheless possible to identify the temporal relationship that exists between the society and the achievements in a more restricted domain.

However, as the contents of these artworks are usually not precisely defined, or even tangible, the problem of their indexing and retrieval is very difficult and almost impossible in general. Current methods for indexing and retrieving using low level features such as colour, shape and texture, are of limited value. So are the metadata about paintings that are manually extracted and recorded. What needed are new techniques that allow high level context be used for indexing, retrieving and archiving purposes, e.g. paintings depicting peasant lifestyle. To this end, we need to develop semantic and context-based models for the paintings, and a language for expressing their symbolic and semantic aspects. Given the diverse nature of artworks, it would be unrealistic to expect that: (i) a generic model can be constructed that would apply to paintings of all genres, styles and subjects; (ii) all contextual information can be extracted automatically from the paintings. Our initial work has therefore been focused on those genres of paintings that exhibit more definite characteristics or patterns, with a limited range of subjects. Notably among these are traditional Chinese paintings, folk paintings, woodcuts and graphic art. We also need to integrate expert knowledge in the model. Such knowledge can be manually annotated or extracted from narratives that accompany the paintings. We have developed a multimedia package that contains a Vietnamese Folk

Proceedings of the Digital Imaging Computing: Techniques and Applications (DICTA 2005)
0-7695-2467-2/05 $20.00 © 2005 IEEE

Paintings Gallery with the ultimate aim to use it for exploring Vietnamese painting styles and identifying high level features and characteristics of these paintings (both artistic and cultural). This Gallery also provides a testbed for indexing and retrieval techniques to be developed.

This paper focuses on the development of an indexing and retrieval scheme for the most well-known Vietnamese folk paintings, *Dong Ho* woodcuts, which came from *Dong Ho* village, 40 km north-east of Hanoi, the capital city of Vietnam. Their simplistic and symbolic nature would facilitate our initial investigation. In section 2, we describe how the Vietnamese Folk Paintings Gallery was constructed, while more detailed information on *Dong Ho* woodcuts is given in section 3. Section 4 discusses the difficulties inherent in the process of indexing and retrieval. We then develop an approach that can address these difficulties in sections 5. The last section presents our future work.

## 2. Vietnamese Folk Paintings Gallery

Vietnamese folk paintings were produced by peasant artists who lived in villages. These paintings are often categorized by localities with distinct artistic styles, printing techniques and material used. Many of these paintings were produced for celebration (e.g. Lunar New Year) while others depict contemporary folk culture through different historical periods of Vietnam. Using Macromedia Director, we have developed a Vietnamese Folk Paintings Gallery which contains four major styles:

- *Dong Ho* woodcuts.
- *Hang Trong* paintings: used imported colour pigments and paper in large formats, with strong colours applied within bold black outlines. These paintings have predominantly cultural themes and were normally displayed in large sitting rooms of city dwellers.
- *Kim Hoang* paintings: done on red or yellow paper imported from China.
- *Sinh* paintings: done in the outskirts of Hue, where the Citadel and Emperors' tombs are located. Their themes reflect Vietnamese ancient primitive thinking about divinities, god and goddesses.

Fig. 1 displays example paintings of each style.

Our aim is to allow a user to browse or search the collection to seek relevant information on these paintings. The Gallery therefore also contains narratives to provide both general information on the whole collection and specific information on individual paintings. The general information are given on the country Vietnam, traditional paintings, symbolism in paintings and customs. Specific information on each painting includes its style, subject, origin and any symbolism conveyed through colours or specific objects (e.g. a peach symbolises longevity, the chrysanthemum is a flower of autumn which symbolises joy, a well-off and secluded life).

Macromedia Director is a good authoring tool for creating interactive multimedia, by simulating the creation of a theatre or movie production. It deploys the concept of cast members displayed on the stage and represented by sprites that determine when and where each cast member should appear. We built the initial database using Microsoft Access, then convert it to a *Valentina* database to make the data more accessible to Macromedia Director. Valentina is a powerful relational database which allows the storing of a large number of data types, including Binary Large Objects (BLOBs). Hence it is the logical choice for our application which involves a large amount of dynamic binary data (images to be processed and manipulated; and image animation sequences). The database structure consists of a collection of tables, each of which involves a particular type of object or a certain relationship. Each column indicates a particular attribute or relationship, and each row represents an instance. We used the *Lingo API* provided by Macromedia Director to develop interactive behaviours.

However, Macromedia does have limited capability for access and search of databases, and for complex computational tasks that are required by our application. We therefore had to develop advanced function modules and connect them to the Gallery by using available software components called *Xtras* which allow plug-in extension to the standard Macromedia Director functionality. We used *V4MD Xtra* to allow Macromedia Director access to the Valentina database via OBDC (Object DataBase Connectivity). In addition, we use *XDIIWrapper Xtra* which allows Macromedia Director access to external programs stored in a Dynamic Link Library (DLL). The DLL is required to provide image manipulation and processing functions and more advanced retrieval functions. The database structure is currently organized by painting styles and subjects to facilitate browsing. Users can navigate through the collection to obtain an overall understanding by following the

narratives. Currently, Macromedia Director only allows a simple retrieval scheme using keywords extracted from the narratives. This limits the users' ability to explore in depth the meanings conveyed by these paintings, and to navigate by following their own lines of thought. We aim to develop appropriate schemes to provide more effective indexing and retrieval methods.

## 3. Vietnamese Traditional Woodcuts

Master craftmen produced *Dong Ho* woodcuts through a process which consists of four steps: drawing, wood carving, printing and colour application. The painting outlines were firstly done on paper before being glued to a wooden block for carving. The woodcut blocks were passed down as family heirlooms through generations. Special paper was made firstly made from mulberry bush and then prepared with a coating made from the powder of mother-of-pearl oyster shells. This process produced supple parallel lines and silvery lighting effects on the paper. Colour agents came from natural vegetable dyes, plants and minerals. Yellow hues are extracted from sophora japonica flowers , red and orange tints from sappan wood, blue from indigo leaves, black from charred bamboo leaves and white from mother-of-pearl oyster shells. The colours used in these paintings have subtle meanings according to a Vietnamese writer: *"The images for the Tet festival (Lunar New Year) that depict chickens and pigs are red like sticky rice coloured with momordica, yellow like ripe rice, green like the young rice plants, or else yellow like the tumeric with which one cooks fish in brine, brown like turned soil, all colours that the Vietnamese are used to and which they have loved for so many generations. These colours lie deep in the heart and spirit of the country folk and passed down from father to son, they ended up becoming the specific colours of our people".*

Ideas and concepts conveyed through symbolism also add to the special characteristics of these paintings. Wishes for good things in life such as honour, wealth, longevity, virtue, good fortune are also expressed by symbols. For example, the peach and the tortoise symbolize longevity while the pomegranate with its great number of seeds, symbolizes numerous posterity. The sow and her babies symbolizes abundance and the rooster strength and good omen. In order to facilitate the understanding of the cultural aspects depicted by these paintings, such knowledge needs to be disseminated.

These woodcuts can be broadly classified into six subject categories:

- *Festivities:* Lunar New Year, mid-autumn harvest time.
- *Culture:* music, dance, theatre, sport, rural lifestyle.
- *Wishes:* wealth, longevity, posterity, virtues, peace, good omen, status.
- *History:* heroes, wars, emperors, dynasties.
- *Beliefs:* religion (Buddhism), philosophy (Taoism, Confucianism).
- *Satire:* on Vietnamese society (teaching style, mandarin greed), on foreign culture (western materialistic civilization, Chinese rituals).

Some woodcuts may be classified in more than one category. Furthermore, one category may appear as a subcategory of another, e.g. a woodcut depicting a festival may contain elements that express some wishes as well as elements that contain religious connotation. Fig. 2 displays 5 examples of images in each of the above categories.

Many of the original *Dong Ho* paintings are currently in Paris, while the only substantial reference source on them is an out-of-print French book by Durand [4] which contained only black-and-white prints. Colour images used in the Gallery came from the colour prints in [1]. This book only provided very scanty comments on these paintings in Vietnamese, English and French text. A multimedia Gallery therefore will facilitate the dissemination of Vietnamese cultural heritage.

## 4. Indexing and Retrieval Issues

Much research on content-based image retrieval by the computer vision community to date has concentrated on retrieval of low level visual features such as colour, texture and shape, or retrieval of images which are similar to a given sketch or image. While this is adequate for applications such as browsing, it is not useful for applications in arts and humanities (e.g. art history) where users wish to illustrate, disseminate, analyse or learn concepts or ideas. In these cases, it is paramount to be able to describe the semantic contents of images and retrieve images with similar meanings. At a higher level, these contents can be logical features such as certain events or scenes, e.g. paintings depicting dancing or religious

rituals. At an even higher level, they might contain some symbolism (e.g. a rooster representing strength); or evoke a particular mood (e.g. happiness, despair); or provide some aesthetic values. Traditionally, most art collections are manually classified and annotated using either free text (in terms of key words) or controlled vocabularies (e.g. a thesaurus organized in hierarchies). Common metadata which provide information such as the creator of the artwork, time, location and subject, has also been used for indexing and retrieval.

A much more systematic classification scheme used by art historians is the *iconographic classification* which is based on the concept of *iconclass* [1]. An iconclass provides a priori definitions of subject material contained in visual images, organized in hierarchies (e.g. definitions of objects, people, events, abstract ideas and relationships). An example of an iconographic system is the documentation of art historical images by Waal published in 17 volumes [8]. More recently, another similar scheme by Yee et al. [10] is to search images by *faceted metadata* which is composed of orthogonal sets of categories. The facets could be themes, artists' names, periods, locations, etc. , while the metadata maybe single- or multi-valued, flat (e.g. about a specific artist) or hierarchical (e.g. about a specific artist, belong to a particular school, belong to a particular country). They designed a faceted category interface and performed a usability study for about 35,000 images in the Thinker collection of the Fine Arts Museum of San Francisco. They concluded that the category-based was strongly preferred by most participants as a more effective query approach over the standard keyword matching approach.

To take into account of the user requirements, the prominent visual effects and the semantic complexity of the Vietnamese folk paintings, we are developing an indexing and retrieval scheme based on the combinations of keywords, visual features and iconclasses.

## 5. Indexing and Retrieval Schemes

### 5.1. Keywords

During the creation of the Gallery, narratives have to be prepared which convey specific information about each individual painting and generic information about the whole collection. Metadata includes artists, location, painting media, period, style, components of painting sets and targeted customers (e.g. for display in government offices, city dwellers, or on special occasion such as Lunar New Year or Autumn festival). Semantic information includes the theme of each painting; objects, people, events, situations; their symbolic meanings and origins; the aesthetic values of each painting and its possible relationships to other paintings. Such rich content of these narratives provide useful resources for keyword-based search by both novice and expert users. One disadvantage of the keyword-based approach is its ad hoc nature. While it is a good tool for expert users who know why they seek the images and how to judge the usefulness of the retrieved images with respect to their goals, the situation is not the same for novice users who need better guidance. Thus, it would be beneficial to extract information from these narratives to construct a more systematic classification of these paintings and develop tools to guide users through specific paths of query or suggest related paths of query. These issues will be dealt with in subsection 5.3.

### 5.2. Visual Features

The Vietnamese folk paintings, especially the *Dong Ho* woodcuts, have some visual features that are very distinct, and should be taken advantage of. The colour pallets are limited and the colours are saturated. Thus, retrieval by colours is a simple task. The paintings are flat with simple and precise outlines for object shapes. They are also "cartoon-like" in appearance, so shape recognition has to be performed in an approximate way, e.g. ducks belong to a family of similar shapes under certain transformations and constraints. We adapt the approach for representation and detection of deformable shapes by Felzenszwalb [5] for this purpose. Templates for objects are firstly constructed using triangulated polygons which are then used to form deformable models. The shape matching algorithm is based on the minimization of energy functions to find a global optimal solution to a non-rigid shape matching problem. This approach achieves a much better performance than those approaches based on deformable shape contours because the contours do not provide information on the content of objects. Using this approach, we have been able to detect classes of common objects in these paintings, e.g. ducks of different shapes, scales, and postures.

Thus, retrieval by visual features such as colour and shapes is possible and this provides specific information for queries such as "*find paintings with ducks; find paintings with temples"*. However, this

approach would not provide information on deeper knowledge such as "*why are there so many paintings with ducks or temples? What do they represent*?". To answer these queries, we need to find efficient ways to construct and represent iconclasses in order to access the semantic content.

### 5.3. Iconclasses

We devise and construct the following iconclasses to be used as a knowledge base for semantic retrieval:

- Painting Categories (Festivities, Culture, Wishes, History, Beliefs, Satire)
- Painting Styles (*Dong Ho, Hang Trong, Kim Hoang, Sinh)*
- Painting Media (Ink, Paper, Production methods)
- Symbolic Thesaurus (to map symbols, characters and colours to abstract concepts)
- Association Rules for symbols and emblems (ie. those appear on the same image need to be interpreted together)
- Cross references

Fig. 3 displays the basic UML class diagrams for our system.

If each individual painting is described iconographically, then a number of advantages will result. Firstly, users will have immediate access to the semantics. Secondly, users are provided with better guidance to important aspects of the study domain. Furthermore, cross references will offer suggestions of related areas to be explored. Thirdly, by recording information on how users navigate through iconclasses, their profiles can be captured and later used to aid the retrieval process to better match their preferences or needs. In order to do these tasks efficiently and effectively, we need to be able to extract such knowledge from the narratives and to represent it in a simple and convenient language. We also need to be able to deal with vague queries which are inherent in arts and humanities.

### 5.4. XML-based Descriptive Language

XML-based languages have been found to be very effective for expressing both logical and semantic knowledge. For example, the ODRL (Open Digital Rights Language) provides a mechanism to express digital rights of digital works. This facilitates the description , analysis, valuation, trading and monitoring of these assets [6]. Using a similar approach, we can develop a XML-based language to express the concepts covering both the visual features and the iconclasses described in the last subsections. This language will provide vocabulary (data dictionary), structure and semantics of expressions, and their relationships. This language needs to be extensible so that additional vocabularies and expressions can be defined as new knowledge emerges.

Once this language is constructed, information can be automatically extracted from the narratives which are written in a natural language. Or conversely, descriptions for each painting may be firstly annotated in the XML-based language, then converted to a natural language to provide better flow for reading.

### 5.3. Vague Queries

Cultural and artistic analysis is inherently subjective, hence it is essential to provide mechanisms to deal with vague queries. We have dealt with a similar problem for GIS (Geographic Information Systems), to retrieve answers to vague queries such as *"I am an artistic person, find me interesting places to visit"* [9,10]. In this case, we constructed a Bayesian Inference Network (BIN) which is capable of incorporating expert knowledge and spatial relationships between data sets. The spatial Bayesian learning algorithms assign causal linkages between datasets in order to record probability relationships between datasets. Given a vague query, the BIN retrieves all relevant datasets, then specific data items that best match the query. The BIN is also capable of learning from relevant feedbacks.

A similar approach can be applied here, where a BIN can be initially constructed based on experts' beliefs on the category or categories that a particular painting belongs. For example, the presence of a temple gives strong evidence that the painting is a religious one, while the presence of a farm animal (pig, duck) provides evidence that the painting could depict a country scene or a symbolic wish (prosperity). The experts' beliefs also can be deduced and extracted from the narratives and the information on the iconclasses. For a vague query such as " *find paintings that depict prosperity*", the BIN will find the paths leading to paintings which contain colours and symbols that convey prosperity.

## 7. Conclusion

Most methods for image indexing and retrieval provided by the computer vision community are not

useful for applications in arts and humanities due to their requirements of high level semantic contents of these images. We have analysed important issues and difficulties encountered while creating a Vietnamese Folk Painting Gallery with the aim to analyse and disseminate cultural aspects of these paintings. We have also presented various methods for addressing these problems. These methods combine computer vision approaches with indexing and cataloguing approaches in arts and humanities to allow painting retrieval based on semantic content. Although some evaluation has been carried out for various specific tasks, some implementation has not yet been done and an overall evaluation has yet to be carried out. However, the success of these methods in other related areas have been proven by other researchers and by our own group. Thus, we believe that our chosen direction for this application is promising. Once the overall evaluation is completed, we also plan to extend these approaches to a broader class of paintings which have more complex characteristics.

**Acknowledgement.** This work is supported by a grant from the Queensland University of Technology, Brisbane, Australia. The author wishes to thank Tim Dunn for interesting discussions on Macromedia Director, and Lyn Lau, Daniella Wriske, David Healy and Karen Brady for image scanning and programming assistance.

## 10. References

[1] *Ancient and TraditionalVietnamese Paintings,* Cultural Information Publishing House, Hanoi, 1995.

[2] J. Berg van den, Subject retrieval in pictorial information systems, *Proc. 18th International Congress of Historical Science, Round Table 34: Electronic Filing, Recording and Communication of Visual Historical Data,* Montreal, 1995, pp. 21-29.

[3] E. Capon, M.A. Pang, *Chinese Paintings of the Ming and Qing Dynasties - 14th - 20th Century*, Wilke & Company Ltd, Victoria, Australia, 1981.

[4] M. Durand, *Imagerie PopulaireVietnamienns,* Ecole francaise d'Extreme-Orient, 1960.

[5] P.F. Felzenszwalb, Representation and detection of deformable shapes, *PhD Thesis,* Massachusetts Institute of Technology, USA.

[6] L. Levitchi, *Classical Chinese Painting*, translated from Romanian "Pictura Chinese Clasica", Murrays Remainder Books Ltd, London, 1979.

[7] *ODRL Open Digital Rights Language*, http://odrl.net/1.1/ODRL-11.pdf.

[8] Waal van de, *Iconclass: An Iconographic Classification System,* L.D. Couprie, R.H.Fuchs, E.M.A.E. Tholen and G.G.J. Vellekoop (eds.), 17 volumes, Amsterdam /Oxford/ New York.

[9] A.R. Walker, B. Pham, M. Moody, A Bayesian framework for automated dataset retrieval in Geographic Information Systems, *Proc.10th Multimedia Modelling Conf. MMM2004,* Melbourne, pp. 138-144.

[10] A.R. Walker, B. Pham, M. Moody, Spatial Bayesian learning algorithms for Geographic Information Systems, *Proc. 13th GIS Research UK Conf.* , Glasgow, 2005.

[11] K.P. Yee, K. Swearingen, K. Li, M. Hearst, Faceted metadata for image search and browsing, *Proc. CHI 2003*, Ft. Lauderdale, Florida, USA, 2003, pp. 401-408.

[12] D. Zhang, Pham B., Li Y., Modelling traditional Chinese paintings for content-based image classification and retrieval, *Proc 10th International Multimedia Modelling Conf. (MMM2004)*, Brisbane 2004, pp. 258-264.

**Fig.1. Example paintings of 4 styles: *Dong Ho, Hang Trong, Kim Hoang and Sing.***

**Fig. 2. Example *Dong Ho* paintings of 6 categories: Festivities (Dragon dance), Culture (Farming), Wishes (Opulence), History (Lady Trieu went into battle), Beliefs (Princess Ba at Perfume Pagoda), Satire (Classroom with catfish and toads).**

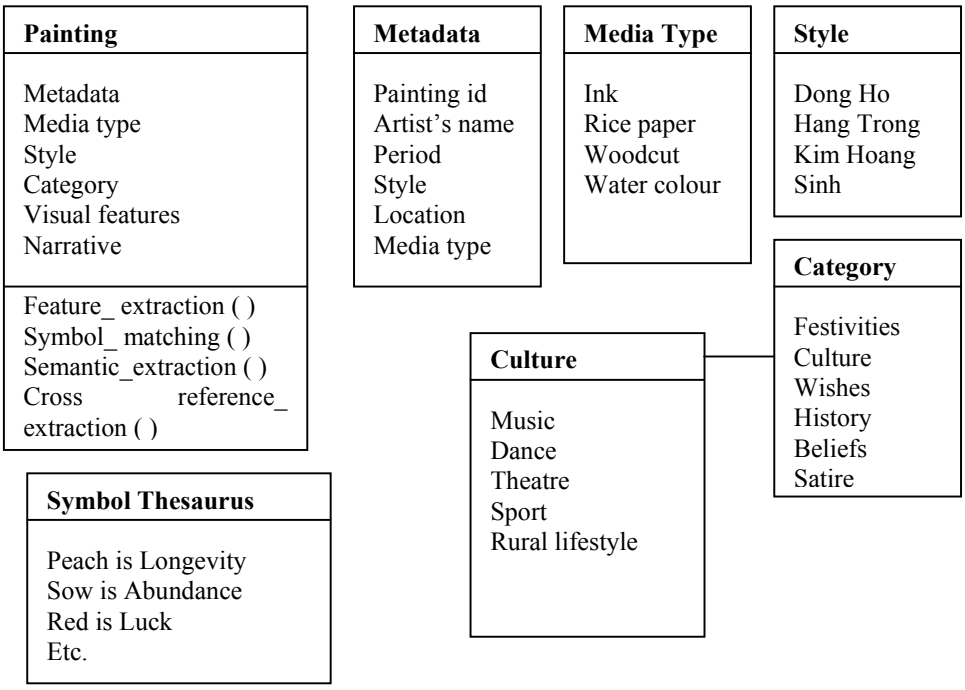


**Fig. 3. UML class diagrams**