Ontology-based Annotation of Paintings using Transductive Inference Framework

Marchenko Yelizaveta, Chua Tat-Seng

National University, Singapore
{marchenk,chuats}@comp.nus.edu.sg

Jain Ramesh

UC Irvine, USA
jain@ics.uci.edu

Abstract. Domain-specific knowledge of paintings defines a wide range of concepts for annotation and flexible retrieval of paintings. In this work, we employ the ontology of artistic concepts that includes visual (or atomic) concepts at the intermediate level and high-level concepts at the application level. Visual-level concepts include artistic color and brushwork concepts that serve as cues for annotating high-level concepts such as the art periods for paintings. To assign artistic color concepts, we utilize inductive inference method based on probabilistic SVM classification. For brushwork annotation, we employ previously developed transductive inference framework that utilizes multi-expert approach, where individual experts implement transductive inference by exploiting both labeled and unlabelled data. In this paper, we combine the color and brushwork concepts with low-level features and utilize a modification of the transductive inference framework to annotate art period concepts to the paintings collection. Our experiments on annotating art period concepts demonstrate that: a) the use of visual-level concepts significantly improves the accuracy as compared to using low-level features only; and b) the proposed framework out-performs the conventional baseline method.

Keywords: Transductive inference, Multi-expert, Concepts Ontology, Paintings.

1 Introduction

Visual characteristics of paintings such as color, brushwork, and composition constitute a large body of artistic concepts that facilitate expert analysis in the paintings domain. They closely relate to high-level semantic information of painting such as the artist names, painting styles and art periods. These concepts have been used for painting analysis to support applications such as brush-stroke detection and image annotation [3, 6, 9, 12, 13]. Several studies [6, 9] performed automatic brushwork analysis for the annotation of paintings with artist names. These methods directly modeled the artist profile based on low-level features. Such approach yields limited accuracy because of two drawbacks. First, it does not incorporate domain-specific knowledge for the disambiguation of results. Second, since visual-level concepts are not represented explicitly, the introduction of other high-level concepts in arts domain will require additional training. To alleviate these problems, in our previous work [12], we proposed a framework for ontology-based annotation of paintings where meta-level artistic concepts such as the color and brushwork are introduced as the basis for annotating higher-level concepts such as the periods of art, artist names and painting styles. In this work, we adopt the proposed framework by utilizing the artistic color and brushwork concepts extracted to support annotation of paintings with the concepts of art periods.

For this task, we first perform annotation of paintings with artistic color concepts based on our earlier proposed method [13] that utilizes color theory of Itten [7] similarly to other studies [3]. We next perform the annotation of brushwork concepts by employing the previously developed framework for brushwork annotation using serial combinations of multiple experts [14, 15]. The paper describes our approach on utilizing the color and brushwork concepts in our ontological and transductive inference framework for the annotation of the high-level concepts of art period.

2 Ontology of Artistic Concepts

In our study we employ the ontology of artistic concepts that includes visual, abstract and application concepts as shown in Figure 1. This ontology is based on external Getty's AAT and ULAN ontologies [16]. It has several advantages. First, the explicit assignment of visual and abstract concepts offers more flexibility for paintings annotation and retrieval. Second, the use of domain-specific ontologies within the proposed framework facilitates concept disambiguation and propagation. Lastly, ontology includes retrieval concepts for both expert and novice user groups.

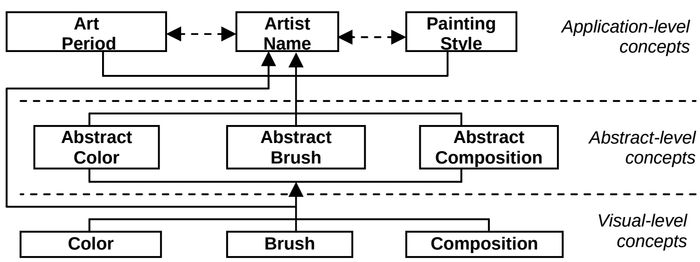


Fig. 1. Three-level ontology of artistic concepts. Double-edged arrows between concepts denote that these concepts are inter-connected.

Concepts of the visual level (atomic concepts) include color, brushwork and composition concepts. In our system, we utilize the visual-level concepts in two ways. First, they represent large vocabulary for retrieval of painting by the expert users. For example, such queries as paintings in *warm* colors, paintings with *temperature* contrast and *impasto* brushwork class are possible. Second, these concepts serve as cues for the annotation of higher-level concepts in abstract and application levels [1, 7]. Abstract-level concepts include concepts defined by artistic theories for the art experts. Application-level concepts denote the widely used concepts for retrieval by novice users in online galleries such as the artist names, painting styles and periods of art etc.

In this paper, we focus on the annotation of paintings with the concepts of art period. Our collection includes paintings by various artists from *Medieval* and *Modern* periods of art. To perform the annotation, we exploit heuristics available in the domain knowledge. For example, paintings of *Medieval* period often exhibit *primary* palette of colors such as red, blue, *light-dark* color contrasts, *mezzapasta*, *glazing* and *shading* brushwork classes. Paintings of *Modern* art often exhibit *complimentary* colors, *temperature* contrasts and variety of brushwork classes such as *scumbling*, *impasto*, *pointillism*, *divisionism* and *grattage* [1]. To account for such heuristics, we utilize the visual-level concepts as mid-level features to assist in the annotation of paintings with high-level concepts. In this section, we also briefly discuss the visual-level concepts.

2. 1 Visual-level Color Concepts

Itten's theory [7] proposes the mapping between colors and artistic color concepts, and is primarily used by artists. Itten defines twelve fundamental hues and arranges them in color circle. Fundamental hues vary through five levels of intensity and three levels of saturation, thus creating their respective subsets of colors. Fundamental colors are arranged along the equatorial circle of sphere, luminance varies along medians and saturation increases as the radius grows. Itten locates the shades of gray colors in the center of the sphere and white and black colors at the poles of the sphere.

Based on the color sphere, Itten defined color temperatures concepts (*warm*, *cold* and *neutral*), color palette concepts (primary, *complimentary* and *tertiary*) and color contrasts (*complimentary*, *light-dark* and *temperature*). We discussed these concepts in detail in our previous work [13].

2. 2 Visual-level Brushwork Concepts

In our study, we employ eight brushwork classes widely used in Medieval and Modern periods of art. Table 1 summarizes information of brushwork. It demonstrates that our brushwork collection includes mostly stochastic textures. They exhibit a variety of properties such as directionality, contrast, regularity etc. In terms of the spatial homogeneity, we can roughly group brushwork patterns as homogeneous (*mezzapasta* and *pointillism*), weakly homogeneous (*divisionism*) and inhomogeneous (*scumbling*, *shading* and *glazing*).

3 Transductive Inference of Concepts using Serial Multi-Expert Approach

To annotate paintings with artistic concepts, we employ previously developed transductive inference framework. We briefly discuss its major components in this section.

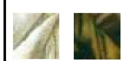
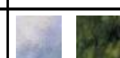
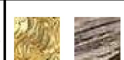
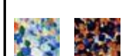
3. 1 Serial Multi-Expert Approach

The decision process within the serial multi-expert framework starts with all classes and the original dataset including both labeled and unlabelled patterns. It progressively reduces the subset of candidate classes to which a pattern might belong to based on the manually pre-defined decision hierarchy, which guides the experts in splitting the input dataset into individual classes.

We denote the subset of candidate classes as *the target set*. We formalize the reduction of the target size as follows. The expert at the *i-th* level has the input vector

(X, S_{i-1}) received from the ancestor node and generates the output vector S_i, where X represents a pattern. S_i represents the set of classes to which the expert of *i-th* level believes the pattern X might belong and the set S_i is a subset of its respective set S_{i-1} $(S_n \subset S_{n-1} \subset S_i \ldots \subset S_0)$. During the annotation process, if the terminal node is reached, then the unlabelled patterns under this node are labeled with a single element of S_i.

Table 1. Visual-level brushwork concepts

Class	Characteristics	Background	Examples
Shading	Depiction of foldings in Medieval Period	Edges and gradients, often directional, intensity contrast, weakly or non-homogeneous	
Glazing	Depiction of nudity/face in Medieval Period	Subset of hues (yellow, red, orange), intensity contrast, gradients, non-homogeneous, may contain edges	
Mezzapasta	Widely used technique in paintings. The color palette used varies with respect to the art period.	Homogeneous, low intensity contrast and small gradients	
Grattage	Depiction of objects and patterns in Fauvism and Expressionism painting styles of Modern Art period	Edges, high gradients, intensity contrast, inhomogeneous	
Scumbling	Depiction of sky, clouds, greenery and atmosphere in various painting styles of Modern art	Soft gradients, low intensity and hue contrast, low directionality, weakly homogeneous	
Impasto	Widely used in Impressionism, Post-impressionism, Pointillism styles of Modern art	Edges, high gradients, often directional, low hue contrast, high intensity contrast	
Pointillism	Often used for depiction of atmosphere/air in Pointillism painting style of Modern art	Medium intensity contrast, medium roughness, no directionality, homogeneous	
Divisionism	Widely used in Pointillism, demonstrates the Color Mixing Principle	High gradients, high roughness, high intensity and hue contrast, no directionality, weakly homogeneous	

We employ Class Set Reduction and Class Set Reevaluation strategies for annotation using the serial multi-expert framework. The Class Set Reduction requires that the experts generate a subset of candidate class labels from the original set of candidate class labels received from the ancestor node. The Class Set Reevaluation extends the intermediate nodes to facilitate additional analysis: if the unlabelled patterns are assigned labels with high confidence, then these assignments become final and the decision process does not evaluates these patterns further.

3. 2 Class Weighted Feature Score

To provide the expert with the feature relevance information, we calculate feature scores with respect to each analyzed class. For this we first calculate tight partitions in the feature space using iterative K-means method. Since the K-means clustering minimizes the intra-cluster distance, the data points within a partition are somewhat close to each other in the feature space and exhibit relatively small variances along some of the feature dimensions. Thus, feature dimension is more likely to be relevant to the partition if the projection of the partition on this dimension has a smaller variance. Second, we employ Chi-square statistics to compare the feature value distributions between this partition and the whole dataset. Intuitively, if the distributions are similar, then the analyzed feature is not representative of the cluster and its Chi-square statistics is comparatively low. We represent the feature distributions using the normalized histograms of each feature in the cluster and the whole dataset. To measure the similarity of distributions, we employ Pearson's Chi-Square test: $X^2=\Sigma(O_i - E_i)^2/E_i$, where we treat the *i-th* histogram bin of the feature distribution in a cluster and the overall dataset as the observed counts O_i, and expected counts E_i respectively. Using the Chi-square statistics we obtain the relevance score of the analyzed feature with respect to a partition. Third, we combine the feature scores of the partitions to calculate the feature scores of the classes. The experts utilize the class weighted feature scores during the model selection step to be discussed in Section 3. 4.

3. 3 Individual Experts

For each individual expert, the decision hierarchy predefines its input target set TS_i and output target sets TS_{O1} and TS_{O2}. To implement individual experts, we train probabilistic mixture model GMM using EM algorithm. This model approximates the patterns of TS_i as k clusters in the feature space using parametric Gaussian distributions $G(\mu_1,\Sigma_1)\ldots G(\mu_K,\Sigma_K)$. Next, the expert maximizes the calculated posterior probabilities $p(x_j,G(\mu_i,\Sigma_i))$ to estimate the cluster membership of each pattern x_j. Using this information, the expert performs annotation of the unlabelled patterns using *the cluster purity* measure. We define pure cluster of class X as the cluster in which more than 75% of the labeled patterns are of that class (or a subset of classes). The cluster purity represents the degree to which the calculated cluster contains labels of class X and is defined as $p(c)=N_X/N_{all}$, where N_X and N_{all} denote the number of labeled patterns of class X and the overall number of patterns in cluster c respectively. The expert measures the purity of clusters based on the class labels in its output target sets. The unlabelled patterns that fall in the pure clusters receive the candidate class label of that cluster. The unlabelled patterns in impure clusters are assigned the label of the biggest labeled class in the input target set.

To perform the model selection step, the system first trains several models using varying input parameters. Next, it select the least erroneous model using Vapnik's combined bound [4] as shown in Figure 3. For each trained model we have its respective hypothesis h, the full sample risk $R(X_{l+u})$, the transduction risk (or test error) $R(X_u)$ and the training error $R(X_l)$. The Vapnik's criterion estimates of the testing error based on training error $R(X_l)$ and on the bounded deviation between the two random variables $R(X_u)$ and $R(X_l)$ around their mean $R(X_{l+u})$.

4 Annotation of Artistic Concepts

To annotate paintings with the concepts of art periods, we perform a three-step procedure. First, we sub-divide paintings in the fixed size blocks and perform iterative

K-means clustering of painting blocks using low-level color and texture features. Second, we perform the analysis of visual color concepts using the method to be discussed in Section 4. 1. Since we perform the analysis of color concepts at the level of fixed-size blocks, we employ the majority vote to assign color concepts to clusters. For the annotation of a cluster with brushwork concepts, we utilize low-level color and texture features of a cluster and employ the transductive inference framework (see Section 4. 2). Using a combination of low-level color and texture features and mid-level color and brushwork concepts, we again employ the transductive inference framework as described in Section 3 to perform the annotation of application-level concepts.

Input:
A full sample set X_{l+u} and training sample set X_l,
Feature weighted scores $F_S(L_j)$ for the candidate class labels L_j,
A maximum number of mixture components K,
A set of cut-off thresholds for the feature ranks T_f
Output:
Candidate class labels of the test set X_u
Algorithm:
1. For each cut-off threshold $tf \in T_f$ and number k of mixture components, $2 \le k \le K$, train GMM on X_{l+u} to generate $(K\text{-}1)\times|T_f|$ number of models $\{M_{k,tf}\}$;
2. Based on the training set, employ the cluster purity measure to generate a set $\{h_{k,tf}\}$ of $(K\text{-}1)\times|T_f|$ hypotheses corresponding to the models;
3. For each hypothesis $\{h_{k,tf}\}$ calculate its training error $R(X_l)$ and its Vapnik combined bound
4. Output a candidate class labels for X_u using $\{h_{k,tf}\}$ with the smallest Vapnik's bound.

Fig. 3. The model selection algorithm

4. 1 Visual-level Color Concepts

For the analysis of color concepts we utilize CIE L*u*v color space. We employ a two-step procedure to assign *warm*, *cold* or *neutral* color temperature concept to a region. First, we model the distribution of various color temperatures within a block. For this, we back-project image colors to the corresponding reference colors in the Itten color space using the following formulae:

$$ref = \arg_{M_c} \min_{1 \le i \le N} dist\,(Rc, Mc(i)) \qquad (\mathbf{1})$$

where *dist* denotes the normalized Euclidean distance, R_c denotes the image colors, $Mc(i)$ denotes the reference color i on the Itten's chromatic sphere, and N denotes the number of Itten colors *($N = 187$,* including 5 shades of grey, black and white colors). The feature vector of a block includes the number of pixels of each color temperature concept, color values of dominant colors extracted from 316-color histogram in HSI color space, spatial coherence of block pixels of each color temperature calculated based on a modification of the color coherence vector. Second, the system utilizes probabilistic SVM [18] and winner takes all strategy to assign color temperature concept to each block. Using the same two-step procedure, we classify blocks with respect to *complimentary*, *primary* and *tertiary* color palette concepts.

To calculate color contrast concepts, we represent each block as a set of color pairs based on its dominant colors [2]. Using formula 1, we calculate the corresponding reference colors. Based on the relative location of reference colors on the chromatic color sphere, we calculate the *complimentary*, *temperature* and *light-dark* contrast values. Lastly, we average the contrast values of all color pairs within the blocks to derive the contrast values for each block.

4. 2 Visual-level Brushwork Concepts

In order to employ the transductive inference framework as described in Section 3 to annotate brushwork concepts, we extract color and texture features and derive the decision hierarchy for the annotation.

We employ variety of feature extraction techniques for adequate representation of brushwork concepts [13] such as major colors [10], directional histograms of image edges and gradients, multi-resolution Gabor Texture features [11], wavelet-based features, Hurst coefficient [8] and Zernike moments [17]. We extract these features based on the fixed-size blocks and average their values to calculate one feature vector per cluster. Figure 4 demonstrates the decision hierarchy for brushwork.

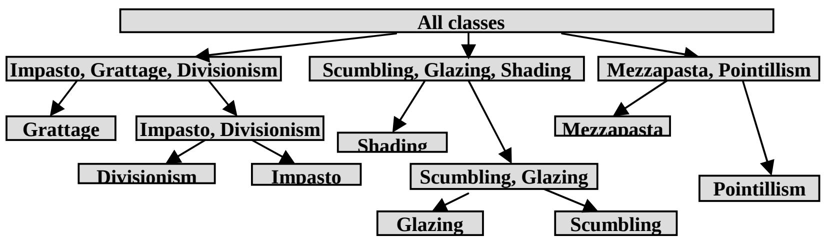


Fig. 4. The decision hierarchy for brushwork annotation

4. 3 Application-level Art Period Concepts

For each image cluster, we now have low-level color and texture features as well as intermediate-level artistic concepts for color and brushwork. We utilize this information to annotate high-level concepts of art periods. Overall, we employ a two-step procedure to perform annotation. First, we annotate the image clusters with high-level concepts. To perform this task, we employ transductive inference framework. However, since our collection includes paintings of only two periods of art, the decision tree has only three nodes: a root node and two leaf nodes. In accord to the decision tree, the framework employs a single expert that annotates the image clusters with one of the two mutually exclusive concepts. To facilitate feature selection, we calculate class weighted feature scores for periods using the method discussed in Section 3. 2. The framework utilizes feature scores during the model selection step as described in Section 3. 3. Second, we back-project clusters onto their respective paintings and employ the majority vote technique to annotate the art period concept to the whole painting.

5 Experiments

For our experiments, we employ 200 and 700 paintings of various artists and painting styles for training and testing respectively. The testing set includes 120 paintings in

Medieval and 580 paintings of Modern period of art. To preserve color and brushwork information, we employ the fixed-size blocks of size 32x32 for the concept analysis.

5. 1 Annotation of Visual-level Color Concepts

To measure the accuracy of labeling with color temperature and color palette concepts we employ 5,000 randomly sampled blocks from the training set. We utilize this dataset to perform training and testing of probabilistic SVM classifiers for annotation of color temperature and color palette concepts respectively. We use 75% of the dataset for training and 25% for testing. We found that we could achieve 91.2% of accuracy in color temperature annotation task and 93.7% in color palette annotation task. We did not evaluate the annotation of blocks with color contrast concepts due to the lack of ground truth, but we have demonstrated its performance for region-based retrieval task [13].

5. 2 Annotation of Visual-level Brushwork Concepts

For this set of experiments, we extract 4880 blocks from 30 paintings of Renaissance, Fauvism, Impressionism, Post-Impressionism, Expressionism and Pointillism painting styles. We randomly select 75% of the dataset for training and use the remaining patterns for testing. Figure 5 demonstrates the distribution of brushwork.

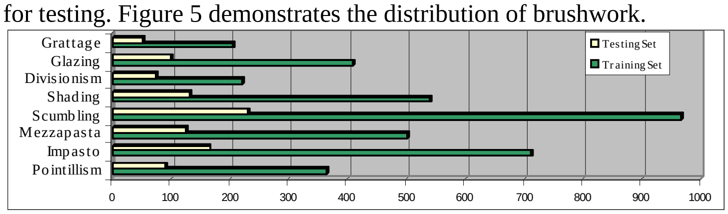


Fig. 5. Distribution of brushwork classes in the training and testing datasets.

Table 2 summarizes the performance of the systems in terms of overall annotation accuracy. We employ a single GMM model as the baseline system for our experiments with brushwork.

Table 2. Performance of the systems for brushwork concepts annotation.

System	Class Reduction	Class Reevaluation
Baseline	**80.07%**	
Baseline with feature selection	**83.6%**	
Multi-expert with model selection	**93.7%**	**87.45%**

Baseline performs the annotation of the unlabelled instances into the brushwork classes on the basis of pure clusters. It can be viewed as a single expert operating on the full feature set. During our experiments, we found that baseline generates the best results using *K=30* mixture components. To evaluate feature selection, we perform another baseline with feature selection as discussed in Section 3. 2. The proposed multi-expert transductive inference framework achieves higher accuracy due to the several reasons. First, it sequentially disambiguates patterns, which yields high annotation accuracy at the leaf nodes. Second, it employs the model selection step that finds most appropriate number of mixture components as well as the cut-off threshold for the features scores with respect to each individual expert.

5. 3 Annotation of Paintings with Art Period Concepts

For our experiments of application-level concept annotation, we perform clustering of blocks from each painting in 60 clusters. The first baseline system (Baseline 1) for our experiments is a binary SVM classification method based on low-level color and texture features. To test the contribution of the visual-level concepts to the overall result, we employ the variation of the baseline system (Baseline 2) that combines visual-level concepts and low-level features with the class weighted feature scores above 0.7. Lastly, we evaluate the proposed transductive inference framework using both low-level features and intermediate-level concepts. Table 3 demonstrates the performance of the systems.

Table 3. Performance of the systems for application-level concepts annotation.

System	Accuracy of cluster annotation, %	Accuracy of image annotation, %
Baseline 1	**68.72%**	**81.48%**
Baseline 2	**79.02%**	**93.56%**
Transductive inference with model selection	**86.84%**	**98.71%**

From these results, we draw the following observations. Baseline 2 results in higher accuracy as compared to Baseline 1 system due to the several reasons. First, the use of visual-level concepts facilitates more accurate mapping from feature vectors to the art period concepts. Second, the use of the weighted feature scores results in the reduction of the noise in the feature space. Next, our proposed method achieves even higher accuracy of 98% at the image-level as compared to Baseline 2 because of several improvements. First, the transductive inference yields higher accuracy due to the use of unlabeled data samples. Second, during the model selection step, the framework finds the parameter values that lead to the least erroneous results in accord to Vapnik's combined bound. Figure 6 illustrates misclassified paintings. All of them belong to Modern art period. However, they all exhibit dark and red colors with large areas of mezzapasta brushwork class similarly to the paintings of Medieval art period.

Fig. 6. Examples of misclassifications by the proposed system

Conclusions

In this paper we proposed a framework for ontology-based annotation of paintings with application-level concepts of art period. Within this framework, we utilize domain-specific knowledge to facilitate annotation. Our experimental results demonstrate that the use of meta-level artistic concepts results in higher annotation accuracy and that the proposed framework outperforms conventional classification approach for annotation of high-level concepts. In our future work, we will focus on several tasks. First, we will perform the annotation of paintings with artist names and painting style concepts. Second, we will develop a methodology to share and integrate the concept ontology used in our study with external ontologies. Third, we will extend the proposed framework to utilize external textual descriptions such as concept definitions in external ontologies and WWW textual information.

References

[1] Canaday J. Mainstreams of Modern Art, Saunders College Publishing, 1981.

[2] Chua T.-S., Lim S.-K., Pung H.-K.. "Content-based retrieval of segmented images". *ACM MM*, 211 – 218, 1994.

[3] Corridoni J. M., Del Bimbo A., and Pala P. Retrieval of Paintings Using Effects Induced by Color Features, CAIVD, pp. 2-11, 1998.

[4] El-Yaniv, R., and Gerzon, L. Effective Transductive Learning via PAC-Bayesian Model Selection. *Technical Report CS-2004-05, IIT*, 2004.

[5] Friedman J. H., An overview of predictive learning and function approximation, From *Statistics to Neural Networks*, Springer Verlag, NATO/ASI, 1-61,1994.

[6] Herik, H.J. van den, Postma, E.O. Discovering the Visual Signature of Painters. In *Future Directions for Intelligent Systems and Information Sciences*, 129-147, 2000.

[7] Itten J. The Art of Color, Reinhold Pub. Corp., NY, 1961

[8] Kaplan L. M. and Kuo C.-C. J., Texture roughness analysis and synthesis via extended self-similar (ESS) model, *IEEE Trans. Pattern Anal. Machine Intell*, 1043–1056, 1995.

[9] Li J., Wang J. Z. Studying Digital Imagery of Ancient Paintings by Mixtures of Stochastic Models, *IEEE Trans. on Image Proc*, vol. 13 (3), 2004.

[10] Low W.-C., Chua T.-S., "Color-Based Relevance Feedback for Image Retrieval". IW-MMDBMS 1998, pp.116-123

[11] Manjunath B. S., Ma W. Y., Texture features for browsing and retrieval of image data. *IEEE Trans. Pattern Anal. Machine Intell* (18), 837–842, 1996.

[12] Marchenko Y., Chua T.-S., Aristarkhova I., Jain R. Representation and Retrieval of Paintings based on Art History Concepts. *IEEE Int'l Conf. on Multimedia and Expo (ICME)*, 2004.

[13] Marchenko Y., Chua T.-S., Aristarkhova I., Analysis of paintings using Color Concepts. *IEEE Int'l Conf. on Mm and Expo (ICME)*, 2005.

[14] Marchenko Y., Chua T.-S., Jain R., Semi-supervised Annotation of Brushwork in Painting Domain using Serial Combinations of Multiple Experts, *ACM Multimedia*, 2006.

[15] Marchenko Y., Chua T.-S., Jain R., Transductive Inference Using Multiple Experts for Brushwork Annotation in Paintings Domain, *ACM Multimedia*, 2006.

[16] Paul Getty Trust. Art and Architecture Thesauri and United List of Artist names. 2000. Available at http://www.getty.edu/research/conducting_research/vocabularies/

[17] Teague, M.R. Image Analysis via the General Theory of Moments, *Journal of the Optical Society of America*, 70 (8), 920-930.

[18] Vapnik, V. Estimation of Dependences Based on Empirical Data. Springer Verlag, New York, 1982.