2014 International Conference on Advanced Robotics and Intelligent Systems (ARIS 2014), June 6- 8 2014,
Nat'l Taiwan Univ. of Sci. and Tech., Taipei, Taiwan, R.O.C.

# A Vision-based Surgical Instruments Classification System

Xian-Heng Liu, Chung-Hung Hsieh, Jiann-Der Lee, Shin-Tseng Lee and Chieh-Tsai Wu

***Abstract*— This paper presents a real-time and automatic online vision-based surgical instruments recognition system, which can be used for surgical instruments monitoring during surgery or robotic applications. The main processes of this system consist of feature extraction and classification. In feature extraction, the image of surgical instruments placed on surgical drape are segmented by using color information. Several shape and contour information of the instruments are extracted as features. A two-stage classification scheme based on naïve Bayesian classifier is then proposed to recognize the surgical instruments according to these features. Experimental results demonstrate that the proposed classification scheme can achieve 90.82% accuracy for classifying 7 instruments.**

**Keywords: Surgical assist systems, surgical instruments monitoring, object classification, Naïve Bayesian Classifier.**

## I. Introduction

With the development in computer science and information technology, application of computer vision and intelligent system in medical field becomes a significant and attractive research topic in engineering. Nowadays, many Surgical Assistance Systems (SAS) have been developed, such as Image-Guided Navigation System (IGNS) or medical robotics [1-3]. These systems or devices can provide useful information or do some actions to help the surgeons or clinical staff to perform surgery or medical treatment more conveniently and stably.

In general, during the surgery, surgeons need someone to pass the surgical instruments in need to them. Therefore, development of a SAS for surgical instruments monitoring is important and necessary in modern surgery. A surgical instruments monitoring system should have the functions to identify and recognize surgical instruments, and further accomplish surgical instrument management or confirm the instrument pass to the surgeon. Therefore, instruments recognition is an important key component for surgical instruments monitoring.

There are several researches which reveal some ideas for surgical instruments recognition. According to the contents and methods they introduced in these articles, these research efforts can be divided into two categories. The first category is radio frequency identification (RFID) based methods [5-6] and the other category is vision based recognition methods [7-9]. Compared to the other category, the greatest advantage of vision based recognition method is that it is non-contact, and it doesn't need markers or tags to help the system to identify and recognize the instruments.

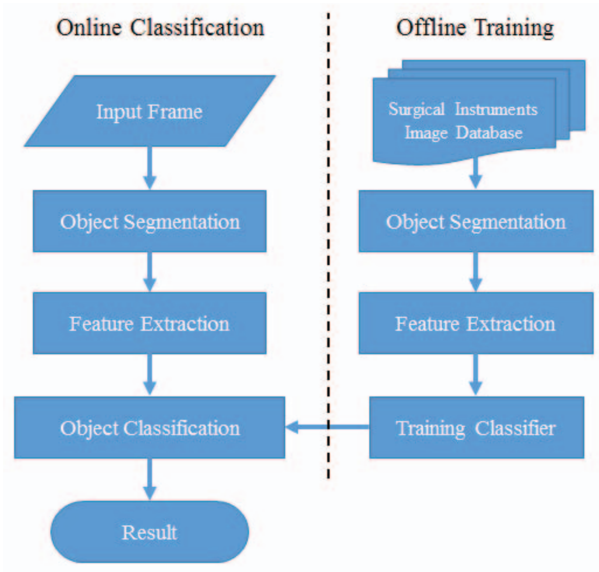


Figure 1. System flowchart

In this paper, we present a real-time and automatic online surgical instruments classification system, which can be used for applications of instruments monitoring or medical robotics. This approach utilize camera to collect a bunch of images of seven surgical instruments to form a surgical instrument image database, and this image database is utilized to train Naïve Bayes classifiers for online instruments classification. The processes of the proposed system can be roughly divided to three parts. First, instruments in camera image are segmented by the color-based segmentation method. The second part is feature extraction. Since the surgical instruments are mainly different on their appearance in substance, some shape and contour information such as aspect ratio, area and Fourier descriptors are chosen as features for classification. The third part is surgical instruments classification. In order to identify the surgical instruments, a two-stage classification scheme which is formed by Naïve Bayes classifiers is designed and applied in our system.

The remaining part of this paper is organized as follows: in section II, an overview of the proposed system is given. Section III describes the methods used in the proposed system. Section III-A reveals the color-based object segmentation method. Section III-B illustrates how to generate and select features for classification. In section III-C, we make a

This work is supported by Chang-Gung Memorial Hospital under Grant CMRPD2C0042 and Ministry of Science and Technology, Republic of China, Taiwan, under Grant NSC102-2221-E-182- 015-MY3.

X. H. Liu and C. H. Hsieh are with the Department of Electrical Engineering, Chang-Gung University, 259 Wen-Hwa 1st Road, Kwei-Shan Tao-Yuan, Taiwan, 333, R.O.C. (e-mail: henryboy0322@hotmail.com, kurenai0413@hotmail.com).
J. D. Lee is with the Department of Electrical Engineering, Chang-Gung University, 259 Wen-Hwa 1st Road, Kwei-Shan Tao-Yuan, Taiwan, 333, R.O.C. (corresponding author, phone: +886-3-2118800#5316; e-mail: jdlee@mail.cgu.edu.tw).
S. T. Lee and C. T. Wu are with Department of Neurosurgery and Medical Augmented Reality Research Center, Chang Gung Memorial Hospital, Tao-Yuan, Taiwan, 333, R.O.C.

978-1-4799-5846-7/14/$31.00 ©2014 IEEE

description of the two-stage classification scheme. Section IV show the experimental results of the proposed system, and finally conclusion about this work is included in in section V.

## II. SYSTEM OVERVIEW

Fig. 1 shows a flowchart of the proposed system. A lot of images of several surgical instruments were collected to form a surgical instruments image database. These images were considered as training data. In the offline training stage, various shape features such as aspect ratio, area and Fourier descriptor were extracted from the object in the training image database. The extracted features were utilized to train naïve Bayes classifiers, and these classifiers were combined to form a two-stage classification scheme for online surgical instruments classification.

In the online stage, the proposed system mainly follows these procedures: object localization, feature extraction and classification. At first, all surgical instruments in camera field of view are segmented by using color-based segmentation method. Then, labeling using connected component concept is applied to the segmented image to localize region of each instrument. Area, aspect ratio and Fourier descriptor of each instrument are extracted as features for classification. Features belong to a single object are then applied to the two-stage classification scheme to determine which instrument it is.

## III. METHODS

### A. Color-Based Object Segmentation

In order to recognize and identify the surgical instruments, at first the instruments should be localized in the camera field of view. In general, surgical instruments are usually put on the table, and the table is fully covered by a sterile surgical drape. Since the surgical drapes usually has uniform color, color information is then used to differentiate background and object, i.e., to separate instruments area from the surgical drape.

In the proposed system, HSV color space is adopted for color-based object segmentation. HSV color segmentation is better than RGB because that HSV space separates intensity from the chromaticity, and it is less affected by the light changes [10]. HSV color space is usually represented as a cylinder or a cone. In the HSV color space, hue describes the main property of a color, which varies from 0° to 360°. Saturation stands for the purity of the color, it varies from 0 to 1. V stands for the intensity of a pixel.

Max is denoted as the maximum value of a single pixel (R, G, B), and Min is the minimum value of (R, G, B). RGB color space can be transfered to HSV space according to following equations:

$$H = \begin{cases} 0^\circ & ,\text{if Max} = \text{Min} \\ 60^\circ \times \frac{G-B}{\text{Max}-\text{Min}} & ,\text{if Max} = \text{R and G} > B \\ 60^\circ \times \frac{G-B}{\text{Max}-\text{Min}} + 360^\circ, & \text{if Max} = \text{R and G} < B \\ 60^\circ \times \frac{B-R}{\text{Max}-\text{Min}} + 120^\circ, & \text{if Max} = \text{G} \\ 60^\circ \times \frac{R-G}{\text{Max}-\text{Min}} + 240^\circ, & \text{if Max} = \text{B} \end{cases} \quad (1)$$

$$S = \begin{cases} 0 & ,\text{if Max} = 0 \\ \frac{\text{Max}-\text{Min}}{\text{Max}} = 1 - \frac{\text{Min}}{\text{Max}}, & \text{otherwise} \end{cases} \quad (2)$$

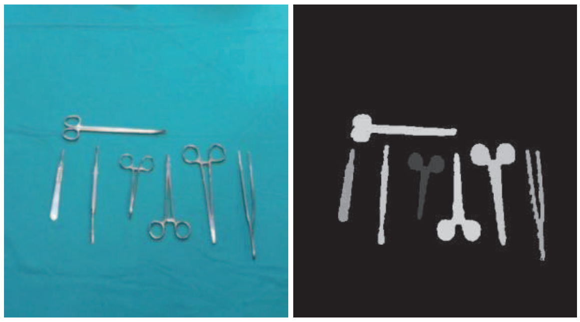

Figure 2. An example of segmentation and labeling result.

$$V = \text{Max} \quad (3)$$

By analyzing the histogram distribution of both H channel and S channel, we can easily set the thresholds to segment objects from the surgical drape in the frame. After the color-based segmentation, connected component labeling [11] is adopted to examine regions in the binary image. Because the camera field of view and distance are fixed during the image capturing, the region of a candidate surgical instrument in the frame will not too small. Therefore, scattered and fragment regions with too small area will be neglected. Fig. 2 shows the result of surgical instruments segmentation, where the segmented regions are labeled in different values.

### B. Feature Extraction

In pattern recognition, features are extracted to describe partial information of a region in the image, and this information is further applied for matching or classification. As mentioned above, it is observed that the appearance is the main diversity between different types of the surgical instruments. Therefore, the classification problem can be simply considered as a shape classification problem. Shape contexts are usually utilized as a feature to describe the appearance characteristic of a shape. Good shape contexts which can be used as a feature for shape classification should have the following characteristics: translation invariant, rotation invariant and scaling invariant. In other words, these features must be affine-invariant.

In [12], various shape features were surveyed and the applicability of them had been listed. In the proposed system, Fourier descriptor was first considered as a feature for classification, since it shows good performance against different kinds of situation, such as affine transformation and noises. Fourier descriptors uses Fourier coefficient to describe characteristic of a shape or a contour [13]. The steps of extracting Fourier descriptors are described as follows:

At first, contours of the object region are extracted by using the border following algorithm [14]. Once the contour of an object is obtained, the distances between contour points and the contour centroid are calculated. Since contours of different object different number of points, contour points need to be sampled to N points to let the descriptors in same dimension. Contour distance the centroid $(x_c, y_c)$ can thus be denote as:

$$d(k) = \sqrt{\left((x_k - x_c) + (y_k - y_c)\right)^2} \quad k = 1,2,3 \ldots N \quad (4)$$

Here, N is set to 28 to maintain the descriptors with the same dimension of 128. Fourier transform is applied to the contour-centroid distances function *d(k)* to get the Fourier coefficients *F(n)*:

$$F(n) = \frac{1}{N}\sum_{k=0}^{N-1} d(k)e^{-\frac{j2\pi nk}{N}} \quad (5)$$

The magnitudes of the Fourier coefficients obtained by Eq. (5) were used to comprise the descriptor of the segmented object, as shown in Eq. (6):

$$FDs = \{|F_0|, |F_1|, \ldots, |F_k|\} \quad (6)$$

In addition, since most of the surgical instruments can be considered as rigid body and they will not deform, area and aspect ratio of the surgical instruments are also chosen as the features for surgical instruments classification. To calculate aspect ratio of the segmented object, considering the surgical instruments will be random positioned with random orientation in the camera field of view, principle component analysis (PCA) is used to find the major and minor axes of a segmented region. The contour points of segmented object are denoted as a 2-dimentional data point group, and the covariance matrix of these contour data points is calculated. Then we find the eigenvalues and corresponding eigenvectors of the covariance matrix. Since these data points are two-dimensional, only two eigenvectors are obtained. The eigenvector corresponding to the bigger eigenvalue is the major axis of the data points, and the eigenvector corresponding to the smaller eigenvalue will be the minor axis of the data points. By finding borders of the object contours along these two new axes, the bounding rectangle of the segmented object can be obtained. Fig. 3 gives some examples of the bounding rectangle after applying PCA to the segmented object region. Aspect ratio of the segmented surgical instrument can be calculated through the bounding rectangle. On the other hand, area of the segmented object can be calculated by simply counting number of points inside the

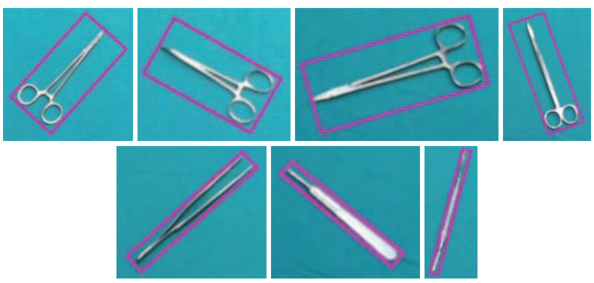

Figure 3. Several examples of PCA bounding rectangle results. The bounding rectangles are utilized to calculate aspect ratio.

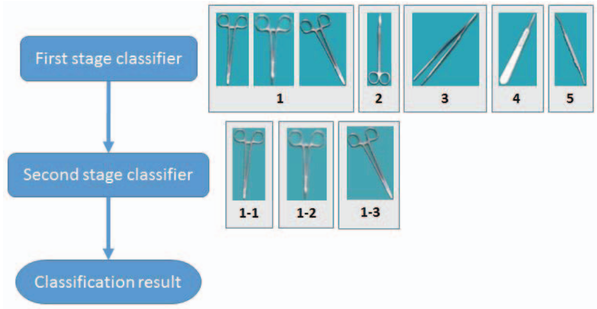


Figure 4. Two-stage classification scheme

segmented object contour.

## C. Classification

In machine learning, Naïve Bayes classifier is one of the classifiers which can be easily implemented and have good performance. Naïve Bayes classifier is mainly based on the Bayes' theorem. It uses the concept of maximum a posteriori estimation [15]. Given an event Y and an evidence X, the posterior probability of event Y under the condition of X happens can be denoted as:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} \quad (7)$$

In the case of sufficient samples, we can estimate the probability of evidence X under happening of event Y, probability of evidence X and probability of event Y. We can thus calculate the probability of event Y under the condition of X. In other words, if we observed a feature X, we can know this feature belong to which class according to the calculated maximum posterior probability.

In the proposed system, these features we utilized is multi-dimensional. Assuming that the evidence feature vector has N dimension, feature vector $\boldsymbol{X}$ can be described as $\boldsymbol{X}=(x_1, x_2, \ldots, x_N)$. Assume there are $K$ instruments used in this study, then we have $K$ classes of event $\{Y_1, Y_2, \ldots, Y_K\}$, where $Y_i$ stands for the event for surgical instrument $i$. According to Bayes' theorem, the a posteriori probability can be formulated as follow:

$$P(Y_i|x_1, x_2, \ldots, x_N) = \frac{P(Y)P(x_1, x_2, \ldots, x_N|Y_i)}{P(x_1, x_2, \ldots, x_N)} \quad (8)$$

By comparing $P(Y_1|\boldsymbol{X}), P(Y_2|\boldsymbol{X}), \ldots, P(Y_K|\boldsymbol{X})$, a class label can be assigned to a feature vector $\boldsymbol{X}$ obtained from the object according to the maximum a posteriori probability, as shown in Eq. (9).

$$C = argmaxP(Y_i|\boldsymbol{X}) \quad (9)$$

Different kinds of feature combinations were examined to find out which feature can most efficiently classify the surgical instruments. Detailed experimental and comparison results are described in section IV. After several testing and observing the results, finally a two-stage classification scheme for the proposed system is derived. Fig. 4 gives a diagram of the two-stage classification scheme. Since some of the instruments are similar in their aspect ratio, these instruments were regarded as a same group in the first stage of classification. According to the experimental results, Fourier descriptor and aspect ratio were combined to construct a 129

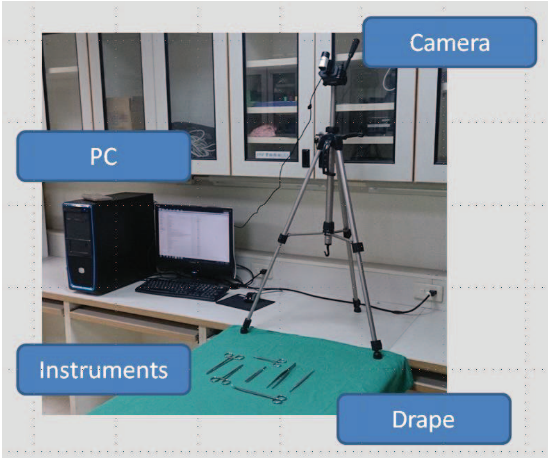


Figure 5. Experimental setting of the proposed system

dimensional feature vector to train the first-stage classifier. As shown in fig. 4, after the first stage if a surgical instrument is identified as group 1, it will be classified by the second-stage classifier. In the second-stage classifier, area is utilized as the feature to classify the grouped instruments.

## IV. Experimental Results

Some experiments had been conducted to evaluate the performance of the proposed system. As mentioned before, different feature combinations had been evaluated, and we finally give a conclusion which feature should be selected for classification of the surgical instruments. This section reveals all of the results

### *A. Hardware, development platform and experiment setting*

The proposed system is running on a PC with Intel i5 3.1GHz CPU and 12GB RAM. An USB camera with 1280x720 resolution is utilized as image acquisition device. Experimental scenario was set as shown in figure 5. The table was cover by a green surgical drape, and surgical instruments were placed on the table. Camera was positioned one meter from the table and with fixed position and field of view. The system had been developed using Microsoft Visual Studio in C++ along with Open Source Computer Vision (OpenCV) [16] library.

### *B. Experiments of surgical instruments classification*

Seven types of surgical instruments are involved in the proposed system. We give each surgical instrument a label number for the convenience on description and programing. Name, label number and appearance of these surgical instruments are listed in Table I. In order to train the classifiers, 100 images for each of the surgical instrument were took to form an image database. Therefore, totally 700 surgical instrument images were utilized to train Naïve Bayes classifiers using different types of feature combinations. Accuracy evaluation was carried out by using 14 images with all instrument in the camera field of view. Thus there are 98 cases for counting the accuracy. For every testing image, the surgical instruments were placed on the surgical drape with random position and orientation. Figure 6 shows two examples of the testing images.

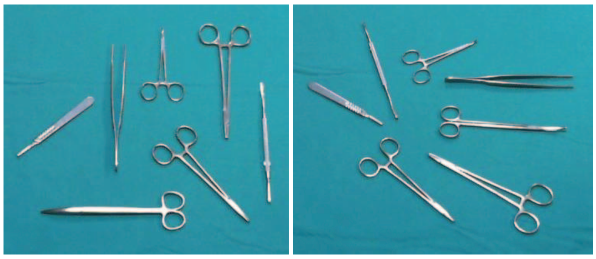

Figure 6. Two examples of the testing image.

In this study, Fourier descriptor (FD), aspect ratio (AR), and area were used for features for classification. First the classification accuracy was evaluated by using single stage classification scheme. In this experiment, each surgical instrument is regarded as a single class. The seven classes of surgical instruments are classified by Naïve Bayes classifier using different feature selection. Four feature combinations were tested: Fourier descriptors only (FD), Fourier descriptors with aspect ratio (FD+AR), Fourier descriptors with area (FD+Area), and all of the three features. Accuracies of every feature combination were calculated and recorded for comparison. Table II reveals the results of the single stage classification scheme. From the results, we see that the accuracy is unsatisfied, i.e., FD+Area is the best among four feature combinations, however, it only gets 82.65% accuracy.

By analyzing result of the first experiment and observing appearance of the surgical instruments, we considered that hemostats are almost similar in the aspect ratio. Difference between the hemostats is only in scaling. Basing on this assumption, we group the hemostats into a class in the first stage of the two-stage classification scheme. Therefore, it becomes a five classes classification problem in the first stage. We examined which combination of the features is best for the grouped instruments, and Table III shows the results of using different feature combinations in the first stage of the

TABLE I. Seven Surgical Instruments which are used in this Study

| Label | Surgical instruments | |
|---|---|---|
| | Name | Appearance |
| 1-1 | Mosquito Hemostat |  |
| 1-2 | Curved Mosquito Hemostat |  |
| 1-3 | Long Mosquito Hemostat |  |
| 2 | Long Scissor |  |
| 3 | Forcep |  |
| 4 | Fine Tip |  |
| 5 | Probe |  |

TABLE II. ACCURACY OF USING SINGLE STAGE CLASSIFICATION SCHEME WITH DIFFERENT FEATURE COMBINATIONS

| | Feature type | | | |
|---|---|---|---|---|
| | ***FD*** | ***FD+AR*** | ***FD+Area*** | ***FD+AR+Area*** |
| Accuracy | 75.51% | 77.55% | 82.65% | 41.8% |

TABLE III. ACCURACY OF GROUPED INSTRUMENTS AFTER FIRST STAGE CLASSIFIER UNDER DIFFERENT FEATURE COMBINATIONS

| | Feature type | | | |
|---|---|---|---|---|
| | ***FD*** | ***FD+AR*** | ***FD+Area*** | ***FD+AR+Area*** |
| Accuracy | 92.70% | 95.92% | 93.87% | 80.61% |

two-stage classification scheme. FD+AR feature can classify the grouped surgical instruments most efficiently.

In summary, with the two-stage classification scheme, FD+AR is first used to train the first-stage classifier, and area is utilized as second-stage feature. The overall accuracy is 90.82%.

## V. CONCLUSION AND FUTURE WORKS

This paper presents an online surgical instruments classification system. This system can be used for applications of surgical instruments monitoring or medical robotics. The classification is accomplished by applying shape and contour features of the object to a two-stage classification scheme which is mainly based on Naïve Bayesian classifier.

In this study, only seven surgical instruments are used to evaluate the proposed system, and the accuracy reach 90.8% in classifying these instruments. Ideally, according to the system and algorithm design, if we want to apply the proposed system to the case with more surgical instruments, we only need to add more training images of new instruments to the training image database. However, it needs to be further approved in practical situation.

From the experimental results, it is observed that to classify hemostats is still a challenge. Tip of the surgical instruments is also a feature to differentiate. In the future work, we will try to include the tip information to improve the system performance. Moreover, we are looking for integrating the proposed system to medical robotics, and testing it in practical situation in the future.

## REFERENCES

[1] G. D. Hager, A. M. Okamura, P. Kazanzides, L. L. Whitcomb, G. Fichtinger and R. H. Taylor, "Surgical and Interventional Robotics: Part III: Surgical Assistance Systems," *IEEE Robot Autom. Mag.*, vol. 15, no. 4, pp. 84-93, Dec. 2008.

[2] K. Cleary and A. Kinsella, "OR2020 workshop repor: The operating room of the future," *CARS 2005 computer assisted radiology and surgery*, Elsevier, Amsterdam, pp. 832–838, May 2005

[3] H. U. Lemke and M. W. Vannier, "The operating room and the need for an IT infrastructure and standards," *Int. J. Comput. Ass. Rad.*, vol. 1, no. 3, pp. 117-121, Nov. 2006.

[4] D. A. Wiegmann, A. W. ElBardissi, J. A. Dearani, R. C. Daly and T. M. Sundt III, "Disruptions in surgical flow and their relationship to surgical errors: an exploratory investigation," *Surgery*, vol. 142, no. 5, pp. 658-665, Nov. 2007.

[5] T. Neumuth and C. Meißner, "Online recognition of surgical instruments by information fusion," *Int. J. Comput. Ass. Rad.*, vol. 7, no. 2, pp. 297-304, Mar. 2012.

[6] K. Yamashita, K. Kusuda, Y. Tokuda, A. Onishi, K. Tanaka, M. Shimizu, R. Kanda, H. Honda, M. Komino, Y. Iwakami, Y. Ohta, and T. Okudo, "Validation of cleaning evaluation of surgical instruments with RFID tags attached based on cleaning appraisal judgment guidelines," in *Proc. IEEE Eng. Med. Biol. Soc.*, pp. 926-929, July 2013.

[7] S. Wang, X. Yin, B. Ge, Y. Gao, H. Xie, and L. Han, "Maschine vision for automated inspection of surgical instruments," *3rd Int. Conf. Bioinf. Biomed. Eng.*, pp. 1-4, June 2009.

[8] E. Carpintero, C. Pérez,, R. Morales, A. Candela, and J. M. Azorín, "Development of a robotic scrub nurse for the operating theatre," *3rd IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics*, pp. 504-509, Sept. 2010.

[9] F. Tan, S. Wang, H. Wang, and J. Zheng, "A new method of surgical instruments automatic identification and counting," *4th Int. Conf. on Image and Signal Processing*, pp. 1797-1800, Oct. 2011

[10] S. Sural, G. Qian, and S. Pramanik, "Segmentation and histogram generation using the HSV color space for image retrieval," *in Proc. on Int. Conf. on Image Processing*, vol. 2, pp. 589-592, 2002

[11] M. Dillencourt, H. Samet, and M. Tamminen, "A general approach to connected-component labeling for arbitrary image representations," *J. ACM*, vol. 39, no.2, pp. 253-280, Apr. 1992

[12] M. Yang, K. Kpalma, and J. Ronsin, "A survey of shape feature extraction techniques," *Pattern Recognition Techniques, Technology and Applications*, 2003

[13] D. Zhang and G. Liu, "Shape-based image retrieval using generic Fourier descriptor," *Signal Process-Image*, vol. 17, no. 10, pp. 825-848, Nov. 2002

[14] T. D. Haig, Y. Attikiouzel, and M. D. Alder, "Border following : new definition gives improved borders," *IEE Proceedings I: Communications, Speech and Vision*, vol. 139, no.2, pp. 206-211, Apr. 1992

[15] M. Aly, "Survey on multiclass classification methods," technical report, Caltech, USA, 2005

[16] OpenCV, "http://opencv.org//"