

# Postage Stamp Recognition System (PSRS)

C. Pornpanomchai and P. Leerasakultham

Faculty of Information and Communication Technology, Mahidol University, Bangkok, Thailand

Email: chomtip.por@mahidol.ac.th, pimprapai.lee@gmail.com

**Abstract**—There are various kinds of pictures on postage stamps, which are very difficult for people and philatelists to identify them. Therefore, the objective of this research is to develop a computer system which is capable of recognizing some postage stamp images by using the digital image processing (DIP) technique. The developing system is called “Postage Stamp Recognition System” or “PSRS”. The PSRS consists of five sub-systems, which are: 1) image acquisition, 2) image preprocessing, 3) feature extraction, 4) image recognition, and 5) result presentation. The experiments were conducted on 3,593 postage stamp images, which are divided into 3,093 images for a training data set, and 500 images for testing the system. The precision of the system is 93.60 percent, with the average access time of 1.43 seconds per image.

**Index Terms**—postage stamp, content-based image retrieval, digital image processing, pattern recognition, euclidean distance

## I. INTRODUCTION

A postage stamp is a small piece of paper, which sticks on an envelope or a parcel post and is used for paying the postal fair. The first postage stamp in the world was called “Penny Black”, which was issued in May 1840 in Britain (as shown in Fig. 1(a)), and the first U.S. postage stamp is in July 1847, which depicted Benjamin Franklin and George Washington (as shown in Fig. 1(b) and Fig. 1(c)). The Ceres series was the first postage stamp series of France, which was issued in 1849 (as shown in Fig. 1(d)). The first Russian postage stamp was issued in January 1858 (as shown in Fig. 1(e)). A pair of dragons facing Japanese characters was the first Japanese postage stamp issued in April 1868 (as shown in Fig. 1(f)). The “Large Dragons” was China’s first postage stamp, issued in 1878 (as shown in Fig. 1(g)), and the first postage stamp in Thailand was called “Solot”, which was issued in August 1883 (as shown in Fig. 1(h)). The first postage stamps were issued around 1840, which was more than 170 years ago, and nowadays, the postage stamps are still used for paying the postal fair.

Normally, the topical stamp collections are classified in many categories, which depend on the collectors. Some popular topical stamp collections are as follows (Hornadge, 2011): [1]

**Famous person’s face:** many postage stamps illustrate the faces of kings, queens, presidents, prime ministers or famous persons. For example, the King Bhumiphon of Thailand is shown in a stamp in Fig. 2(a)

**Map or flag:** many countries show their country map and flag in the postage stamp. The Thai flag and map stamp image is shown in Fig. 2(b).

**Animals:** there are many kinds of animals, which are very beautiful and represent the local animal, shown in the postage stamp. A tiger, an open-billed stork bird, a clown knife fish and a butterfly (Papilio Palinulus) stamp image are shown in Fig. 2(c)-Fig. 2(f), respectively.

**Flowers:** not only the local animals but also the beautiful flowers are also illustrated in postage stamps. The golden-shower flower stamp is shown in Fig. 2(g).

**Cultural heritage:** there are many cultural heritages and important festivals displayed on stamp pictures. The grand palace stamp image is shown in Fig. 2(h), including Song-Kran festival stamp image shown in Fig. 2(i).

**Historical events :** the important events, namely: Olympic Games, Asian Games, etc. are frequently illustrated on postage stamp images. For example, the Asian-games V in Bangkok 1966 is shown in Fig. 2(j) and this figure also demonstrates a sport image on a postage stamp.



Figure 1. The first postage stamp of various countries, (a) Britain, (b) and (c) U.S.A., (d) France, (e) Russia, (f) Japan, (g) China, and (h) Thailand (Pictures from the Internet)



Figure 2. The various kinds of postage stamp images. (a) King Bhumiphon’s Face, (b) Thai Flag and map, (c) tiger, (d) open-billed stork bird, (e) a clown knife fish

Manuscript received April 3, 2014; revised June 26, 2014.

## II. LITERATURE REVIEWS

There are various kinds of stamp images and all the images represent some important people, cultural heritage, festivals, animals, plants and flowers, etc. It is very difficult for people and stamp collectors to identify and understand images on a postage stamp. Therefore, the main objective of this research is to recognize some stamp images including their series, series name and issuance date. There are many researchers who have applied various techniques to recognize a small picture on a stamp. The briefly summary of each technique is given below.

### A. The Content Based Image Retrieval (CBIR)

The CBIR technique segments an image into many regions and applies image processing methods to recognize each region. Afterwards each region is identified and then combined with the whole image result. The researchers who used the CBIR technique are the following:

Choras *et al.*, 2007 applied color, texture and shape with CBIR to identify a postage stamp image. The experiment was conducted on 1,300 images in a system database and tested on 500 postage stamp images [2].

Takala *et al.*, 2005 used the CBIR method and many image features (color, texture, motive, size and aspect ratio) to retrieve about 900 German stamp images [3].

Goel *et al.*, 2013 demonstrated the CBIR with soft dipole vector to recognize 200 stamp images [4].

### B. Fuzzy Logic Technique

The fuzzy logic theory is a method of solving the artificial intelligent problems. The fuzzy set method works by grading the membership function of the interesting problems and then classifying them. Some postage stamp image recognition by a fuzzy logic set includes the following.

Soft Pro 2010 established the stamp recognition software or SRS. The system applied fuzzy logic engine to search for the stamp image. The SRS consists of more than 18,000 unique US postage stamps [5].

Sigelkow and Burkhardt, 2002 developed a fuzzy set histogram to retrieve and classify postage stamp images. The experiment was conducted on 12,000 stamp images in the system database [6].

Herendeen *et al.*, 2011 implemented a fuzzy logic set with color shading to discriminate postage stamps issued in Slovenia from 1919-1921 (Scott Yugoslavia 3LJ1-7 and 3LJ8-14) [7].

### C. Euclidean Distance Technique

The Euclidean distance is a mathematical method to find the distance or similarity between two objects. The Euclidean distance equation is shown in (1):

$$ED = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2} \quad (1)$$

where *ED* is the Euclidean distance value, *n* is number of features, *X<sub>i</sub>* is a value of feature *i<sup>th</sup>* in the system database, *Y<sub>i</sub>* is a value of feature *i<sup>th</sup>* of an unknown stamp image. The researchers who applied the Euclidean distance technique to recognize postage stamps are listed below.

Miciak 2006 used the color information to recognize postage stamps by using the Euclidean distance method. The experiment was conducted on 192 postage stamp images [8].

Kumar and Sharan, 2007 employed the Garbor wavelet grid to find the region of an image and used PCA and Euclidean distance to recognize human faces, postage stamps, typing characters and currency banknotes [9].

Based on the previous researches, the PSRS adopts the CBIR technique to identify a sophisticate stamp image by dividing it into four regions, and then finds the features of each region. Finally, the PSRS uses all features in the four regions to recognize a stamp by using the Euclidean distance technique. The PSRS system design and implementation are presented in the next section.

## III. METHODOLOGY

The experiment was conducted using the following computer hardware specifications: 1) CPU Intel® Pentium (R) D 3.4GHz, 2) Memory DDR3 2GB and 3) Harddisk 600GB. The computer software used Microsoft Windows XP (Microsoft Corp.; Redmond, WA, USA) as the operating system and MATLAB R2013a (The MathWorks Inc.; Natick, MA, USA) for the development tool.

### A. Conceptual Diagram

The PSRS system conceptual diagram is illustrated in Fig. 3, which starts from a stamp corrector scanning an unknown stamp image and then importing to the PSRS. After that, the PSRS recognizes the stamp image by using DIP. Finally, the recognition results include a set of stamp images displayed on the user screen.

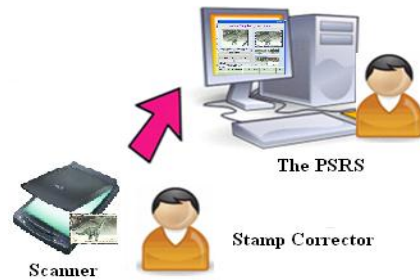


Figure 3. The PSRS conceptual diagram

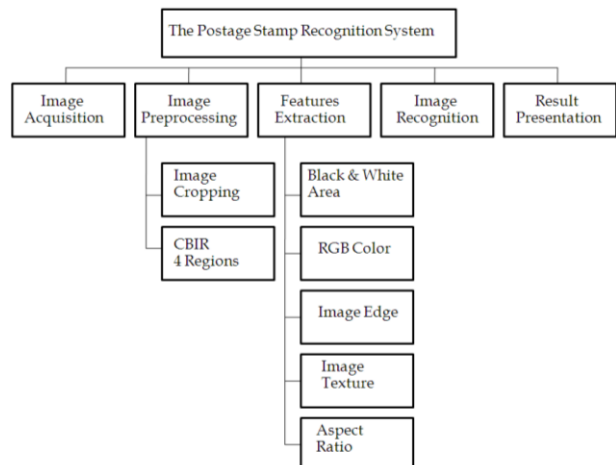


Figure 4. The PSRS structure chart

**B. Structure Chart**

For more understanding, the system structure chart, which maps to the PSRS conceptual diagram, is shown in Fig. 4. The PSRS consists of 5 sub-systems, which are: 1) image acquisition, 2) image preprocessing, 3) feature extraction, 4) image recognition, and 5) result presentation. Each sub-system has the following details.

1) *Image acquisition*: this sub-system captures a postage stamp image by using a scanner machine. The image should be in the proper orientation. A skew image below 5 - 10 degrees is acceptable (as shown in Fig. 5(a)).

2) *Image preprocessing*: this sub-system consists of two modules, which are image cropping and image separating into 4 CBIR regions. Each module has the following details.

*Image cropping*: this module reduces stamp perforation out and crops only the stamp image (as shown in Fig. 5 (b)).

*CBIR 4 Regions*: this module separates a cropping image into 4 regions (as shown in Fig. 5 (c)) and the (x, y) coordination of each region calculates by using (2) to (5):

$$\text{Region number 1} = [(15, 15), (x, y)] \quad (2)$$

$$\text{Region number 2} = [(x-15, 15), (x, y)] \quad (3)$$

$$\text{Region number 3} = [(15, y-15), (x, y)] \quad (4)$$

$$\text{Region number 4} = [(x-15, y-15), (x, y)] \quad (5)$$

where  $x$ =image width/3,  $y$ =image height/3 (in pixel). The first pair variables in the first square bracket is the starting coordination, and the second pair variables is extended value from the starting point to  $x$ -pixel in horizontal axis and  $y$ -pixel in vertical axis.

3) *Feature extraction*: there are five modules in this sub-system, which have the following details.

*Black and White area*: this feature counts the number of black pixels in a black & white image. This module starts by converting RGB image to grayscale image by using (6):

$$GR = 0.299 * R + 0.587 * G + 0.114 * B \quad (6)$$

where  $GR$  is grayscale color,  $R$  is red color,  $G$  is green color and  $B$  is blue color. After obtaining the grayscale image, the system converts it into a black and white image using the highest histogram to be a threshold value. The grayscale and Black and White image are shown in Fig. 5(b) and Fig. 5(d), respectively. The black area of the dinosaur stamp image in Fig. 5 (a) is equal to 12,143.75.

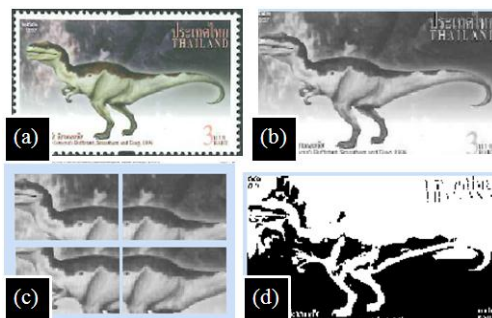


Figure 5. The postage stamp image after passing image preprocessing process (a) An original image (b) a cropping image in grayscale, and (c) the CBIR 4 regions image (d) a Black and White image

*RGB color*: this feature measures the mean value of a red, green and blue color of a stamp image. The (7)-(9) use for calculating the mean values of a red, green and blue color, respectively.

$$Mr = r/N \quad (7)$$

$$Mg = g/N \quad (8)$$

$$Mb = b/N \quad (9)$$

where  $Mr$  is a mean of red color,  $Mg$  is a mean of green color,  $Mb$  is a mean of blue color,  $r$  is a total of red pixels,  $g$  is a total of green pixels,  $b$  is a total of blue pixels, and  $N$  is a total number of image pixels. The  $Mr$ ,  $Mg$  and  $Mb$  of dinosaur stamp image in Fig. 5 (a) are equal to 166.43, 169.87 and 162.22, respectively.

*Image edge*: the PSRS used 2 edge detection techniques, which are the Sobel edge detection and the Canny edge detection, which have the following details:

(a) *Sobel edge*: there are two  $3 \times 3$  templates used in the Sobel edge detection technique. Template "a" (Fig. 6(a)) is applied to determine the difference in the horizontal axis and template "b" (Fig. 6(b)) is used for the vertical axis. The result from the Sobel edge detection process is shown in Fig. 6(c). The white color of Sobel edge in Fig. 6(c) is equal to 3,080.75 (Rebecca and Folorunso, 2009) [10].

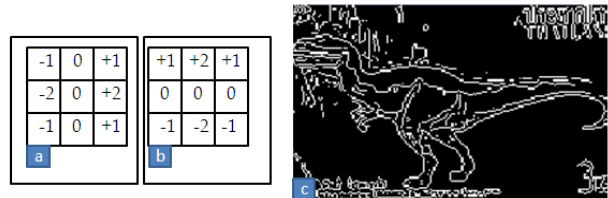


Figure 6. The postage stamp image after passing the Sobel edge detection (a) a vertical  $3 \times 3$  mask, (b) a horizontal  $3 \times 3$  mask, and (c) the Sobel edge detection result

(b) *Canny edge*: there are 5 steps in Canny edge detection technique, which are 1) smoothing image by using  $5 \times 5$  template as shown in Fig. 7(a), 2) finding gradients, 3) mark only local maximum edge, 4) determine potential edge by thresholding method, and 5) edge tracking. The result from the Canny edge detection process is shown in Fig. 7(b). The white color of Canny edge in Fig. 7(b) is equal to 3,814.38 (Canny, 1986 and Green, 2002) [11], [12].

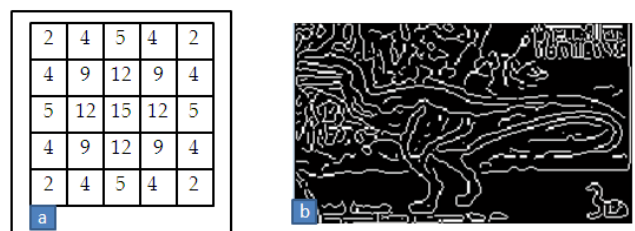


Figure 7. The postage stamp image after passing the Canny edge detection (a) a smoothing  $5 \times 5$  mask and (b) the Canny edge detection result

*Image Texture*: The entropy is a statistical measure of randomness that can be used to characterize the texture of the postage stamp image. The entropy can be found by



(10). The entropy texture of dinosaur stamp image in Fig. 5(a) is equal to 630.92 (Beliakov, *et al.*, 2008) [13].

$$\sum_{i,j=0}^{N-1} P_{i,j} (-\ln P_{i,j}) \quad (10)$$

where  $P_{i,j}$  is entry in a normalized gray-tone spatial-dependence matrix and  $N$  is a number of distinct gray levels in the quantized image.

*Aspect ratio:* this feature calculates by dividing image width by image height. The width and height of dinosaur stamp image in Fig. 5(b) are 603 and 120, respectively and the aspect ratio is equal to 5.03.

4) *Image recognition:* the PSRS employs the Euclidean distance method to recognize the postage stamp image by using 36 features in the feature extraction sub-system. Each stamp image consists of 36 features and 5 regions (one cropping image region and four CBIR regions). Each region contains 8 features, namely: 1) Black and White area in an image, 2) mean of red color, 3) mean of green color, 4) mean of blue color, 5) Sobel edge, 6) Canny edge, 7) entropy texture, and 8) aspect ratio. The aspect ratio applies only to an image cropping region but does not apply to in the CBIR regions because each CBIR region has the same size.

5) *Result presentation:* after the PSRS recognizes the stamp image, it will display the recognition result on the graphic user interface (GUI), including the stamp recognition details. The GUI of the PSRS is depicted in Fig. 8, which has the following details.

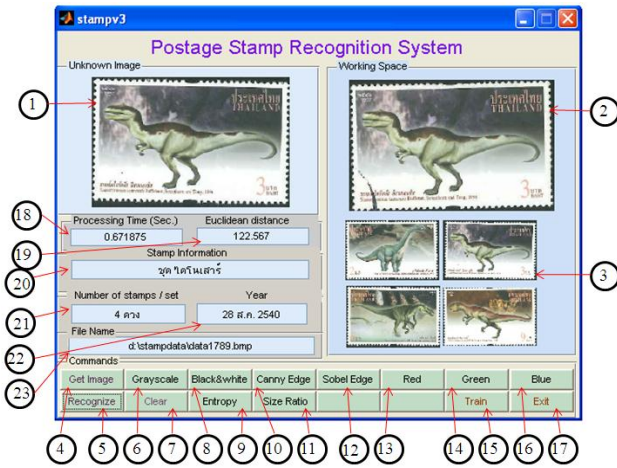


Figure 8. The GUI of the PSRS

*Image box:* there are three image boxes, namely: 1) input postage stamp image box (label number 1), 2) recognition image box (label number 2), and 3) stamp series image box (label number 3).

*Command button:* There are 14 command buttons, namely: 1) Get image file button (label number 4), 2) Recognition button (label number 5), 3) Grayscale button (label number 6), 4) Clear button (label number 7), 5) Black and White button (label number 8), 6) Entropy button (label number 9), 7) Canny edge button (label number 10), 8) Aspect ratio button (label number 11), 9) Sobel edge button (label number 12), 10) Red-color button (label number 13), 11) Green-color button (label number 14), 12) Blue-color button (label number 15), 13) Train button (label number 16), and 14) Exit button (label number 17).

*Text box:* there are six text boxes, which are: 1) processing time (label number 18), 2) Euclidean distance value (label number 19), 3) recognition stamp series name (label number 20), 4) number of stamps in the series (label number 21), 5) issuance year (label number 22), and 6) recognition stamp filename in the system database (label number 23).

#### IV. EXPERIMENTAL RESULTS

The experiment was conducted on 3,593 postage stamp images, which are divided into 3,093 images for a training data set, and 500 images for testing the system. The experiment was also conducted on 8 Euclidean threshold values which vary from 100 to 800, and the experimental results are shown in Table I. Based on the experimental result in Table I, the applicable threshold value is 400 because it gives a high matching rate including acceptable mismatch values. The precision rates of the system with the threshold value 400 are 93.6 percent, 2.6 percent and 3.8 percent for match, mismatch and unknown, respectively. The average recognition speed is 1.43 seconds per image.

In previous study, the stamp recognition software (SRS) developed by the Soft Pro 2001 can search an unknown postage stamp image and display the possible matching results. But in this research, the PSRS can not only matching and displaying the postage stamp images but also giving the stamp collectors more information. The PSRS provided the recognition stamp series information, which include: 1) images of all postage stamp in the series, 2) series name, 3) a total number of stamps in the series, and 4) the issuance year. Moreover, the PSRS can match between two different stamp images but they are in the same series, for examples, the PSRS can matching between used and unused stamp image, as shown in Fig. 9(a) and Fig. 9(b) and the PSRS can matching between two stamp images with different postmarks, as shown in Fig. 9(c) and Fig. 9(d). However, there are many similar stamp images, which give the mismatch results for examples, in Fig. 9(e) and Fig. 9(f), show two stamps consist of the King's face, which have a very similar color but they have a different price and Fig. 9(g) and Fig. 9(h) have the same designing pattern but they have different color and price.

#### V. CONCLUSION

In this research, the PSRS fulfilled the research objective by extracting four main stamp image features, namely: 1) RGB color, 2) edge, 3) texture, and 4) aspect ratio, for the recognition of stamp image samples. The experiment was conducted on 3,593 postage stamp images, which are divided into 3,093 images for a training data set, and 500 images for testing the system. The experiment was also conducted on the Euclidean distance method with the precision rates of the system are 93.6 percent, 2.6 percent and 3.8 percent for match, mismatch and unknown, respectively. The average recognition speed is 1.43 seconds per image.

TABLE I. THE PSRS VARIOUS THRESHOLD VALUES AND RESULTS

Threshold	#Testing	Match	Mismatch	Unknown
100	500	103	0	397
200	500	352	1	147
300	500	438	4	58
400	500	468	13	19
500	500	470	22	8
600	500	472	25	3
700	500	474	25	1
800	500	475	25	25



Figure 9. Some experimental results, (a) match with (b), (c) match with (d), (e) mismatch with (f), and (g) mismatch with (h)

REFERENCES

[1] B. Hornadge. (Apr. 22, 2013). Topical stamp collecting. *Philatelic Database*. [Online]. Available: <http://www.philatelicdatabase.com/topicals-thematics/topical-stamp-collecting/>

[2] R. S. Choras, T. Andrysiak, and M. Choras, "Integrated color, texture and shape information for content-based image retrieval," *J. Pattern Anal Applic.*, vol. 10, pp. 333-343, 2007.

[3] V. Takala, T. Ahonen, and M. Pietikainen, "Block-Based method for image retrieval using local binary patterns," *LNCS.*, vol. 3540, pp. 882-891, 2005.

[4] A. Goel, P. Saxena, A. Srivastava, and C. Sheth, "Study of dissociated dipole histogram for content based image retrieval using stamp recognition system," in *Proc. 4th International Conference on Intelligent Systems, Modeling and Simulation*, Bangkok, Thailand, 2013, pp. 246-250.

[5] *Stamp Recognition Software User's Manual*, SoftPro 2001 118 Palomino Drive, Sault Ste. Marie, Ontario, Canada, P6A 6K4, 2010

[6] S. Siggelkow and H. Burkhardt, "Improvement of histogram-based image retrieval and classification," in *Proc. 16th International Conference on Pattern Recognition*, Quebec city, Canada, 2002, pp. 367-370.

[7] D. L. Herendeen, J. A. Allen, and T. Lera, "Philatelic shade discrimination based on measured color," *London Philatelist*, vol. 120, pp. 105-117, 2011.

[8] M. Miciak, "The color information as a feature for postage stamp recognition," in *Biometrics, Computer Security Systems and Artificial Intelligence Applications*, Springer US, 2006, pp. 69-78.

[9] V. Kumar and S. Sharan, "Pattern recognition with localized gabor wavelet grids," in *Proc. International Conference on Computational Intelligence and Multimedia Applications*, Tamil Nadu, India, 2007, pp. 517-521.

[10] O. R. Vincent and O. Folorunso, "A descriptive algorithm for Sobel image edge detection," in *Proc. Informing Science and IT Education Conference*, Macon, GA, USA, 2009, pp. 97-107.

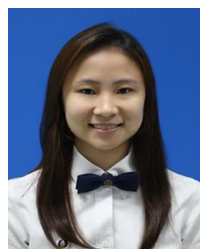
[11] J. Canny, "A computational approach to edge detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 679-698, 1986.

[12] B. Green. (Nov. 2013). Canny edge detection tutorial. [Online]. Available: [http://dasl.mem.drexel.edu/alumni/bGreen/www.pages.drexel.edu/\\_weg22/can\\_tut.html](http://dasl.mem.drexel.edu/alumni/bGreen/www.pages.drexel.edu/_weg22/can_tut.html)

[13] B. Gleb, J. Simon, and T. Luigi, "Texture recognition by using GLCM and various aggregation functions," in *Proc. International Conference on Fuzzy System*, Hong Kong, China, 2008, pp. 1472-1476.



**Chomtip Pornpanomchai** received his B.S. in general science from Kasetsart University, M.S. in computer science from Chulalongkorn University and Ph.D. in computer science from Asian Institute of Technology. He is currently an assistant professor in the faculty of Information and Communication Technology, Mahidol University, Bangkok, Thailand. His research interests include artificial intelligence and pattern recognition.



**Pimprapai Leerasakultham** was born in Bangkok Thailand. She received her B.S. in Information and Communication Technology, Mahidol University. She is currently a master degree student in the faculty of Information and Communication Technology, Mahidol University, Bangkok, Thailand.