Pill Recognition Using Imprint Information by Two-step Sampling Distance Sets

Jiye Yu
Graduate School of Information, Production and System
Waseda University
Kitakyushu, Japan
troyspur@fuji.waseda.jp

Zhiyuan Chen
School of Electronic Information and Electrical Engineering
Shanghai Jiao Tong University
Shanghai, China
shchen-zhiyuan@sjtu.edu.cn

Sei-ichiro Kamata
Graduate School of Information, Production and System
Waseda University
Kitakyushu, Japan
kam@waseda.jp

***Abstract*— Huge variety of medicine cures diseases. But unlabeled pills sometimes confuse people, even causing adverse drug events. This paper introduces a high accuracy automatic pill recognition method based on pill imprint which is a main discriminative factor between different pills. To describe the imprint information clearly, we propose a Two-step Sampling Distance Sets (TSDS) descriptor based on Distance Sets (DS) using a two-step sampling strategy. The two-step sampling strategy applies a resampling according to imprint segmentation, which divides an imprint into separated strokes, fragments and noise points. The TSDS is able to take control over the selection of feature points, aiming to cut down the noise points and unwished fragments generated by imprint extraction which will cause disturbance on recognition. In the aspect of the imprint extraction, we preprocess the pill image by dynamic contrast adjustment to cope with the exposure problem. Modified Stroke Width Transform (MSWT) is used to extract the imprint by detecting the coherent strokes on the pill. Finally, several experimental results have shown 86.01%, rank-1 matching accuracy, and 93.64%, within top 5 ranks, when classifying pills into 2500 categories.**

Keywords— pill recognition; imprint extraction; Two-step Sampling Distance Sets (TSDS); image retrieval

I. INTRODUCTION

Huge variety of medicine has been manufactured to benefits human beings. But it brings a problem that errors happen while classifying the unlabeled pills. Patients always have no idea to distinguish the unwrapped pills, which possibly even leads to the adverse drug events. To avoid these, pill identification websites are created to help people distinguish different kinds of pills. In U.S., Food and Drug Administration (FDA) issued the regulation code 21CFR206 [1] in order to enforce the unique look for every prescription pill on the market in terms of size, shape, color and imprint. These four features become the common marks of one pill, which can be used to distinguish different kinds of pills either by human eyes or by machines.

Constructing a stable system to accurately identify pills based on the pills database is what we focus on. Besides contour shape and color, imprint is the most discriminative feature to distinguish from one kind of pill from another. How to utilize the information of imprint is the key to realize a successful identification. Then two steps, extracting the imprint and describing the imprint, are used to implement the imprint information type-in system. To the imprint extracted from a pill image, flaws are hard to prevent. In other words, imprints extracted from different pill images of the same category pill always possess tiny differences. Most of the differences are just noise and unwished fragments caused by variance of luminance, which will cause disturbance on pill retrieval. Good imprint extraction method can do some improvement on this problem, but that is not enough. So when describing an imprint, this interference should be suppressed as much as possible. In this paper, based on Distance Sets (DS) [2], we propose a descriptor named Two-step Sampling Distance Sets (TSDS) using the two-step sampling strategy which can exclude the disturbance of the noise and unwished fragments by resampling the feature points according to imprint segmentation. And as regard extracting the imprint, we start with the preprocessing of image's contrast, improving the exposure problem to provide Modified Stroke Width Transform (MSWT) [3] a better environment to extract clear imprint patterns. At last, recognition system is set combining the effort of contour shape, color and the main factor, imprint, to achieve high accuracy and efficient image retrieval. Here contour shape and color features are used to filter pill categories.

The rest of the paper is organized as follows. Section II reviews the previous work. In Section III, we introduce the concept of TSDS. Section IV gives the introduction of recognition system. Section V presents the experimental results. Finally Section VI concludes the paper.

II. PREVIOUS WORK

In recent years, more and more medical authorities and research institutions pay their attention to the development of medicine identification system. According to the working manners, existing pill recognition systems can be divided into two categories: manual input method and automatic recognition method.

Manual input method can be used individually because of its usability. On the internet, we can find several websites offering the pill recognition tools, such as WebMD Pill Identification Tool [4], RxList Pill Identification Tool [5], Healthline Pill Identifier [6], Pillbox [7] and so on. These identification tools require people to input pill's shape, color and imprint or brand and then output the proper result. When

This work is partially funded by NSFC key project NO. 61133009, Techniques for Complicated Scene Modeling and Super-high Resolution Rendering

1051-4651/14 $31.00 © 2014 IEEE
DOI 10.1109/ICPR.2014.544

disposing large amount of pills, manual input method is not viable because it is time-consuming and costly in manpower.

Different from the manual input method, automatic method uses image process algorithm to get the required information while doing the identification. Process using such information to classify the inputs into categories does not need manual picking and typing in operation; that makes batch processing possible. Many algorithms have been developed to realize this process. Andreas Hartl [8] presented a mobile computer vision system, which can take pill size, color and contour shape into consideration. Young-Beom Lee's algorithm [9] succeeded to use the imprint as key information, in which imprint shape is extracted by means of edge detection method and feature vectors are generated based on edge values using Hu invariant moments [10]. Shape Distribution is introduced in [11] to measure the similarity between 3D shapes. Reference [12] and [13] makes use of this sense and applies it to deal with pill images. Just last year, we introduced the Weighted Shape Context to describe the pills' imprints, which are extracted by Modified Stroke Width Transform [3]. That algorithm has been proved to be effective for dealing with both debossed imprint and printed imprint. But regarding on various illumination and exposure conditions, it still displays a little bit lack of robustness.

III. TWO-STEP SAMPLING DISTANCE SETS

Unlike shapes in the usual sense, imprint images, may be constituted by several regions, not by a cohesive whole. But it still can be treated as a shape and applied to kinds of shape descriptors. Descriptor TSDS is proposed for imprint images, but it can also be used on other shape images with the similar characteristic such as trademarks and road signs.

Distance Sets [2] introduced by C. Grigorescu, et al is a kind of rich local descriptor. As a local descriptor, distance sets describe the distance between a given point and its K neighbor points on the shape's contour. After applied this descriptor to every sampling points for a pill image, set of distance sets can be got by just assembling these distance sets. In summary, distance set describes a local arrangement of points around the specified point, while set of distance sets describing the global spatial arrangement of the whole image.

Generally, a shape contour needs sampling before applying a descriptor and uniform sampling (Fig. 1 shows an example) is commonly used. However, different from single closed curve shape, the imprint shapes might be more complex and irregular. And there could be some noise points caused by the extraction process on the imprint image. So we suggest the two-step sampling strategy. That can be comprehended as resampling on the sampled imprint image to get the feature points, whose

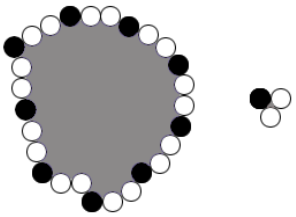

Fig. 1. Example of Uniform Sampling

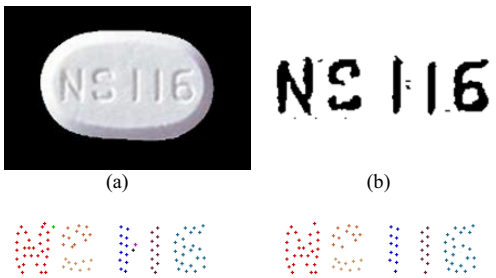

Fig. 2. (a) pill image; (b) the extracted imprint image; (c) first sampling points (9 regions); (d) second step sampling results; we can find that noise points are removed after second step sampling.

local descriptors (distance sets) should be used to constitute the set of distance sets to describe the whole imprint image.

As an imprint usually consists of some letters, symbols and other kinds of marks, it can always be divided into several separated regions, which can be regarded as imprint segmentation. Here, regions are independent strokes on the imprint. And if noise points appear, they can be regarded as regions, too. As a result, whether two points are in the same connected stroke shown in Fig. 2 (b) is regarded as the criterion of whether two points belong to one region or not. In that imprint, there are 6 regions in 5 letters ("N", "S", "1", "1", "6"; "S" is constituted by two regions because there is a breakage in the middle of the stroke) and several noise regions. Fig.2 (c) is the uniform sampling result of Fig. 2 (b). In this figure, these regions are shown in different colors (totally there are 9 regions in this figure).

Regions are constituted by the sampling points, so the number of sampling points can be used to evaluate the size of a region. Generally, these regions are always not even sized, and noise regions are relatively much smaller. According to the imprint segmentation, we can execute the two-step sampling strategy. The basic idea of second step sampling is:

1) The smaller a region is, the more points it should be cut.

2) A region relatively larger than other regions cuts fewer points.

By this way, noise regions, whose region size is always small, will be cut down.

Following is the proposed method in detail. First, we apply uniform sampling and get N_1 sampling points. Points in k-th regions can be shown as $R_k = \{p_1^{(K)}, p_2^{(K)}, ..., p_{\#R_k}^{(K)}\}$, where $\#R_k$ is the numbers of first sampling points in k-th region. Then we have the following equation,

$$N_1 = \sum_{k=1}^{T} \#R_k, \qquad k \in [1, T]. \tag{1}$$

Here, T is the number of regions.

Then we need to decide the number of second step sampling points $N_2, (N_2 < N_1)$. The difference $N_1 - N_2$ needs to be larger than the number of noise points. And the number

of points in each region after the second step sampling is $\#A_k$ ($\#A_k \le \#R_k$), which should meet $N_2 = \sum_{k=1}^{T} \#A_k$ ($\#A_k \le \#R$). Points in k -th region after second step sampling can be shown as $A_k = \{\tilde{p}_1^{(K)}, \tilde{p}_2^{(K)}, \cdots, \tilde{p}_{\#A_K}^{(K)}\}$. Based on the basic idea of second step sampling, $\#A_k$ is decided by

$$\#A_k = \#R_k - \left(\frac{\#R_k}{N_1}\right)^{\gamma} \cdot \frac{1}{r} \cdot (N_1 - N_2), \quad k \in [1, T], \tag{2}$$

where $r = \sum_{k=1}^{T} \left(\frac{\#R_k}{N_1}\right)^{\gamma}$, $\frac{1}{r}$ is a normalization factor, which aims to ensure $N_2 = \sum_{k=1}^{T} \#A_k, k \in [1, T]$, and $\gamma \in (0,1]$ is the curvature parameter. The smaller γ is, the more complete elimination of small regions will be done. The choice of γ is decided by experiments shown in section V. Note that when γ is close to 0, sometimes we can get an exceptional case that $\#A_k < 0$ happens. In that case, we can just set $\#A_k = 0$, and other regions apportion the lacking points. Fig. 2 (d) shows the second step sampling result, with $N_1 = 100, N_2 = 90, \gamma = 0.1$. We can find that the noise points are cut down completely, while the number-fixed feature points are in a quite uniform distribution.

From [2], we get the expression of local descriptor DS of point p to its K nearest neighbors points within N_1 sampling points of shape $S = \{R_1, R_2, ..., R_T\}$ like

$$DS_{S,K}(p) = \{d_1(p), d_2(p), ..., d_i(p), ..., d_K(p)\}. \tag{3}$$

Here, $d_i(p)$ is the distance between point p and its i -th nearest neighbor from shape S, $1 \le i \le K$. Then after using the two-step sampling strategy, TSDS can be expressed as

$$RDS(A_k) = \{DS_{S,K}(\tilde{p}_1^{(K)}), DS_{S,K}(\tilde{p}_2^{(K)}), \cdots, DS_{S,K}(\tilde{p}_{\#A_k}^{(K)})\}, \tag{4}$$

$$TSDS_K(S) = \{RDS(A_1), RDS(A_2), ..., RDS(A_k)\}, \tag{5}$$

where A_k is the k -th region after second step sampling. And distance measure between TSDS descriptors is the distinction between two shapes with their respective N_2 feature points. Let $\psi : S_1 \mapsto S_2$ be a point-to-point mapping from S_1 to S_2 and let Ψ be the set of all such mappings. Then a cost of the mapping $\psi \in \Psi$ is defined as follows,

$$C_K^{(\psi)}(S_1, S_2) = \frac{1}{N_2}\left(\sum_{i=1}^{N_2} D_{S_1,K,S_2,K}(\tilde{p}_i, \psi(\tilde{p}_i))\right), \tag{6}$$

where $D_{S_1,K;S_2,K}(p,q)$ is a dissimilarity between two distance sets $DS_{S_1,K}(p)$ and $DS_{S_2,K}(q)$. And a dissimilarity between $TSDS_K(S_1)$ and $TSDS_K(S_2)$ is defined by

$$\varepsilon_K(S_1, S_2) = \min\{C_K^{(\psi)}(S_1, S_2) \mid \psi \in \Psi\}. \tag{7}$$

We describe the process of imprint description by the diagram shown on Fig. 3. TSDS descriptor is applied to describe the imprint images. In the TSDS description process, second step sampling is done after calculating the distance sets for every first step sampling points. The reason why we do not calculate the distance sets after the second step sampling is that the noise points are fewer and dispersive, that affect the construction of distance sets little. And in another aspect, first step sampling (uniform sampling) preserves the primitive structure of imprint better.

Computation of the dissimilarity between two sets of distance sets can be reformulated in terms of minimum weight assignment problem in bipartite graph and solved efficiently in $O(v \times (e + v\log(v)))$ [14], (v and e are the number of vertices and edges of the associated graphs, respectively), which is still a little bit time-consuming when cardinality of set of distance sets is large. When processing high resolution images in a large dataset, a number of sampling points are needed. At that time, two-step sampling strategy can also be used to shorten the running time when keeping the precision well.

IV. Construction of Recognition System

Fig. 4 shows a diagram of our recognition system. It contains several parts, including features extraction and representation. Features of a pill include contour shape, color and imprint. In order to organize an efficient recognition system, contour shape and color features should also been taken into account. In our system, contour shape and color features are used to select pill categories before applying the recognition by imprint, which is the main part of the recognition system.

A. Contour Shape Feature

The feature for contour shape in our method is in a vector. By exploiting the fact that pills are convex objects, we use a vector V_i to represent the contour shape of pill i :

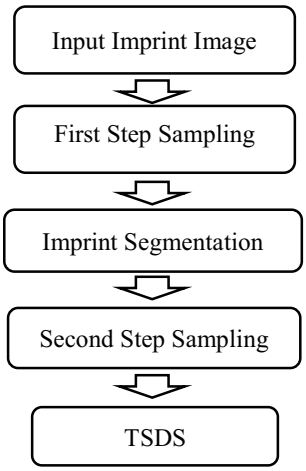

Fig. 3. Diagram of TSDS Process

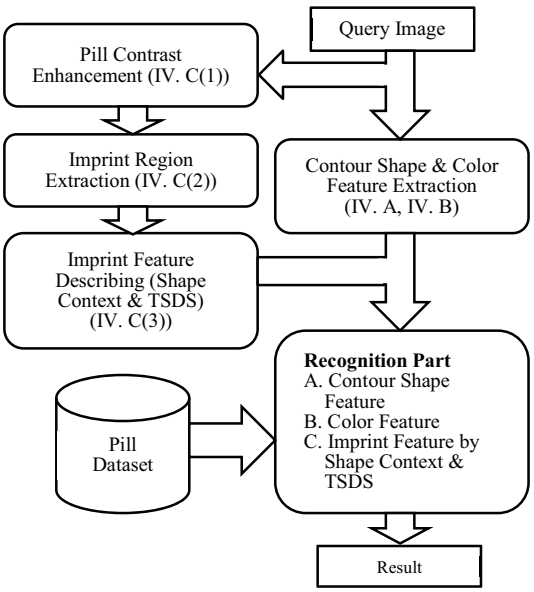

Fig. 4. Diagram of the Recognition System

$$V_i = (c_{i,1}, c_{i,2}, ..., c_{i,M}). \tag{8}$$

We need to uniformly space M points on the external boundary, and $c_{i,j}$ in (8) is just the distance between the center of pill and the j-th point. Fig. 5 shows an example. Here the distance can be measured in Euclidean distance.

Cross-correlation $r(V_a, V_b)$ is used to determine the degree to which two features V_a and V_b are correlated. The higher the score is, the more similar two contour shapes are

$$r(V_a, V_b) = \max(\sum_{j=1}^{M} V_{a,j} V_{b,(j+k) \bmod M}) \quad (0 \le k < M). \tag{9}$$

Here the subscript $(j+k) \bmod M$ introduces the cyclic shifting property. The Maximum value is obtained at the position when two shape are aligned.

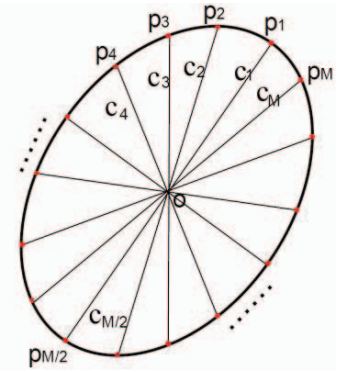

Fig. 5. Uniformly sampling M points on the boundary, construct the shape feature by computing the distance between center O and each point.

B. Color Feature

In order to eliminate the disturbance of luminance, we convert the pill images into HSV color model. V channel can be removed for its no contribution to the color information. Another reason of using HSV model not RGB model is that color in HSV model is more similar to the way human eye perceive color.

We build color histogram to construct the color feature. In color feature comparing, the metric for histogram matching uses intersection as the following formula [15], [16] for the sake of quick comparison:

$$d_{\text{intersection}}(H_1, H_2) = \sum_i \min(H_1(i), H_2(i)). \tag{10}$$

Here H_1 and H_2 are two histograms built from different pill images.

C. Imprint Feature

As the key information of one pill, imprint region should be treated seriously, including extracting the imprint as intact as possible and describing the imprint in detail.

1) Image Preprocessing

As a matter of fact, many pill images have such problems as overexposure or underexposure. Those conditions will interfere with the working of imprint extraction, and result in creating incomplete or even blank imprint images. In order to eliminate this problem, one method is to enhance the contrast as preprocess before applying any imprint extraction algorithm.

About contrast enhancement, there are several existed methods. Adjusting the histogram is one series of these methods. Global Histogram equalization is the most commonly used that can improve the image's contrast globally. But it will cause the loss of detail and import noise. Local Histogram Equalization [17] also cannot solve the noise problem. As a further optimization, Dynamic Histogram Equalization has been proposed in [18]. In this algorithm, histogram is partitioned by local minima. Based on the partitions, specific gray level ranges are applied. Then these partitions are equalized respectively. When the input image histogram has already spanned almost the full spectrum of the grayscale, this method widens the range of dominant grayscale partition, while shortening the non-dominant grayscale partition's range. Unfortunately, regarding to pill images, the imprint part is obviously not the dominant part, so noise will appear on the non-imprint area of the pill.

Another kind of method to enhance the contrast is just applying the contrast formula on V channel of pill images under the HSV color model commonly used:

$$I_{m,n}^{(out)} = \bar{I} + \left(I_{m,n}^{(in)} - \bar{I}\right) \times \omega. \tag{11}$$

Here, $I_{m,n}^{(in)}$ is the input pill pixel locating at (m,n), $I_{m,n}^{(out)}$ is its output pill pixel, $\bar{I}$ is the average value of whole input pill pixels, and ω is the rate of enhancement. According to this formula we can get a smooth enhance result without changing the entire image luminance.

There is another problem that if the rate is fixed to a constant value, excessive enhancement may happen to some pill images, which will influence the extraction. One solution is using fluctuating rate:

$$\omega = \frac{\alpha}{\ln \sigma^2} + \beta. \tag{12}$$

Here, $\sigma^2 = \frac{\sum_{m,n} \left(I_{m,n}^{(in)} - \bar{I} \right)^2}{N}$, α and β are adjustable parameters, N is the sum of input pill pixels. Rate fluctuates according to the variance of the image. In this system, we try to let rate fluctuate from 1.2 to 1.4 by selecting the proper parameters α and β.

2) Imprint Region Extraction

In [3], we proposed the Modified Stroke Width Transform (MSWT), which contained Stroke Width Transform [19] and Switching Function. Stroke width Transform extracts stroke which is contiguous, and owns a nearly constant width while Switching Function can locate the rough position of imprint before applying the Stroke Width Transform.

3) Imprint Feature Description

At the recognition by imprint part, two-step recognition is used. First step, we apply the Shape Context [20] and make a rough ranking on all candidate categories. The using of two-step recognition aims to reduce the time consumption. We make the recognition only by Shape Context with 50 sampling points, and result shown in Table I. According to Table I, we just reserve the 50 most similar categories to ensure sufficient accuracy. Next, we can apply the TSDS to do the further retrieval.

V. Experimental Results

We have built an image-capturing device in order to construct the pill dataset in order to evaluate the performance of our pill recognition system. In our database, there are 2500 different pill categories, each category providing at least one image for corresponding pill. Fig. 6 shows the samples of pill images in the dataset. Query images are generated by randomly rotating the images, as well as randomly changing the brightness from -30% to 0, in order to verify the robustness on rotation and illumination. For each category, we prepared 5 query images so that totally, there are 12500 query images in the dataset. About the resolution, all the pill images are normalized into the size of 200×200.

TABLE I. Recognition Result by Shape Context

Shape Context	**Rank**					
	1st	2nd	…	40th	50th	60th
Accuracy (%)	55.56	65.89	…	96.59	97.31	97.84

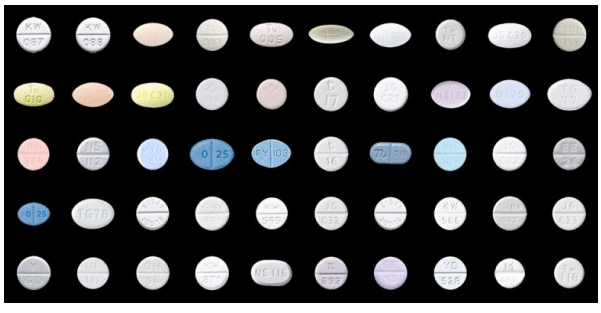

Fig. 6. Pill image samples in the dataset

A. Parameter Estimation

In order to evaluate the various selections of parameters, we just pick out the 500 pills images with similar imprints from the dataset to do the experiments. In (2), we show that how resampling strategy works. Here different selection of curvature parameter γ will affect the final recognition result. $\gamma \in (0,1]$ and $\gamma = 1$ means that uniform sampling is applied. The smaller γ is, the more complete elimination of small regions will be done. That will decrease the influence of noise.

Table II shows the results under various curvature parameter γ selections. ($N_1 = 100$, first step sampling points, $N_2 = 75$, second step sampling points) We can find that a small γ leads to a better result than uniform sampling ($\gamma = 1$). That means the elimination of noise works.

B. Comparison with Other Pill Recognition Algorithms

This time we use the whole 2500 categories and 12500 query images to do the comprehensive experiments. And compare the result among different algorithms. This dataset is challenging because pills in it have similar scale and shape; same series of pills even have the similar imprints.

Table III depicts the comparison of results among our proposed method and other methods. Parameter selection of our method is $N_1 = 100$, $K = 50$ (nearest neighbor points), $N_2 = 85$, and $\gamma = 0.3$. Comparing to MSWT+WSC (exponential) [2], proposed method works better, for its robust on imparity brought by luminous and exposure variation. Besides the difference of descriptor, there are two factors that cause the improvement. First, preprocessing of the pill image can help imprint extraction algorithm work better. Second, TSDS puts more attention on noise disturbance.

The processing speed is about 1.56 second per query pill under the condition of Core(TM)2 Duo CPU E7200, 2.53Hz.

C. Analysis on Failed Imprint Extraction

Fig. 7 shows some pill images which failed to extract fairly

TABLE II. Recognition Result Under Various γ

γ	**Rank** (Accuracy %)		
	1st	2nd	3rd
1	89.6	94.7	96.7
0.7	89.6	95.2	96.8
0.5	90.2	95.5	97.1
0.3	90.5	95.8	97.6
0.1	90.3	95.7	97.4

TABLE III. COMPARISON OF RETRIEVAL RESULTS AT TOP 5 RANKS

Method	Rank (Accuracy %)				
	1st	2nd	3rd	4th	5th
MSWT+WSC(exponential) [2]	79.62	86.47	89.27	90.80	91.91
proposed method	86.01	90.42	91.93	92.98	93.64

complete imprints by our system. There are four reasons: (a) Excessively overexposure makes imprint hard to identify even by human eyes, (b) Excessively blurry printed imprint, which is hard to mark out the boundary, (c) Rough pill surface disturbs the contour extraction of imprints, (d) Imprint fades making the strokes incoherence. Fig. 7 (b) and (c) need sharpness adjustment when doing preprocess. Problems shown in Fig. 7 (a) and Fig. 7 (d) are much tougher problems. More sensitive and robust imprint extraction system is needed to conquer these problems.

VI. CONCLUSION

In this paper, we proposed a pill recognition system, which can identify the pills with high accuracy. This system combines the effort of pill's contour shape, color and imprint features, and imprint plays the decisive role among these three. In the process of imprint extraction, we keep imprint strokes contours coherent by enhancing the contrast of pill image, which makes the modified stroke width transform work better. In imprint description part, we modified the Distance Sets by two-step sampling strategy to build the Two-step Sampling Distance Sets descriptor which can reduce the disturbance of noise and unwished fragments on imprint images caused by variation of luminance and exposure. Even in the condition of relatively low resolution, our system still achieved 93.64% accuracy among top 5 matches when testing with 12500 query images. This pill recognition system can be applied to the mass production in pharmaceutical factories as doing quality inspection. It can also be used to confiscate the illicit drugs.

REFERENCES

[1] "CFR – code of federal regulations title 21," U.S. Food and Drug Adminstration, http://www.fda.gov.

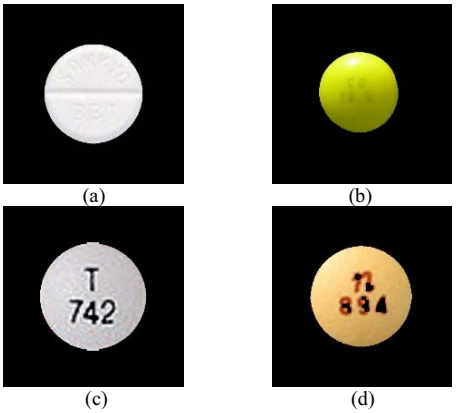

Fig.7. Pill images which failed to extract fairly complete imprints for these reasons. (a) excessively overexposure, (b) excessively blurry printed imprint, (c) rough pill surface, (d) imprint fading.

[2] Grigorescu, Cosmin, and N. Petkov. "Distance sets for shape filters and shape recognition," Image Processing, IEEE Transactions on 12.10 (2003): 1274-1286.

[3] Z. Chen and S. Kamata, "A new accurate pill recognition system using imprint information," Machine Vision (ICMV), 2013 6th International Conference on, 2013

[4] "WebMD pill identification tool," WebMD, http://www.webmd.com/pill-identification/.

[5] "RxList pill identification tool," RxList, http://www.rxlist.com/pill-identification-tool/article.htm.

[6] "Healthline pill identifer," Healthline.com, http://www.healthline.com/pill-identifier.

[7] "Pillbox," National Library of Medicine, National Institutes of Health, United States, http://pillbox.nlm.nih.gov/.

[8] Hartl, Andreas, "Computer-vision based pharmaceutical pill recognition on mobile phones," Proceedings of CESCG 2010: The 14th Central European Seminar on Computer Graphics, May 2010.

[9] Y. Lee, U. Park, and A. K. Jain, "Pill-ID: matching and retrieval of drug pill imprint images," Pattern Recognition (ICPR), 2010 20th International Conference on, IEEE, 2010.

[10] Hu, Ming-Kuei, "Visual pattern recognition by moment invariants," Information Theory, IRE Transactions on, 1962, 8(2): 179-187.

[11] Osada, Robert, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape distributions," ACM Transactions on Graphics (TOG) 21.4 (2002): 807-832.

[12] J. Caban, A. Rosebrock, and T. S. Yoo, "Automatic identification of prescription drugs using shape distribution models," Image Processing (ICIP), 2012 19th IEEE International Conference on. IEEE, 2012.

[13] M. Bukovec, Ž. Špiclin, F. Pernuš, B. Likar, "Automated visual inspection of imprinted pharmaceutical tablets," Measurement Science and Technology, 2007, 18(9): 2921.

[14] Mehlhorn, Kurt, LEDA: a platform for combinatorial and geometric computing, Cambridge University Press, 1999.

[15] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain "Content-based image retrieval at the end of the early years," Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2000, 22(12): 1349-1380.

[16] C. W. Niblack, R. Barber, W. Equitz, M. D. Flickner, E. H. Glasman, et al, "QBIC project: querying images by content, using color, texture, and shape," IS&T/SPIE's Symposium on Electronic Imaging: Science and Technology, International Society for Optics and Photonics, 1993: 173-187.

[17] T. K. Kim, Paik J K, B. S. Kang, "Contrast enhancement system using spatially adaptive histogram equalization with temporal filtering," Consumer Electronics, IEEE Transactions on, 1998, 44(1): 82-87.

[18] M. Abdullah-Al-Wadud, Md. Hasanul Kabir, M. Ali Akber Dewan, and Oksam Chae, "A dynamic histogram equalization for image contrast enhancement," Consumer Electronics, IEEE Transactions on 53.2 (2007): 593-600.

[19] Epshtein, Boris, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform." Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, IEEE, 2010.

[20] Belongie, Serge, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," Pattern Analysis and Machine Intelligence, IEEE Transactions on 24.4 (2002): 509-522.