13-1 MVA2013 IAPR International Conference on Machine Vision Applications, May 20-23, 2013, Kyoto, JAPAN

Pharmaceutical Blister Pack Recognition using Local Features

Koki Saitoh[1,2]
[1]TeamLab Inc, Japan
[2]The University of Tokyo, Japan
saito@team-lab.com

Abstract

In this paper, we propose a high-accuracy, high-speed method for recognizing pharmaceutical blister packs. In the proposed method, we first rank reference images of blister packs stored in a database for a given query image by using a simple voting process based on a nearest-neighbor search for local features. Next, we evaluate the results in ranking order by similarity of shape and color with reference to a template image for each reference image. In the evaluation process, we utilize an arrangement of local features to specify the area in the query image that is similar to the template image. By doing so, we reduce the computational time compared to simple template matching methods. To test our proposed method, we constructed a prototype and tested it at a pharmacy. As a result of this experiment, we demonstrate that the proposed method offers a practical method of recognizing blister packs accurately and quickly.

1 Introduction

Although pharmacists try their best to avoid dispensing errors, it is difficult to do in practice. Pharmacists have many kinds of blister packs at their disposal, of which many have similar designs, as shown in Fig.1. This means that they have to pay careful attention while dispensing medication to avoid selecting an incorrect blister pack.

A study has reported that the rate of dispensing errors in a pharmacy was substantially decreased by identifying medications using barcodes prior to dispensing[11]. However, most blister packs in Japan come without barcodes. Therefore, an automated system that can recognize a blister pack with a high degree of accuracy would reduce dispensing errors and lessen the burden on pharmacy staff.

To the best of our knowledge, there are no studies that tackle the task of recognizing a pharmaceutical blister pack from a captured image. Although there are systems that implement image processing techniques for preventing dispensing errors[1][3], such systems inspect the query image with reference to prescription data to confirm that no mistake has been made. In this study, the task is to search for the correct blister pack in an image database, but not by comparing a query image against reference images by using prescription data.

In order to recognize a blister pack with a high degree of accuracy and thereby prevent dispensing errors, we must consider the following points. First, the number of different blister packs handled by a pharmacy in Japan can range from several hundred to several thousand, in which many similar designs exist (see Fig.1). Second, the query image may be a blister pack that has been cut or has a rubber band around it, thus obscuring some details (see Fig.2). Third, to prevent dispensing errors, high reliability is required for the recognition result.

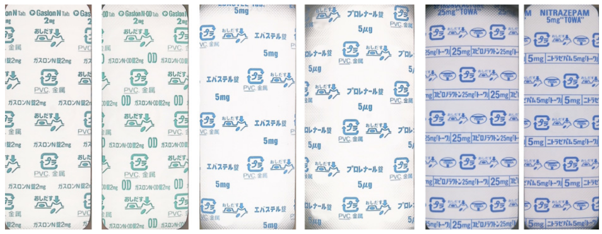

Figure 1. Examples of similar blister pack designs such as those found at pharmacies.

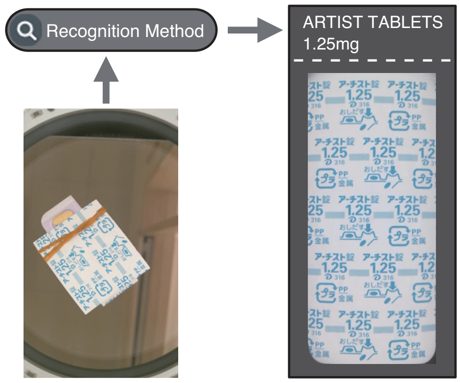

Figure 2. We propose a high-accuracy, high-speed method for recognizing blister packs. The image to the left is a query image, which was captured by our prototype. The image to the right is a reference image, which is in the image database.

In this paper, we propose a high-accuracy, high-speed method of recognizing a blister pack from a captured image as shown in Fig.2. To achieve this, we utilize a recognition technique based on local features, a method that is popular in the computer vision community. The proposed method is a two-step process. First, to obtain a ranked list for a given query image, we use a voting process based on a nearest-neighbor search for local features, which we call the *ranking process*. Second, to improve recognition accuracy of the results, we inspect the results in ranking order according to shape and color similarity by using a template image, which is the part of each reference images that was specified manually beforehand. This process is called the *inspection process*.

In the *inspection process*, we utilize an arrangement of local features to specify the area in the query image that is similar to the specified template image. By

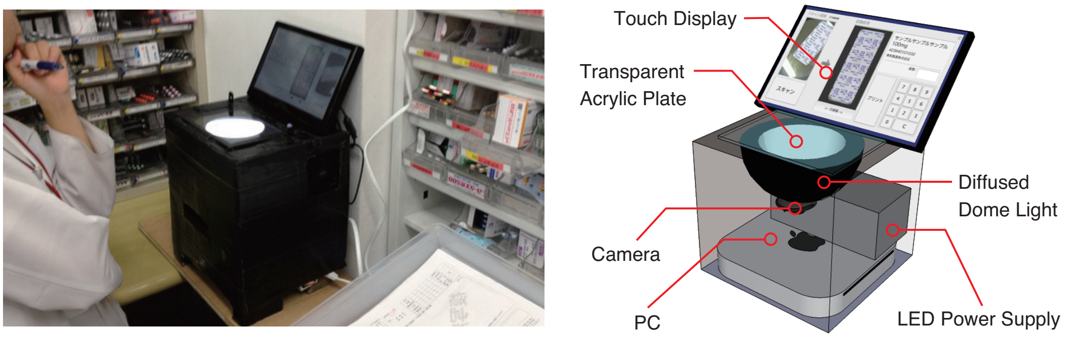

Figure 3. The prototype for our blister-pack recognition system. The image to the left is a photgraph of the prototype, and the image to the right is a schematic of its inner-structure.

doing so, we reduce the computational time compared to simple template matching methods. It should be noted that we make use of local features in both the *ranking process* and the *inspection process* in order to recognize a blister pack accurately and quickly.

2 Prototype Overview

Fig.3 shows the prototype of our blister-pack recognition system. We use a diffused dome light to suppress diffused reflection from the blister packs, most of which are made of laminated aluminum. Note that the distance from the camera to the transparent acrylic plate is fixed. Thus, the size of captured images for all blister packs is the same.

When a blister pack is placed on the plate, the recognition result is displayed on the touch display. There are no restrictions on the direction of the blister pack, so all one has to do is simply place a blister pack on the plate. This makes the system easier to use than a barcode reader. To automatically detect whether a blister pack has been placed on the plate, a background subtraction technique is used for every frame.

From the viewpoint of preventing dispensing errors, it is essential for the output of the system to maintain high reliability. Our system guarantees the required reliability by outputting the recognition result only if a reference image passes the *inspection process* for a given query image. If the reliability is not enough, the system displays a screen that alerts the pharmacist to visually verify the image.

The aim of our system is to prevent dispensing errors and reduce the burden on pharmacy staff, so we make two assumptions for the system: 1) The recognition algorithm must maintain a high pass ratio for the *inspection process*; and, 2) If a reference image passes the *inspection process*, the recognition accuracy must be very close to 100%.

3 Blister-Pack Recognition

Many object-recognition technologies based on local features such as SIFT[9] and SURF[5] have been proposed. The advantage of using local features in object recognition is robust for deformation, occlusion, and rotational change of the target. This advantage is appropriate for our recognition task because a blister pack that is captured as a query image could be

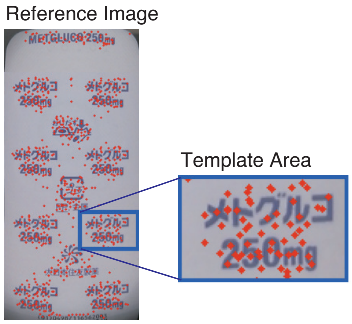

Figure 4. An example of keypoints for local features. The image is a reference image in the database and the red dots represent keypoints where descriptors are computed. The blue rectangular area represents the template area used on the *inspection process*.

cut or partially occluded by a rubber band, and its orientation is not known.

Although a nearest-neighbor search for local features is a simple technique, it can achieve high recognition rates for object-recognition tasks[6]. In addition, this method does not require a learning stage. However, it is not clear whether such a method would be effective for blister packs as they have a man-made design, for which there exist many similar designs.

Through an experiment we conducted, we found that a simple voting process based on a nearest-neighbor search for local features, which we call the *ranking process*, can recognize a blister pack accurately in most cases. To improve the recognition accuracy and guarantee the result, the *inspection process* is performed where the results are inspected in ranking order by shape and color similarity using a template image for each reference images. Note that the template image is the part of the reference image that was specified manually beforehand, as shown in Fig.4. Below, we describe the details of each process.

3.1 Ranking Process

Local features are retrieved by a two-stage process: 1) keypoint detection, which finds the positions for effective recognition; and, 2) keypoint description, which describe the features at the detected keypoints. We use

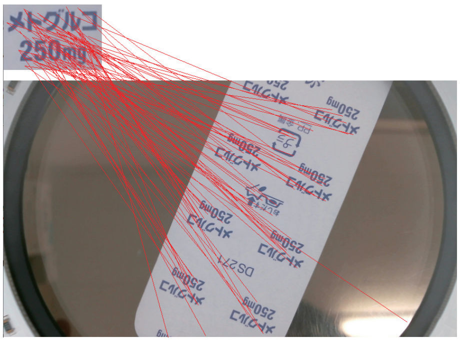

Figure 5. An example of corresponding points by local features. Red lines are drawn in the corresponding points of local features between the query and the template images.

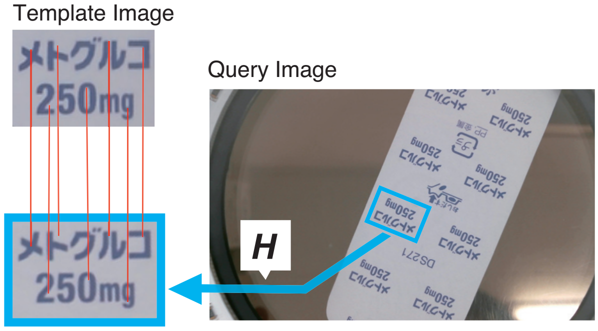

Figure 6. The query image is projected by the matrix $\boldsymbol{H}$ and trimmed to the lower left image. The red lines shows the corresponding points between the query and template images.

Fast-Hessian[5] for keypoint detection and FREAK[4] for keypoint descriptions. FREAK is described by a binary vector, and it has been reported that FREAK outperforms more-recent state-of-the-art keypoint descriptors while remaining faster[4].

The local features of the reference images are extracted beforehand and added to a database. When given a query image, first the local features for the query image are extracted, and then, by a voting process that uses an approximate nearest-neighbor search on local feature descriptors, a ranked list of the reference images in the database is obtained. A hierarchical k-means tree[10] is used for the approximate nearest-neighbor search. And to reduce the computational time, we use 300 local features of the query image, which is randomly selected, for the voting process.

The above process can be performed quickly, and in many cases, the correct blister pack ranks highest. However, as there are many similar images in the database; as a result, sometimes similar but incorrect images get the most votes. Therefore, an inspection is performed sequentially from the top-ranked reference image.

3.2 Inspection Process

Ichimura[7] and Iwamura et al.[8] each proposed a method for specifying the template area for a query image using local features when more than one similar template image is found in the query image. While these methods are useful for detecting the area, our aim is to ensure the high reliability of the results. Therefore, after specifying similar template image areas that might exist in the query image, we conduct a precise inspection by calculating the similarity in the shape and color of the specified area.

The algorithm for inspection is described using the following expression. Local features extracted from the query image are described as $\boldsymbol{f}_i^q = \{\boldsymbol{p}_i^q, \sigma_i^q, \boldsymbol{d}_i^q\}$, where $\boldsymbol{p}_i^q$ is the position where a keypoint is detected, σ_i^q is the scale of the local feature, and $\boldsymbol{d}_i^q$ is the descriptor. Local features extracted from the reference image within the template area are described as $\boldsymbol{f}_j^t = \{\boldsymbol{p}_j^t, \sigma_j^t, \boldsymbol{d}_j^t\}$.

To obtain an appropriate transform matrix, we first find the corresponding points of the local features between the query and the template images that satisfy the following constraints:

$$||\boldsymbol{d}_i^q - \boldsymbol{d}_j^t|| \quad < \quad T_d, \tag{1}$$

$$||\sigma_i^q - \sigma_j^t|| \quad \leq \quad T_s, \tag{2}$$

where $||\boldsymbol{d}_i^q - \boldsymbol{d}_j^t||$ is the Hamming distance, and T_d is the threshold to be set in advance. In addition, the difference between the scale of the keypoints must be the same, because the reference and query images were captured at the same size. Thus we set the constraint (2). From this, we obtain a set of corresponding points $C = \{\boldsymbol{p}_k^q, \boldsymbol{p}_k^t\}$. An example of corresponding points is shown in Fig.5 (we set as $T_d = 160$, $T_s = 1$).

In many cases, the corresponding set contains outliers. In such a case, a transformation matrix between the two images can be estimated robustly by RANSAC. However, in the query image there could be more than one area that is similar to the template image. Therefore, we uses the RANSAC with local restriction.

From a set of corresponding points C, we randomly select one set of corresponding points. We describe the corresponding points as $\{\boldsymbol{p}_{l0}^q, \boldsymbol{p}_{l0}^t\}$. Next, we extract a set of corresponding points from C where each distance between the points in the query image and $\boldsymbol{p}_{l0}^q$ is less than r, and describe it as $C_r = \{\boldsymbol{p}_l^q, \boldsymbol{p}_l^t\}$. We set r as the length of the diagonal of the template image. Next, we randomly select two corresponding points from C_r; after which we calculate the transformation matrix $\boldsymbol{H}$ with the affine transformation restriction from the correspondence among the three corresponding points. And as an appropriate transformation is a rigid transformation, we confirm whether $\boldsymbol{H}$ satisfies the following condition, where $\boldsymbol{H} = [\boldsymbol{R}|\boldsymbol{t}]_{2\times 3}$ and we set T_{det} as 0.05:

$$1.0 - T_{det} < \det(\boldsymbol{R}) < 1.0 + T_{det}. \tag{3}$$

If matrix $\boldsymbol{H}$ satisfies the condition (3), the set of points $\boldsymbol{p}_l^q$ within set C_r are projected by $\boldsymbol{H}$ onto the template image. Thus the total number of points that satisfy the condition where the distance between the projected and corresponding points is less than ϵ are obtained as inliers (we set ϵ as 5.0). The above process is repeated multiple times, and a set of corresponding points is selected where the total number of inliers is the highest. Finally, the transform matrix $\boldsymbol{H}$ is calculated from the inliers. Then the query image is projected by $\boldsymbol{H}$ and trim the projected image by the size

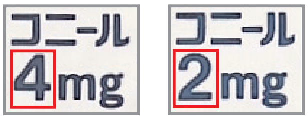

Figure 7. If there are similar refrence images in the database, we specified an area for detecting the similar images.

of the template image (see Fig.6).

Next, we calculate the shape and color similarity between the transformed query and template images. Shape similarity is calculated by normalized cross correlation (NCC), and described as S_{shape}. For the color similarity, which is described as S_{color}, we generate a color histogram of both images in the HSV color space and calculate the Bhattacharyya distance between both histograms. When S_{shape} and S_{color} satisfy the following conditions, we consider the reference image to have passed the *inspection process*:

$$S_{shape} > T_{shape} \quad \text{and} \quad S_{color} > T_{color}. \tag{4}$$

However, there could be similar reference images, like those shown in Fig.7. To deal with this problem, we specify the area for detecting similar images beforehand. For a reference image that satisfies the condition (4), if it is known that there are similar images for the reference image, we calculate the NCC in the specified area. We described the NCC as S_{local} and confirm whether it satisfies the following conditions (we set as T_{shape}=0.80, T_{color}=0.70, T_{local}=0.90.):

$$S_{local} \quad > \quad T_{local}. \tag{5}$$

If a reference image passes the inspection described above, the image is output as a result. If it does not, the next reference image in the ranked order is inspected. If there is nothing that passes the *inspection process* for the top T_n, which we set as $T_n = 5$, by the sum of S_{shape}, S_{color} and the rate of the vote obtained by the *ranking process* for each reference image, we sort the images in descending order and output the top-ranked reference image as a result. In such a case, as we cannot guarantee the reliability, an alert screen is displayed on the touch display.

4 Experiment

The recognition algorithm was implemented in C++ using OpenCV 2.4.2[2] on a 2.7GHz machine. We constructed a prototype and tested it at a pharmacy. The number of blister packs handled at the pharmacy was 497. We used 1169 images which were captured during the test period for recognition accuracy.

The recognition results are shown in Table 1. As shown, the recognition algorithm using the *ranking process* only can correctly recognize 95.5% of test images. This indicates that local features are effective for recognizing blister packs. But it also indicates that this process alone is not enough to guarantee the reliability.

Our proposed method improve the recognition accuracy as shown in Table 1. Note that if no reference image passes the *inspection process* within the top five reference images, we re-rank the top five by the inspection score. Even more importantly, in our proposed method, 1038 test images out of 1169 (88.8%) passed the *inspection process*, and for the test images that passed the *inspection process*, the recognition accuracy is 100%. Thus we have proved that our proposed method contributes to preventing dispensing errors.

Table 1. Recognition result by *ranking process* only and by our proposed method.

Correct blister pack within	Top1	Top5
ranking process only	95.5%	99.6%
Our proposed method	99.1%	99.6%

The proposed system is also very fast. The average processing time is 350ms, while the *ranking process* takes 170ms. This means that the recognition result appears on the display as soon as a blister pack is placed on the prototype.

5 Conclusion

In this paper we proposed a method for recognizing pharmaceutical blister packs with a high degree of accuracy and at a high speed. Experiments showed that although a simple voting process using local features can correctly recognize most query images, but the process alone is not enough to guarantee the reliability. Therefore, we inspected the reference images in ranking order. To streamline the *inspection process*, we utilized an arrangement of local features, which enabled us to identify the template image area in the query image. Note that we utilized local features in the both processes. The results of the experiment verified the effectiveness of the proposed method.

Acknowledgement

The authors would like to thank Mr. Kawaguchi, Mr. Kataoka and Mr. Takeshita of Medical-Polygon Ltd. for their support.

References

[1] Okura Information System ltd. http://www.ois92.co.jp.
[2] willowgarage. http://opencv.willowgarage.com/.
[3] Windy ltd. http://www.e-windy.co.jp.
[4] A. Alahi, R. Ortiz, and P. Vandergheynst: "FREAK: Fast Retina Keypoint," In *CVPR*, 2012.
[5] H. Bay, T. Tuytelaars, L. V. Gool: "SURF: Speeded Up Robust Features," In *ECCV*, 2006.
[6] O. Boiman, E. Shechtman, and M. Irani: "In defense of nearest-neighbor based image classification," In *CVPR*, 2008.
[7] N. Ichimura: "Recognizing multiple billboard advertisements in videos," In *PSIVT*, 2006.
[8] M. Iwamura, T. Kobayashi, and K. Kise: "Recognition of multiple characters in a scene image using arrangement of local features," In *ICDAR*, 2011.
[9] D. Lowe: "Distinctive image features from scale-invariant keypoints," In *IJCV*, 2004.
[10] D. Nistér and H. Stewénius: "Scalable recognition with a vocabulary tree," In *CVPR*, 2006.
[11] EG Poon, et al: "Medication dispensing errors and potential adverse drug events before and after implementing bar code technology in the pharmacy," In *Annals of Internal Medicine*, 2006.