Optical character verification of print on pharmaceutical capsules

Miha Možina[1], Dejan Tomaževič[1,2], Franjo Pernuš[1,2] and Boštjan Likar[1,2]

[1]Sensum, Computer Vision Systems
Tehnološki park 21, 1000 Ljubljana, Slovenia

[2]University of Ljubljana, Faculty of Electrical Engineering,
Laboratory of Imaging Technologies, Tržaška 25, 1000 Ljubljana, Slovenia

E-mail: miha.mozina@sensum.eu

Abstract

In this paper, the problem of print quality inspection on pharmaceutical capsules is addressed. Good print quality, a property that makes the print legible, is of utmost importance for identification and preventing mix-ups among various types of capsules. Spatial distortion obtained at image formation process, where a 3-D surface gets projected onto a 2-D image plane, causes print appearance variations. Elimination or compensation of print appearance variations is necessary for improving defect detection performance of inspection systems. For that reason, a novel appearance-based method for print quality inspection on capsules is proposed. The performance of the proposed method was evaluated on a real capsule image database of printed capsules, where a "gold standard" was established by manually classifying the capsules into non-defective and defective class. The obtained results indicate that the proposed method is more effective than the standard defect detection method.

1 Introduction

In order to protect the customer from mix-ups among various types of pharmaceutical products, a regulation code 21CFR206 [1], issued by the Food and Drug Administration, enforces the pharmaceutical companies to produce products with unambiguous identification of products' active substance and dosage by size, shape, texture, imprint, etc. Besides the identification [2], visual appearance also plays an important role in marketing, because imperfect appearance of a single product in a package can raise serious doubts about integrity and quality of the product. Since manual visual inspection of large batches is subjective, unreliable, slow, tedious and costly, automated visual inspection systems are more and more commonly used.

Various automated visual inspection systems for tablets and capsules are being produced by companies, such as Ackley, Eisai, Ikegami, Mutual, Proditec, Seidenader, Sensum and Viswill. These systems have to be robust and general enough to be able to inspect the vast amount of all possible products and defects. However, the robustness comes with the price of sensitivity. The robustness-sensitivity trade-off means that these inspection systems in general perform well in detecting large defects such as cracks, broken products, contrast stains, dots and size variations on products without imprints or texture, but have suboptimal performance for some specific products, such as tablets and capsules with imprints. In case of products with imprints, defects in the vicinity of the imprint can have similar contrast to imprint and therefore, before the classification, the system should distinguish defects from imprint. In addition, imprint quality has to be inspected. Good imprint quality, a print property that makes the imprint legible, is crucial for the identification. The imprint quality is reduced by imprint degradations such as partly missing imprint, blurred imprint, low contrast imprint, color and size variations of imprint, etc.

The field most related to the imprint inspection of pharmaceutical products is optical character verification (OCV), where OCV is considered as a machine vision software tool used to inspect a string and confirm its legibility. In addition to checking that the content of the presented text string is correct, it will also inspect for contrast and sharpness, and will flag or reject poor quality samples. Most OCV applications are found for integrated circuit (IC) and text based codes inspection on various products made by automated printing machines [3][4][5]. In the most recent work in the field of automated visual

Figure 1. Images of the same capsule with print at different positions

inspection of pharmaceutical products, Bukovec *et al* [6], proposed a method for automated visual inspection of imprinted tablets. However, print inspection on capsules has proven to be an additionally challenging task due to two reasons. Firstly, the print is not on flat surface but is on tube-shaped capsules. At image formation process, where a 3-D surface is projected onto a 2-D image plane, the tube-shaped capsules cause spatial distortion of print on the images. Secondly, capsules can rotate around their axis (Figure 1), therefore, only part of the print may be visible on the image. If the quantity of the visible print is very low, the print might seem nothing more than a dark dot defect. Capsule inspection was considered by Karloff *et al* [7], however, they considered capsules without the print and inspected trivial defects, such as capsule length, missing cap and large dents. On the other hand, Špiclin *et al* [8] proposed print localization on capsules for the purpose of distinguishing defects from print. They also attempted to eliminate the spatial distortion of the print caused capsul's shape by transforming the capsule image into cylindrical coordinates. While the transformation eliminated the shape distortion of the print, it introduced an additional contrast distortion of the print in the images (Figure 2). The contrast distortion is similar to the shape distortion in that it is the strongest when the print is on the border of a capsule.

The aim of this paper is to improve sensitivity of the automated visual inspection systems for capsules by improving the capsule print inspection. Although the state of the art systems have reliable product manipulation mechanisms, where capsules can be held in reproducible position, the rotation of capsules around their axis is still random. Since the print appearance depends on capsule position at image acquisition, we propose exploiting the position information of the capsule for generation of an accurate print appearance model.
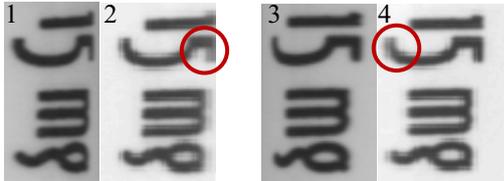
Figure 2. Print images with spatial distortion (1,3) and corresponding print images after the transformation into cylindrical coordinates (2,4) with marked contrast distortion.

2 Method

The proposed method is based on comparison of an inspected print to a print appearance model of non-defective print. Before the comparison, segmentation, i.e. partitioning the image into capsule region and the background, is done using the border tracking algorithm [9]. Then, shape normalization [8] of the capsule is performed in order to eliminate the spatial distortion caused by capsules' shape. Shape normalization consists of capsule rotation into vertical position and transformation of the image cartesian coordinates into cylindrical coordinates $(x, y) \rightarrow (\varphi, y)$ with a constant step of φ. The last preprocessing step is print localization [8], which determines the region of interest (ROI), i.e. the print region,
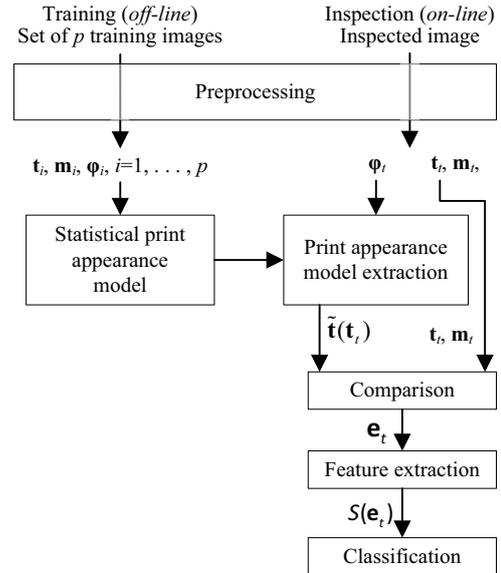
Figure 3. Framework of the proposed method for visual quality inspection of print appearance

and sets the prints into spatial correspondence, enabling the generation of the print appearance model and comparison of the inspected print to the print appearance model.

The method consists of two stages (Figure 3). In the first, training stage, the print appearance model is generated from the training images, i.e. images of print without defects. In the second, on-line (verification) stage, a given input image is compared to the corresponding print appearance model and then accordingly classified.

Firstly, notations are introduced. Let $\mathbf{t} = [t_1, \ldots, t_n]^T$ be vector of intensities of an individual image with n pixels of a full capsule after preprocessing and let $\boldsymbol{\varphi} = [\varphi_1, \ldots, \varphi_n]^T$ be the vector of the first cylindrical coordinates φ for the corresponding individual image $\mathbf{t}$. The capsule appearance $\mathbf{t}$ can be modeled by an appearance model $\tilde{\mathbf{t}}$:

$$\mathbf{t} \approx \tilde{\mathbf{t}} = \overline{\mathbf{t}} + \mathbf{c}(\boldsymbol{\varphi}) \quad (1)$$

consisting of an average intensity model $\overline{\mathbf{t}} = [\overline{t}_1, \ldots, \overline{t}_n]^T$ and intensity correction vector $\mathbf{c}(\boldsymbol{\varphi}) = [c_1(\varphi_1), \ldots, c_n(\varphi_n)]^T$, which shifts the average intensity model $\overline{\mathbf{t}}$ according to $\boldsymbol{\varphi}$. The intensity correction vector $\mathbf{c}(\boldsymbol{\varphi})$ is obtained as a linear function:

$$c_i(\varphi_i) = r_i \cdot \left(q_i(\varphi_i) - \overline{q}_i \right), \quad i = 1,...,n \quad (2)$$

with $\mathbf{r} = [r_1, \ldots, r_n]$ denoting regression coefficients, while $\mathbf{q}(\boldsymbol{\varphi})$ and $\overline{\mathbf{q}}$ represent pose vector and average pose model, respectively. The pose vector $\mathbf{q}(\boldsymbol{\varphi})$ describes the correlation of contrast distortion, which is obtained at the shape normalization, and position $\boldsymbol{\varphi}$:

$$\mathbf{q} = \mathbf{q}(\boldsymbol{\varphi}) = \frac{1}{\cos(\boldsymbol{\varphi})} \quad (3)$$

Since the focus of this paper is print only, the appearance model of capsule $\tilde{\mathbf{t}}$ will only be calculated for the print region and therefore $\tilde{\mathbf{t}}$ will be notated as print appearance model. In order to limit the model to print region only, we are introducing the masks to define the region of interest (ROI). Let $\mathbf{m} = [m_1, \ldots, m_n]^T$ be mask defining ROI for $\mathbf{t}$. Vector $\mathbf{m}$, where elements with value of one ($m = 1$) belong to ROI and those with value of zero ($m = 0$) do not, was obtained in the preprocess step of print localization. The print appearance model is obtained statistically

in the training phase from a set of p training images $T = [\mathbf{t}_1, \ldots, \mathbf{t}_p]$, i.e. images without defects, the corresponding set of masks $M = [\mathbf{m}_1, \ldots, \mathbf{m}_p]$ and the set of pose vectors $Q = [\mathbf{q}_1, \ldots, \mathbf{q}_p]$. The average intensity model $\bar{\mathbf{t}}$ and the average pose model $\bar{\mathbf{q}}$ were calculated using the weighted average:

$$\bar{t}_i = \begin{cases} \dfrac{\sum_{j=1}^{p} m_{ij} \cdot t_{ij}}{\sum_{j=1}^{p} m_{ij}}, & \sum_{j=1}^{p} m_{ij} > 0 \\[4mm] 0 & \sum_{j=1}^{p} m_{ij} = 0 \end{cases}, i = 1,...,n \quad (4)$$

where for the average pose model $\bar{\mathbf{q}}$ set Q instead of set T was used. The regression coefficients $\mathbf{r}$ were calculated as weighted regression of intensity t and the corresponding pose value q:

$$r_i = \begin{cases} \dfrac{\sum_{j=1}^{p} m_{ij}\left(t_{ij} - \bar{t}_i\right)\left(q_{ij} - \bar{q}_i\right)}{\sum_{j=1}^{p} m_{ij}\left(q_{ij} - \bar{q}_i\right)^2}, & \sum_{j=1}^{p} m_{ij} > 0 \\[4mm] 0 & \sum_{j=1}^{p} m_{ij} = 0 \end{cases}, i = 1,...,n \quad (5)$$

In the verification stage, let $\mathbf{t}_t$ be an inspected capsule after preprocessing with mask $\mathbf{m}_t$ and corresponding coordinate vector $\boldsymbol{\varphi}_t$. By equation (1), the print appearance model $\tilde{\mathbf{t}}(\mathbf{t}_t)$ is calculated as:

$$\tilde{\mathbf{t}}(\mathbf{t}_t) = \bar{\mathbf{t}} + \mathbf{c}(\boldsymbol{\varphi}_t) \quad (6)$$

The defect detection feature was chosen as the maximum difference between the print appearance model $\tilde{\mathbf{t}}(\mathbf{t}_t)$ and inspected capsule $\mathbf{t}_t$ and was calculated in equations (7) and (8). Firstly, image of error $\mathbf{e}_t$ is defined as the absolute difference of $\mathbf{t}_t$ and $\tilde{\mathbf{t}}(\mathbf{t}_t)$:

$$\mathbf{e}_t = \left| \mathbf{t}_t - \tilde{\mathbf{t}}(\mathbf{t}_t) \right| \quad (7)$$

The defect classification feature $S(\mathbf{t}_t)$ is then calculated by finding the maximum value $\mathbf{e}_t$ over the ROI of the inspected capsule:

$$S(\mathbf{e}_t) = \max_{\mathbf{m}_t} \left| \mathbf{e}_t \right| \quad (8)$$

Classification of inspected capsule $\mathbf{t}_t$ to defective or non-defective is then done by thresholding the feature $S(\mathbf{e}_t)$.

3 Experiment and results

The proposed method was qualitatively evaluated by comparison to "Average" method, which uses the average intensity model for print appearance model, i.e. $\tilde{\mathbf{t}} = \bar{\mathbf{t}}$. The methods were compared on a set of real images of capsules with print. Details about "gold standard" database, model analysis and results are given in the following subsections.

3.1 Image database with "gold standard"

The experimental database of images of capsules, which consisted of 312 images with print, was acquired with a Sensum SPINE (Figure 4) automatic tablet and capsule inspection machine. The database included 216 non-defective images and 96 defective images of capsules with print. Some of the examples of defective print images are given in Figure 5. The training dataset consisted

Figure 4. Sensum SPINE visual quality inspection machine (left), which automatically inspects and sorts pharmaceutical capsules or tablets (right); the speed for capsules is up to 180,000 capsules per hour.

of 104 non-defective images, while the rest were used to evaluate the defect detection performance of the proposed method. A "gold standard" of the database was established by a careful and manual classification of the capsules as non-defective and defective.

3.2 Model analysis and results

The accuracy of the print appearance model and the average intensity model for a randomly chosen pixel on print is given in Figure 6. The obtained results for the given datasets are given in Figure 7, where separate box-whisker diagrams for the distribution of defect detection feature $S(\mathbf{e}_t)$ of non-defective and defective print are given for both methods. Furthermore, the normalized distance D between the distribution of non-defective $\mathcal{N}(\mu_n, \sigma_n)$ and distribution of defective $\mathcal{N}(\mu_d, \sigma_d)$ was calculated for each method:

$$D = \frac{\left| \mu_n - \mu_d \right|}{\frac{1}{2}\left(\sigma_n + \sigma_d\right)} \quad (9)$$

The normalized distances D obtained for the proposed method and the "Average" method are given in Table 1.

Table 1. Distance between the distribution of non-defective and defective print defect detection feature $S(\mathbf{e}_t)$ for both methods

	$\mathcal{N}(\mu_n,\sigma_n)$	$\mathcal{N}(\mu_d, \sigma_d)$	D
Proposed method	(3.92, 1.27)	(17.01, 5.77)	3.72
"Average"	(3.89, 1.44)	(13.42, 4.29)	3.33
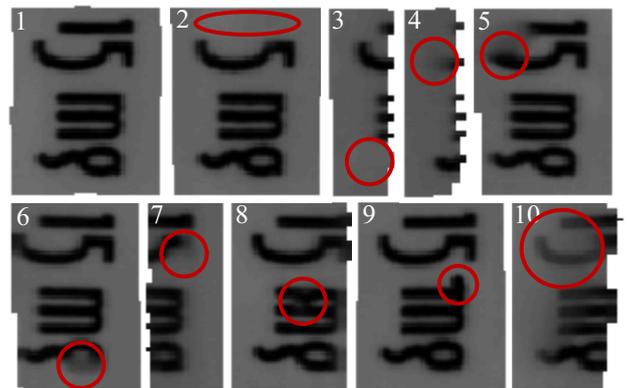
Figure 5. An example of non-defective print (1), partly missing print (2,3), defective and/or blurred print (4-7), spot/dot on print (8,9), low contrast print (10)
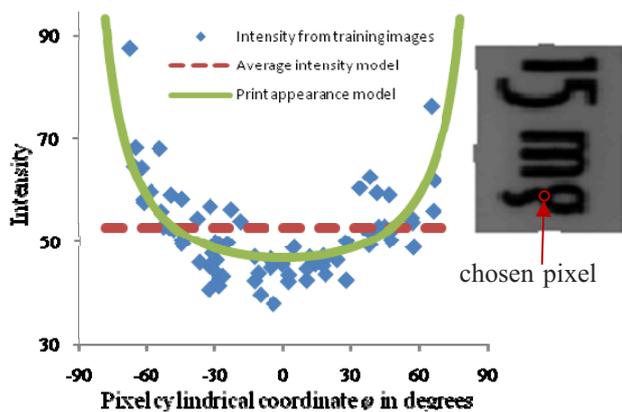
Figure 6. Modelling accuracy of print appearance model and average intensity model for a randomly chosen pixel on print
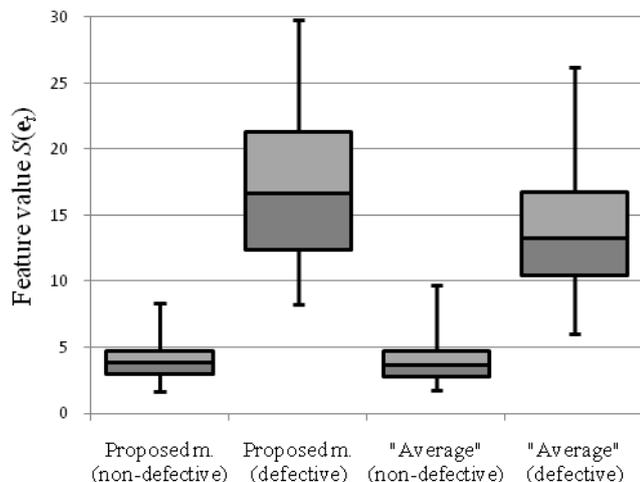
Figure 7. Results for the given dataset with separate whisker-box diagrams for non-defective and defective diagrams

4 Discussion and conclusion

Image formation process of automated visual inspection systems induces object-dependent appearance variations. In the case of capsules with print, print appearance variation is a consequence of spatial distortion caused by a 3-D surface projection onto a 2-D image plane. Elimination or compensation of print appearance variations is necessary for improving defect detection performance of inspections systems.

An alternative modelling strategy of print appearance variations on cylindrical surfaces is proposed in a novel image analysis method. The method uses automatically obtained statistical *a priori* knowledge to model the print appearance variations and is based on comparison of an inspected print to a corresponding print appearance model of non-defective print.

The proposed method was evaluated by analyzing the print appearance model on a randomly chosen pixel on the print. The analysis showed the proposed method's print appearance model accurately fits the training data (Figure 6). Furthermore, the method was evaluated on a database of real images acquired on a real industrial machine vision system Sensum SPINE by comparison to the "Average"

method. The defect detection performance was assessed with separate box-whisker diagrams for the distribution of the non-defective and defective print for both methods (Figure 7), where the proposed method showed significantly better detection of defects. The distance between the non-defective and defective distributions is larger for the proposed method, which was confirmed by calculation of distance D. The distance D was 3.72 and 3.33 for the proposed method and the "Average" method, respectively. The reason for the larger distance of the proposed method is because of better accuracy of the print appearance model, which resulted in a smaller standard deviation of non-defective distribution and a higher average value of the defective distribution. In terms of speed and computational load, the proposed method only requires one extra multiplication per pixel in ROI compared to the "Average" method, which is negligible even for real-time applications.

In conclusion, a novel method for defect detection and print quality inspection of print on capsules has been proposed. The method's modelling approach accurately models the print appearance variations, which improves the defect detection performance. The method has been evaluated on real images and showed to be effective for defect detection performance and print quality inspection.

Acknowledgement

References

[1] FDA, "FDA 21CFR206, Imprinting of solid oral dosage form drug products for human use," Revised. 2009.

[2] A. Berman, "Reducing medication errors through naming, labeling, and packaging," *Journal of medical systems*, vol. 28, no. 1, pp. 9–29, 2004.

[3] Y. Lu, "Machine Printed Character Segmentation - An Overview," *Pattern Recognition*, vol. 28, no. 1, pp. 67-80, 1995.

[4] B. Jiang, et al.: "Machine vision-based gray relational theory applied to IC marking inspection," *IEEE Transactions on Semiconductor Manufacturing*, vol. 15, no. 4, pp. 531-539, 2002.

[5] R. Nagarajan, et al.: "A real time marking inspection scheme for semiconductor industries," *The International Journal of Advanced Manufacturing Technology*, vol. 34, no. 9, pp. 926-932, 2006.

[6] M. Bukovec, et al.: "Automated visual inspection of imprinted pharmaceutical tablets," *Measurement Science and Technology*, vol. 18, no. 9, pp. 2921-2930, 2007.

[7] C. Karloff, et al.: "A Flexible Design for a Cost Effective, High Troughput Inspection System for Pharmaceutical Capsules," *Proc. Int. Conf. on Industrial Technology ICIT*, Chengdu, pp. 1-4, 2008.

[8] Ž. Špiclin, et al.: "Real-time print localization on pharmaceutical capsules for automatic visual inspection," *International Conference in Industrial Technology (ICIT), 2010 IEEE*, pp. 279–284, 2010.

[9] M. Možina, et al.: "Real-time image segmentation for visual inspection of pharmaceutical tablets," *Machine Vision and Applications*, 2009