

# Automated Machine Vision System for Liquid Particle Inspection of Pharmaceutical Injection

Hui Zhang<sup>✉</sup>, Xuanlun Li, Hang Zhong, Yimin Yang, *Member, IEEE*, Q. M. Jonathan Wu, *Senior Member, IEEE*, Ji Ge, and Yaonan Wang

**Abstract**—The particle matter inspection for pharmaceutical injection is inevitable in the field of pharmaceutical manufacturing, as it has the direct impact on the quality of the drugs. It is a challenge to inspect the contaminated injection online using an imaging system. This paper introduces a novel and effective inspection machine consisting of three modules, a mechanical system with 120 carousel grips, an image acquisition system with multihigh resolution cameras and a multilight sources station, and a distributed industrial electrical computer control system. Particle visual inspection machine first acquires image sequence using the high-speed image acquisition system. The image capture process at each camera module is alternately synchronized with different LED illumination techniques (light transmission method and light reflection method), enabling independent capture of particle images from the same container. Then, a set of novel algorithms for image registration and fast segmentation are proposed to minimize false rejections even in sensitive conditions, which enable the identification of all the tiny potential defects.

Manuscript received August 4, 2017; revised December 11, 2017; accepted December 17, 2017. Date of publication February 19, 2018; date of current version May 10, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61401046 and Grant 61473049, in part by the National Technology Support Project under Grant 2015BAF11B01, in part by the Scientific Research Fund of Hunan Provincial Education Department under Grant 17C0046, in part by the Hunan Province Key Laboratory of Videometric and Vision Navigation under Grant TXCL-KF2013-001, and in part by the Changsha University of Science and Technology Innovation Project under Grant CX2016SS07. The Associate Editor coordinating the review process was Dr. Emanuele Zappa. (*Corresponding author: Hui Zhang.*)

H. Zhang is with the College of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha 410114, China, and also with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada (e-mail: zhanghuihb@126.com; zhanghui@uwindsor.ca).

X. Li is with the College of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha 410114, China (e-mail: csd1174412@163.com).

H. Zhong and Y. Wang are with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China (e-mail: zhonghang@hnu.edu.cn; yaonan@hnu.cn).

Y. Yang is with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada (e-mail: yangyi\_min@126.com).

Q. M. J. Wu is with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada, and also with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China, and also with the Key Laboratory of Shanghai Education Commission for Intelligent Interaction and Cognitive Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: jwu@uwindsor.ca).

J. Ge is with the Advanced Micro and Nanosystems Laboratory, University of Toronto, Toronto, ON M5S 3G8, Canada (e-mail: jige@mie.utoronto.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIM.2018.2800258

Finally, a particle tracking and classification algorithm based on an adaptive local weighted-collaborative sparse model is also presented. The experiments demonstrate that the proposed inspection system can effectively detect the particles in the pharmaceutical infusion solution online, and achieve a performance rate of above 97% average accuracy.

**Index Terms**—Adaptive local weighted, collaborative sparse model, image segmentation, liquid particle tracking, pharmaceutical inspection, visual inspection.

## I. INTRODUCTION

THE quality of pharmaceutical is very important for the patients, and receiving increasing attention, so the inspection of that has become crucial and urgent. Generally, foreign matters, such as glass, fiber, white spot, metal, and coating fragment, which adhered to the bottom or the sidewall of the container, can be seen in the liquid pharmaceutical if there is a certain deficiency in the process of filtration and sealing. It is reported that when the injection solution contains foreign matters (substances), the patients who use the medicine may catch a fever or chill, and in some case, even face death [1]. Therefore, it is a significant responsibility for pharmaceutical manufacturers to improve the standard of inspection [2].

### A. Techniques for Pharmaceutical Impurity Inspection

Pharmaceutical product defect detection has highly concerned pharmaceutical manufacturers, and related techniques have been improved greatly over the last few decades. Traditionally, the manual visual inspection system (VIS) was adopted earlier as a guaranteed method for the final quality of parenteral drugs produced by the pharmaceutical manufactures. The pharmaceutical factories employ many human inspectors, who are in charge of sorting defective products from qualified products. When performing the manual visual inspection, the operators shake the liquid content of container by means of a circular movement; and then hold the container still, because of the movement inertia of the liquid, possible particles will move together with it. Using a dark background and a bottom-light source, bright particles will become clearly visible, on the other hand, a white background will highlight the presence of black particles [2]. The workers are required to inspect about 3000 bottles in 8 h each day, they tend to get tired easily and cannot assure the detection accuracy. Hence, the manual inspection of pharmaceutical products is an inefficient, subjective, which may cause a second contamination.

Then, the need arises to keep the inspection quality achieved by human inspection while increasing the production rate. At the beginning, the choice is set to develop a semiautomatic inspection system [3], which is able to facilitate the manual inspection procedures when handling the containers to be inspected, but this method only changes the handling process.

The limitation of human inspection has led to many advanced nondestructive testing (NDT) techniques, which acquires the condition of the filling pharmaceutical container by certain sensors (visual and laser sensors) and then detected defects using sophisticated software. Nowadays, available NDT techniques for pharmaceutical inspection include the use of visual cameras [4], laser sensors, light-resistance method, opt-electronic method, static division (SD) sensor, etc.

Opt-electronic inspection method [5] has the best performance in detecting the internal particles of the liquid. According to the Chinese Pharmacopoeia, U.S. FDA, and European Pharmacopoeia, this method is always chosen with the regular sampling inspection in the manufacture testing laboratory. However, its inspection speed is significantly low, thus it cannot be used in the production line. Furthermore, the test results are easily disturbed by external light, and the detection of repeatability is poor. Several improved opt-electronic techniques are proposed to increase the speed and accuracy, such as dual-light-path optical structure inspecting method and the laser method. But they cannot satisfy the demand of online detection for particle defects.

Visual inspection [6], [7] has been developed in recent years with a great progress in computer vision techniques [8]–[10], which is commonly used for inspection in fabric [11], semiconductor wafers [12], rail heads [13], and concrete materials crack detection [14]. To reduce these manual and semiautomatic harmful effects, some factories have planned to develop an automatic inspection equipment for pharmaceutical, with higher speed and accuracy. VIS has the advantages of high speed, low cost, and appealing performance, which is regarded as the most attractive technique for liquid particles detection [15]. The work mainly focuses on the key techniques of VIS.

### B. Challenges of Online Particles Visual Inspection

The online inspection of particles [16], [17] is a main issue in the quality of many pharmaceutical products. For bottling pharmaceutical, the structure of an automatic inspection machine and optic illumination is very complex [18]. A system with multiple high-resolution cameras and high-speed capturing and processing is used to increase the performance (accuracy and efficiency) of the inspection process. Due to the difference in transparency and dimensional tolerances of pharmaceutical packaging materials, the inspection machines often have high rates of false. Consequently, in our country, there are about 6000 enterprises producing the bottling pharmaceutical, but less than 100 of them use the foreign machines to inspect the pharmaceutical foreign particles. In summary, particle defects are difficult to be inspected for a particle visual inspection machine (PVIM) because of the following factors.

- 1) Complexity and precision of mechanism structure.  
Particles lie on the bottom of the container or float

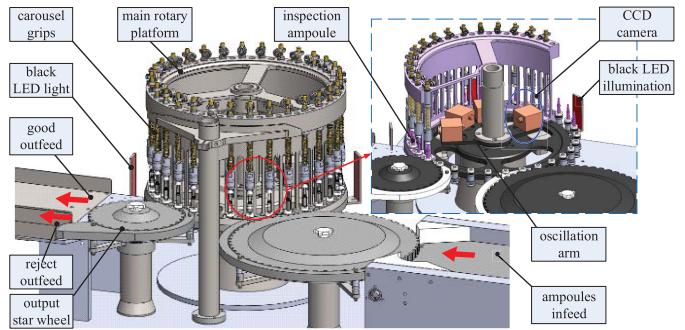


Fig. 1. Prototype of our proposed pharmaceutical inspection system.

in the liquid. They cannot be distinguished from dust or scratches on the outside. The container should spin and rotate by some specific machine. The liquid inside the container, including the particles, start to rotate. The camera captures a series of images. Each camera inspects each individual product. Consequently, the requirements of a mechanical system for online inspection are high-speed rotation and synchronous tracking, vibration-free, and high precision.

- 2) Illumination inequality. Pharmaceutical container images are captured by a camera installed under a moving tracking apparatus in a synchronous operation circumstance [19]. The setup of the image acquisition system used in the inspect machine is seen in Fig. 1. The illumination in an image is ready to be changed because of ambient light and the shake of the machine.
- 3) Limited features are available for the particle tracking and recognition of particles. The particle targets are small, and the interference as well as noise, which are similar to the particles, are complex and diverse. The propensity to form microbubbles [20] or protein particles can also interfere with the inspection. The particles category, shape, and size are uncertain and very familiar with the bottle scratches. Foreign objects in the sequence image, motion trajectories and image deformation, as well as the identification and classification of foreign objects under high-speed movement, these uncertain conditions make particles difficult to be detected.
- 4) The requirement of high detection speed, accuracy, and reliability. Inspection systems are expected to be in real time in high-speed pharmaceutical systems, so a preferable inspection system should have a low computational complexity [21]. The detection resolution of reflective and opaque particles inside the products is up to  $50 \mu\text{m}$  of dimension. Automated inspection offers a consistent and fast process for visual inspection of parenteral products.

### C. Outline of Our Work

This paper presents a real-time PVIM for pharmaceutical product defects, which comprises three subsystems: the mechanical rotatory and tracking system, the image acquisition system, and the image analysis system. The mechanical system

with 120 carousels grip the containers from above keeping it clear for best illumination and inspection of all parts. Image acquisition subsystem acquires a set of gray image sequences of the container, and the images with possible defects can be detected by related algorithm, which can identify the particles from interferences. In this paper, we focus on the followings: design of an inspection machine for particle detection, a new registration method based on improved binary local descriptor [22], [23] for pharmaceutical image sequences, a defect segmentation algorithm based on fuzzy cellular neural networks (FCNNs) [24]–[26], and a robust object tracking algorithm using a collaborative model [27] named as adaptive local weighted-collaborative sparse model (ALW-CSM). The PVIM has the following advantages.

- 1) The improved binary local descriptor for imaging registration can greatly overcome the expensive computation of the traditional feature descriptor. We use the nonmaximal suppression, Hessian matrix, the entropy difference and distance constraint to select stable feature points. A horizontal direction search strategy is used to find the best matching parameters for image matching.
- 2) The particles image segmentation algorithm based on improved FCNNs (IFCNNs) can effectively solve the problem appearing in the small edge detection, which may not be solved by the existing methods, that achieves a better approximation to the original images. It can also well capture the particle variations due to itself spin or drift in the liquid.
- 3) The ALW-CSM is fast and robust to target appearance change, because it relies on the collaborated mode which considers both the latest observations and the original template. When the appearance of the target changes, the model can update the discriminant dictionary in real time, thereby deal with the appearance change effectively and alleviate the drift problem.

This paper is organized as follows. The particle inspection system and its principle are introduced in Section II. Sections III and IV describe the proposed inspection algorithms. Experimental results obtained by applying the proposed method are given in Section V. Finally, Section VI provides the conclusion and suggests the possible future investigations.

## II. OVERVIEW OF PVIM

### A. Mechanical Structure of The PVIM

The proposed automatic inspection machine simulates the manual visual inspection process, the pharmaceutical container caught by a smart robot arm is rotated before the inspection takes place and abruptly stopped in order to have the particles in movement when the container is in front of the camera. According to the material of particles, as well as the manual inspection, different kinds of light sources can be used. The machines use a combination of all types of the lighting system in order to offer the best inspection efficiency regardless of particles nature. When the container stops in front of the camera, a sequence of images are acquired at high speed. Then, the image processing system detected the possible defects by means of special mathematics algorithms

specifically developed depending on the feature of the defects themselves. According to the analysis results, the machine sorts the product for good and rejects it.

The pharmaceutical inspection system is made up of several parts, including the infeed, 120 gripping rotation system (main rotary platform, carousel grips), image acquisition system (12 charge-coupled devices (CCD) and LED light illumination), and outfeed, as shown in Fig. 1.

- 1) The containers are loaded onto the loading table and then sent to the turret by means of a loading star wheel. The center main rotary turret is equipped with independent grips which can catch the container one by one from the loading star wheel.
- 2) The container and liquid are rotated at high speed by a servo drive in each position of the inspection carousel.
- 3) The oscillation arm combined with the CCD and LED moves synchronized with the transport of the container, where the imaging acquisition system trigger camera and LED light, and capture a sequence of images.
- 4) The images are sent to the image analysis system to recognize the particles. In the main rotary turret, the machine sets several light sources and different lighting techniques. Therefore, the process involves 13 cameras divided into 7 analysis stations for glass and level analysis, bottom analysis, and contents analysis.
- 5) The turret output is according to the inspection results, where the container is addressed toward two outfeed levers and the rejected products are automatically collected in rejected trays.

### B. Illumination System and Optical Structure

Dedicated illumination and optical system play a critical role in machine vision applications. Stable and reliable light is an important factor for obtaining an excellent image. To properly visualize any particulate matters that may be present in the liquid, first, we sort the particulates into two types: extraneous reflecting particles (glass, fiber, etc.) and no-reflecting particles (fibers, whiter spots, coating fragments, etc.). According to the reflection and transmission of two types of particles, two different illumination techniques are set in the system. Some appear as black spots against diffused illumination [Fig. 2(a)], the others appear as luminous spots when illuminated by collimated light [Fig. 2(b)]. The typical particles to be detected are shown in Fig. 3. In pharmaceutical particle inspection, LED light is the first choice due to its high efficiency, excellent performance, and easiness of control. Fig. 2 shows the optical structure used for different particle inspection of the liquid.

**Light reflection method (bottom-light illumination):** a single-spot white LED from advanced illumination is sent through the container bottom as it passes along the production line. The camera looking from the side detects light reflections scattered by the particle against a dark background. **Light transmission method (back-light illumination):** a flat red LED is sent through the container side as it passes along the production line. The camera looking from the other side of the container detects shadows created by the particle.

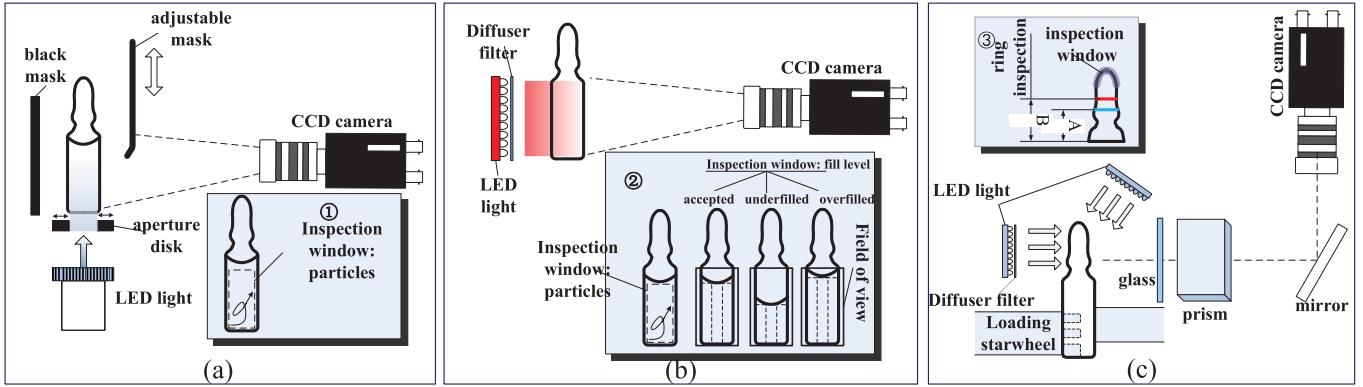


Fig. 2. Light illumination system. (a) Light reflection method for particles inspection (bottom-light illumination): concentrated light is sent through the container bottom. The camera looking from the side detects light reflections scattered by the reflection particle against a dark background. (b) Light transmission method for no-reflection particles inspection (back-light illumination): light is sent through the container side. The camera looking from the other side of the container detects shadows created by the particle. (c) Light illumination for container seal inspection: LED light is sent from back and above with white LED lighting systems and diffuser filters to inspect ampoule tip defects including burned deposit, scratches, stains, defective shape, and identification of color code.

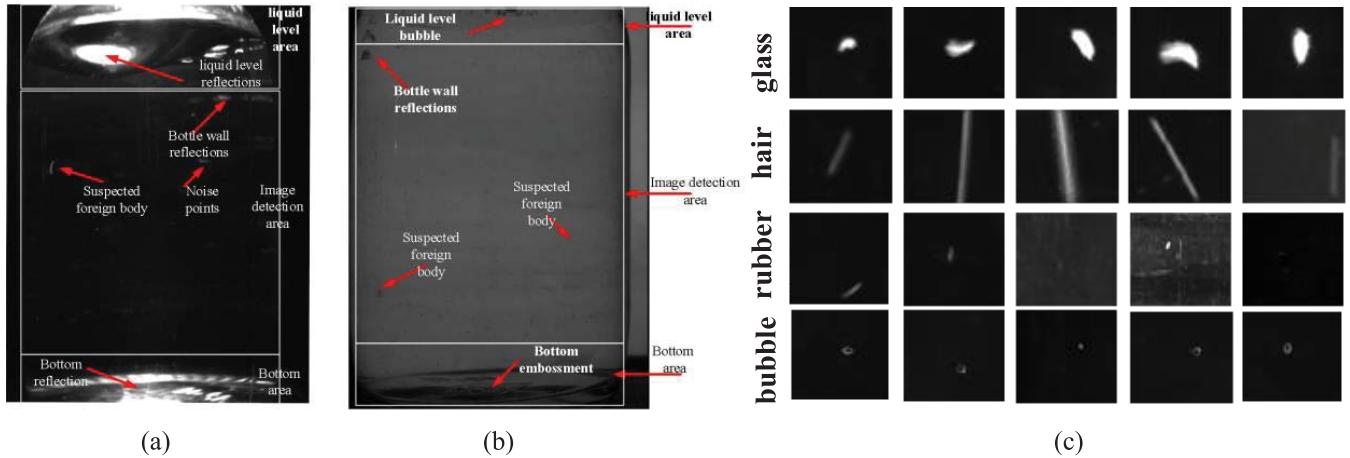


Fig. 3. Example of (a) container image under light reflection method for particles, (b) container image under light transmission method for no-reflection particles, and (c) typical defects of different foreign particles.

### C. Image Acquisition System of Inspection Station

Due to the processing of large images of the container at very high speed and high precision, especially in liquid foreign matter inspection, particles and fill level inspection are formed by 6 detecting stations. One station is responsible for the tip-shape and ring color inspection, two are responsible for fill level and no-reflection particles inspection, and the last three are responsible for reflection particle inspection. Each inspection station consists of two cameras and their dedicated LED light source is located at the bottom or at the back of the bottle. The two CCD cameras are arranged to inspect two containers labeled with an odd and even number, respectively. For the camera, our system uses a Gigabit Ethernet vision CCD camera with an image resolution of  $1600 \times 1200$  pixels. The camera is mounted with an oscillation arm allowing the tracking of the container movement. Accompanied by the camera, we use an M1214-MP2 lens. The related LED light and CCD will be triggered when the drug containers are conveyed through different inspection positions. At the same time, the industry touch PC will start the image capture and complete the real-time inspection. The inspection results will

be transferred to the power line carrier, which will control the ejector to reject the bottles.

### III. PARTICLES DETECTION ALGORITHMS

The pharmaceutical inspector is one of the typical applications of machine vision and a digital image processing technology in industrial production. The most important module of the software is the inspection algorithms, which must be capable of high-speed and accurate application [28]–[30]. In the liquid particle inspection, the foreign matter of liquid varies in size, shape, and position. And there are many factors that will cause a disturbance, such as the bubbles in the liquid and the flicker spots of the bottle itself, surrounding light, and so on. Moreover, the fast-moving mechanism system shakes the bottles randomly, which may result in the image deviation. It is more difficult to deal with the sequence of images. Hence, a very ideal and stable image is often unavailable while utilizing a dedicated light and image capture system.

A specific algorithm flowchart is shown in Fig. 4, for pharmaceutical foreign matter inspection. The data processing module is composed of three major components:

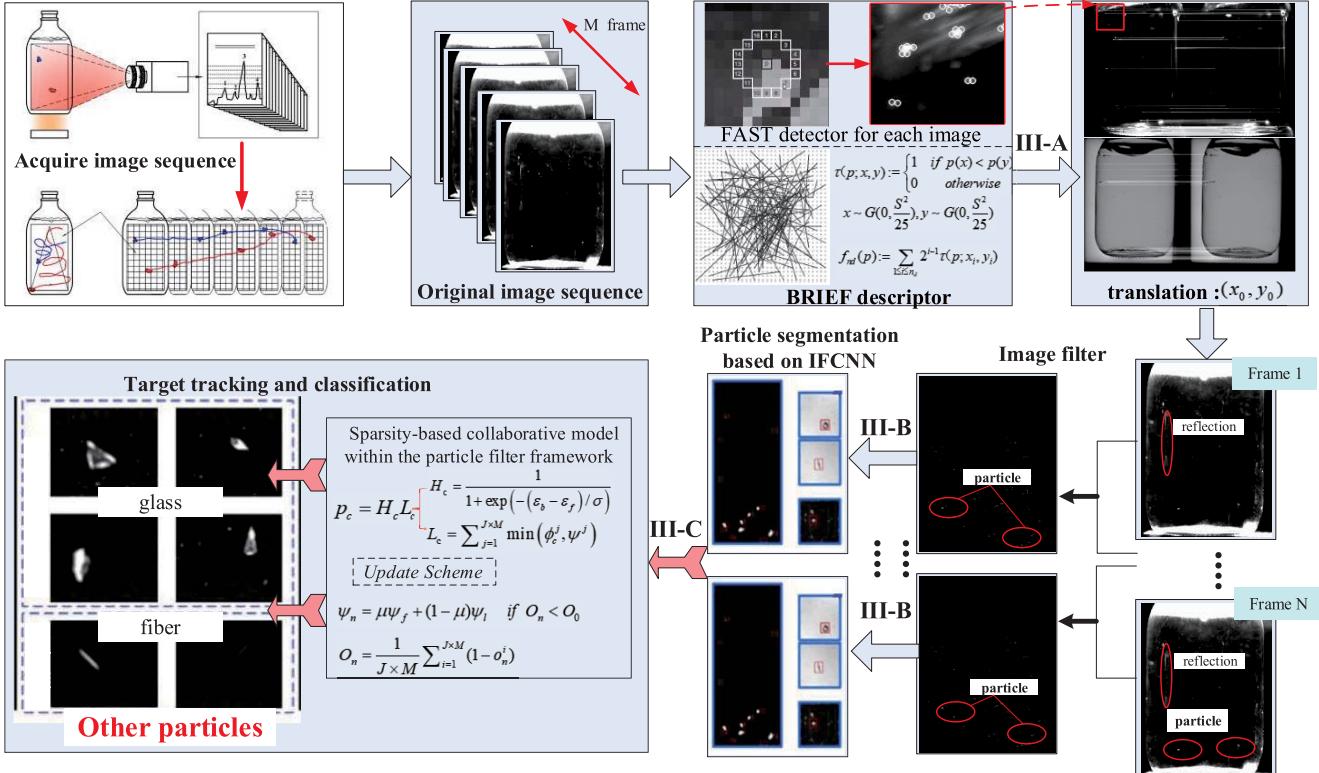


Fig. 4. Flowchart illustrating the steps of particle recognition algorithm.

1) registration and deviation position adjustment; 2) faster filter and segmentation; and 3) particles tracking and classification. First, the raw sequences of images are acquired by the camera. A fast acceleration split test (FAST), high-quality image registration method using FAST detector [31], and BRIEF descriptor [32] are proposed. To detect the particles in movement, a segmentation algorithm [33] based on the sequences image difference using an FCNN is used. Next, the moving tracking and classification method can finally distinguish bubbles and particles from the images sequence. If a defect is successively detected, it will be stored and labeled in the final image. In this algorithm procedure, only moving particles are detected, other elements of the image even though they are very bright or reflecting, such as particles, will not affect the inspection result because it does not produce luminosity variations.

#### A. Registration and Deviation Position Adjustment

Due to the interference of the image displacement deviation, a misjudgment phenomenon occurred when an intelligent inspection machine extracted the foreign body in the infusion liquid using a frame difference method. In order to solve the problem, a new method of binary descriptor block matching is proposed based on features of the accelerated segment test (FAST) [31]. First, the feature points are detected by accelerating the segment test for different scales of the image, and the best feature point is chosen using nonmaximal suppression and the entropy difference. Then, the improved template is used for sampling around the feature point, which forms the new binary descriptor with strong robustness to scale changes,

noise interference, and illumination changes. The dimension of the new descriptor is further reduced. Finally, by using the block matching and threshold methods, the two frame images are matched quickly and accurately, and the displacement deviation is solved and compensated. The detailed description is as follows.

*Step 1:* Using FAST for fast feature point detection. For each location on the circle  $x \in \{1 \dots 16\}$ , the pixel at that position relative to  $p$  (denoted by  $p \rightarrow x$ ) can have one of three states

$$S_{p \rightarrow x} = \begin{cases} d, & I_{p \rightarrow x} \leq I_p - t \quad (\text{darker}) \\ s, & I_p - t < I_{p \rightarrow x} < I_p + t \quad (\text{similar}) \\ b, & I_p + t \leq I_{p \rightarrow x} \quad (\text{brighter}). \end{cases} \quad (1)$$

Choosing an  $x$  and computing  $S_{p \rightarrow x}$  for all  $p \in P$  (the set of all pixels in all training images) partitions  $P$  into three subsets,  $P_d$ ,  $P_s$ , and  $P_b$ , where each  $p$  is assigned to  $P_{S_{p \rightarrow x}}$  and  $t$  is a threshold. Intensity values (I1–I16) are examined by the above method. After the comparison using (1), if a number of points whose intensity is brighter or darker more than 12, it would be a candidate keypoint. In the nonmaximal suppression step, it selects final keypoints that take the maximum value among neighboring keypoints using the following equation:

$$V = \max \left( \sum_{x \in S_{bright}} |I_{p \rightarrow x} - I_p| - t, \sum_{x \in S_{dark}} |I_p - I_{p \rightarrow x}| - t \right). \quad (2)$$

*Step 2:* Using BRIEF algorithm for describing and generating the descriptor. The sampling point gray value sum of the

three adjacent neighborhoods is used to replace the gray value of the single point, which improves the antinoise ability.

The BRIEF algorithm generates bit string description, which consists of 0 and 1

$$\tau(p_{\Sigma}; x, y) = \begin{cases} 0, & p_{\Sigma}(x) < p_{\Sigma}(y) \\ 1, & p_{\Sigma}(x) \geq p_{\Sigma}(y). \end{cases} \quad (3)$$

Binary test  $\tau$  is defined as (3). In (3),  $p_{\Sigma}(x)$  and  $p_{\Sigma}(y)$  indicate the random point, respectively, on the midpoint and the location of the  $3 \times 3$  subwindow pixels gray value and the sum. When it compares intensity values at the two  $3 \times 3$  subwindows, if the  $p_{\Sigma}(x)$  value is less than  $p_{\Sigma}(y)$ , then 0, otherwise 1 of the binary test is assigned

$$f_n(p_{\Sigma}) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(p_{\Sigma}; x_i, y_i). \quad (4)$$

In (4),  $n$  is the number of keypoint pairs, which is the dimension of the descriptor. It can be seen that the BRIEF descriptor will significantly reduce the effect of the single noise point on the descriptor and will be more robust to the noise.

*Step 3:* Using the Hamming distance algorithm to measure the distance between the images. For vectors  $(x_1, x_2, \dots, x_{256})$  and  $(y_1, y_2, \dots, y_{256})$ , using the following equation to calculate the Hamming distance

$$\text{dist} = \sum_1^s (x_i \oplus y_i) \quad (5)$$

where  $i = 1, 2, \dots, s$ ,  $\oplus$  is XOR calculation. Using (6) to determine whether the feature points are matching points, which is suitable for the calculation of the current descriptor distance,  $\text{dist}$ , and  $n$  is a multiple threshold. The minimum distance is the best match point in the window

$$\text{dist} < n \cdot \text{dist}_{\max}. \quad (6)$$

*Step 4:* Calculate the displacement deviation compensation amount. According to the  $i$ th point in the best matches point, the match point coordinate  $P_{1,i}(x_{1,i}, y_{1,i})$  of the front frame image is subtracted from the match point coordinate  $P_{2,i}(x_{2,i}, y_{2,i})$  of the next frame image, where the horizontal and vertical displacement deviation in pixels are got.  $\Delta x_1, \Delta x_2, \dots, \Delta x_m$  and  $\Delta y_1, \Delta y_2, \dots, \Delta y_m$  are obtained by  $m$  match points. The histogram of the distribution of  $\Delta x_1$  and  $\Delta y_1$  is counted, respectively, and then the statistical law is used to remove the small error term that deviates from the real value. The error of the deviation is completely eliminated by the error matching. Finally, the value of the peak is the deviation of the motion displacement.

## B. Target Detection and Segmentation

Before segmentation, the targets are compared with the sequence multiframe image [34], and the motion object is extracted from the sequence image [35] by using the method of interframe difference. In the case of difference, the current frame is different from the first frame of the sequence image. The calculation of the interframe difference method can be expressed as

$$d(x, y, k) = |f(x, y, k) - f(x, y, 1)| \quad (7)$$

where  $d(x, y, k)$  denotes the gray value of the difference result image at the point  $(x, y)$ ,  $f(x, y, k)$  denotes the gray value of any point  $(x, y)$  on the image at  $k$ th frame,  $f(x, y, 1)$  denotes the gray value of any point  $(x, y)$  on the image at first frame,  $k$  is the index of frames of the image sequence,  $k = 2, 3, \dots, 8$ , this paper takes 8 frames. After the above calculation, a sequence difference image is obtained. The goal of image segmentation is to extract important information from the difference images. Many algorithms for weak liquid target segmentation have been developed for blood cell image [25] and leukocyte nucleus [36]. Segmentation in this area [37], [38] is performed for extracting the edge of the small object by using the image histogram analysis method, color-based clustering. Unfortunately, the threshold is not a good approach when SNR is low.

In real pharmaceutical liquid, however, the images used for experiments are rarely noiseless, and the literature does not support a reliable technique that is equally effective in segmenting images corrupted by noise. One of the advantages in the white blood cell detection research is the algorithm proposed by Shitong *et al.* [25], which is based on the FCNN. On the basis of the literature [39], this paper presents an improved FCNN algorithm for the particles detection of pharmaceutical images.

An  $4 \times 4$  IFCNN is similar to the CNN, which is a locally connected network. The output of a neuron is connected to the inputs of every neuron in its  $r \times r$  neighborhood, and similarly, the inputs of a neuron are only connected to the outputs of every neuron in its  $r \times r$  neighborhood. It is apparent that feedback connections will be presented. Considering an  $M \times N$  FCNN, each neuron in the FCNN performs as follows. The state equation of a cell  $C(ij)$  is given as

$$\begin{aligned} C \frac{dx_{ij}}{dt} = & -\frac{1}{R_x} x_{ij} + \sum_{c_{kl} \in N_r(i,j)} A(i, j; k, l) y_{kl} \\ & + \sum_{c_{kl} \in N_r(i,j)} B(i, j; k, l) u_{kl} + I_{ij} \\ & + \tilde{\bigwedge}_{c_{kl} \in N_r(i,j)} (A_{f_{\min}}(i, j; k, l) + y_{kl}) \\ & + \tilde{\bigvee}_{c_{kl} \in N_r(i,j)} (A_{f_{\max}}(i, j; k, l) + y_{kl}) \\ & + \tilde{\bigwedge}_{c_{kl} \in N_r(i,j)} (B_{f_{\min}}(i, j; k, l) + u_{kl}) \\ & + \tilde{\bigvee}_{c_{kl} \in N_r(i,j)} (B_{f_{\max}}(i, j; k, l) + u_{kl}) \\ & + \tilde{\bigwedge}_{c_{kl} \in N_r(i,j)} (F_{f_{\min}}(i, j; k, l) + x_{kl}) \\ & + \tilde{\bigvee}_{c_{kl} \in N_r(i,j)} (F_{f_{\max}}(i, j; k, l) + x_{kl}). \end{aligned} \quad (8)$$

The input equation of  $C(ij)$  is given as

$$u_{ij} = E_{ij}, \quad 1 \leq i \leq M, \quad 1 \leq j \leq N. \quad (9)$$

The output equation of  $C(ij)$  is given as

$$y_{ij} = f(x_{ij}) = \frac{1}{2}(|x_{ij} + 1| - |x_{ij} - 1|). \quad (10)$$

The constraint conditions are

$$\begin{aligned} |x_{ij}(0)| &\leq 1, \quad 1 \leq i \leq M, \quad 1 \leq j \leq N \\ |u_{ij}| &\leq 1, \quad 1 \leq i \leq M, \quad 1 \leq j \leq N \end{aligned} \quad (11)$$

where  $C > 0$ ,  $R_x > 0$ .

The symmetric conditions are

$$\begin{aligned} A(i, j; k, l) &= A(k, l; i, j) \\ A_{f \min}(i, j; k, l) &= A_{f \min}(k, l; i, j) \\ A_{f \max}(i, j; k, l) &= A_{f \max}(k, l; i, j) \\ F_{f \max}(i, j; k, l) &= F_{f \max}(k, l; i, j) \\ F_{f \max}(i, j; k, l) &= F_{f \max}(k, l; i, j) \\ 1 \leq i \leq M, \quad 1 \leq j \leq N \end{aligned} \quad (12)$$

where  $u$ ,  $x$ , and  $y$  denote the input variable, state variable, and output variable, respectively.  $A_{f \min}(i, j; k, l)$ ,  $A_{f \max}(i, j; k, l)$ ,  $B_{f \min}(i, j; k, l)$ , and  $B_{f \max}(i, j; k, l)$  are elements of the fuzzy feedback minimum (MIN) template, fuzzy feedback maximum (MAX) template, fuzzy feed-forward MIN template, and fuzzy feed-forward MAX template, respectively.  $F_{f \max}(i, j; k, l)$  and  $F_{f \min}(i, j; k, l)$  denote the connected weights between neuron  $C(ij)$  and  $C(kl)$ , respectively.  $A(i, j; k, l)$  and  $B(i, j; k, l)$  are elements of feedback template and feed-forward template, respectively, and denote the fuzzy AND and fuzzy OR, respectively. Obviously, the set of controlling parameters of FCNN is  $A$ ,  $B$ ,  $A_{f \min}$ ,  $A_{f \max}$ ,  $B_{f \min}$ ,  $B_{f \max}$ ,  $F_{f \min}$ ,  $F_{f \max}$ ,  $C$ ,  $R_x$ ,  $I(ij) = I$ . Thus, the key to IFCNN is how to set the first eight elements.  $C$  and  $R_x$  determine the recall time. In this paper, we fix  $R_x$  at 1. The output range of IFCNN is  $[-1 \ 1]$  representing a binary image  $(-1 \ 1)$  or a gray image  $[0 \ 1]$ .

#### IV. PARTICLES TRACKING USING ADAPTIVE LOCAL WEIGHTED-COLLABORATIVE SPARSE MODEL

Although recently there has been a surge in the use of object tracking algorithms [40], [41] that include adaptive mean-shift sparse representation, those literature always focus on human surveillance, vehicle tracking, or intelligent video surveillance, and so far little research has been conducted in the study of medical liquid foreign particles to identify the types of particles under different environmental influences by tracking particles. The process is not only faster and reliable but also has less previously unobserved experience.

Based on these considerations, we proposed an ALW-CSM-based objective tracking method within a particle filter framework, which is inspired by the similar techniques for video tracking [42], [43]. The algorithm adopts an adaptive trajectory to obtain candidate image and histogram intersection function by using the optimization function to optimize the sparse discriminant classifier based on the sampling operator, to achieve a better separation of foreground and background. The 2-D Fourier transform is then performed on the candidate image in the weighted sparse generative model (WSGM).

In the WSGM module, the spatial information and the contribution to the local appearance model of each transformed local block are included in the target tracking using the weighted histogram. Furthermore, a weighted histogram intersection function is defined to measure the similarities between candidates and the target model. Finally, the target state is estimated by the maximum posterior probability. The updated scheme considers both the dictionary of the sparse discriminant classifier and the local patch of the WSGM to effectively deal with appearance changes and alleviate the tracking drift problem.

##### A. Tracking Framework

Visual tracking algorithm has been commonly formulated within the Bayesian filtering framework with a goal to determine a posteriori probability, as defined by the following equations:

$$p(\mathbf{x}_t | \mathbf{z}_{1:t-1}) = \int p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \mathbf{z}_{1:t-1}) d\mathbf{x}_{t-1} \quad (13)$$

$$p(\mathbf{x}_t | \mathbf{z}_{1:t}) \propto p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{z}_{1:t-1}) \quad (14)$$

where  $\mathbf{x}_t$  is the object state, and  $\mathbf{z}_t$  is the observation at time  $t$ . Let  $\mathbf{x}_t = [l_x, l_y, \theta, s, \alpha, \varphi]^T$ , where  $l_x, l_y, \theta, s, \alpha, \varphi$  denote  $x, y$  translations, rotation angle, scale, aspect ratio, and skew, respectively. The affine parameters are assumed to be independent and modeled by six scalar Gaussian distributions. The motion model  $p(\mathbf{x}_t | \mathbf{x}_{t-1})$  predicts the state at  $t$ , which employs the immediately previous state, and the observation model  $p(\mathbf{z}_t | \mathbf{x}_t)$  describes the likelihood of observing  $\mathbf{z}_t$  at state  $\mathbf{x}_t$ . The particle filter is an effective realization of the Bayesian filtering, which is able to predict the state regardless of the underlying distribution. The optimal state is obtained by the maximum a posteriori estimation (MAP) over a set of  $N$  samples

$$\hat{\mathbf{x}}_t = \arg_{\mathbf{x}_t^i} \max p(\mathbf{z}_t | \mathbf{x}_t^i) p(\mathbf{x}_t^i | \mathbf{x}_{t-1}) \quad (15)$$

where  $\mathbf{x}_t^i$  is the  $i$ th sample at frame  $t$ . In Section IV-B and IV-C, a tracking algorithm within the particle filter framework is presented. The motion model of the particle filter will adopt sparse representation, and the observation model will utilize WSGM.

##### B. Motion Model and Adaptive Sampling

Suppose the size of the sampling image block is  $b$ , and that  $S = [s_1(x_1, y_1), s_2(x_2, y_2), \dots, s_t(x_t, y_t)]$  is denoted as the target position estimated in the previous  $t$  frame ( $s_t$  is denoted as the target position in the  $t$  frame). The sampling strategy of the candidate image block adopted by conventional collaborative sparse representation is: in  $t+1$  frame, centering the foreign particles position estimated in  $t$  frame with a radius of  $r$ . The candidate image block can be obtained as

$$\begin{aligned} C &= \{c_n(x_n, y_n) | x_n = -r + x_t + 2r \cdot \text{rand}([0, 1]), \\ y_n &= -r + y_t + 2r \cdot \text{rand}([0, 1]), n = 1, 2, 3, \dots, N\} \end{aligned} \quad (16)$$

where  $c_n$  is the candidate image block,  $x_n$  and  $y_n$  are coordinates of  $c_n$ ,  $x_t$  and  $y_t$  are denoted as coordinates of  $s_t$ ,  $\text{rand}([0, 1])$  is a function that generates a random number

TABLE I  
ADAPTIVE SAMPLING STRATEGY

| The image trajectory of the $t$ -th frame                             |                       |                       |                      |        |
|---|-----------------------|-----------------------|----------------------|--------|
| $x_1 - x_t > \delta$  | $x_1 - x_t < -\delta$ | $x_1 - x_t < -\delta$ | $x_1 - x_t > \delta$ | Others |
| The distribution of candidate image blocks in the $t + 1$ frame image |                       |                       |                      |        |
|   |                       |                       |                      |        |
|   |                       |                       |                      |        |

between 0 and 1, and  $N$  denotes the number of candidate image blocks ( $N = 220$ ). However, such a sampling method does not take the target's past trajectory into consideration during candidate image blocks selection. Since the pharmaceutical sequence images under tracking are images of foreign particles' regular centrifugal whirligig, it is necessary to consider the past trajectory when selecting candidate blocks.

To improve the sampling method of candidate image blocks, the sampling strategy in the current frame is adaptively adjusted according to the past trajectory of the foreign particles, as shown in Table I. Thereby, the candidate image blocks assembled in the  $t + 1$  frame is obtained as follows:

$$\begin{aligned} C' = \{c'_n(x_n, y_n) | x'_n &= x_t + \omega_x \cdot 2r \cdot \text{rand}([0, 1]), \\ y'_n &= y_t + \omega_y \cdot 2r \cdot \text{rand}([0, 1]), n = 1, 2, 3, \dots, N\} \end{aligned} \quad (17)$$

where  $c'_n$  is the candidate image block,  $x'_n$ ,  $y'_n$  are the coordinates of  $c'_n$ ,  $x_n$  and  $y_n$  are coordinates of  $s_n$ , and  $\omega_x$  and  $\omega_y$  are defined as

$$\begin{aligned} \omega_x &= \begin{cases} 1, & x_1 - x_t < -\delta \\ -1, & x_1 - x_t > \delta \end{cases} \\ \omega_y &= \begin{cases} 1, & y_1 - y_t < -\delta \\ -1, & y_1 - y_t > \delta. \end{cases} \end{aligned} \quad (18)$$

For improving the sampling method of candidate image blocks, the robustness is enhanced to the interference caused by other foreign particles' motion.

### C. Observation Model

Most tracking methods exploit rectangular image regions to represent targets, and thereby, it is inevitable for background pixels to be included as a part of the foreground objects. Consequently, when the background patch is used as an updated foreground, the classifier based on the local representation will have a serious impact. In this paper, we propose a collaborative observation model that integrates an optimized discriminative

classifier based on holistic templates and a weighted generative model using local representations.

1) *Holistic Appearance Model*: Motivated by the literature [42], we propose a sparse discriminative classifier (SDC) based on the optimized dictionary for liquid particles object tracking using the following steps.

Step 1: To obtain templates. The training image set is composed of  $n_p$  positive templates and  $n_n$  negative templates.  $n_p$  sample images are obtained around the target location within a radius of a few pixels and the selected images are downsampled to a canonical size ( $64 \times 64$ ). Each downsampled image is stacked together to form the set of positive templates. Similarly, the negative training set is composed of images further away from the target location. Thus, the negative training set consists of both the background and the images with parts of the target object.  $n_p$  positive templates and  $n_n$  negative templates are vectorized separately to get the matrix  $A_+$  and  $A_-$ , which are stacked together to obtain matrix  $A$ .

Step 2: To calculate histogram matrix. We calculate the histogram of  $n_p$  positive template separately and expand each histogram to a vector (each histogram is expanded into a  $256 \times 1$  dimensional vector in our experiments). In this way, we get a matrix  $H_p$  that contains the positive template histogram. Similarly, a matrix  $H_n$  can be obtained.

Step 3: To define a filter function. In order to improve the discriminative property of the holistic model, we define a filter function

$$Ch(A_-) = \begin{cases} A_-(: , i) & S(H_n(: , i), H_p(: , 1)) \leq M \\ A_-(: , i)\text{delete} & S(H_n(: , i), H_p(: , 1)) > M \end{cases} \quad i = 1, 2, 3, \dots, n_p \quad (19)$$

where  $S()$  is a histogram cross function computed by

$$S() = \sum_{j=1}^{256} \min(H_n(j, i), H_p(j, i)). \quad (20)$$

And  $M$  is denoted as the maximum similarity between images in the foreground dictionary, computed by

$$M = \max \left\{ \sum_{j=1}^{256} \min(H_P(j, i), H_P(j, 1)) \right\} \\ i = 1, 2, 3, \dots, n_p. \quad (21)$$

In this way, a dictionary  $A^c$  is improved for integrating an optimized background dictionary exploiting filter function  $S()$  and foreground dictionary, which can eliminate all kinds of interferences that may be included in the background.

*Step 4:* To select determinative features. Since the matrix  $A^c$  mentioned before is rich yet redundant, we design a classifier to extract determinative ones that best distinguish the foreground objects from the background. The classifier is defined as

$$\min_s \| (A^c)^T s - p \|_2^2 + \lambda_1 \| s \|_1 \quad (22)$$

where vector  $p \in R^{(n_p+n'_n) \times 1}$  represents the property of each template in the training set  $A^c$ , i.e.,  $+1$  for positive templates and  $-1$  for negative templates. The solution of (22) is the sparse vector  $s$ , whose nonzero elements are denoted as the most discriminative features selected from the K-D feature space. The feature selection scheme adaptively chooses a suitable number of discriminative features in dynamic environments via the  $l_1$ -norm constraints. We project the features to a subspace via a projection matrix  $S$ , which is formed by removing all-zero rows from a diagonal matrix and the elements are determined by

$$S'_{ii} = \begin{cases} 0, & s_i = 0 \\ 1, & \text{otherwise.} \end{cases} \quad (23)$$

Both the training template set and the candidates sampled by a particle filter are projected to the discriminative feature space. Thus, the training template set and the candidates in the projected space are  $A' = SA^c$  and  $x' = Sx$ .

*Step 5:* To measure confidence. The proposed SDC method has been developed based on the assumption that a target image region can be better represented by the sparse combination of positive templates, while a background patch can be better represented by the span of negative templates. An optimized matrix  $A'$  with a dimensionality reduction of the training set is employed to be more discriminative. The SDC is implemented by a sparse decomposition of the dictionary. Given a candidate region  $x$ , it is represented by the training template set with the coefficients computed using

$$\min_{\alpha} \|x' - A'\alpha\|_2^2 + \lambda_2 \|\alpha\|_1 \quad (24)$$

where  $x'$  is the projected vector of  $x$  and  $\lambda_2$  is a weight parameter. A candidate region with smaller reconstruction error using the foreground template set indicates that it is more likely to be a target object and vice versa. Thus, we formulate the confidence value  $H_c$  of the candidate  $x$  by

$$H_c = \frac{1}{1 + \exp(-(\varepsilon_b - \varepsilon_f)/\delta)} \quad (25)$$

where  $\varepsilon_f = \|x' - A'\alpha'\|_2^2$  is the reconstruction error of the candidate  $x$  with the foreground template set, and

is the corresponding sparse coefficient vector. Similarly,  $\varepsilon_b = \|x' - A'_-\alpha_-\|_2^2$  is the reconstruction error of the candidate  $x$  using the background template set and  $\alpha_-$  is the corresponding sparse coefficient vector.

With the holistic appearance proposed in [42], the reconstruction error is computed on the basis of the dictionary of optimized positive and negative template sets. However, we adopt the histogram cross function as the optimization function to eliminate similar images in positive and negative templates, which is helpful to improve the discriminative ability of the holistic appearance consequently.

*2) Local Appearance Model:* In this section, we present a weighted generative model based on local appearance, which not only takes the gray-scale features of local image blocks into consideration but also the location. Generally speaking, a local image located at the center of the candidate consists of more information of the object, because they are likely to have a greater contribution. Therefore, we design a contributing factor of the image for the generative model. The specific steps are shown as follows.

*Step 1:* Spectrogram transformation. Because the gray characteristics of the object are not obvious in the experiment. That is to say, the amount of information is not enough. Consequently, we transform the gray-scale features of the objects into spectrogram through 2-D Fourier transformation.

*Step 2:* Sparse coefficient vector computing. We exploit the spectrogram to represent the local features of an object where each image is normalized to  $32 \times 32$  pixels. We use overlapped sliding windows on the normalized images to obtain  $M$  patches, and each patch is converted to a vector  $y_i \in R^{G \times 1}$ , where  $G$  denotes the size of the patch. The sparse coefficient vector of each patch is computed by

$$\min_{\beta_i} \|y_i - D\beta_i\|_2^2 + \lambda_3 \|\beta_i\|_1 \quad \text{s.t. } \beta_i \geq 0 \quad (26)$$

where the dictionary  $D \in R^{G \times J}$  is generated from the  $J$  cluster centers using the  $k$ -means algorithm on the  $M$  patches from the first frame and  $\lambda_3$  is a weight parameter.

The contribution of these local patches is not taken into account in the literature [42], which results in the tracking drift caused by neglecting background information. Therefore, we formulate a contribution  $w_k$  factor to give different contributions to different patches. And the contribution factor  $w_k$  of the  $k$ th patch is defined as follows:

$$w_k = 1 + e^{-\beta(|i - \frac{1+r}{2}| + |j - \frac{1+c}{2}|)}, \quad i = 1, 2, \dots, r, \quad j = 1, 2, \dots, c \quad (27)$$

where  $\beta$  is a constant,  $r$  denotes the number of row of a patch in the whole image, and  $c$  is the number of columns in the image of the local image block.  $e^{-\beta(|i - ((1+r)/2)| + |j - ((1+c)/2)|)}$  represents those patches from different positions having different contributions. More specifically, the patch at the center of the image [ $i = ((1+r)/2)$ ,  $j = ((1+c)/2)$ ] has a maximum contribution factor of 2, while the contribution factor of the patches on the edge of image ranges from 1 to 2.

In this paper, we have constructed a weighted histogram based on the sparse coefficient vector  $\beta_i \in R^{J \times 1}$  and

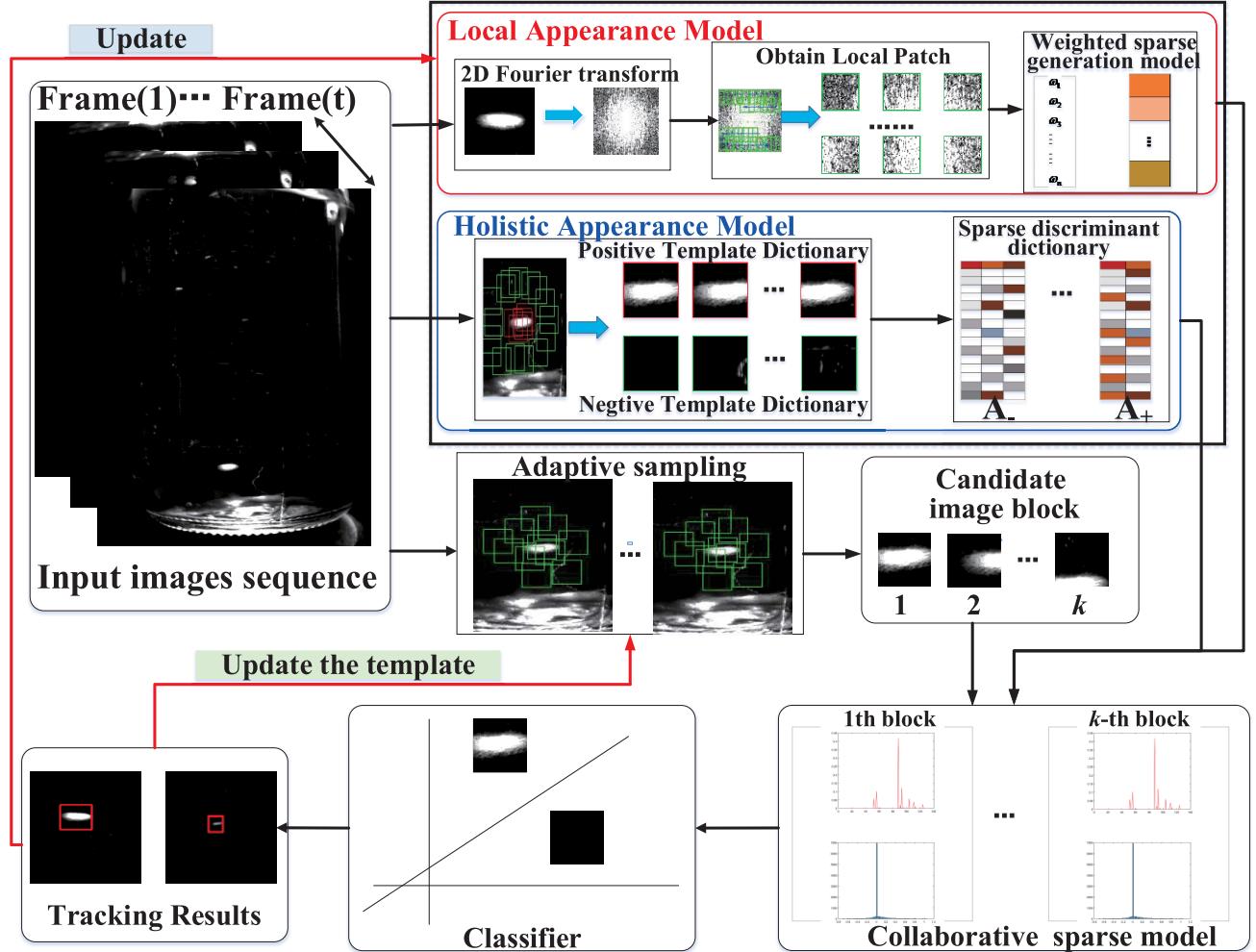


Fig. 5. Flowchart of ALW-CSM algorithm.

contribution factor  $\omega_k$

$$\rho = [\omega_1 \beta_1^T, \omega_2 \beta_2^T, \dots, \omega_M \beta_M^T]^T \quad (28)$$

where  $\rho \in R^{(J \times M) \times 1}$  is the weighted histogram for one candidate region.

*Step 3: Similarity function.* We exploit the histogram intersection function to compute the similarity of histograms between the candidate and the template by

$$L_c = \sum_{i=1}^M \omega_i \sum_{j=1}^J \min(\phi_c^{ji}, \psi^{ji}) \quad (29)$$

where  $\phi_c$  and  $\psi$  are the histograms for the  $c$ th candidate and the template, respectively. The histogram of the template  $\phi_c$  is generated by (23)–(26).

*3) Collaborative Sparse Model:* We propose a collaborative sparse model by adopting a discriminative classifier based on an optimized dictionary and a weighted generative model based on spectrogram within the particle filter framework. In our tracking algorithm, both the confidence value based on the holistic templates and the similarity measure based on the local patches contribute to an effective and robust probabilistic

appearance model. The likelihood function of the  $c$ th candidate region is computed by

$$p(z_t | x_t^c) = H_c L_c \quad (30)$$

where  $H_c$  is obtained by (25) and  $L_c$  is obtained by (29). What's more, each tracking result is the candidate with the MAP.

*4) Update Scheme:* It is important and necessary to the update scheme due to the huge appearance change of objects while tracking. In this paper, we have updated the local appearance model and the holistic appearance model independently.

For holistic appearance model, we have updated the dictionary every two frames due to a huge appearance change and fast variation of objects. The SDC is used to distinguish target foreground from the background, so it is vital to ensure that the dictionary is correct and discriminative. For the local appearance model, the dictionary remains the same when tracking. Therefore, to account for the change of the local appearance model, the new template histogram is calculated as follows:

$$\psi_n = \mu \psi_f + (1 - \mu) \psi_l \quad (31)$$

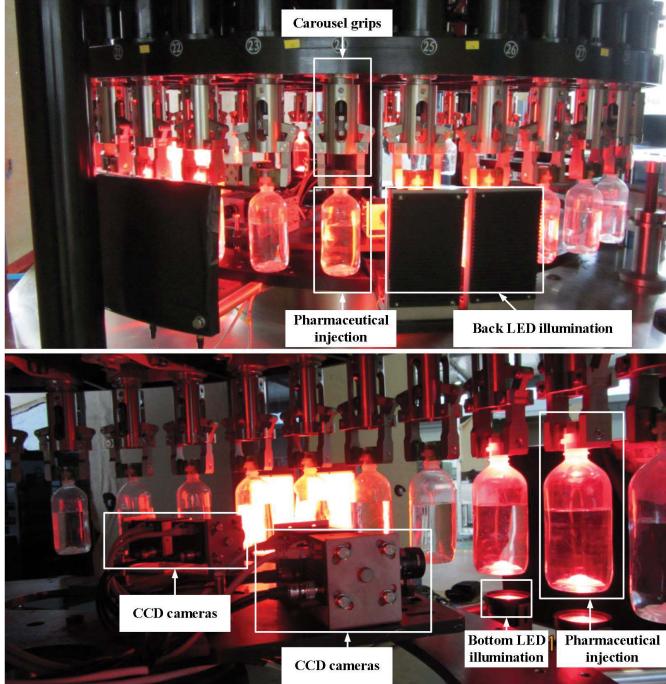


Fig. 6. Pharmaceutical VIS based on multicameras.

where  $\psi_f$  is the histogram in the first frame,  $\psi_l$  is the histogram in the latest frame, and  $\mu$  is the weight. Our appearance model is updated every two frames, and the update frequency is the same for the holistic appearance model and the local appearance model. To make the ALW-CSM algorithm more clearer, we give a flowchart of the ALW-CSM tracking algorithm, as shown in Fig. 5.

## V. EXPERIMENTAL RESULTS AND DISCUSSIONS

### A. Experiment System

To evaluate the performance of the proposed method, a pharmaceutical VIS (as shown in Fig. 6) is established according to the scheme given in Section II. The VIS mainly consisted of several vision stations with PointGrey cameras equipped with 16-mm lens and 5-V/5-W red LED point light source or area source. The camera resolution is  $1600 \times 1200$  pixels, and we use the CCD capture  $80 \text{ mm} \times 100 \text{ mm}$  250-ml infusion solution. The cameras are placed on the motion platform. The inspection vision systems are mounted on an oscillation arm to capture the image sequence continuously. The CPU of the host computer is Intel Core i7 processor and 8-GB RAM using VS2013+OPENCV2.4.9. After building a machine vision detection platform, we obtained the images from different light sources, as shown in Fig. 6.

### B. Registration and Positioning Adjustment Experiments

According to the registration and deviation position adjustment method presented in Section III. There is no scale or illumination changes, mechanical jitter caused by the image deflection is only a small offset, the rotation will not exceed 5. The parameters in (1)–(6) are the threshold value  $t = 5, 10, 15, 20$ , the match subwindow  $3 \times 3$ , back-light

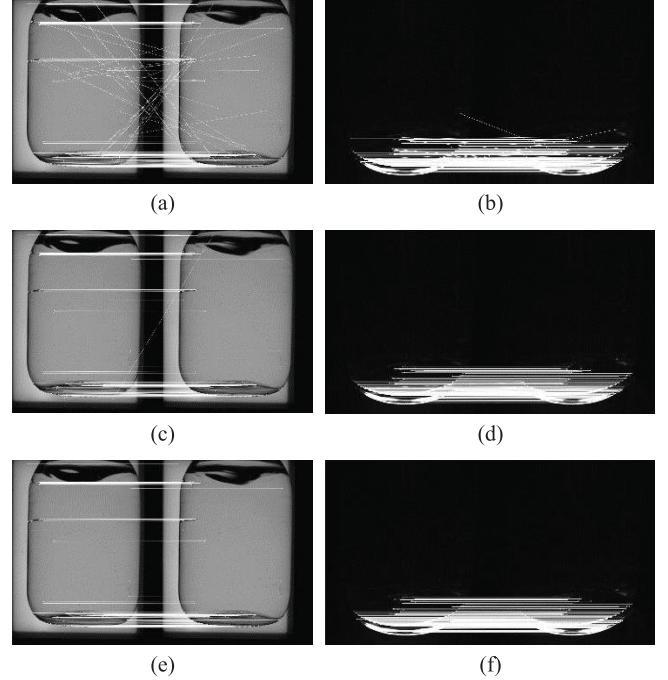


Fig. 7. Matching results under proposed method. (a) and (b) No matching distance threshold of back-light and bottom-light illumination images. (c) Matching distance threshold  $n = 0.4$ . (d) Matching distance threshold  $n = 0.3$ . (e) Distance threshold  $n = 0.4$ ,  $3 \times 3$  subwindow block matching. (f) Distance threshold  $n = 0.3$ ,  $3 \times 3$  subwindow block matching.

matching  $n = 0.4$ , and bottom-light matching  $n = 0.3$ . The experimental results of our method with different parameters are reported in Table II. In the back-light station, our method with the distance threshold  $n = 0.4$ ,  $3 \times 3$  subwindow block matching can get a matching error rate of 3.2%, which is significantly better than other parameter methods. The feature point detection, generated descriptor, and matching time is 125, 78, and 15 ms, respectively, there are 125 matching points. The method for the bottom-light station matching is similar to the situation at the back-light station.

According to the matching point pairs, the calculated motion offset results are shown in Fig. 7. The results showed that the offset in the horizontal  $X$  direction is much larger than the offset in the vertical  $Y$  direction, and the fluctuation range of the data is much larger than that of the vertical  $Y$  direction, both in the back-light and bottom-light image sequences. In the horizontal  $X$  direction, the back-light fluctuates within 12 pixels of “0–12,” and the bottom light fluctuates within 18 pixels of “−5–13”; for the vertical  $Y$  direction, the back-light fluctuates within 4 pixels of “−3–1,” and the bottom light fluctuates within 5 pixels of “−2–3.”

In Fig. 8, the motion deviation vector of the back-light image used in the experiment is  $\Delta U_B = (\Delta x_B, \Delta y_B)$ , the number of occurrences of 9-pixel appearances in the horizontal  $X$ -direction offset  $\Delta x_B$  is 75 times, and the number of occurrences of −1 pixel in the vertical  $Y$ -direction offset  $\Delta y_B$  is 59 times. Here, the image motion vector displacement  $\Delta U_B = (9, -1)$ . We use the same method to calculate the bottom-light deviation  $\Delta U_D = (9, 0)$ , which is shown in Fig. 8(b).

TABLE II  
RUNNING TIME AND TOTAL MATCHING NUMBER AND MATCHING ERROR RATE FROM TWO IMAGE SEQUENCES

| Station      | Method                                 | Running Time(ms)        |                     |          | Matching Number | Matching Error Rate |
|--------------|--|-------------------------|---------------------|----------|-----------------|---------------------|
|              |  | Feature Point Detection | Generate Descriptor | Matching |                 |                     |
| back-light   | No matching distance threshold         | 120                     | 93                  | 15       | 194             | 34.54%              |
|              | distance threshold n=0.4               | 125                     | 78                  | 15       | 141             | 14.18%              |
|              | n=0.4, 3 × 3 sub-window block matching | 125                     | 78                  | 15       | 125             | 3.20%               |
| bottom-light | No matching distance threshold         | 131                     | 95                  | 16       | 275             | 27.27%              |
|              | distance threshold n=0.3               | 124                     | 94                  | 15       | 243             | 20.57%              |
|              | n=0.3, 3 × 3 sub-window block matching | 121                     | 83                  | 16       | 222             | 13.06%              |

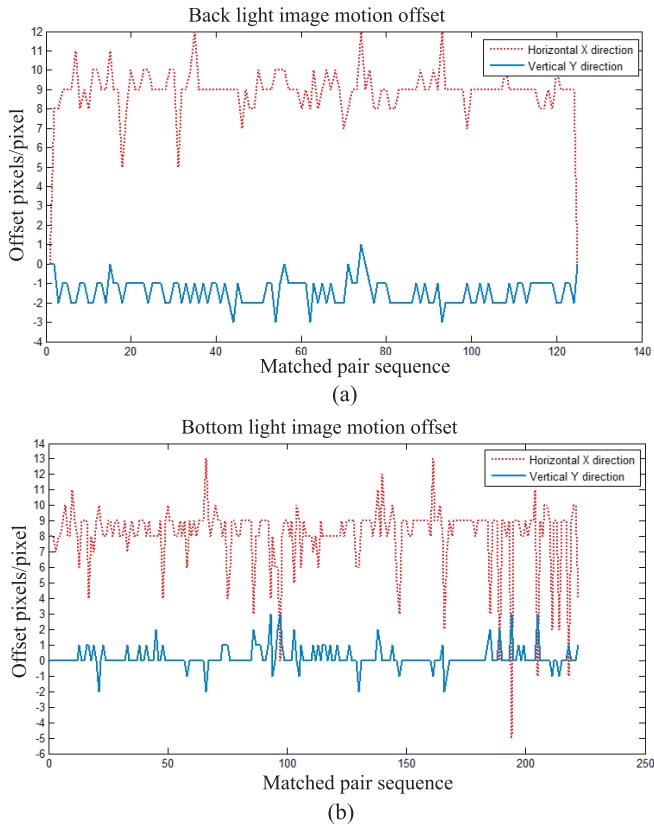


Fig. 8. Deviation computation results of (a) back-light sequence deviation and (b) bottom-light sequence deviation.

### C. Performance of IFCNN for Image Segmentation

In the following, we evaluate the parameter templates performance of IFCNN. However, we cannot determine how to choose them well using a theoretical means. So, a feasible trial strategy is taken in this paper based on a lot of experimental results. Three examples of the segmentation results are shown in Fig. 9, which represents over dilation, over corrosion, and suitable. From Fig. 9, we can see that the proposed method can accurately extract the contour lines of the foreign particle images that include hair, rubber flolage, and glass [Fig. 9(a)].

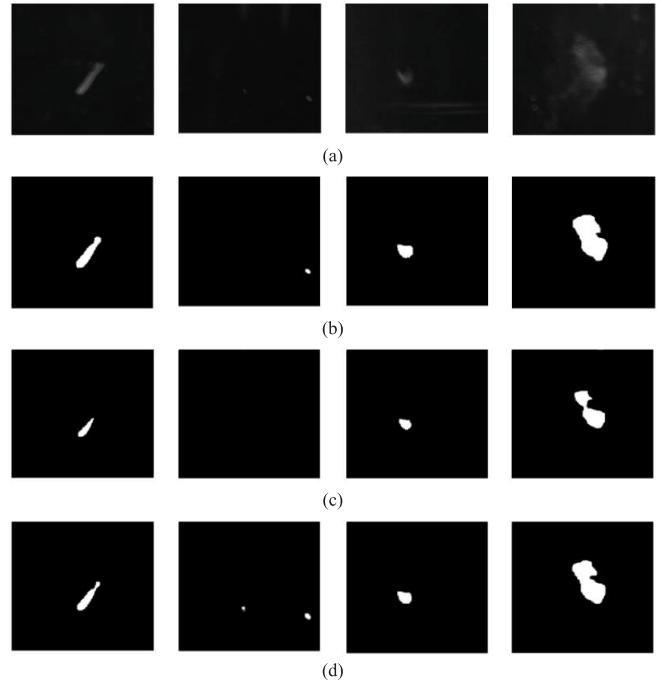


Fig. 9. Foreign substances image processing results of IFCNN. (a) Original image. (b) Template A (over dilation). (c) Template B (over corrosion). (d) Template C.

#### 1) Template A:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A_f \min = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 2/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}, \quad A_f \max = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$B_f \min = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}, \quad B_f \max = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$F_f \min = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad F_f \max = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$R_x = 1, I = 0, u = X, x_0 = \text{undefined}, y = Y.$$

2) Template B:

$$A = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & -3/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -2 & 2 \\ 1 & 2 & -1 \end{bmatrix}$$

$$A_f \min = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad A_f \max = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$B_f \min = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad B_f \max = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$

$$F_f \min = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 2/9 & 1/9 \\ 1/9 & 1/9 & 5/9 \end{bmatrix}, \quad F_f \max = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$R_x = 1, I = 0, u = X, x_0 = \text{undefined}, y = Y.$$

3) Template C:

$$A = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A_f \min = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 2/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$

$$A_f \max = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 0 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$

$$B_f \min = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 2/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}, \quad B_f \max = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$F_f \min = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 2/9 & 1/9 \\ 1/9 & 1/9 & 5/9 \end{bmatrix}$$

$$F_f \max = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 2/9 & 1/9 \\ 1/9 & 1/9 & 3/9 \end{bmatrix}$$

$$R_x = 1, I = 0, u = X, x_0 = \text{undefined}, y = Y.$$

From the experimental results shown in Fig. 9, we can find that with the template parameter A, the split map we got is too inflated, which cannot effectively correspond to the original image of the foreign body [Fig. 9(b)]. The segmentation result obtained by using the template parameter B is too corroded [Fig. 9(c)]. After several tests, the template parameter C can be used to realize the approximate perfect segmentation of the foreign object [Fig. 9(c)]. Then, according to the image preprocessing method and the image differential motion foreign body extraction process, we tested the large infusion sequence images and used this IFCNN method to segment the moving foreign body. It can be seen from the experimental results in Fig. 10, whose image difference [Fig. 10(c)] and cumulative operation adopted in this paper, have achieved good results in the extraction of moving objects [Fig. 10(d)]. Comparing the IFCNN [Fig. 10(f)] with CNN [Fig. 10(e)] segmentation methods, it was found that IFCNN was superior to the traditional CNN method in the process of foreign body segmentation, and the visible foreign body can be detected by complete segmentation.

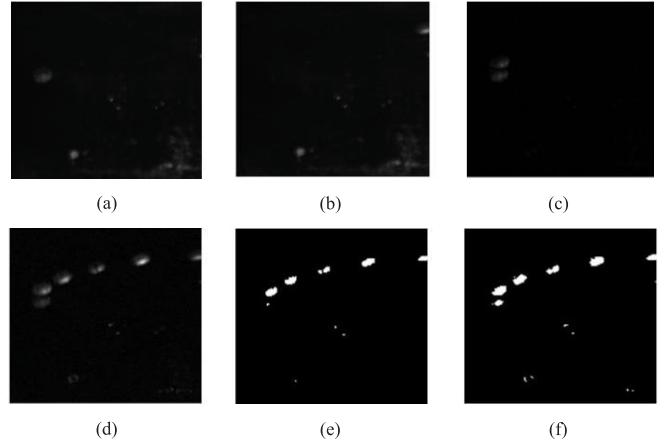


Fig. 10. Image segmentation experiment of back illumination. (a) First frame image. (b) Sixth frame. (c) Difference imaging. (d) Difference image accumulation. (e) CNN imaging segmentation. (f) IFCNN imaging segmentation.

TABLE III  
IMAGE TRACKING SEQUENCES USED IN OUR EXPERIMENTS

| Test sequence | #Frames | Main challenges                                     |
|---------------|---------|---|
| Foreign1      | 8       | Fiber foreign particles tracking                    |
| Foreign2      | 8       | Glass foreign particles tracking                    |
| Foreign3      | 8       | Rubber foreign particles tracking                   |
| Occlusion     | 8       | Occluded foreign particles tracking                 |
| Cross1        | 8       | Subject to another movement of foreign interference |
| Cross2        | 8       | Subject to another movement of foreign interference |

#### D. Particle Tracking Results and Analysis

In this section, to evaluate the performance of ALW-CSM algorithm, we conducted experiments and analyzed the results both qualitatively and quantitatively. The detail of the selected image sequence is listed in Table III. The tested images are acquired for fusion medicine during production via intelligent inspection machine, which is shown in Fig. 6. The selected test sequence covers most of the challenges in object tracking, such as appearance changes, occlusion, scale variation, motion blur, and so on. We have chosen six state-of-the-art tracking algorithms and compared their tracking performance. The selected tracking methods are L1 [44], adaptive local appearance model (ALSA) [45], sparse collaborative appearance model (SCAM) [42], partial least squares analysis (PLS) [46], structured multi-task sparse learning (MTT) [47], and local distribution fields [48], all of which are proved to perform well in object tracking. All tracking algorithms exploit the application codes provided by us and are in the same initial state.

The parameters of the proposed tracking algorithm are fixed in all experiments. The numbers of positive templates  $n_p$  and  $n_n$  negative templates are 50 and 200, respectively. The weight parameters  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  of (22), (24), and (26) are set to 0.01. In all experiments, the size of sampled patch  $M$  of our algorithm in holistic appearance model is  $64 \times 64$ , and the size of local patch  $m$  that is sampled of our algorithm in local appearance model is  $32 \times 32$ . The weight parameter  $\mu$  in (31) is 0.5. In all experiments, the parameters  $\theta$ ,  $s$ ,  $\alpha$ , and  $\varphi$  of the object state  $x_t$  are set to be 0.

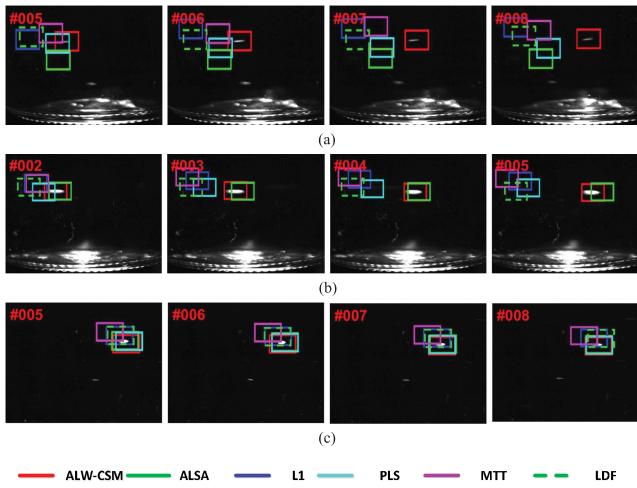


Fig. 11. Sample tracking result of fiber, skins, and rubber dander foreign particles. (a) Fiber foreign particle. (b) Skins foreign particle. (c) Rubber dander foreign particle.

*1) Qualitative Evaluation:* In this section, we evaluate the ALW-CSM algorithm qualitatively under the circumstances of interference, and present tracking results on each image sequence. In order to display the tracking results of the ALW-CSM algorithm more clearly in the image, the test results of six methods of each test sequence are given.

*a) Experiment 1:* When the target is drastically changed in undisturbed condition, the tracking effect of the ALW-CSM algorithm is shown in Fig. 11. In the fiber foreign particle sequence [Fig. 11(a)], almost all algorithms, except ALW-CSM, began to fail to track the target as the scale of the fiber changed dramatically. In addition, except for the ALW-CSM algorithm, other algorithms have completely lost the target in the last two frames. As shown in Fig. 11(b), in the glass foreign particle sequence, both the ALW-CSM algorithm and the MTT algorithm perform effectively in the tracking of the target trajectory due to quite a small appearance change. Moreover, the ALW-CSM algorithm performed better with smaller errors. As for rubber chips sequence shown in Fig. 11(c), all algorithms could track the object successfully. However, relatively, the ALW-CSM achieved better tracking accuracy and smaller error. Above all from the aspect of qualitative evaluation, the ALW-CSM algorithm can not only track several typical foreign objects effectively but also obtain a better tracking accuracy.

*b) Experiment 2:* As shown in Fig. 12, the appearance of the object will change because of occlusion caused by scratches of the bottle, which has a great impact on the tracking effect. It can be seen that the object has undergone deformation twice throughout this sequence. The first time occurred in frame 1 and frame 4, and the second time in frame 5 and frame 6. At the first occlusion, the MTT, PLS, and L1 algorithms failed to track the object due to varied scale. At the second occlusion, the tracking method based on the traditional sparse representation gradually fail to locate the target because of deformation resulted by occlusion. In contrast, the method based on ALSA and ALW-CSM algorithms performed effectively. However, because ALSA does

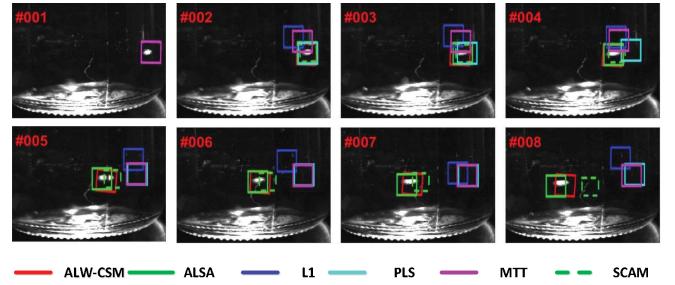


Fig. 12. Sample tracking result of foreign particle on image sequences with occlusion.

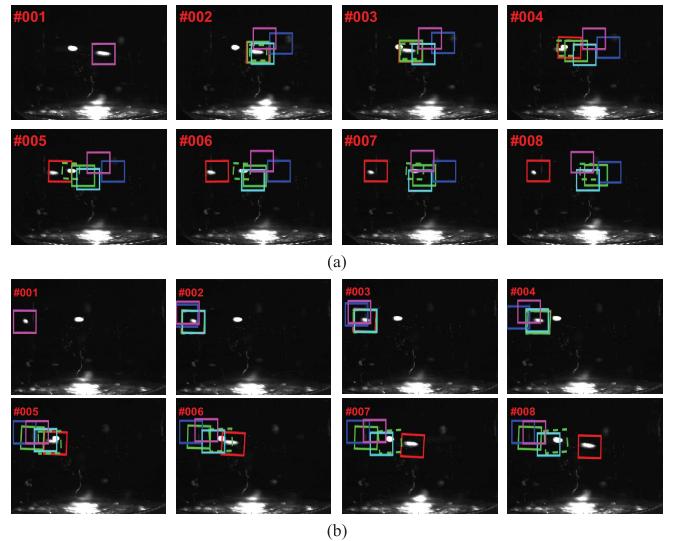


Fig. 13. Sample tracking result of foreign particle on two image sequences with other similar foreign matter interference.

not update the target template in time, its tracking error is larger, while the ALW-CSM algorithm is not only able to track the trajectory of the target but also achieves less tracking error.

*c) Experiment 3:* As shown in Fig. 13, the tracking is made vulnerable by other objects' motion when the target is moving, resulting in mistakenly locating the target. In the fourth frame of the sequence, it can be seen that the target overlapped with other moving objects in the moving process, which made it impossible to distinguish it from other objects. At the beginning, the L1 and PLS methods failed to track the target right away in the second frame due to scale variation of the target. However, the ALSA, SCAM, MTT, and ALW-CSM methods can track the target with the interference of other objects. Once the interference occurs, only the ALW-CSM method is able to keep the tracking of the target. The ALSA, SCAM, and MTT methods are unable to handle a situation where the motion trajectories of the target and other objects are in crosstalk, because they do not take the past motion trajectory of the target into consideration. In contrast, the ALW-CSM algorithm performs quite robustly in crosstalk by using the past motion trajectory of the target to eliminate the errors.

*2) Quantitative Evaluation:* We evaluate the above algorithm quantitatively by using the center position error.

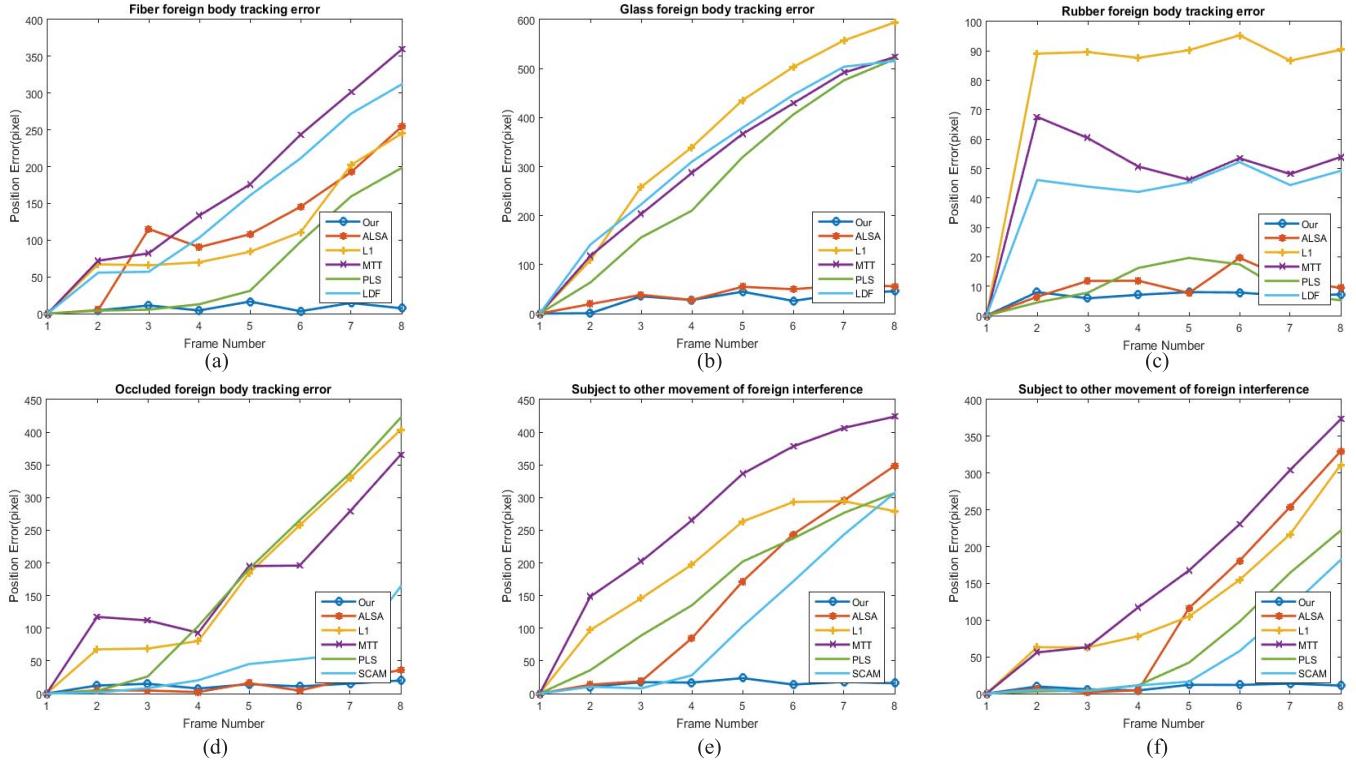


Fig. 14. Center location error plots for our tracker and the five state-of-the-art trackers. (a) Foreign 1 sequence of fiber foreign particle. (b) Foreign 2 sequence of skins foreign particle. (c) Foreign 3 sequence of rubber dander foreign particle. (d) Occlusion sequence. (e) Cross-interference sequence 1. (f) Cross-interference sequence 2.

The center position error is defined as

$$PosErr = \sqrt{(c_x^g - c_x^t)^2 + (c_y^g - c_y^t)^2} \quad (32)$$

where  $c_x^t$  and  $c_y^t$  are denoted as the tracking result of location coordinates, and  $c_x^g$  and  $c_y^g$  denote as the references tracking the result of coordinates.

The average location error is defined as

$$AvrErr = \left( \sum_{i=1}^n PosErr_i \right) / n \quad (33)$$

where  $n$  is the number of frames of the test sequence and  $PosErr_i$  is denoted as the center position error of the target in frame  $i$ .

Fig. 14 and Table IV show the center position error of tracking the fiber object, glass object, and rubber object, respectively, of which the center position error of ALW-CSM is in a smaller range all the time. In addition, Fig. 14(d) shows the center position error of tracking an object with occlusion, which makes any great difference in tracking the effect of ALW-CSM. At last, Fig. 14(e) show the center position error of the tracking object under interference, which is still very small due to the aid of the past motion trajectory.

For the pharmaceutical foreign particles in liquid infusion, in this case, tracking the target trajectory becomes challenging, especially considering the appearance of the target is not stationary under the industrial environment, e.g., significant pose varying, the variations of illumination and reflection,

TABLE IV  
AVERAGE CENTER LOCATION ERROR (IN PIXEL)

| Sequence  | Ours | ALSA  | L1    | PLS   | MTT   | LDF   | SCAM  |
|-----------|------|-------|-------|-------|-------|-------|-------|
| Foreign1  | 7.6  | 113.8 | 105.5 | 170.9 | 63.5  | 146.5 | -     |
| Foreign2  | 8.5  | 37.9  | 349.6 | 302.3 | 268.7 | 314.8 | -     |
| Foreign3  | 7.1  | 9.9   | 78.6  | 47.6  | 9.7   | 40.4  | -     |
| Occlusion | 11.8 | 11.9  | 173.9 | 169.8 | 168.7 | -     | 44.1  |
| Cross1    | 26.1 | 147.1 | 196.3 | 270.1 | 160.3 | -     | 108.9 |
| Cross2    | 8.7  | 112.5 | 169.7 | 68.8  | 131.3 | -     | 50.2  |

partial occlusion, motion blur, low contrast, and random noise. Since the proposed tracking method consists of SDC module and WSGM, for this specific application, we consider the particle motion trajectory prior to improving the adaptive sampling operator and updating per two frames to modify the particle appearance. On the other hand, we use the 2-D Fourier transform and weighed histogram to extremely measure the similarities between the candidate image and the target model. The experiment demonstrated the merits of the several accordingly improvements. The tracking results are presented in Fig. 14.

In most cases, as shown in Fig. 14 and Table IV, our algorithm performed better than the state-of-the-art methods. The traditional tracking method is less effective. The main reason is the SDC module used in the traditional tracking method that has poor performance in separating the foreground object from the background and is easily influenced

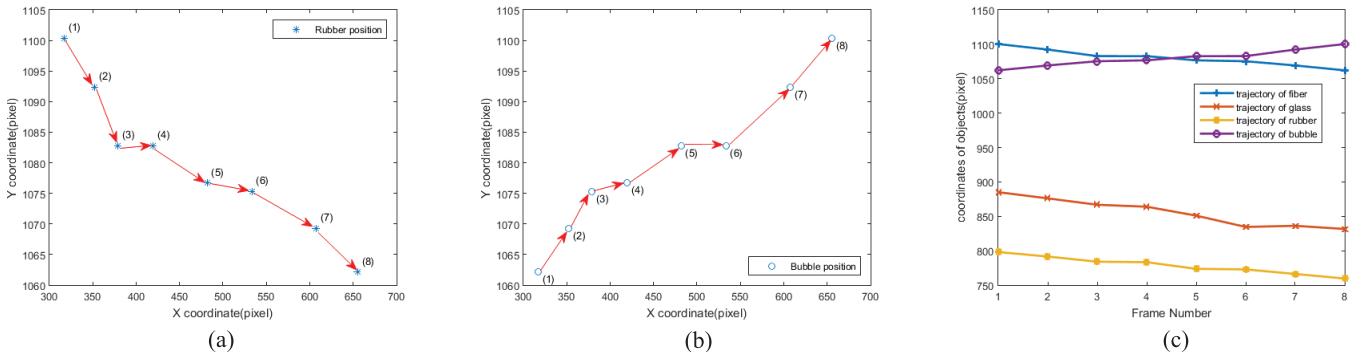


Fig. 15. (a) and (b) Experimental trajectories of rubber and air bubble. Serial numbers (1)–(8) represent the image frame number. (c) Trajectories of glass, random noise, fiber, and an air bubble constructed through connecting their coordinates frame by frame.

by heavy occlusions. In contrast, our local appearance model optimizes were designed to consider the occlusion; thus, WSGM performs well. In some cases (e.g., the Cross1 and Cross2 sequences), those sequences contain several objects with similar appearance shape to the target, which have low contrast to the background. The proposed tracking algorithm based on the improved weighted collaborative sparse model achieves good results.

#### E. Trajectory-Based Particle Recognition and Analysis

Eight captured sequential images of pharmaceutical infusion solutions, containing fiber, glass, and rubber foreign body particles are randomly selected as an example to evaluate recognition accuracy [49]. Besides foreign particles, air bubble and random noise also appeared in the selected sequential images. As shown in Fig. 15, each target is extracted and labeled with a red square after image preprocessing (i.e., image registration and positioning adjustment, background subtraction, object segmentation, and labeling suspected targets). On this basis, the tracking algorithm was used to identify the target in every image, and then to determine whether the multiple-suspected targets include bubble according to the abscissa of the target.

As is illustrated in Fig. 15, moving object trajectories in the sequence are constructed by connecting their coordinates frame by frame. The experimental trajectory of a foreign particle with larger density (e.g., fiber, glass, and rubber) is downward, while the air bubbles trajectory is upward. Under the rotation speed of 1200 r/min, particles can also be observed that the rotation radius of the target becomes smaller with the particles falling down or air bubbles floating up due to the pharmaceutical liquid viscosity. In the experiments described in Section V-D, particles trajectory is assumed to be consecutive and directional. Although different particle shapes can lead to a differences motion curve, the movement law of particle, air bubble, and random disturbance is used to recognize different types of impurities.

Several trajectories of particle and air bubble are experimentally proven effective for identifying particle, as shown in Fig. 15(a) and (b). In Fig. 15(c), the random noise disappeared in some images (see frames 4, 6, and 9) and its trajectory displays random fluctuations. Consequently, foreign

particles in pharmaceutical infusion solutions can be differentiated from air bubbles and random noise according to their distinct moving trajectory trends.

For our particle visual inspection machine, this paper used the proposed image processing approach, with several parts as shown in Fig. 4: 1) in the sequential image acquisition step, we used the industrial GigE CCD(BFLY-PGE-20E4C) which has a frame rate of 47 f/s, capture eight frames of each container need 170 ms in one image inspection station, where the tracking mechanism returned to start position 120 ms. Thus, the rate of our machine can achieve about 200 bottles/min (in this paper, we set the inspection speed 150 bottle per minute); 2) in the registration and deviation position adjustment, the running time for the back-light illumination station is 218 ms, the running time for the bottom-light illumination station is 220 ms; 3) in the image segmentation step using IFCNNs, the running time is 546 ms; and 4) in the particle tracking result and analysis step, we used our proposed ALW-CSM algorithm, where the running time is 5.32 s; thus, the total implement time costs is about 6.25 s for one drug container.

#### F. Knapp Test

The Knapp test [50], [51] is carried out to ensure that the manual inspection is similar to the automated inspection. The Knapp test designed by the U.S. FDA and European Pharmacopoeia is used for evaluating the performance of an inspection system [50]. Julius Knapp present a methodology to validate a semiautomated inspection method or fully automated inspection machines. In this section, after collecting  $X$  numbers of rejected bottles from various batches, these bottles can be blindly placed into  $Y$  number of uninspected bottles from various batches. Now with  $X+Y$ , have inspectors (let us consider five different inspectors) manually inspect these bottles 10 times each (5 inspectors times 10, times each bottle, is 50 inspections times each bottle). Using the Knapp probabilistic methodology, anything with a probability of 0.7 is classified as a reject.

Thus from the  $X+Y$ , any bottle that has been rejected 44 times (five technologists rejection times are 8, 10, 10, 6, and 10) out of 50 times would be rejected. Using these rejected bottles classified as bad ones and the bottles that

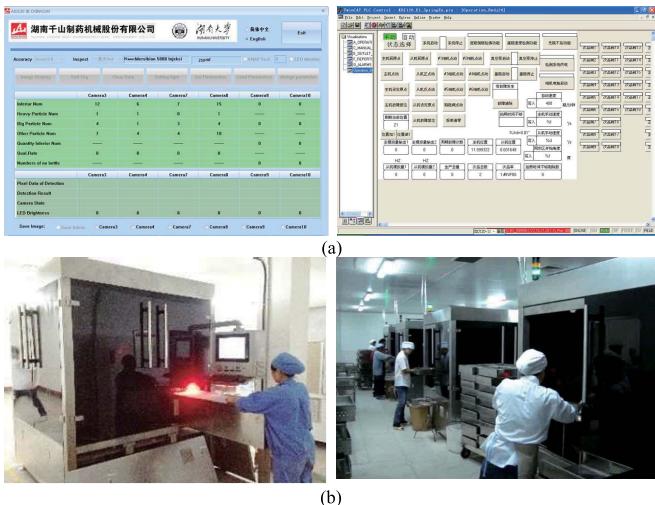


Fig. 16. Developed software user interface and visual inspection machine for pharmaceutical manufacture. (a) Software user interfaces for image capture and inspection. (b) Pharmaceutical manufacture application and test site.

are classified as good bottles (anything with a probability of 0.3 and less), a standard qualification set can be created. The quality factor of manual inspection for an  $i$ th bottle is  $FQA_i = (44/50) \times 10 = 8.8 \approx 9$ . According to the Knapp test protocol, only  $FQA_i > 7$  is added. The quality factor of automatic inspection machine for an  $i$ th bottle  $FQB_i$  is calculated using the same procedure. For 300 bottles of infusion, the sum over  $FQA_i$  and  $FQB_i$  ( $i = 1, \dots, 300$ ), if the ratio (i.e.,  $\sum FQB_i / \sum FQA_i$ ) is equal or greater than 1, it means that the automated inspection system is equivalent or better than the manual inspection method [51]. As shown in Fig. 16, our inspection machine for the pharmaceutical manufacturers on-site testing, quality factors of the inspection system and manual method are 1037 and 830, respectively. The ratio of 1.24 proves that the automated inspection system is superior to manual inspection.

To more quantitatively compare the inspection accuracy between the manual method and our automated system, four batches of 250 ml 10% glucose and 0.9% sodium chloride infusion solution are further tested. In each batch, 500 bottles containing glass, rubber, fiber, and hair (125 bottles with each type of particle) are randomly selected and mixed with 500 qualified bottles. Every container is labeled, then evaluated and confirmed by the commercial particle counter (e.g., Puluody PLD-0203, U.K.). Inspection accuracy is defined as

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (34)$$

where true positive (TP), true negative, false positive (FP), and false negative (FN) corresponded to the count of defective unqualified, defect-free qualified, defective qualified, defect-free unqualified products, respectively.

The detection results of flaws are listed in Table V. The accuracy of 10% glucose injection infusion solution is 96.3%, and the accuracy of 0.9% sodium chloride injection infusion solution is 98.2%. Thus the average accuracy is 97.25%.

TABLE V  
DETECTION RESULT OF FLAWS

| Medicine type | Particle type | Number of true positive | Number of false positive | accuracy |
|---------------|---------------|-------------------------|--------------------------|----------|
| 10%           | glass         | 11                      | 5                        | 96.30%   |
|               | rubber        | 3                       | 2                        |          |
|               | fiber         | 8                       | 3                        |          |
|               | hair          | 5                       | 0                        |          |
| 0.90%         | glass         | 8                       | 2                        | 98.20%   |
|               | rubber        | 1                       | 0                        |          |
|               | fiber         | 2                       | 1                        |          |
|               | hair          | 4                       | 0                        |          |

Based on our system, from a statistical point of view, two years application in the pharmaceutical production factories such as Xinjiang China Shidan Pharmaceutical Company Ltd., Jiangsu Dahong Eagle Hengshun Pharmaceutical Company Ltd., etc., the average inspection accuracy is above 97%. Moreover, the average inspection accuracy is higher than 90% for the manual inspection method. In addition, the standard deviation of automated inspection system is significantly lower than that of manual inspection, which indicated a higher precision or repeatability achieved by our automated inspection system. Scrutinizing the false detected situations revealed that misjudgment stemmed from random heavy lighting reflection, coincidental floating in the liquid level, and shape variations of foreign particles. For instance, in a few cases, the reflection spot of glass was not directly captured by the CCD camera. Therefore, the glass could disappear in some frames. Among fast rotation of the liquid, very few fiber or hair particles stay and stop by chance in regions such as the sidewall and neck or liquid level.

#### G. Comparison of the Proposed System and Other Existing Detection Systems

There are several other related inspection systems and methods, for example 1000-ml bottle infusion solutions inspection system [15], ampoule particle inspection systems with feature space clustering algorithm [52], frame difference and particle filter algorithm [53], online sequential extreme learning machine (OS-ELM) [54], wine production particle inspection system using shape recognition algorithm [55], 250-ml infusion transfusion single experimental detection device by morphology and mean-shift tracker [17], and unknown liquid particle detection system by fuzzy C-means method with fuzzy support vector machines [56]. The performances of some methods are greatly affected by the inspection electromechanical system architecture and the specific algorithm. For instance, the OS-ELM particle classification method requires feature extraction including area, shape, mean gray value, and statistical properties accurate. Due to different kinds of inspection objects, the comparison results of the other strategies are shown in Table VI.

According to Table VI, it can be observed that the proposed inspection system based on the ALW-CSM method

TABLE VI  
PERFORMANCE COMPARISON OF THE PROPOSED SYSTEM AND OTHER DETECTION SYSTEMS

| Inspection System     | Inspection object                           | Algorithm and description  | Accuracy |
|-----------------------|---|--|----------|
| Ishii A's system [12] | 1000ml infusion solutions with 500 samples  | Traces of foreign substances are obtained after minimum or maximum value selection   | 90.00%   |
| Lu G's system [49]    | ampoule with 200 samples                    | Moving-object clustering and frame difference  | 97.18%   |
| Xiao F's system [50]  | ampoule with 3500 samples                   | The target and background differences, particle filter algorithm   | 96.00%   |
| Ge J's system [51]    | ampoule with 1000 samples                   | Online sequential extreme learning machine (OS-ELM)  | 94.14%   |
| Wang Y's system [52]  | wine production with 200 samples            | The target and background differences, Shape recognition   | 96.60%   |
| Lu J's system [14]    | 250ml infusion transfusion with 180 samples | Grayscale dilation and image thresholding, mean-shift tracker. Single machine detection device   | 91.80%   |
| Liu H's system [53]   | unknown liquid production 400 samples       | Fuzzy C-means method with fuzzy support vector machines  | 96.80%   |
| Our system            | 250ml solutions with 500 samples            | Automatic inspection system with multi-carousel grips, adaptive local weighted-collaborative sparse model (ALW-CSM) tracking algorithm | 97.25%   |

outperformed other methods in most of the samples and the proposed method achieved the highest average accuracy in all the methods. This proved that the proposed online vision-based inspection system has a better performance when compared with those other methods. Due to a large production volume of medicine (production speed from 9000 to 46 000 bottles/per h), an improvement of 1% in accuracy could lead to a significant savings of the cost incurred by potential defects.

## VI. CONCLUSION

In this paper, we present automatic liquid particles inspect machine using cameras. The mechanical machine structure, multitype multistation light source, and image acquisition system are described in detail. Eight sequential images for each injection container are captured. A fast image registration and positioning adjustment method are utilized to compensate a mechanical error. IFCNNs are then used for precise segmentation and to get the exact initial tracking position. Moreover, the adaptive tracking method based on the adaptive local weighted-collaborative sparse model is proposed and applied to construct targets moving trajectory, which is used as a foreign particle judging criterion. These tracking algorithm and movement law are combined to recognize the true particles from the trajectory. The qualitative and quantitative results indicated that the method and automatic machine achieved high accuracy and repeatability in foreign particle detection.

After more than two years of application in the pharmaceutical production line, our system can be used for various kinds of medicine. Its accuracy and robustness are sufficient to satisfy the pharmaceutical industrial requirements. The proposed particle inspection system can be applied in the following aspects.

- 1) Pharmaceutical process monitor: the proposed system shows good detection accuracy in particle matter defects. If the categorized defective drugs for a high proportion of the total amount, it is likely that some problems may exist in the manufacturing process before the inspection. So the PVIM can be considered as a real-time monitoring system for the pharmaceutical.

- 2) Particle classification: the proposed method can be applied for the classification and recognition of the particle types by taking advantage of the multifeatures according to the nature of the application.
- 3) It can be extended to other pharmaceutical drugs packaging forms, such as ampoules, syringes, powder, and lyophilized drugs. In this kind of new application, how to adjust the parameters of the algorithm is worth studying.

## ACKNOWLEDGMENT

The authors would like to thank Hunan Chinasun Pharmaceutical Machinery Company, Ltd., Hunan, China, for inspection machine joint development and manual inspection by their certified technologists.

## REFERENCES

- [1] S. E. Langille, "Particulate matter in injectable drug products," *PDA J. Pharmaceutical Sci. Technol.*, vol. 67, no. 3, pp. 186–200, 2013.
- [2] *Chinese Pharmacopoeia: 2015 Edition*, Chinese Pharmacopoeia Commission, Beijing, China, 2015.
- [3] H. Grindinger, H. Neusser, N. Seidenader, and V. Wedershoven, "Product testing apparatus," U.S. Patent 0117149 A1, Jun. 2, 2005. [Online]. Available: <https://www.google.com/patents/US20050117149>
- [4] Y. Wang, J. Ge, H. Zhang, and B. Zhou, "Intelligent injection liquid particle inspection machine based on two-dimensional Tsallis entropy with modified pulse-coupled neural networks," *Eng. Appl. Artif. Intell.*, vol. 24, no. 4, pp. 625–637, 2011.
- [5] Z. Taoa, H. Mei-Zhena, L. Ming-Lianga, J. Jun-Bob, L. Tian-Yuanb, and L. Pei-Huab, "Signal processing methods for detecting impurities in ampoules," *J. Shanghai Jiaotong Univ.*, vol. 47, no. 6, pp. 998–1002, 2013.
- [6] S. Shirmohammadi and A. Ferrero, "Camera as the instrument: The rising trend of vision based measurement," *IEEE Instrum. Meas. Mag.*, vol. 17, no. 3, pp. 41–47, Jun. 2014.
- [7] Z. Zhou, Y. Wang, Q. M. J. Wu, C.-N. Yang, and X. Sun, "Effective and efficient global context verification for image copy detection," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 1, pp. 48–63, Jan. 2017.
- [8] H. Liu, J. Qin, F. Sun, and D. Guo, "Extreme kernel sparse learning for tactile object recognition," *IEEE Trans. Cybern.*, vol. 47, no. 12, pp. 4509–4520, Dec. 2017.
- [9] J. Wang, T. Li, Y.-Q. Shi, S. Lian, and J. Ye, "Forensics feature analysis in quaternion wavelet domain for distinguishing photographic images and computer graphics," *Multimedia Tools Appl.*, vol. 76, no. 22, pp. 23721–23737, 2017.

- [10] B. Gu, V. S. Sheng, K. Y. Tay, W. Romano, and S. Li, "Incremental support vector learning for ordinal regression," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 7, pp. 1403–1416, Jul. 2015.
- [11] C.-S. Cho, B.-M. Chung, and M.-J. Park, "Development of real-time vision-based fabric inspection system," *IEEE Trans. Ind. Electron.*, vol. 52, no. 4, pp. 1073–1079, Aug. 2005.
- [12] C.-T. Su, T. Yang, and C.-M. Ke, "A neural-network approach for semiconductor wafer post-sawing inspection," *IEEE Trans. Semicond. Manuf.*, vol. 15, no. 2, pp. 260–266, May 2002.
- [13] Q. Li and S. Ren, "A real-time visual inspection system for discrete surface defects of rail heads," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 8, pp. 2189–2199, Aug. 2012.
- [14] D. Zhang, Q. Li, Y. Chen, M. Cao, L. He, and B. Zhang, "An efficient and reliable coarse-to-fine approach for asphalt pavement crack detection," *Image Vis. Comput.*, vol. 57, pp. 130–146, Jun. 2017.
- [15] A. Ishii, T. Mizuta, and S. Todo, "Detection of foreign substances mixed in a plastic bottle of medicinal solution using real-time video image processing," in *Proc. 14th Int. Conf. Pattern Recognit.*, vol. 2, Aug. 1998, pp. 1646–1650.
- [16] F. Zhou, Z. Su, X. Chai, and L. Chen, "Detection of foreign matter in transfusion solution based on Gaussian background modeling and an optimized BP neural network," *Sensors*, vol. 14, no. 11, pp. 19945–19962, 2014.
- [17] J. Lu, Y.-N. Wang, J. Zhang, and B.-W. Zhou, "On-line detection of foreign substances in glass bottles filled with transfusion solution through computer vision," in *Proc. IEEE Int. Conf. Inf. Autom. (ICIA)*, Jun. 2008, pp. 424–429.
- [18] F. Duan, Y.-N. Wang, H.-J. Liu, and Y.-G. Li, "A machine vision inspector for beer bottle," *Eng. Appl. Artif. Intell.*, vol. 20, no. 7, pp. 1013–1021, 2007.
- [19] M. Carrasco, L. Pizarro, and D. Mery, "Image acquisition and automated inspection of wine bottlenecks by tracking in multiple views," in *Proc. 8th Int. Conf. Signal Process., Comput. Geometry Artif. Vis. (ISCGAV)*, 2008, pp. 82–89.
- [20] I. Szatmári, A. Schultz, C. Rekeczky, T. Kozek, T. Roska, and L. O. Chua, "Morphology and autowave metric on CNN applied to bubble-debris classification," *IEEE Trans. Neural Netw.*, vol. 11, no. 6, pp. 1385–1393, Nov. 2000.
- [21] T. Akilan, Q. M. J. Wu, and Y. Yang, "Fusion-based foreground enhancement for background subtraction using multivariate multi-model Gaussian distribution," *Inf. Sci.*, vols. 430–431, pp. 414–431, Mar. 2018.
- [22] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, pp. 977–1000, Oct. 2003.
- [23] M. Calonder, V. Lepetit, M. Ozysal, T. Trzcinski, C. Strecha, and P. Fua, "BRIEF: Computing a local binary descriptor very fast," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 7, pp. 1281–1298, Jul. 2012.
- [24] G. Vagliasindi, A. Murari, P. Arena, L. Fortuna, and G. Mazzitelli, "Cellular neural network algorithms for real-time image analysis in plasma fusion," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 8, pp. 2417–2425, Aug. 2009.
- [25] W. Shitong, K. F. L. Chung, and F. Duan, "Applying the improved fuzzy cellular neural network IFCNN to white blood cell detection," *Neurocomputing*, vol. 70, nos. 7–9, pp. 1348–1359, 2007.
- [26] S. Wang, D. Fu, M. Xu, and D. Hu, "Advanced fuzzy cellular neural network: Application to CT liver images," *Artif. Intell. Med.*, vol. 39, no. 1, pp. 65–77, 2007.
- [27] C. Xie, J. Tan, P. Chen, J. Zhang, and L. He, "Collaborative object tracking model with local sparse representation," *J. Vis. Communun. Image Represent.*, vol. 25, no. 2, pp. 423–434, 2014.
- [28] Y. Zhang, D. Lefebvre, and Q. Li, "Automatic detection of defects in tire radiographic images," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 3, pp. 1378–1386, Jul. 2017.
- [29] H. Feng, Z. Jiang, F. Xie, P. Yang, J. Shi, and L. Chen, "Automatic fastener classification and defect detection in vision-based railway inspection systems," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 4, pp. 877–888, Apr. 2014.
- [30] X. Tao, Z. Zhang, F. Zhang, and D. Xu, "A novel and effective surface flaw inspection instrument for large-aperture optical elements," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 9, pp. 2530–2540, Sep. 2015.
- [31] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *Computer Vision—ECCV*. Berlin, Germany: Springer, 2006, pp. 430–443.
- [32] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "BRIEF: Binary robust independent elementary features," in *Computer Vision—ECCV*. Berlin, Germany: Springer, 2010, pp. 778–792.
- [33] A. R. J. Fredo, R. S. Abilash, and C. S. Kumar, "Segmentation and analysis of damages in composite images using multi-level threshold methods and geometrical features," *Measurement*, vol. 100, pp. 270–278, Mar. 2017.
- [34] B. Porat and B. Friedlander, "A frequency domain algorithm for multiframe detection and estimation of dim targets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 4, pp. 398–401, Apr. 1990.
- [35] J. Xu, H.-W. Chang, S. Yang, and M. Wang, "Fast feature-based video stabilization without accumulative global motion estimation," *IEEE Trans. Consum. Electron.*, vol. 58, no. 3, pp. 993–999, Aug. 2012.
- [36] A. Jati *et al.*, "Automatic leukocyte nucleus segmentation by intuitionistic fuzzy divergence based thresholding," *Micron*, vol. 58, pp. 55–65, Mar. 2014.
- [37] A. Basset, J. Boulanger, J. Salamero, P. Bouthemy, and C. Kervrann, "Adaptive spot detection with optimal scale selection in fluorescence microscopy images," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4512–4527, Nov. 2015.
- [38] S. Agaian, M. Madhukar, and A. T. Chronopoulos, "Automated screening system for acute myelogenous leukemia detection in blood microscopic images," *IEEE Syst. J.*, vol. 8, no. 3, pp. 995–1004, Sep. 2014.
- [39] W. Shitong and W. Min, "A new detection algorithm (NDA) based on fuzzy cellular neural networks for white blood cell detection," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 5–10, Jan. 2006.
- [40] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A survey of sparse representation: Algorithms and applications," *IEEE Access*, vol. 3, no. 1, pp. 490–530, May 2015.
- [41] C. Li, X. Sun, X. Wang, L. Zhang, and J. Tang, "Grayscale-thermal object tracking via multitask laplacian sparse representation," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 4, pp. 673–681, Apr. 2017.
- [42] W. Zhong, H. Lu, and M.-H. Yang, "Robust object tracking via sparse collaborative appearance model," *IEEE Trans. Image Process.*, vol. 23, no. 5, pp. 2356–2368, May 2014.
- [43] Y. Xie, W. Zhang, C. Li, S. Lin, Y. Qu, and Y. Zhang, "Discriminative object tracking via sparse representation and online dictionary learning," *IEEE Trans. Cybern.*, vol. 44, no. 4, pp. 539–553, Apr. 2014.
- [44] X. Mei, H. Ling, Y. Wu, E. P. Blasch, and L. Bai, "Efficient minimum error bounded particle resampling L1 tracker with occlusion detection," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2661–2675, Jul. 2013.
- [45] J. Yang, R. Xu, J. Cui, and Z. Ding, "Robust visual tracking using adaptive local appearance model for smart transportation," *Multimedia Tools Appl.*, vol. 75, no. 24, pp. 17487–17500, 2016.
- [46] Q. Wang, F. Chen, W. Xu, and M.-H. Yang, "Object tracking via partial least squares analysis," *IEEE Trans. Image Process.*, vol. 21, no. 10, pp. 4454–4465, Oct. 2012.
- [47] T. Zhang, B. Ghanem, S. Liu, and N. Ahuja, "Robust visual tracking via structured multi-task sparse learning," *Int. J. Comput. Vis.*, vol. 101, no. 2, pp. 367–383, Jan. 2013.
- [48] W. Shi, M. Lin, W. Wang, and W. Jian, "Local distribution fields for tracking in complex scenarios," *J. Comput.-Aided Des. Comput. Graph.*, vol. 26, no. 11, pp. 2031–2038, 2014.
- [49] M. Feng, Y. Wang, and C. Wu, "Foreign particle inspection for infusion fluids via robust dictionary learning," in *Proc. 10th Int. Conf. Intell. Syst. Knowl. Eng. (ISKE)*, Nov. 2015, pp. 571–577.
- [50] J. Z. Knapp, "The scientific basis for visible particle inspection," *PDA J. Pharmaceutical Sci. Technol.*, vol. 53, no. 6, pp. 291–302, 1999.
- [51] J. Z. Knapp and H. K. Kushner, "Generalized methodology for evaluation of parental inspection procedures," *PDA J. Pharmaceutical Sci. Technol.*, vol. 34, no. 1, pp. 14–61, 1980.
- [52] G. Lu, Y. Zhou, Y. Yu, and S. Du, "A novel approach for foreign substances detection in injection using clustering and frame difference," *Sensors*, vol. 11, no. 10, pp. 9121–9135, 2011.
- [53] F. Xiao, Y. Wang, J. Zhang, and L. Cao, "Research on an on-line detection system for foreign substances in ampoule," *Comput. Meas. Control*, vol. 18, no. 2, pp. 295–298, 2010.
- [54] J. Ge *et al.*, "A system for automated detection of ampoule injection impurities," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 2, pp. 1119–1128, Apr. 2017.
- [55] Y. Wang, B. Zhou, H. Zhang, and J. Ge, "A vision-based intelligent inspector for wine production," *Int. J. Mach. Learn. Cybern.*, vol. 3, no. 3, pp. 193–203, Sep. 2012.
- [56] H.-J. Liu, "Development of intelligent liquid inspection system based on machine vision," in *Advances in Mechanical and Electronic Engineering*. Berlin, Germany: Springer, 2012, pp. 209–215.



**Hui Zhang** received the B.S., M.S., and Ph.D. degrees in pattern recognition and intelligent system from Hunan University, Changsha, China, in 2004, 2007, and 2012, respectively.

He is currently an Assistant Professor with the College of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha. He is also a Visiting Scholar with the CVSS Laboratory, Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada. His current research interests include machine vision, sparse representation, and visual tracking.



**Q. M. Jonathan Wu** (M'92–SM'09) received the Ph.D. degree in electrical engineering from the University of Wales, Swansea, U.K., in 1990.

In 1995, he was with the National Research Council of Canada, Ottawa, ON, Canada, where he was a Senior Research Officer and a Group Leader. He is currently a Professor with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON. He is also a Visiting Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. He has authored more than 300 peer-reviewed papers in computer vision, image processing, intelligent systems, robotics, and integrated microsystems. His current research interests include 3-D computer vision, active video object tracking and extraction, interactive multimedia, sensor analysis and fusion, and visual sensor networks.

Dr. Wu was the Tier 1 Canada Research Chair in Automotive Sensors and Information Systems. He is an Associate Editor of the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS and *Cognitive Computation*. He has served on technical program committees and international advisory committees for many prestigious conferences.



**Xuanlun Li** received the bachelor's degree from the Hunan University of Arts and Science, Changde, China, in 2010. He is currently pursuing the master's degree in control science and control engineering with the Changsha University of Science and Technology, Changsha, China.

His current research interests include machine vision, sparse representation, visual tracking, and so on.



**Hang Zhong** received the B.S. and M.S. degrees from Hunan University, Changsha, China, in 2012 and 2015, respectively, where he is currently pursuing the Ph.D. degree in control science and control engineering.

He is currently involved in computer vision, robotics, and related problems.



**Ji Ge** received the B.S. and M.S. degrees in mechatronics engineering from the Jiangxi University of Science and Technology, Ganzhou, China, in 2002 and 2005, respectively, and the Ph.D. degree in pattern recognition and intelligent system from Hunan University, Changsha, China, in 2012.

He is currently a Post-Doctoral Fellow with the Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, ON, Canada. His current research interests include machine learning and computer vision.



**Yimin Yang** (S'10–M'13) received the M.Sc. and Ph.D. degrees in electrical engineering from the College of Electrical and Information Engineering, Hunan University, Changsha, China, in 2009 and 2013, respectively.

He is currently a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada. He has authored or co-authored over 20 refereed papers. His current research interests include computer vision, robotics, and machine learning.



**Yaonan Wang** received the Ph.D. degree in electrical engineering from Hunan University, Changsha, China, in 1994.

From 1994 to 1995, he was a Post-Doctoral Research Fellow with the Normal University of Defense Technology, Changsha. From 2001 to 2004, he was a Visiting Professor with the University of Bremen, Bremen, Germany. Since 1995, he has been a Professor with the College of Electrical and Information Engineering, Hunan University, Changsha. His current research interests include robotics and image processing.

Dr. Wang was a Senior Humboldt Fellow in Germany, from 1998 to 2000.