An Efficient Parts Counting Method based on Intensity Distribution Analysis for Industrial Vision Systems

Qiaochu ZHAO  Ittetsu TANIGUCHI  Makoto NAKAMURA  Takao ONOYE
Osaka University  Osaka University  Laboratory of Hi-Think Corporation  Osaka University
Suita 565-0871, Japan  Suita 565-0871, Japan  Kyoto 600-8815, Japan  Suita 565-0871, Japan
zhao.qiaochu@ist.osaka-u.ac.jp  i-tanigu@ist.osaka-u.ac.jp  makoto@lhcc.co.jp  onoye@ist.osaka-u.ac.jp

Abstract—In this paper, we proposed an efficient parts counting method based on intensity distribution analysis for industrial vision system. Counting productions, as a preliminary operation in assemble line, is essential for calculating many industrial index such as deficiency rate. Conventional approach for counting problem is based on template matching, which we consider it as both stiff and time-consuming. In the proposed approach, counting problem is converted into an equivalent classification problem, in which a trained classifier is used to classify whether a specific line segment region belongs to parts or not. While parts flow through this line segment, number of the flowed parts can be effectively counted according to the interlace of different classified results. Experiments revealed that the proposed method superior to conventional template-matching method by being capable of counting with significant improvement of speed as well as with higher accuracy and stronger robustness. We also considered the proposed method can be readily extended to data with similar properties.

I. INTRODUCTION

With developments of manufacture techniques, significant improvements of performance as well as constantly lowering price have made sensors more and more easily available. Especially in recent years, with popularity of the concepts of Industry 4.0 and IoT, more and more sensors have been deployed in factories hoping for accelerating both automation and quality assurance. Industrial vision system, in particular, is playing an essential role in such process. It has been widely utilized to monitor the quality of products as well as collecting products’ information.

Modern industrial camera in vision system, aiming for capture products in real time, varies from civilian use camera mainly by its high speed, in some cases capture one frame for industrial camera can only cost about 2ms. However, this advantage also comes with a price — while conventional image processing method such as template matching or feature extraction has been fully explored in research field[2][3][4], they often become inapplicable in real industrial scene for its limitation on processing speed.

Specifically, counting products problem, in which products conveyed on high speed assembly line are needed to be counted in real time, is one of the most basic requirements in industry and plays an essential role to calculate some index such as non-defective ratio. The most common approach to solve counting problem can be considered as template matching[3], however, since in a plain version of template matching method, the template should be scanned over the entire area, it often turns out to be time-consuming as well as deficient accuracy for real time processing. Some researches have been reported to circumvent these drawbacks including fast template matching using Fourier transform, or rather than matching on the whole template, only matching on some extracted features[5] in favor of dimensionality reduction. Moreover, in recent years, with the development of GPGPU(general purpose computing on graphic processing unit), speed issue can also be considered alleviating through a hardware approach by allotting computation intensive job on GPU[6].

Despite speed requirement in industry is stricter than general occasion, properties of industry data are also represents to be diverge from ordinary image data mainly in the following three respects:

1. Color and texture simplicity: Unlike general image data like crowded people or scenery view which often contains complex color and texture information, image data in industrial field often appears to have simpler texture and relatively monotone color infor-
mation.

2. Streaming property: Due to manufacture operations are often lined up in factory, which means that objects will flow through the whole production line following a determined path, industry data always manifests a streaming property.

3. Individual homogeneous: Volume production in factories means that there will be only a slight variants among each productions, and depends on the situations, defective products can be significantly dissimilar to a normal one.

All these properties enable us to come up with more efficient counting methods for industry applications.

On the other hand, intensity distribution, as one of the most common statistical variable of an image, is often utilized to characterized images’ features. It has been applied to various field, such as HoG[2] for pedestrian detection, SIFT[4] for classification. In this paper, we also considered intensity distribution as a way of representing images’ content and further using it to resolve counting problem.

The rest of this paper is organized as following: as a background knowledge, mechanism of feeder machine and parts counting problem is introduced in section 2, conventional solution to this issue and its drawbacks are represented in section 3, details of our proposed method are explained in section 4 and experiment results of our proposed method and comparison results with conventional method are summarized in section 5, conclusions and future works are given in the last section.

II. MECHANISM OF FEEDER MACHINE AND PARTS COUNTING PROBLEM

A. An overview of feeder machine’s operation

Feeder machines are common devices used to transport individual component parts for assembly on industrial production lines. They are used when bulk package of small components must be fed into another machine one-by-one in correct direction. Fig. 1 illustrates an example of the vibratory bowl feeder, one kind of the feeder machine family, which is also the main object in this research.

A conceptual diagram of feeder machine is illustrated in Fig.2. Parts are first dumped into the bowl from the funnel. While dumped parts spin around in the "bowl", conveyor chute which is shaped to fit the part gradually shaken so that the parts are sent to the feeder’s conveyor one-by-one. Then, images of the dispatched parts are taken by high speed industrial cameras deployed over the conveyor to check each part. Specifically, size of the dispatched parts in our research are constrained about 2.8mm[1].

Although dispatched parts are required in a same orientation in this case, there is no guarantees about it, several parts in different orientations during checking are illustrated in Fig.4. Black color marks are attached to parts in order to show their orientations. From Fig.4 it can be inspected that either parts can be flipped along the transverse axis, or it is also very common that parts rotated from horizontal to vertical.
Parts in wrong orientations will be sent back to correct its directions, while the left parts sending to the next assembly procedure needed to be counted.

B. Counting numbers of parts and its challenges

Due to these special requirements of counting parts, two major challenges spontaneously raised as following.

1. High speed requirement: Since parts are passing through conveyor very rapidly (speed of conveyor is about 4m/min to 6m/min) and industrial camera also works in high speed, counting method should also be efficient enough to meet the speed.

2. Orientation and texture robustness: As it has been shown in previous, parts are not necessarily in a same orientation when counting, and depends on the illumination conditions as well as position of marks, textures of parts are also varying. Therefore, robustness for orientation and texture is also desired.

As it will be presented later, these challenges lead to the primary drawbacks of conventional template-matching method.

III. Conventional template-matching method and its drawbacks

Template matching method, in which objects are searched by matching the registered template over entire image region, can be intuitively applied in counting problem. An example of using template-matching for counting in our research is illustrated in Fig.5. Noted that matching process is only conducted in a constrained region in this case, for it is enough for counting purpose. It is also worth to notice that even though the matched part is slightly varied from the template in this scene, matching result can still be correctly obtained by properly setting the acceptance threshold.

While conventional template matching method retains robustness to some extent, it is not sufficient dealing with all cases. Orientations and textures switched from parts to parts can leads to failure matching. A specific examples of failing to match is given in Fig.6. It can be observed that this issue usually arise when vertical parts are continuously coming so that registered horizontal template cannot recognize both.

Another issue of template matching is, whereas, repetitive hit, which means that a same object matched in different frames, an example is given in Fig.7. This issue typically occurs when a single vertical part following another normal horizontal part in conveyor.

It is straight to circumvent miss-hit problem by registering more templates e.g. parts in vertical direction. However, countermeasure as such will also causes aggravation of repetitive hits as well as prolonging match processing time.

IV. The proposed method

Taking into account of the properties as described in section I, as well as drawbacks of the conventional template-matching method explained in section III. We proposed a novel counting method which is more accurate as well as time efficient.

A. Profiles of part and interstice of neighbored parts

Noted that parts’ streaming property means that parts are conveyed in a determined route over time. Therefore, two dimensional spatial information actually can be
effectively expressed in one dimensional spatial information with variations along time. Thereby, it is natural to consider that operation like template matching which conducted in two dimension spatial area is actually redundant.

Give a close observation of parts’ components in this case.

![Components of parts](image)

**Fig. 8. Components of parts**

As illustrated in Fig.8, part in our research is mainly composed by three components: two ends of part are made of metal, while ceramic constitutes the part’s torso, mark in black is used to indicate the orientation of parts.

Viewing parts in one dimension, we sampled several vertical lines in part as well as the interstice between two neighbored parts. Sampled lines in various regions are illustrated in Fig.9

![Sampled lines in parts and interstice regions](image)

**Fig. 9. Sampled lines in parts and interstice regions**

Correspondingly, intensity profiles of each sampled line is illustrated in Fig.10. Horizontal axis represents indexes of pixels, vertical axis represents intensities of the corresponding pixel.

It can be observed that in parts (a), intensities in ceramic region has consistent high value appearance, while in mark region, whose color is significantly darker than the ceramic region, a lopsided peak can be observed in the depicted profile. Profile in metal region has the most complicate appearance whose intensities is overall high with abundant variances. While in the case of vertical part as illustrated in (b), profiles actually is divided into three segments: highlight segment for top metal and ceramic region, dark segment for mark region, highlight segment again for bottom metal region. Sample line (v) represents interstice between two neighbored parts, it is obvious that profile has consistent low intensities appearance because of dark background. Situation where two parts which are interlapped to each other are excluded in this research for its rare incidence.

B. Convert counting into classification problem

Following the streaming property, motion of parts passing through conveyor can be effectively expressed as a one dimension line segment scanning over the entire parts with its profile varying. Therefore, counting number of parts’ problem is equivalent to counting number of switches between parts’ profile and parts interstice’s profile. Specifically, counter will increase itself by one when profiles belong to parts like Fig.9(a) or (b) is switched into non-part profile like (c), which means that one part has finished and another part will start.

C. Training classifier using neural network

In order to discriminate part’s profile from interstice’s profile, a trained classifier is necessary. As illustrated in Fig.10, although dissimilarities of interclass is remarkable, intraclass dissimilarities is also substantial. Therefore, a qualified classifier should be capable of discriminating different classes of obvious different appearances, meanwhile with enough tolerance for relatively large dissimilarity within a same class.

There are several alternatives when training a profile classifier, e.g. SVM[7], Random Forest[8] as well as Neural Network[9]. SVM is argued to be especially effective to small dataset, but it is also considered being sensitive to outliers contained in training data[10]. Because intraclass variance is also substantial in our research, effect of SVM is not promising. Another choice is Random Forest method, in which decisions are made by ensemble result of each decision tree. This approach is efficient as well as accuracy in training phase, but its speed is inversely proportional to the amount of trees[11] in inference phase, which made it unrealistic to train a high accuracy and efficient classifier for real-time application.

In this research, we adopted artificial neural network to obtain an efficient and high-accuracy classifier. By concatenating multiple layers with predecessor layer’s output as successor layer’s input, both flexibility and accuracy can be achieved by neural network. Neural network contains several types such as convolution neural network(CNN)[12], recurrent neural network(RNN)[13] for different purposes, while in our case, fully connected(FC) neural network is used since the input of network is one dimensional profile of intensities. Dimension of the input is fixed to the line segment’s length, and output dimension is two which represents the part-interstice binary classes.
V. EXPERIMENT

A. Experimental setups

Proposed method as well as its comparison experiment with conventional template matching method are verified on a same video of parts.

Video specification: Video with duration of 1'43" which contains 3000 frames is used as the experiment data. Although Original frame rate is 500 fps which means that each frame is taken in 2ms, in this experiment, frame rate is slowed down to 29 fps in purpose due to the consideration of visually verified the results of the proposed method easily.

In proposed method, a line segment in a fixed interval from [145,190] to [145,230] is put in every frame to sample the current profile of intensities. An example is illustrated in Fig. 11.

Fig. 11. Line segment of fixed position is posed to sample intensities distribution

We manually labeled the line segment in each frame. Number of each class (parts or interstice) is summarized in Table I.

<table>
<thead>
<tr>
<th>Classes</th>
<th>#Frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parts</td>
<td>2,545</td>
</tr>
<tr>
<td>Interstice</td>
<td>455</td>
</tr>
</tbody>
</table>

Among which, 400 samples of interstice frames are taken as negative training data, in consideration of the balancing of positive and negative samples, equal frames of parts are taken as positive samples. Temporally, testing is performed on the same video sequence as training dataset because of the conservation and limited availability of industrial data.

An architecture of three hidden-layer fully connected neural network is used for training, its settings are summarized in Table II.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>40 neurons</td>
</tr>
<tr>
<td>Hidden layers</td>
<td></td>
</tr>
<tr>
<td>First layer</td>
<td>800 neurons</td>
</tr>
<tr>
<td>Second layer</td>
<td>500 neurons</td>
</tr>
<tr>
<td>Third layer</td>
<td>100 neurons</td>
</tr>
<tr>
<td>Output</td>
<td>2 neurons</td>
</tr>
<tr>
<td>Learning rate</td>
<td>0.001</td>
</tr>
<tr>
<td>Batch size</td>
<td>100</td>
</tr>
<tr>
<td>Epochs</td>
<td>10</td>
</tr>
<tr>
<td>Optimizer</td>
<td>Adam</td>
</tr>
</tbody>
</table>

Trained model with the best outcome, which achieved an accuracy of 94.9% on training set, is picked from several trials.

B. Experiment results

We compared the proposed method with conventional template-matching based method, results are summarized in Table III. “TM” is short for conventional template matching method and “Proposed” represents our proposed method.

Noted that counting precision is different from the precision in training by it is the number of parts which are
TABLE III
Comparison of proposed method and template matching method

<table>
<thead>
<tr>
<th></th>
<th>Precision [%]</th>
<th>Time [ms]</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM</td>
<td>92.91</td>
<td>2.186</td>
</tr>
<tr>
<td>Proposed</td>
<td>97.72</td>
<td>0.276</td>
</tr>
</tbody>
</table>

The processing time is evaluated on a PC with 2.7 GHz Intel Core i5 CPU, template matching method costs about 2.186ms per frame, because each frame is taken shorter than 2ms by high speed camera, it is actually not applicable to count using template matching method in a real scene. Comparatively, the proposed method demonstrated both more efficiency and accuracy.

Deficiency of the current counting precision mainly comes from parts that are concatenate with each other, which will become unrecognizable for line segment classifier and is expected to be improved in our future work.

The proposed method achieved higher accuracy than template matching method attribute to its robustness of direction varieties.

VI. Conclusions and Future work

In this paper, properties of industrial data are fully explored and a novel counting method for industrial vision system is proposed. Experiments revealed that the proposed method can achieve higher accuracy as well as less time consumption than conventional template matching approach.

It is worth to be noted that the proposed method still classifying in a frame-by-frame manner and not yet taking temporal relationship among frames into consideration. So, our future work includes also taking temporal relationship among frames into consideration for further improving the accuracy of proposed method.

ACKNOWLEDGEMENT

We would like to express our gratitude for Daishin Co.,Ltd. The research could not be possible without the experiment data they provided as well as their precious advice.

REFERENCES