Proceedings of The 2016 IEEE International
Conference on Real-time Computing and Robotics
June 6-9, 2016, Angkor Wat, Cambodia

Defects Detecting of Gloves Based on Machine Vision

Xu Sun and Qijun Chen, *Senior Member, IEEE*

***Abstract*— This paper is focused on solving problems in production lines of gloves. To solve those problems we will use some image processing algorithm including contour extraction, color space changing and visual tracking. Using Kalman filter track gloves in production lines is the first step to process images. In industrial production, it makes request on the time of image processing. In order to process the image as soon as possible, we should improve the ratio of area that glove covers in the whole image , that is to say we should cut the image until it just contain the glove, so we used Canny edge detector to extract the contour of glove. With the same contour extraction method, we can also detect rip on gloves. As to the glove with oil, we changed color space from RGB to HSV to get a more accuracy detecting result.**

I. Introduction

Machine vision has be used in defects detecting for such a long time. From the very beginning, Machine vision successfully detected defects on the surface of steels. Up to now, defects detecting based on machine vision has plenty of applying examples in industrial production. However, products that are detected are always with fixed geometry in production lines, like steels and glasses [1], even fruits [2]. Gloves are so soft that their shapes are different from each other in production lines, which makes it a great challenge to do defects detecting on this kind of products and it can be seen as a new attempt in defects detecting with machine vision.

Defects detecting in production lines that produce gloves is done by human eyes now. It causes problems for both factory workers and enterprise. On one hand, because of materials and productive technology to produce gloves, the environment of production lines are not so health-friendly. On the other hand, workers have to work only a few hours and then have a long time to rest both for their health and detecting quality, so the enterprise has to employ a huge number of workers. To sum up, it is such a problem that needs to be solved in both technology and economy.

In this paper, in the respect of the particularity production lines of gloves, specific detecting algorithm should be used based on ordinary image processing algorithms. There are three main steps to do defects detecting in production lines of gloves: track gloves in production lines, cut the images until just contain one glove and detect defects.

Xu Sun is with College of Electronic and Information Engineering, Tongji University, Shanghai, China (e-mail: 1433353@ tongji.edu.cn).

Qijun Chen, is with College of Electronic and Information Engineering, Tongji University, Shanghai, China (e-mail: qjchen@tongji.edu.cn).

II. Visual Tracking

A. Background

Gloves in production lines have two kinds of motion modes: horizontal movement and rotation. In order to get images of one glove from every direction, there will not be only one glove in the view of a camera. With constraints of parameters of production lines, translational velocity and rotational speed, there are 6 to 8 gloves in the view of a camera (Figure 1). In this situation, tracking gloves is necessary.

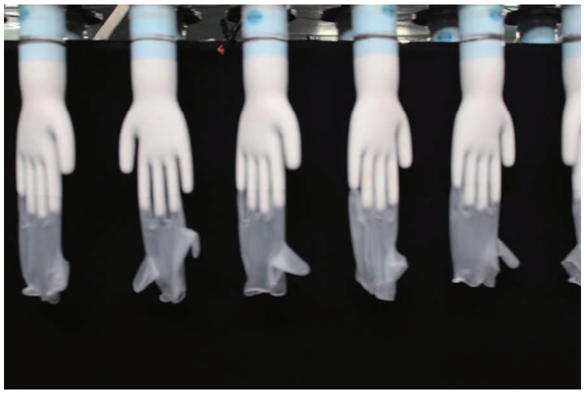

Figure 1. Image of production line catched by camera.

B. Kalman Filter

The Kalman filter is a set of mathematical equations that provides an efficient computational means to estimate the state of a process, in a way that minimizes the mean of squared error. The filter is very powerful in several aspects. It supports estimations of past, present and even future states.

Kalman filter estimates a process by using a feedback control. In other words, Kalman filter estimate a state at some time and then obtains feedback in the form of measurements [3]. In order to implement the feedback, there are two kinds of equations in Kalman filter: time update equations and measurement update equations. The time update equations:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} \tag{1}$$

$$P_k^- = AP_{k-1}A^T + Q \tag{2}$$

The matrix A relates the state at previous time step $k-1$ to the state at the current step k. The matrix B relates the optional control input u to the state x. The process noise covariance Q might change with each time step or measurement. The $\hat{x}_k^-$ is a prior estimate and the $\hat{x}_k$ is a posterior state estimate. The P_k^- is a prior estimate error covariance and P_k is a posterior estimate error covariance. Measurement update equations:

$$K_k = P_k^- H^T \left(H P_k^- H^T + R \right)^{-1} \quad (3)$$

$$\hat{x}_k = \hat{x}_k^- + K_k \left(z_k - H \hat{x}_k^- \right) \quad (4)$$

$$P_k = \left(I - K_k H \right) P_k^- \quad (5)$$

The K_k is Kalman gain. The matrix H relates the state to the measurement z_k .

Time update equations are responsible for projecting forward the current state and error covariance estimates to obtain a prior estimates for the next step and measurement update equations are responsible for the feedback.

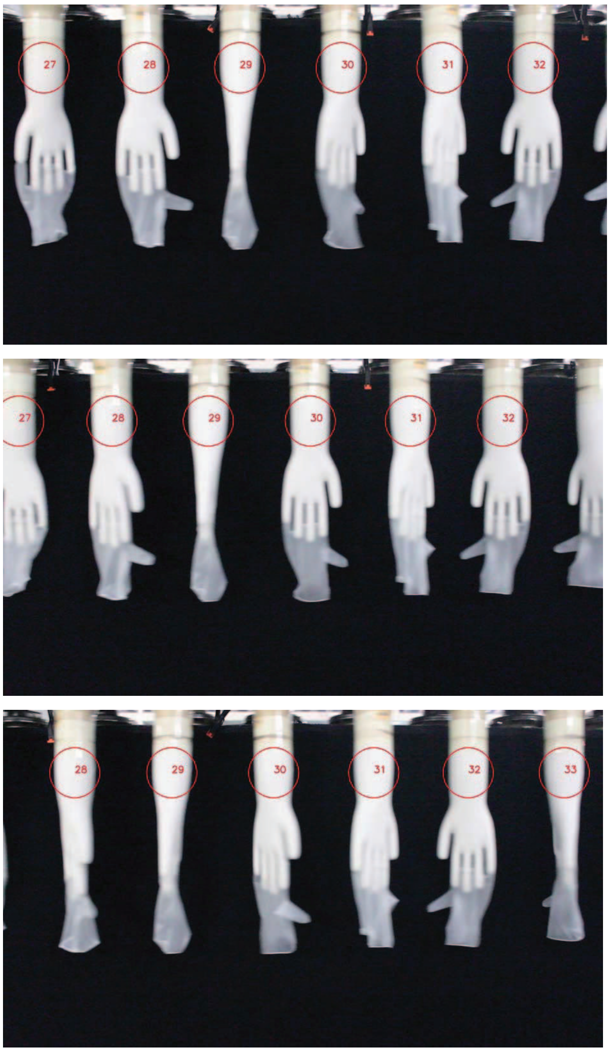

Figure 2. Kalman filter tracking gloves. *(each number corresponding to one glove to distinguish them and the center of circle s)*

C. Kalman Filter in Visual Tracking

Kalman filter have become a standard approach for reducing errors in a least squares sense and using measurements from different sources. Among its many applications, the Kalman filter is most essentially a part of vision development in robots. The purpose of the filter is to use visual measurements that contain noises and uncertainties captured by video cameras over time. Using a Kalman filter can produce values that tend to be closer to the true spatial measurements of targets [4][5][6]. In industrial producing environment, there exists many uncertain situations which will lead to noise in images that cameras captured. Using Kalman filter to track gloves can get a more accurate result, after all visual tracking is the first step in this application. In practice, each detected target is consist of one glove and one hand-model. The whole target approximately can be seen as a rectangle and, go a step further, using the center point of the rectangle stands for the tracking target (Figure 2). Generally speaking, we can use the glove position in current frame to predict the same glove's position in next frame by using Kalman filter, so we can get a more accurate result and resist unpredictable noise.

III. CUTTING IMAGES

After positions of each detected target have been determined, the whole image should be cut into 6 to 8 smaller images, number of smaller images is depend on number of tracked targets in the image, and each smaller image contains one target. In order to improve images processing efficiency, smaller images should be as small as possible, so the best smaller image is the one just contain a glove.

A. Canny Edge Detector

Canny suggested that edge detectors should optimize two specific performance criteria. First, the edge detector should have a good signal to noise ratio, in case image quality is poor. Second, edge detectors should localize the edges accurately, to support subsequent visual processes that need a high degree of positional accuracy. In addition, Canny also found it is necessary to constrain the smoothness of the edge detector [7]. The resulting constrained optimal filter was similar to a derivative of a gaussian.

Consider a linear filter $f(t)$ designed to detect an edge $g(x)$, located at $x=0$.

$$h(x) = \int_{-r}^{+r} g(x-t) f(t) dt + \int_{-r}^{+r} n(x-t) f(t) dt \quad (6)$$

Where $h(x)$ is the filter response and edges are marked by peaks in the filter response. The $n(x)$ is white noise. The filter is zero outside the interval $[-r, r]$. The $h(x)$ is consist of two parts. One is the filter responses to edge $h_g(x)$. The other is the filter responses to noise $h_n(x)$.

The signal to noise ratio of the filter is its response to the edge $h_g(0)$ divided by the r.m.s. response to noise $E[h_n(0)^2]^{1/2}$ and n_0 is the r.m.s. amplitude of the noise:

$$SNR(f) = \int_{-r}^{+r} g(-x) f(x) dx / n_0 \sqrt{\int_{-r}^{+r} f(x)^2 dx} \tag{7}$$

Canny defined the localization to be $E[x_{\max}^2]^{-1/2}$ where at some point $x_{\max}$ the response $h_g(x)$ has a peak. Close to the true location, the filter response can be approximated by a Taylor expansion:

$$\begin{aligned} h(x) &\approx h(0) + xh'(0) + x^2 h''(0)/2 \\ &= h(0) + xh_g'(0) + x^2 h_g''(0)/2 \\ &\quad + xh_n'(0) + x^2 h_n''(0)/2 \end{aligned} \tag{8}$$

Then do the derivative with respect to x and substituting $x_{\max}$

$$h'(x_{\max}) = h_n'(0) + x_{\max}(h_g''(0) + h_n''(0)) \tag{9}$$

When $x = x_{\max}$, $h'(x_{\max})$ is zero, then

$$x_{\max} = -h_n'(0) / (h_g''(0) + h_n''(0)) \tag{10}$$

Canny assumed $h_n''(0) = 0$ and the localization is

$$L_C(f) \approx \left| h_g''(0) \right| / E[h_n'(0)^2]^{1/2} \tag{11}$$

In order to optimize both $SNR(f)$ and $L_C(f)$, one way is to optimize their product, then we can get the optimal edge detector. However, there must has a constrain on the average distance between the noise peaks, in case noise peaks are in the vicinity of the edge [8][9]

$$Z(f) = \sqrt{\int_{-r}^{+r} f'(x)^2 dx / \int_{-r}^{+r} f''(x)^2 dx} \tag{12}$$

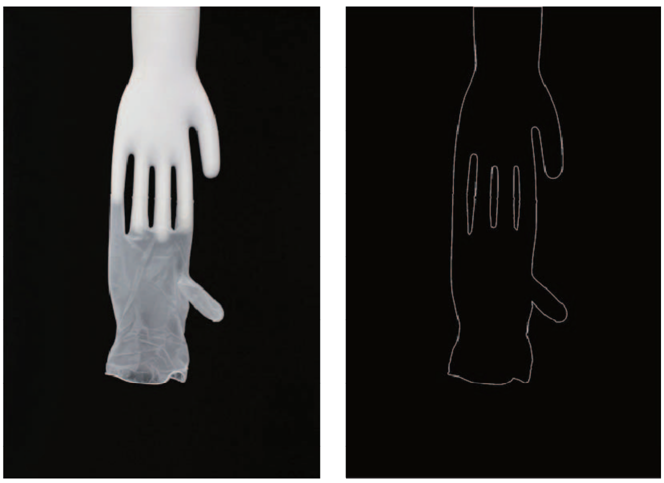

Figure 3. Canny edge detector apply to glove and hand-model.

B. Cutting Images Based on Contour

In this step, images will be further narrowed down. Canny edge detector algorithm has return a set of point of the target which is consist of glove and hand-model. Generally speaking, those points can form four closed graph. The largest one, called Contour A, is the outer contour of the target. The second largest one, called Contour B, is olive-shaped and it between the ring finger and little finger of hand-model, whose the lowest point is lower than any other two olive-shaped contours' lowest points in vertical. Use the vertical coordinate of this point as the upper base of a rectangle and use the vertical coordinate of the lowest point in Contour A as the lower base, then find the very left and very right point in Contour A and use their horizontal coordinates as two sides. Now a rectangle has been got and cut it out of the original image, then we get the image that we want (Figure 4). Next step we will start image processing in the image that we got.

Figure 4. Original image and image after cut.

IV. RIP ON GLOVES

After we have got images focused on glove itself, we begin to detect defects on gloves. The most common defect is rip on gloves. There is a procedure in production line that using fixture separate glove from hand-model. Because of error or happenstance, fixture may rip the glove. Rip defect is the main defect of production lines. A glove with rip will has another contour except the outer contour, called inner contour.

A. Dilate and Erode

Because, in the images we got, the background is black and the glove is white, it is a good way to do binaryzation in images in order to detect rip. In ideal condition, there will only two contours in the image. One is outer contour which is the contour of glove and the other is inner contour which is the contour of rip. Unfortunately the number of contours we get is always more than two, because images contain some noise points that come from dirty camera lens or dirty background. To restrain this, we have to do dilation and erosion.

Dilation is a method to delete some small noise points in images. It is especially appropriated to glove contour extraction situation. To do dilate, there will be a structural element. It may has different shapes, but here we choose rectangle. Dilation algorithm will scan every pixel in the image and do “and” operation between structural element and the pixels it covered. If all results are zero, then the value of this pixel is zero or it will be one. So if the size of noise point is only one pixel, after dilation, the noise point will disappear from the image.

As just said before, we can delete noise points in images. However, if the rip is so small that dilation will make the rip smaller which makes it hard to detect, we should use erosion algorithm. Similar with dilation algorithm, erosion algorithm also need a structural element with different shape. After some experiments using different structural element shape, we choose rectangle again. A little be different from dilation algorithm, erosion algorithm scan every pixel in image and also do “and” operation between structural element and the pixels it covered. But if all results are one, the value of this pixel is one or it will be zero. Now the image delete noise points and hold the size of rip in a reasonable scope.

B. *Detect Rip*

Rip detecting algorithm is based on contours. Generally speaking, if there is an inner contour in the outer contour of glove, rip can be found. But, after many experiments, we can find there are two kinds of rip which is depend on their positions. First, if the rip is close to the bottom of glove enough, the color in the rip is the same with background (left image of Figure 5). Second, if the rip is not so close to the bottom of glove, the color in the rip is gray (right image of Figure 5), because the color of glove can be seen as two layers of glove material overlay on the background, however, the color of rip can be seen as only one layer of glove material overlay on the background. So with a threshold in grayscale image, we can distinguish those two kinds of rip from image, actually we detect first kind of rip in this way. As to the second kind of rip, there exists some noise. Because of the random shape of glove in production lines or the angle of camera installed, the part near the bottom of gloves may similar to rip which also can be seen as one layer of glove material overlay on the background(Figure 6).

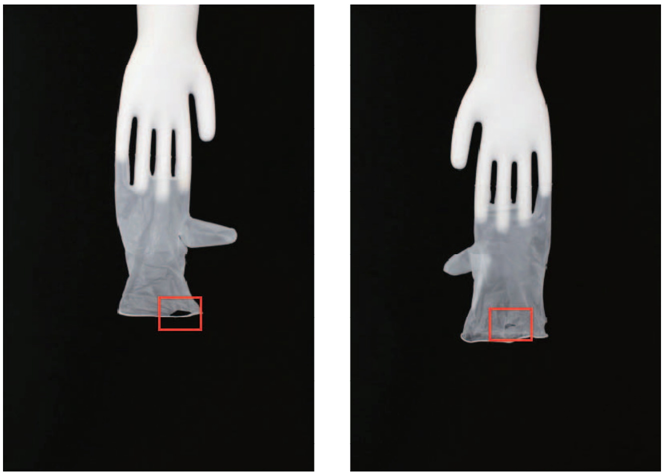

Figure 5. Two kinds of rip on gloves. *(left image shows the first kind of rip and second image shows the second kind of rip)*

Take that situation into account, we can estimate the contour of bottom part of glove (Figure 6), because if rip is in this region, the color in rip will the same with background; if not, we are sure this could be the bottom part. So if we find that there is another contour over the contour of bottom part of glove, it should be the second kind of rip.

Figure 6. Bottom part of glove that may be seen as rip. *(the part of glove in red rectangle has only one layer glove material)*

V. OIL ON GLOVES

In addition to rip, oil on gloves is another problem in defects detecting. Oil in the mechanical component of production lines and hand-model may pollute gloves. Oil on gloves are always light yellow and it is hard for human eyes to see it(left image of Figure 7), so the detecting method used now, detected by human eyes, can’t solve this problem.

A. *HSV Color Space*

HSV color space include three dimension: Hue, Saturation and Value. A three dimensional representation of the HSV color space is a hexacone, whose central vertical axis represents the intensity [10]. Hue is defined as an angle in the range $[0, 2\pi]$, specifically, red at angle 0, green at $2\pi/3$ and blue at $4\pi/3$. Saturation is the depth or purity of the color and it is measured by a radial distance from the central axis with value between 0 at the center to 1 at the outer surface in hexacone. Value stands for the brightness of color and if value equals to 0, it means that all the colors are approximated as black whatever be the Hue or the Saturation [11][12].

In practice, we always catch images from cameras with RGB color space, so we have to do the job to change color space from RGB to HSV:

$$h = \begin{cases} 0^o, & if \ max = min \\ 60^o \times \dfrac{g-b}{max - min} + 0^o, & if \ max = r \ and \ g \geq b \\ 60^o \times \dfrac{g-b}{max - min} + 360^o, & if \ max = r \ and \ g < b \\ 60^o \times \dfrac{b-r}{max - min} + 120^o, & if \ max = g \\ 60^o \times \dfrac{r-g}{max - min} + 240^o, & if \ max = b \end{cases} \quad (13)$$

$$s = \begin{cases} 0, & if \ max = 0 \\ \dfrac{max - min}{max} = 1 - \dfrac{min}{max}, & otherwise \end{cases} \quad (14)$$

$$v = max \quad (15)$$

Where “r”, “g” and “b” are got from original images, then change them into the region $[0,1]$. The “max” is the maximum value among “r”, “g” and “b”, correspondingly “min” is the minimum value.

B. Detecting Oil in HSV Color Space

In RGB color space, one way to detect oil on gloves is to divide color space. But the RGB value of oil point is always close to the RGB value of clean point in at least one of dimension of RGB, which makes it hard to distinguish oil point on gloves. In HSV color space, it can be considered that oil color and clean glove color are in different dimensions. In addition, oil on gloves is always light color, so with constrain on S dimension will improve the accuracy of detecting.

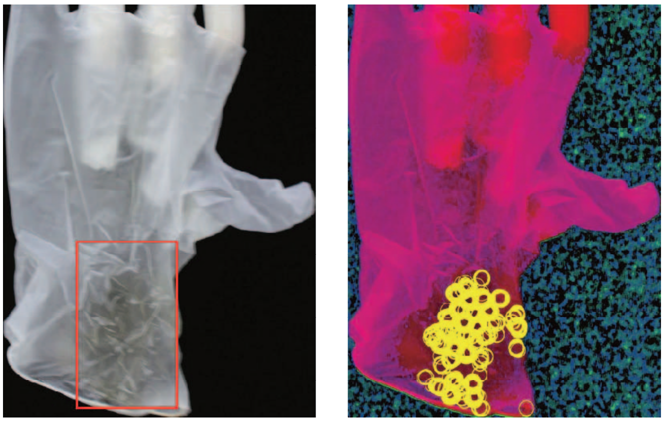

Figure 7. Change color space from RGB to HSV and find oiled pixel in HSV image. *(left image is in RGB color space, right image is in HSV color space and the center of each yellow circle is the oiled pixel found in HSV image)*

VI. CONCLUSION

We got 2126 samples with three types of gloves in laboratory (1026 qualified gloves, 568 gloves with rip and 532 gloves with oil) to do defects detecting.

TABLE I. DEFECTS DETECTION RESULT

Glove Type	Detected as		
	Qualified	***ripped***	***oiled***
Qualified	100%	0	0
ripped	4.25%	95.75%	0
oiled	3.68%	0	96.32%

As we can see, when input gloves with rip or gloves with oil, the algorithm may determine them to be qualified gloves. One Reason of this is that ripped area or oiled area on glove is too small that we take those pixels as noise point for granted. The other reason is that the threshold in HSV color space to distinguish oiled pixels and clean pixels is not well-adapted enough, when oil color changes in some way, it will ignore the oil pixels. We can use a dynamic threshold with algorithm that can decide the threshold automatically, but that will a little be time-consuming. We might compromise between efficiency and accuracy.

In practice, production lines rarely have defective gloves. If we change the data which is provided by the company into the experiment situation, miss ratio is nearly 10%, so the method we are now used is a kind of effective one.

ACKNOWLEDGMENT

Thanks my friends in my laboratory for helping me collecting samples of gloves and building experiment environment for defects detecting.

REFERENCES

[1] Peng X, Chen Y, Yu W, et al. An online defects inspection method for float glass fabrication based on machine vision[J]. International Journal of Advanced Manufacturing Technology, 2008, 39(11):1180-1189.
[2] Li J B, Huang W Q, Zhao C J. Machine vision technology for detecting the external defects of fruits — a review[J]. Imaging Science Journal the, 2015, 63(5):241-251.
[3] Welch G, Bishop G. An Introduction to the Kalman Filter[J]. University of North Carolina at Chapel Hill, 1995(7):127-132.
[4] Chen S Y. Kalman Filter for Robot Vision: A Survey[J]. IEEE Transactions on Industrial Electronics, 2012, 59(11):4409-4420.
[5] Funk N. A Study of the Kalman Filter applied to Visual Tracking[J]. University of Alberta, 2003.
[6] Ontiveros-Gallardo S E, Kober V. Objects tracking with adaptive correlation filters and Kalman filtering[C]// SPIE Optical Engineering + Applications. International Society for Optics and Photonics, 2015.
[7] Canny J ,. A computational approach to edge detection.[J]. Pattern Analysis & Machine Intelligence IEEE Transactions on, 1986, PAMI-8(6):679-698.
[8] Mcilhagga W. The Canny Edge Detector Revisited[J]. International Journal of Computer Vision, 2011, 91(3):251-261.
[9] Cope R K, Rockett P I. Efficacy of Gaussian smoothing in Canny edge detector[J]. Electronics Letters, 2000, 36(19):1615-1617.
[10] Stockman G, Shapiro L G. Computer Vision[M]// Computer vision /. Prentice Hall,, 2001:vii–viii.
[11] Sural S, Qian G, Pramanik S. Segmentation and histogram generation using the HSV color space for image retrieval[C]// In International Conference on Image Processing: P Viith Digital Image Computing: Techniques & Applications. 2002:589--592.
[12] Burdescu D D, Brezovan M, Ganea E, et al. A New Method for Segmentation of Images Represented in a HSV Color Space[M]// Advanced Concepts for Intelligent Vision Systems. Springer Berlin Heidelberg, 2009:606-617.