Vending Machine Based on Optical Flow Method

Xingxing Li[a], Chao Duan[b], Yan Zhi[c], Panpan Yin[d]

Department of Electronics and Information Engineering, Guangzhou College of Technology and Business, Foshan, 528100, China

[a]lixingxing_Edooy@163.com, [b]MDL0857a@126.com, [c]sparks666888@163.com, [d]abc0604@126.com

Keywords: Moving target tracking, Shi-Tomasi corner detection, Kanade-Lucas-Tomasi (KLT), Vending machine

Abstract: Moving target tracking plays a vital role in vending machines. This paper proposes a method of applying the optical flow method to a vending machine, which can detect and extract the corner features of moving objects and identify objects that track motion. Firstly, the image features in the video stream are extracted by Shi-Tomasi corner detection, and then the detected corner features are input to the Kanade-Lucas-Tomasi (KLT) algorithm to estimate the positional relationship of the feature points in the adjacent two frames. Finally, the method can determine the moving trajectory of the object according to the positional relationship, thereby determining whether there is a moving object in the vending machine, and providing the information to the upper computer for automatic judgment.

1. Introduction

Along with the modern quality of life, the vending machine has entered public life, a vending machine can automatically determine whether the shopping behavior. The vending machine is stationary in the normal state. When people use their hands to take the goods in the vending machine, the hands and articles will move. Therefore, the vending machine needs to automatically judge whether there are moving objects in it.

Optical flow method is a classical target tracking algorithm. It processes a continuous video frame sequence. For each video sequence, a certain target detection method is used to detect possible foreground targets. The advantage of optical flow method is that optical flow not only carries the motion information of moving objects, but also carries rich information about the three-dimensional structure of the scene [1]. It can detect moving objects without knowing any information of the scene. With the improvement of modern computer performance, the problems of time-consuming and poor real-time performance of optical flow method have been solved. Therefore, the algorithm can be applied to vending machines, so that the machine can automatically judge whether someone has taken out the goods.

2. Principle of optical flow method

2.1 Shi-Tomasi corner point measurement

The Shi-Tomasi algorithm is an improvement of the Harris algorithm. In the Harris algorithm, it is determined whether the corner point is based on the combination of two eigenvalues of the covariance matrix M [2]. In the Shi-Tomasi algorithm, it is determined whether the corner point is based on whether the smaller eigenvalue is greater than the threshold. This judgment is based on the fact that a smaller eigenvalue indicates that the variance in the direction of the eigenvalue is smaller, and if the smaller variance can be greater than the threshold, the change in this direction satisfies the judgment requirement of the corner point.

The covariance matrix M is expressed as follows: $\omega(x, y)$ represents a window function and is usually not used. I_x, I_y represent the difference in the x and y directions, respectively. $\mathrm{M} =$

$\sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$, solve the eigenvalues, generally solve such a λ, so that:

$$\det(\lambda I - A) = 0 \tag{1}$$

For a 2×2 matrix like A, we can break it down into:

$$\det\left(\lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - A\right) = \det\left(\begin{vmatrix} \lambda & 0 \\ 0 & \lambda \end{vmatrix} - \begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = 0 \tag{2}$$

$(\lambda - a)(\lambda - c) - bd = \lambda^2 - (a + c)\lambda + ac - bd = 0$. According to the root formula, there is

$$\lambda = \frac{(a+c)\pm\sqrt{(a+c)^2-4(ac-bd)}}{2} = \frac{(a+c)\pm\sqrt{(a-c)^2+4bd}}{2} \tag{3}$$

Bringing a specific M, the smaller eigenvalues are:

$$\lambda = \frac{(\sum I_x^2+\sum I_y^2)-\sqrt{(\sum I_x^2-\sum I_y^2)^2+4(\sum I_x I_y)^2}}{2} \tag{4}$$

ShiTomasi corner extraction is to obtain the strong corner point in the harris corner [3], that is, only the smaller one of the two eigenvalues is selected to construct the feature expression, if the smaller eigenvalue can satisfy the set threshold Condition, then the corner point is considered a strong corner point.

2.2 Kanade-Lucas-Tomasi (KLT)

KLT corner tracking algorithm is called Kanade-Lucas-Tomasi Tracking, also known as LK tracking algorithm. The algorithm assumes that the target is in the video stream, only produces a consistent small displacement, and the target's gray level does not change much. Then the algorithm must play a good effect on the premise that the following three assumptions are established [2]:

(1) The brightness is constant.

(2) The time is continuous or the motion displacement is small.

(3) Spatial consistency, adjacent points have similar motion and remain adjacent.

Satisfy hypothesis 1 to ensure that the target is not affected by the brightness; satisfy the hypothesis 2 to ensure that the feature points can correspond in the target domain; satisfy the hypothesis 3 to ensure that the displacements of all the points are the same in the same window [4]. Defining the same target appears in two frames of images I, J. If two points in the image match, then centering on these two points, W is the window with a very small gray-squared difference ε, defined as:

$$\epsilon = \iint_w [J(x) - I(x-d)^2]\omega(x)dx \tag{5}$$

where coordinate $x=[x, y]^T$ offset $d=[d_x, d_y]^T$, generally set the weight function ω(x)=1, and [J(x)-I(x-d)] in equation (1) is replaced by the symmetrical form [J(x+d/2)-I(x-d/2)], rewritten as:

$$\epsilon = \iint_w [J\left(x + \frac{d}{2}\right) - I\left(x - \frac{d}{2}\right)]^2 \omega(x)dx \tag{6}$$

In order to get the best match, let ε be the smallest, let the derivative of the above formula be 0, and find the minimum value. The above formula is used to perform Taylor series expansion at point $a=[a_x, a_y]^T$, and the high-order reserved linear term is truncated.

$$J(\varepsilon) \approx J(a) + (\varepsilon_x - \alpha_x)\frac{\partial J}{\partial x}(a) + (\varepsilon_y - \alpha_y)\frac{\partial J}{\partial y}(a) \tag{7}$$

where $\varepsilon=[\varepsilon_x, \varepsilon_y]^T$.

Using the formulas above, we bring x+d/2=ε and x=a into the formulas above, and x-d/2=ε and x=a into the formulas above, respectively.

$$J\left(x + \frac{d}{2}\right) \approx J(x) + \frac{d_x}{2}\frac{\partial J}{\partial x}(x) + \frac{d_y}{2}\frac{\partial J}{\partial y}(y) \tag{8}$$

So there is.

$$\frac{\partial \epsilon}{\partial d} = 2\iint_w \left[J\left(x+\frac{d}{2}\right) - I\left(x+\frac{d}{2}\right)\right]\left[\frac{\partial J\left(x+\frac{d}{2}\right)}{\partial d} - \frac{\partial J\left(x-\frac{d}{2}\right)}{\partial d}\right]\omega(x)dx \quad (9)$$

According to (6), there are approximately:

$$\frac{\partial \epsilon}{\partial d} \approx \iint [J(x) - I(x) + g^T d] g(x) w(x) dx \quad (10)$$

Among them are

$$g = \left[\frac{\partial}{\partial x}\left(\frac{I+J}{2}\right) \frac{\partial}{\partial y}\left(\frac{I+J}{2}\right)\right]^T \quad (11)$$

In order to find the offset d, set the formula (8) to be 0 and get the minimum value.

$$\frac{\partial \epsilon}{\partial d} \approx \iint [J(x) - I(x) + g^T d] g(x) w(x) dx = 0 \quad (12)$$

Equivalent to solving equation:

$$Zd = e \quad (13)$$

Z is a matrix of 2×2

$$Z = \iint g(x) g(x)^T w(x) dx \quad (14)$$

e is a vector of 2×1, which is the residue of calculation.

$$e = \iint [I_x - J_y] g(x) w(x) dx \quad (15)$$

When the residual error is less than a certain threshold, the approximate solution can be obtained. For the target in the image, KLT does not seek the offset of all points in the target frame and tracking frame, but chooses some corners with invariant features, and can use different corners with invariant features as tracking points, such as SIFT, SURF, FAST, SUSAN, HARRIS, etc.

Here shi-tomasi proposes a corner of Good Features. They believe that the symmetric matrix Z in formula (9) contains noise and good conditions. When the two eigenvalues of matrix Z are small, it means that the window centered on the point W is flat; two eigenvalues, large or small, correspond to undirected texture features; and two larger eigenvalues represent corner or salt-pepper texture. Therefore, when the two eigenvalues are larger than a certain threshold, this point is chosen as the corner point.

$$\min(\lambda_1 - \lambda_2) > \lambda \quad (16)$$

where lambda is the set threshold. Therefore, if the residual e obtained by equation (9) is small enough (set threshold), it is considered as a corner to be tracked, and the offset d of the corner is obtained.

2.3 Experiment

(1) Image preprocessing: gray image, the input color image into gray image.

(2) Corner detection: Shi-tomasi algorithm is used to detect the corner of the image, the effect is shown in Figure 2.

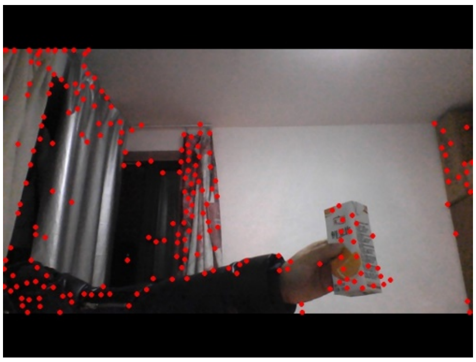

Figure 1. Shi-tomasi corner detection effect

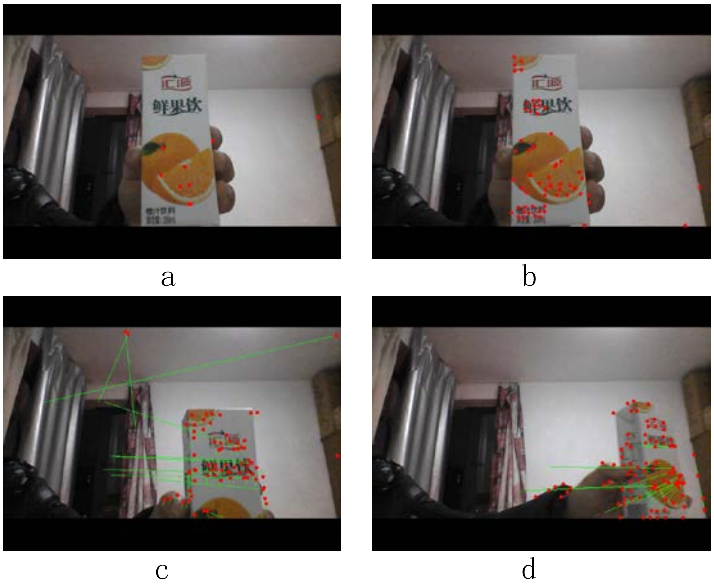

Figure 2. KLT algorithm tracking process

All corners in the image are detected by shi-tomasi corner detection, and they are regarded as feature points. If an object moves in the image, such as the hand and the beverage in the hand in Figure 3, these corners will follow and change. Therefore, these feature points need to be updated in the second frame image.

(3) Use Kanade Lucas Tomasi (klt) to track the corner features, the tracking process is shown in Figures 3a-3d. From the second time, some static corners will not be updated compared with the previous one. As shown in Figure 3a. In the house, the corner point of the wardrobe has been removed, only there are only a few corner points on the material. When the material is removed from the field of view, as shown in Fig. 3b, a large number of feature points begin to appear on the material shown by 3b. When moving to a certain extent, as shown in Fig. 3c, the calculation determines that the number of moving feature points is greater than a certain value. The value of the KLT calculation begins to track these feature points, resulting in a certain movement trajectory, as shown by the green line in Figure 3c. Eventually, when the hand and material are about to move out of the field of view, these feature points are continuously tracked as shown in Figure 3d.

3. Conclusion

The optical flow method is based on the three assumptions of illumination invariance, motion continuity and spatial consistency. The applicable scene of the vending machine just meets these three conditions. The vending machine has a fixed light source. The cargo movements are continuous, and the space is also consistent. Using the optical flow method, the shi-tomasi algorithm is used to detect the corner features in the image, and they are used as the feature points of the tracking. Then the KLT is used to track the features of the corner points, and the phenomenon of missing feature points will appear during the tracking process. By constantly updating the new feature points to keep track of the state, the motion trajectory of the object is drawn and an automatic judgment is made to the vending machine.

References

[1] Wu Y, Lim J, Yang M, et al. Online Object Tracking: A Benchmark [C]. Computer Vision and Pattern Recognition, 2013: 2411-2418.

[2] Lima M V, Kurka P R, Silva Y, et al. Indoor visual localization of a wheelchair using Shi-Tomasi and KLT [C]. Canadian Conference on Electrical and Computer Engineering, 2017: 1-4.

[3] Li H, Lu J, Shi G, et al. Tracking features in retinal images of adaptive optics confocal scanning laser ophthalmoscope using KLT-SIFT algorithm [J]. Biomedical Optics Express, 2010, 1(1): 31-40.

[4] Mstafa R J, Elleithy K M. A video steganography algorithm based on Kanade-Lucas-Tomasi tracking algorithm and error correcting codes [J]. Multimedia Tools and Applications, 2016, 75(17): 10311-10333.