ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 13, No. 2, 2018, pp.089-099

On a New Technique in Laser Spot Detection and Tracking using Optical Flow and Kalman Filtering

Yang Bai,[1, a] Xiuli Wang,[2,b] Wen Liu,[3,c] Ming Yang,[4,d]
[1] Chongqing University of Technology, Chongqing, China, 400054
[2] Anhui University, Hefei, Anhui, China, 230039
[3] Department of Mathematics, Lamar University, Beaumont, TX, US 77710
[4] Department of Computer Science Southern Illinois University Carbondale, IL, US,62901
a kgit@cqut.edu.cn, b wangxiuli@ahu.edu.cn, c wliu@lamar.edu, d ming.yang@siu.edu

(Received January 02 2018, accepted March 18 2018)

Abstract. Laser spot emitted by laser pointers were widely used in numerous fields such as pointing during presentations, guiding robots in human machine interaction (HCI), and guiding munitions to targets in military, etc. Our interest is to develop an effective method for laser spot detection and tracking with exploiting fewer features. In this paper, we jointly combined pyramid Lucas–Kanade (PLK) optical flow for detection and extended Kalman filter for accurately tracking laser spot in low-resolution and varying background video frames. Using this new technique, we only need to deploy intensity of laser spot with displacement of laser spot and get a very good result. This technique could be embedded into multi-core processing devices with inexpensive cost for potential high benefits in the future.

Keywords: spot detection and tracking, PLK optical flow, extended Kalman filter.

1. Introduction

Laser comes from the initials of Light Amplification by Stimulated Emission of Radiation. It has many remarkable features such as monochromatic, highly concentrated, highly bright, etc. It is well known that the laser beam is produced in limited range frequency, which means the wavelength distribution is very narrow. In other words, laser could be easily avoided by polluting other colorful noise. The high centralization characteristic makes it emit narrow beam with low divergence over long distances. And high intensity means high energy density within the injection space. Due to these excellent features, laser could be practically exploited in different fields.

Today, there are many different varieties of laser products for industrial and commercial purpose. Thus, we can easily choose the right laser equipment for guidance in HCI. In [13], the authors used the librarian robots to identify the position of a target relying on the laser point. And for accurate tracking in launching munitions to targets, the system could seek and lock aim point without operator intervention by exploiting a beam of laser energy, as in [14]. Other applications are about surveillance, mapping, remote sensing, etc. So, there are many interesting studies focused on laser. Our interest is to develop an attainable algorithm to detect and track laser spot precisely.

It is a challenge to detect laser spot since it is very tiny in shape. And the spot would deform to ellipse, strip, or other irregular shapes in movement. So, it is hard to detect the whole spot in general conditions. Instead, we just detect the center of spot. The laser equipment could emit red, blue, green or other possible colored beam as well. Besides, the luminance of laser spot depends on the output power and background intensity. For safety, the output power of laser equipment is restricted in most jurisdictions. It cannot be very high to cause the moving laser spot changing in brightness and tending to be unstable in noisy environment. So, it is difficult to extract general features or segment from all possible laser spots under real environmental background. That means, it is very easy to miss detecting laser spots or false detecting noisy points. In the tracking parts, the moving speed and direction of laser spot are unknown. Obviously, the movement is nonlinear which makes tracking more difficult. The spots may drown in noisy obstructions in frames or even just move out of frames. It is a big challenge to track those missing spots correctly.

For now, there are already different methods proposed for object detection and tracking. These methods have deficiencies and their performance are limited. In [7], they reported success rate is low when using template matching technique, because their technique is not suitable for tiny object tracking. While in [6], color segmentation and brightness features were applied in color frames that increase the complexity of computation and could not work well when meeting similar color noises. In [10], the authors reported good result, but still used template matching.

Our algorithm mainly addresses the problem under some conditions of weakening the difficulties. Grayscale frames are processed to get rid of color limitation. At present, we only consider locate one spot each frame and leave multi-spots location in the future. We assume that the intensity of laser spot is brighter than the local background and there exists a displacement difference between the laser spot and background. The proposed method makes use of the relative ratio of intensity between the spot and local background. Besides, we take into account the movement trajectory of laser spot to detect small spot in low-resolution images in time. The detailed algorithm will be described in section II. The experiment demonstrates that this new method is fast attainable and very effective.

The rest of this paper is organized as follows. In section II, we present our algorithm framework in detail. In section III, the Pyramid Lucas- Kanade (PLK) optical flow technique is introduced. Next, in section IV, we describe the extended Kalman filter (EKF). And we present the experimental results in section V. In the last, section VI is the conclusion.

2. Algorithm Framework

In this section, we would like to introduce our method in detail. Before dealing with data, the color video frames are converted to grayscale images and to double precision [0, 1]. First, we make two adjacent images (fpre, fnex) subtraction to get the difference frame and set all negative pixels to zero. This adjacent frame subtraction processing is for removing noise and most of the background. Then, we apply the image gradient technique to detect the candidates of laser spot in the difference frame. The gradient, denoted by ∇f , is defined as the vector g_x in horizontal and gy in vertical direction in frame f ,

$$\nabla f \equiv grad(f) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{bmatrix} \tag{1}$$

where $\partial f / \partial x$ and $\partial f / \partial y$ are partial derivatives of frame f , correspondingly. They could be obtained by kinds of approximation operators in digital image processing, like Sobel operators, Prewitt operators, etc see [3]. Then, after we got the derivatives, the magnitude of vector ∇f , $M(x, y)$, is computed by

$$M(x, y) = mag(\nabla f) = \sqrt{g_x^2 + g_y^2} \tag{2}$$

It yields the gradient image.

Next, we can binary the gradient image by choosing suitable threshold and use dilation to get the backup areas. Because the gradient just stresses the variation of intensity, when the intensity of some areas changes from bright to dark, the gradient is still large. For removing some dark areas, we also binarize the frame f_{pre} by using the high intensity as threshold. Then we compare the areas extracted from the binary frame f_{pre} and the backup areas from the gradient image of difference frame. Then the spot candidates will be picked up if they belong to the two areas.

Assume the spot could not be too tiny or too large, we only pick k regions with middle size as laser spot candidates. For each region R_k, the shape is different, so only the region center is computed as

$$x_{kc} = \frac{1}{m} \sum_{j=1}^{m} x_{kj} \tag{3}$$

$$y_{kc} = \frac{1}{m} \sum_{j=1}^{m} y_{kj} \tag{4}$$

where m is the number of pixels in each region (generally 5< m < 150) and (x_{kj}, y_{kj}) represents the position of each pixel in region R_k. The (x_{kc}, y_{kc}) is a candidate location of the laser spot center.

After all the candidates are captured, PLK are exploited to estimate serial new positions $\{(\tilde{x}_1, \tilde{y}_1), \ldots, (\tilde{x}_{kc}, \tilde{y}_{kc})\}$ of tracked candidates in the adjacent frame f_{next}. Thus, the displacement d_k could be obtained from the difference of the two positions in the two adjacent frames.

$$d_k = \sqrt{(x_{kc} - \tilde{x}_{kc})^2 + (y_{kc} - \tilde{y}_{kc})^2} \quad (5)$$

Then, we sort the displacement and get a set $D = \{d_1, d_2, ..., d_k\}$. Because our method is under the assumption of existing largest displacement difference between spot and background, the median d_m from D is chosen as background motion to calculate the displacement difference di_k,

$$di_k = abs(d_k - d_m) \quad (6)$$

We rank the candidates again with displacement difference and mark the ranked first (x_{1c}, y_{1c}) as the center of laser spot. As well, the corresponding $(\tilde{x}_{1c}, \tilde{y}_{1c})$ is merged in the candidates of adjacent frames. If there are more than one candidate with the highest displacement difference and gradients, we mark the laser spot missed in the frame and remain this problem in tracking parts. This technique could detect the laser spot no matter if it moves faster than background or fixes in stationary. But this technique may not be efficient when laser spot has speed similar or equal to background's.

After the position of laser spot are detected, the EKF will be applied for position fine tune and missing spot tracking. The whole processing structure is shown in Figure 1.

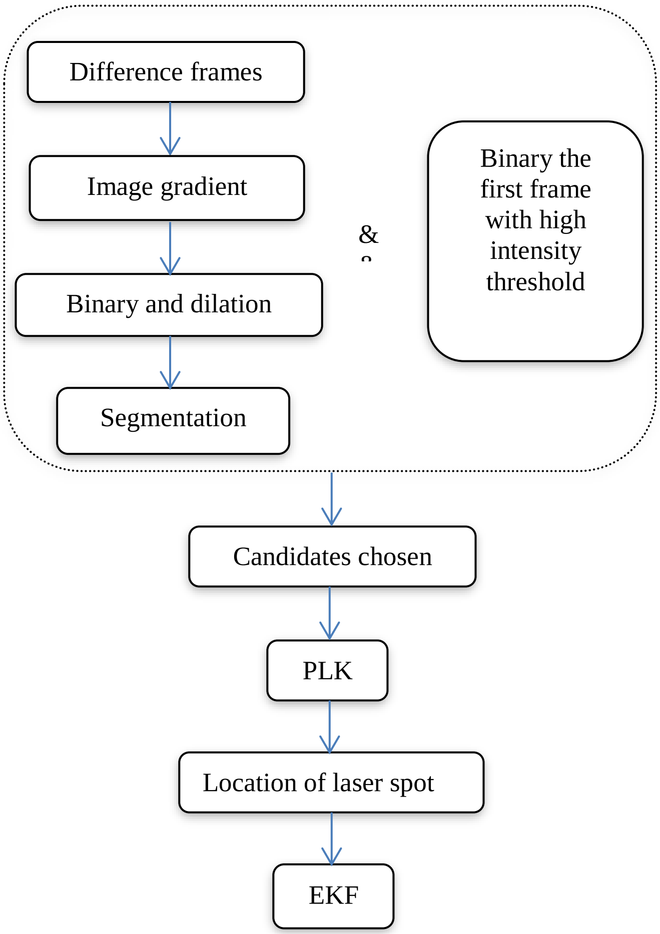


Fig (1) the whole structure of laser spot detection and tracking

3. PLK optical flow

Optical flow is the distribution of instantaneous velocity of apparent motion observed in visual system. It expresses the variation of brightness in images that contain important spatial information of relative motions. So, it could be used to determine the movement of targets. Optical flow is wildly applied in many fields like object detection and tracking, robot navigation, background extraction and separation, etc. Traditional optical flow computation methods for motions are differential techniques like Horn-Schunck (HS) global algorithm, Lucas-Kanade (LK) local algorithm [2][8], region-based matching [4][8], energy-based methods[2][8], phase-based techniques, etc [2][8]. They all have their own advantages and disadvantages. To improve the performance, many methods are introduced in [1][11]. PLK is famous for quickly computing sparse optical flow. For our own situation, only the candidates which composite the sparse feature matrix need optical flow to be estimated, so we could use PLK technique to tackle it.

The computation of optical flow is based on two assumptions. First, the luminance does not change or change very slightly over time between adjacent frames. In other words, the brightness is constrained by constancy:

$$I(x+u, y+v, t+1) \approx I(x, y, t) \tag{7}$$

where the flow velocity (u, v) is defined as optic flow at frame t. Second, the movement is very small, which means that the displacement does not change rapidly. So using Taylor expansion, we have

$$I(x+u, y+v, t+1) = I(x, y, t) + I_x u + I_y v + I_t + H.O.T \tag{8}$$

For small displacement, after omitting the higher order terms (H.O.T), we have:

$$I_x u + I_y v + I_t = 0 \tag{9}$$

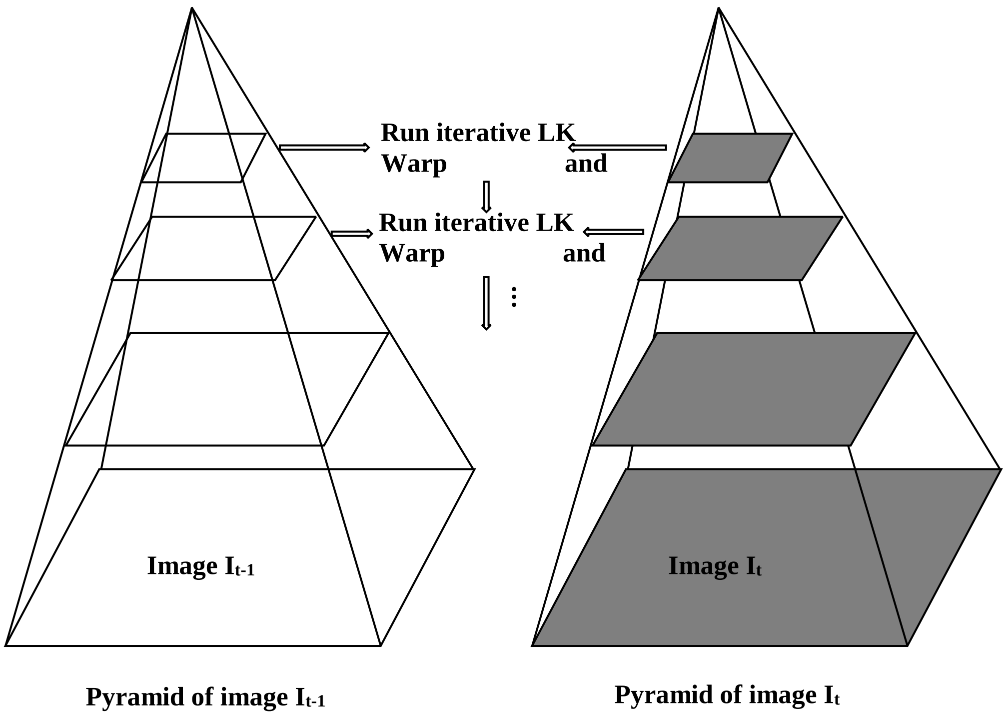
where the subscripts denote partial derivations. Just (9) is not enough for us to compute the two unknown velocity components (u, v) uniquely. To solve this problem, we should introduce another extra assumpt-ion to constrain the equations. By assuming that the moving pixels displace in a small neighborhood, we could use the LK algorithm, which is a local differential technique with a weighted least square fit to mi-nimize the squared error function

$$E_{LK}(u, v) = \sum_{p \in \Omega} W^2(p) * (I_x u + I_y v + I_t)^2 \tag{10}$$

Here $W(p)$ represent window weight function for pixel (x, y) associate with spatial neighborhood Ω of size p. Then we can easily compute the motion of the central pixel using the neighboring spatial and temporal gradients $\partial_u E_{LK}=0$ and $\partial_v E_{LK}=0$. This sets up a linear system.

$$\begin{pmatrix} W(p)^2 * (I_x^2) & W(p)^2 * (I_x I_y) \\ W(p)^2 * (I_x I_y) & W(p)^2 * (I_y^2) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -W(p)^2 * (I_x I_t) \\ -W(p)^2 * (I_y I_t) \end{pmatrix} \tag{11}$$

This system is not only easy and fast for computation but also robust under noise. But in more general situation, large and non-coherent displacements are typical. To ensure the last constraint, we recommend PLK method to circumvent this assumption. The detail is described in [1]. That means the motion could be estimated using an image pyramid over the largest spatial scales at the top layer and then the initial motion velocity could be iteratively refined down layer by layer to the raw image pixels at the bottom layer. In short, it is a coarse to fine optical flow. The structure is shown as follows.



Fig(2) coarse-to- fine optical flow estimation

From Fig (2), we can see that the optical flow run over the top layer to estimate the motions. Then for the next each layer, the resulting estimation from the previous up layer are used as the starting points for continue estimation until reach the bottom layer. While combining with the image pyramids, it could be implemented to estimate faster motions from image data sequences, as well, reduce lots of the computational cost. So, it is very suitable for our case.

4. EKF

A Kalman filter (KF) is a highly effective recursive data processing algorithm that could produce an optimal estimation of unknown variables underlying system state. It is widely applied for navigation, tracking and estimation because of its simplicity, optimality and robustness.

Let us define a linear dynamic system as follows,

$$x_{k+1} = Ax_k + w_k \tag{12}$$

$$y_k = Hx_k + v_k \tag{13}$$

where x_k is the state variable at time k, y_k is the measurement at time k, A is the state transition matrix and H is the observation matrix in linear dynamic system, w_k and v_k are system noise, respectively.

The basic Kalman filter works on two steps in linear system. In the prediction step, the Kalman filter gives an uncertain estimation of the current state variables using the formula

$$\hat{x}_k^- = A\hat{x}_{k-1} \tag{14}$$

$$P_k^- = AP_{k-1}A^T + Q \tag{15}$$

Where x_k^- is the prediction of the state variables at time k, x_{k-1} means the estimate of the state variables at time $k-1$, P_k^- represents the prediction of error covariance at time k, P_{k-1} is the estimate of the error covariance at time $k-1$, Q is the covariance matrix of wk.

In the correction process, the estimates are updated by the compensation of the difference between measurement from the system and the previous prediction.

$$K_k = P_k^- H^T \left(H P_k^- H^T + R \right)^{-1} \tag{16}$$

$$P_k = P_k^- - K_k H P_k^- \tag{17}$$

$$\hat{x}_k = \hat{x}_k^- + K_k \left(y_k - H \hat{x}_k^- \right) \tag{18}$$

where K_k is called Kalman gain, R is the covariance matrix of v_k, x_k means the updated estimates of state variables.

From the mathematical model above, we can see the Kalman filter works on linear dynamic system very well. But it is difficult to apply on nonlinear systems in practical engineering applications. In our experiment, with the initial position, instant velocity, acceleration and direction all unknown, we only wave the laser pointer aimlessly. So, the trajectory of laser spot is obviously nonlinear. Therefore, the traditional KF should be adjusted to fit for the nonlinear situation.

Many kinds of extension and generalization algorithms for Kalman filter are already developed to mitigate the effects of nonlinearities. The EKF could simply approximate a linear model for the nonlinear dynamic system and exploits the Kalman filter theory at each state [5][9][12]. In the general nonlinear system model, the nonlinear functions replace the parameters (A, H) of linear system.

$$x_{k+1} = f(x_k) + w_k \tag{19}$$

$$y_k = h(x_k) + v_k \tag{20}$$

So, in the prediction and correction steps, the EKF estimates the state variables by

$$\hat{x}_k^- = f(\hat{x}_{k-1}) \tag{21}$$

$$\hat{x}_k = \hat{x}_k^- + K_k \left(y_k - h(\hat{x}_k^-) \right) \tag{22}$$

Also, instead of A and H in the two steps, the Jacobian matrix of nonlinear function $f(\bullet)$ and $h(\bullet)$ are computed below,

$$A \equiv \frac{\partial f}{\partial x}\bigg|_{\hat{x}_k} \tag{23}$$

$$H \equiv \frac{\partial h}{\partial x}\bigg|_{\hat{x}_k^-} \tag{24}$$

where $\partial / \partial x$ denotes partial derivatives with x. When the consecutive linearization approximates smoothly, the EKF could converge well. And the EKF works recursively, so we only need the last estimate state rather than the whole process. And the computation consumption is affordable, which is very suitable for our case.

5. Experiments

In this section, we demonstrate the performance of the technique using live video frames. A web camera is used to record some short videos including one laser spot moving in different backgrounds at recording speed 25 frames per sec. Each frame has 720×1280 pixels and a few unstable frames at the beginning and the last parts are discarded, 300 frames are picked up in each video. The spot is projected by a 640-660 nm red laser pointer with a tiny output power (<1 mW). The video is recorded from 1 to 3-meter distance of the spot to the camera. The size of spots becomes bigger when zooming in the camera distance, otherwise, it become smaller. Those backgrounds have some noisy strips or brightly patch regions in the similar shapes of laser spot as obstacles. Fig (3) shows video frame samples about laser spot moving in grass, floor, bricks and wall background up to down, as below:

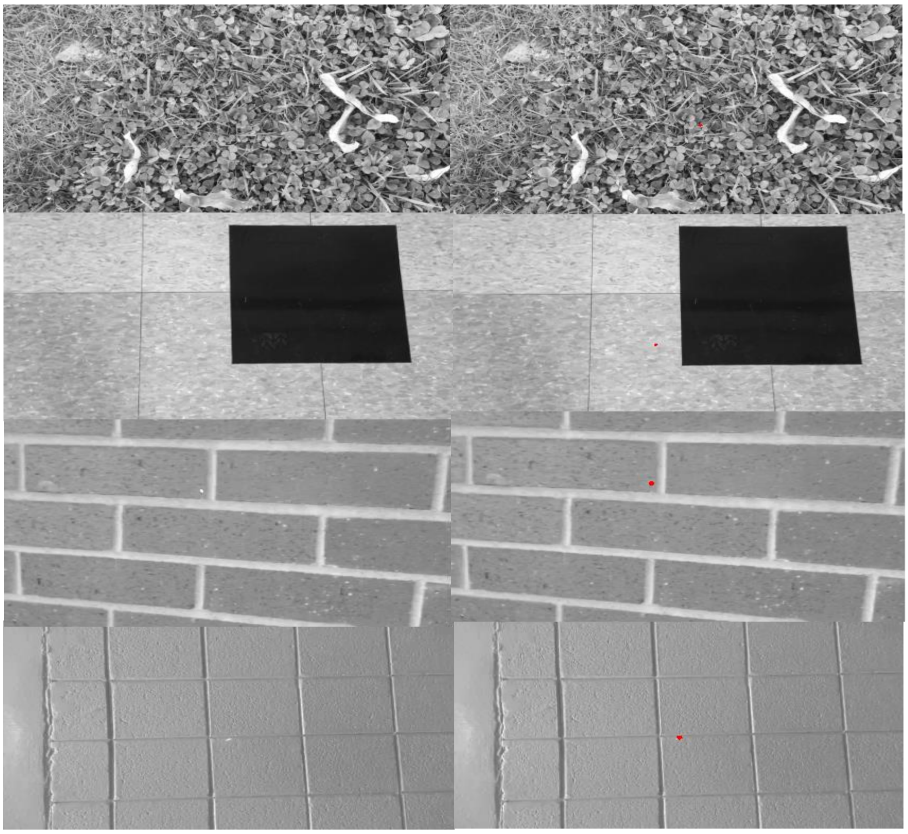


Fig (3) the laser spot with different background sample

In the left lane, there are original samples with the background from noisy to noiseless. In the right lane, we mark the laser spot position by red dot manually. With different background light condition and distance, the size of laser spot also changes. For convenience, we magnify the spot by cutting a size of 31*31 local windows in Fig (4). (up left: grass, up right: floor, down left: bricks, down right, wall). Obviously, the size of spot in grass is the smallest, and the size of spot in bricks is the largest. The shapes of spots in grass and wall are more strip and the shapes of spots in floor and bricks are more circular. Also, the spot in bricks and wall are brighter than in grass and floor.

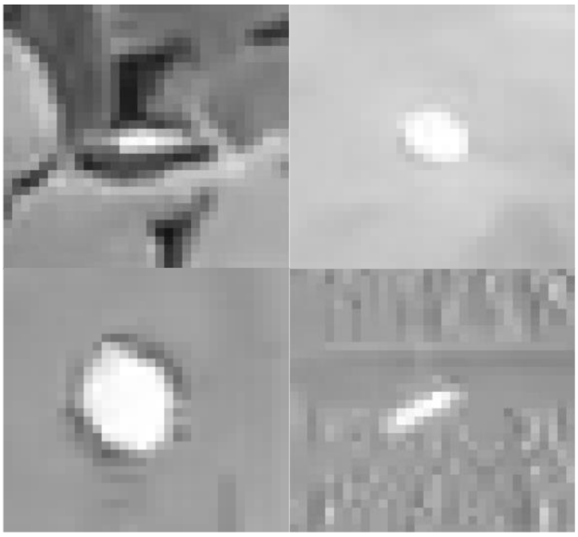


Fig (4) the laser spot with local window in different background

To determine the influence of brightness and size for detecting the spot, we calculate the ratio of average intensity of local widow to the laser spot (ratio A), the ratio of average intensity of the whole background to the laser spot (ratio B), the size of laser spot in each video. Table I gives the range of ratios and size of laser spot.

Table 1: the range of ratios and size of laser spot

	grass	floor	bricks	wall
Ratio A	0.67~0.88	0.35~0.87	0.61~0.78	0.62~0.82
Ratio B	0.68~0.78	0.58~0.64	0.62~0.68	0.61~0.66
Number of laser spot pixels	6~10	20~35	45~120	10~20

From the experiment, we find that the ratio A has more contribution than ratio B. The intensity contrast of laser spot and local background is the key factor for detection. Generally, when ratio A and ratio B are higher, the size of the spot should be larger to detect the laser spot. Also, the high intensity ratio and small size of spot increase the difficulty for detection. If the ratio A is greater than 0.85 or the number of spot pixels are less than 8, it is almost impossible to detect the laser spot. Fig 5 shows the relationship of the ratio A and the number of laser spot pixels.

We also add Gaussian noise artificially onto the video frames to evaluate the effects of noise for detection. While the variance of Gaussian noise changes from 0.001 to 0.1, the difficulty of detection increases correspondingly. It is easier to detect when the ratio A is low and the size of spot pixels is large. If the variance is greater than 0.1, the spot is completely submerged in noise and could not be detected any more. It also clearly shows in Fig 5.

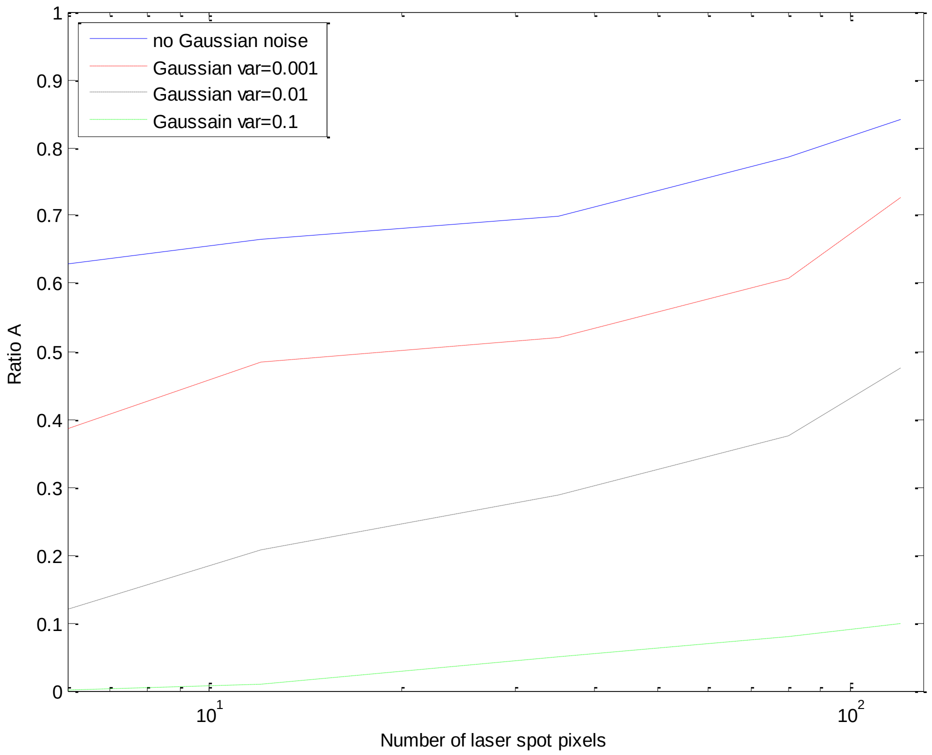


Fig (5) relationship between the ratio A and number of laser spot pixels

We use PLK to calculate the optical flow of each spot candidate in two adjacent frames. Then, we can compute the displacement difference between the two adjacent frames and pick up the unique candidate with the largest displacement difference as spot. If the number of candidates with the largest displacement is bigger than 1, then we think the spot is missing in current frame and left this problem in tracking part. Fig (6) shows the detected spot marked by red dot. For convenience, we only cut a 300*200 window and magnify the window to display.

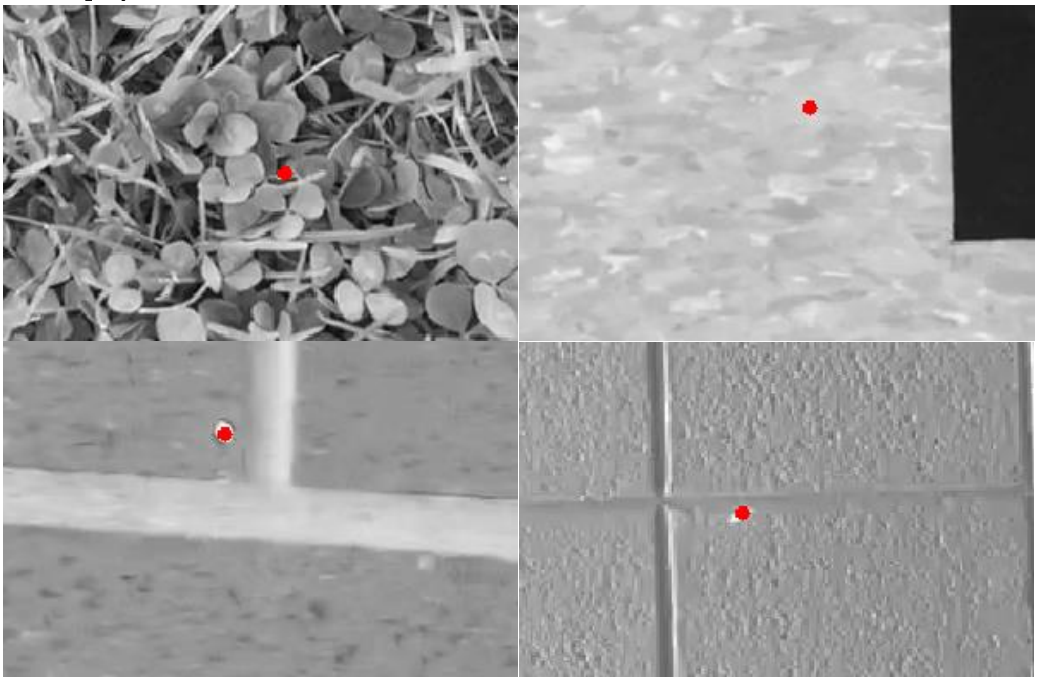


Fig (6) the detection samples of laser spot by PLK samples

After capturing the initial measurement, if the number of detected laser spot is more than one in a frame, we marked that frame missing spot. Then EKF is exploited to fine adjust and tracking the missing spot. Samples shown in Fig 7 indicate that the tracking is successful. For convenience, we only cut a 120*100 window to display. The red dot is remarked as detection by PLK and the blue circle is denoted as EKF tracking. But when the wrong point is far from the real laser spot detected, the EKF cannot work well on the lock of the right position.

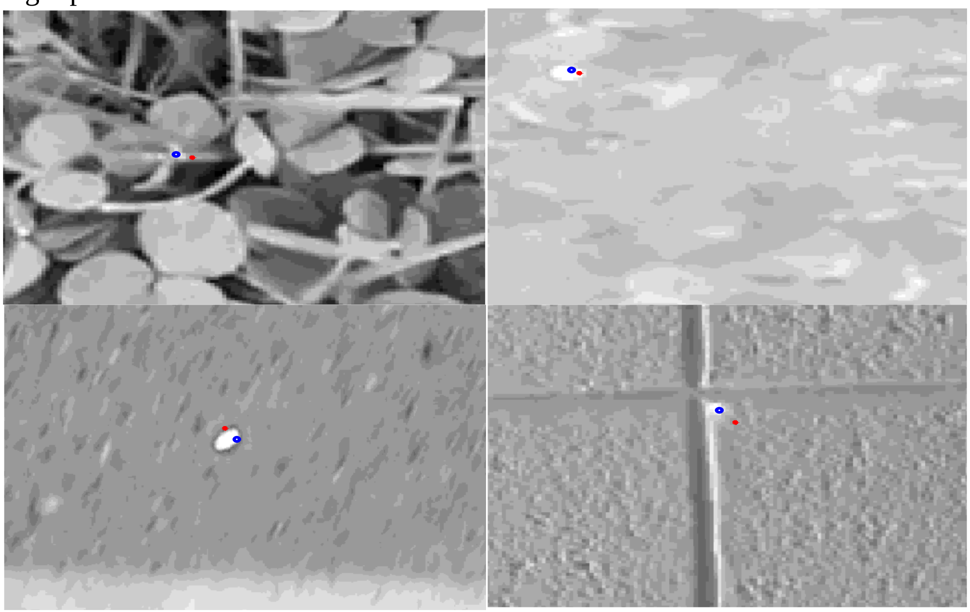


Fig (7) tracking sample

For the 4 videos, the accuracy of detection and tracking is shown in Table II .

Table 2： the accuracy of detection and tracking

	grass	floor	bricks	wall
accuracy	40.3%	81.7%	93.3%	92.1%

From these results, the accuracy of video with bricks background is the best, because the spot is large, and the obstacle is rare. Conversely, the accuracy of video with grass is the worst, because the spot is very small and weak compared to the background.

6. Conclusion

This technique is attainable in laser spot fast detection and tracking. Now, we could also work on single spot case with color, size and shape invariant. It is much easier to detect and track when the laser spot is brighter, bigger or moving faster. When the background is noisy, the processing time will become longer. For further improvements, we would like to do some optimization and to solve multiple spots case, when it will be much more difficult to determine the candidates and suitable threshold values. And it is a challenge to solve the missing and false detection problem.

7. References

[1] Brox, Thomas, et al. "High accuracy optical flow estimation based on a theory for warping." Computer Vision-ECCV 2004. Springer Berlin Heidelberg, 2004. 25-36.

[2] Bruhn, Andrés, Joachim Weickert, and Christoph Schnörr. "Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods." International Journal of Computer Vision 61.3 (2005): 211-231.

[3] Gonzalez, Rafael C. Digital image processing. Pearson Education India, 2009.

[4] Brox, Thomas, and Jitendra Malik. "Large displacement optical flow: descriptor matching in variational motion estimation." Pattern Analysis and Machine Intelligence, IEEE Transactions on 33.3 (2011): 500-513..

[5] Grewal, Mohinder S., and Angus P. Andrews. Kalman filtering: theory and practice using MATLAB. John Wiley & Sons, 2011.

[6] Meško, Matej, and Štefan Toth. "Laser spot detection." Journal of Information, Control and Management Systems 11.1 (2013).

[7] Chávez, F., et al. "Automatic laser pointer detection algorithm for environment control device systems based on template matching and genetic tuning of fuzzy rule-based systems." International Journal of Computational Intelligence Systems 5.2 (2012): 368-386.

[8] Beauchemin, Steven S., and John L. Barron. "The computation of optical flow." ACM Computing Surveys (CSUR) 27.3 (1995): 433-466.

[9] Ribeiro, Maria Isabel. "Kalman and extended kalman filters: Concept, derivation and properties." Institute for Systems and Robotics (2004): 43.

[10] Sun, Yanpeng, Bo Wang, and Yingshuo Zhang. "A Method of Detection and Tracking for Laser Spot." Measuring Technology and Mechatronics Automation in Electrical Engineering. Springer New York, 2012. 43-49.

[11] Bouguet, Jean-Yves. "Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm." Intel Corporation 5 (2001): 1-10.

[12] Julier, Simon J., and Jeffrey K. Uhlmann. "New extension of the Kalman filter to nonlinear systems." AeroSense'97. International Society for Optics and Photonics, 1997.

[13] Mikawa, Masahiko, Yuya Morimoto, and Kazuyo Tanaka. "Guidance method using laser pointer and gestures for

librarian robot." RO-MAN, 2010 IEEE. IEEE, 2010.

[14] Livingston, Peter M. "Target-tracking laser designation." U.S. Patent No. 5,973,309. 26 Oct. 1999