IMPLEMENTATION OF A PROGRESSIVE TARGET RECOGNITION SYSTEM FOR ALL THE ANGLES

LI Juan，MA Huimin *，ZHOU Luyao

*Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
Tel: 86-10-62781432 E-mail: mhmpub@tsinghua.edu.cn

Keywords: the scan image; 3D target; FPGA; progressive recognition system.

Abstract

3D target recognition is a popular research field in the laser near-field detection, where the scan image of the target is obtained line by line. The arbitrary spatial position between the detector and the target, along with the interference caused by the clouds, make it difficult to realize accurate real-time identification for the detector. This paper presents a method to generate gray-scale plane images under any viewpoint by calculating the echo signals from the surface of 3D planes. Besides, considering both scan imaging and signal processing in hardware, a local feature based on improved and simplified SURF algorithm is implemented in hardware to describe the target. Then a progressive 3D target recognition system is built on the FPGA, and the experiments performed on the Modelsim simulation software show that the system can reach a high recognition rate.

1 Introduction

The scan imaging is an important method to acquire target information for the laser detector [1]. On the one hand, it is relatively easy to access the target image line by line, which makes it suitable for the pattern of progressive local recognition, namely, the laser detector performs the target recognition while obtaining the scan image. On the other hand, it is a big challenge for the hardware to generate the scan image, especially when the high-speed detectors try to get the push-broom image. In addition, choosing an appropriate local feature to describe the line scan image can greatly contribute to the improvement of target recognition accuracy. Therefore, it is of great practical value to study the simulation of the method of the laser scanning imaging and the description of the target characteristics.

In the field of high-speed probe, Yu *et al* [2] propose a method to acquire the binary image, which is convenient to implement in hardware and attains a high processing speed. But arbitrary intersection angles and positions make a serious loss of target information in the binary image. As shown in the left side of Figure 1, the binary image cannot express a distorted target appropriately, and it is hard to extract target feature from such a binary image. Besides, shape features of the blob-like cloud and the target are similar, so it is easy to arouse false recognition as shown in the right side of Figure 1. Kou and Liu [3] give a method to get the gray-scale image in the infrared detection field. This method achieves the local recognition, but it adopts the passive imaging system, so its false recognition rate is relatively high. Generating gray-scale image in the laser detection not only overcomes the lack of information in the binary image, but also employs the active detection. Therefore, it can greatly improve the local recognition rate for the scan image. Besides, Liu *et al* [4] employ the global feature to describe the target, and its false rate turns out to be relatively high, so a local improved SURF feature can reduces the recognition rate obviously.

Fig.1 The binary image with the laser scan

In the laser imaging system, the intensity of the echo information reflects the gray information of the target [5]. Therefore, it is feasible to get the target reflection intensity information by reasonable quantification of the echo signal. Thus we can obtain the intensity image of target different from that of the background and clouds [1]. The intersection process between the detector and the target is quite short and there is only one intersection process. Therefore, it is necessary to adopt a progressive target image recognition method, that is, once N lines of image data are obtained, the recognition is performed. Taking into account the characteristics of the laser scan imaging and the realizability in hardware, we use the simplified SURF algorithm [6] and implement it on the FPGA to extract the feature of the target for a laser linear scan image. The feature of the target has good invariance in the translation, rotation and scale [7]. Besides, it maintains a certain degree of stability in visual changes and affine transformation. So it can better express the gray changes characteristics of the target.

2 Gray-scale Image Simulation

According to the principle of the laser probe, the echo signal can be approximatively calculated [8]. The dynamic range of its value depends on the laser detection range. First, the statistical maximum and minimum values of the echo signal are computed. Then, the target and background can be separated by the predetermined threshold. Finally, the echo signal which is above the threshold is quantified to form the gray-scale image. Figure 2 shows the gray-scale image obtained in the simulation platform.

Fig.2 A gray-scale image simulated with the echo power

However, it is difficult to implement the qualifiation in the hardware platform. First, the n-channel input analog signal needs to be converted into the digital signal by an AD converter. In accordance with 12-bit conversion accuracy, 12n input ports in the chip are required. When the line array element n is very large, it is impossible to realize so many ports in hardware. Therefore, the power changes of the echo signal are simulated in the size of the width changes. In fact, the value of the laser transimitter beam follows a Gaussian distribution [5]. After the receiver performs photoelectric conversion and amplification, the distribution of the amplified voltage waveform varies in the value of the echo power. If the echo power is in the linear region of the receiver amplifier, the amplified voltage waveform continues to show a Gaussian distribution. And if it falls in the saturated zone, the top of the waveform would be cut down [9]. But the width of waveform is still increasing as shown in Figure 3.

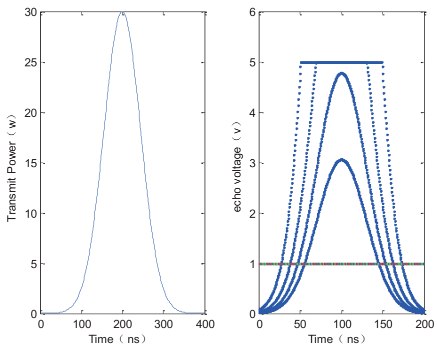

Fig.3 Simulation the height with the width

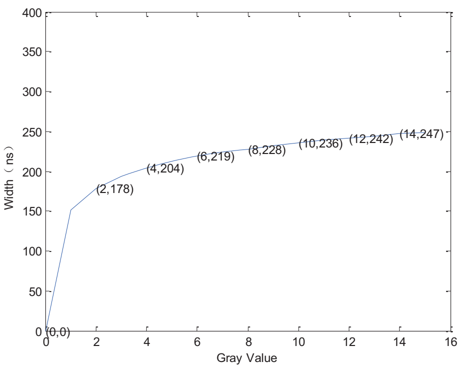

Fig.4 The non-uniform quantization curve

Therefore, the first step during the implementation in hardware is to count the sampling pulse,then map it to corresponding gray value.Considering the value of the echo power follows Gaussian distribution, instead of uniform quantization based on the value of echo, non-uniform quantization base on the log of width is used to get the gray value. Figure 4 shows the non-uniform quantization curve. Figure 5 shows the gray-scale image simulated with the pulse width.

Fig.5 A gray-scale image simulated with the pulse width

3 Local Feature Extraction Algorithm

In order to realize the real-time recognition for the gray-scale image obtained by the line scan, we propose a progressive local recognition, as shown in Figure 6.

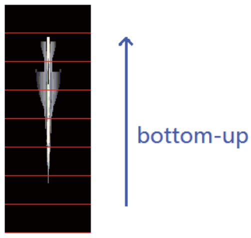

Fig.6 The progressive local recognition

The image is obtained by the bottom-up scan. We easily know that share of the target in the scan image is 30 lines or so. Taking into account the content and the achievability of the image, 16 lines are chosen as the number of the update lines. Once 16 lines of image data are ready, the feature is calculated and the recognition is performed. If the recognition result turns out to be successful, we give a success flag and stop calculating, or continue to wait for another 16 lines and calculate the feature again.The local feature extraction process works in the following steps:

3.1 Build Scale Space

It is optimal to use Gaussian to anlasis the degree of changes in the image [10]. Formula (1) shows Hessian matrix,which is represented by the convolution of the second order patial derivatives of Gaussian function and the image. Lxx, as shown in Formula (2), is the convolution of Gaussian kernel and the image, in which σ represents Gaussian variance.

$$H(X,\sigma)=\begin{bmatrix} L_{xx}(x,\sigma) & \mathrm{L}_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & \mathrm{L}_{yy}(x,\sigma) \end{bmatrix} \quad , \quad (1)$$

$$\mathrm{L}_{xx}(\mathrm{x},\sigma)=G(\sigma)*I(X) \quad , \ (2)$$

In order to reduce the computational complexity, integral image [11] and box filter [6] are used to approximate the Gaussian Filter. Figure 7 shows the 9×9 box filter. Then the space of $\sigma=1$ is formed by the determinant of Hessian matrix at each pixel in the image. The specific method of calculation is shown in Formula(3) below.

$$\det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad , (3)$$

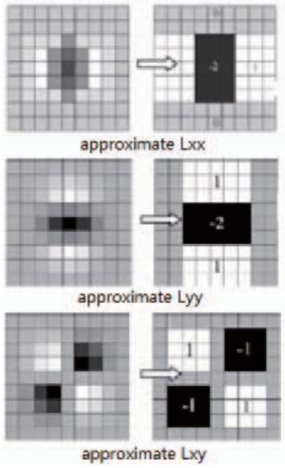

Fig.7 Approximate the second order Gaussian derivatives

The first step in the calculation is to obtain the integral image, and Figure 8 shows the integral results simulated with Modelsim. For a 128*48 image with 16-level gray scale, each pixel needs 16-bit after integral, so a block of 16 rows and 16 columns occupies 4K bits. In order to make full use of 4KRAM resouces inside the chip[12], every 16×16 integral image are saved in a 4KRAM, so the whole integral image needs 24 pieces of 4KRAM. According to the position in the image, the correct RAM number and address in the new RAM is calculated to achieve data reading and writing as shown in Figure 8.

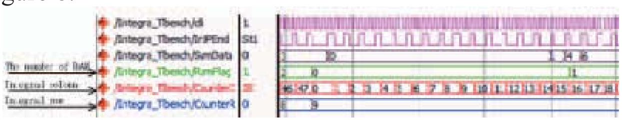
Fig.8 The simulation result of the integral module

Scale space represents the smoothness of the image under a different structure information. Details of the image can be gradually simplified from a small scale to a large-scale. As for a target with a simple structure, After many scales smooth,the too many scales would simplify the information out. Our objective is to find changes in the key points, so three scales are chosen to find the right key. The sizes of box filter under the three scales are 9×9,15×15,21×21. Through the filter of different sizes, we get a scale space{H0，H1，H2} from the integral image.

3.2 Non -maxima Suppression

In order to find out the final key point, the non-maxima suppression is performed in the scale space. Specifically, first, the value in the scale H1 is compared with a constant threshold. If the value is larger than the threshold ,then it is chosen to be a candidate for the maximum point. By doing so we obtain a preliminary determination, then the non-maxima suppression is performed in the 3×3×3 scale space.

Floating point calculation brings the equation of determinant in the scale space. In this case, to simplify the computational complexity, an improved measure is designed to select the key point as shown in Figure 9. "×" represent the points to be suppressed, "☆"represent the point appeared earlier than "×", and "△"represent the point later than "×". As for the positions equal to the suppressed point, if located in the "☆"area, the suppressed point is to be neglected. But if it is located in the "△" area, the suppressed point is to be reserved. It ensures the key points to be non-adjacent.

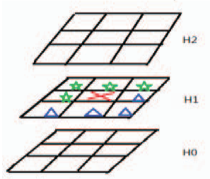

Fig.9 The non-maxima suppression

If the non-maxima suppression is conducted until the calculation of the whole scale space is finished, it is really a waste of time and storage space. Therefore, we present a parallel design, namely, the non-maxima suppression and Hessian matrix are calculated at the same time. Five rows of Hessian data occupy about 4KRAM , so in order to make full use of the storage space in the chip [12], five rows of Hessian data are sent to the non-maxima suppression module while the next five rows of Hessian data are calculated. Ping-pong RAM is an appropriate technique to achieve the parallel computing of non-maxima suppression and scale space calculation. The flow chart is shown in Figure 10 .

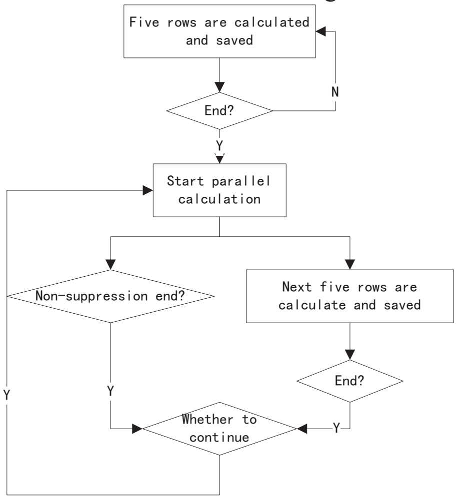

Fig.10 The flow chart of the parallel computing

3.3 Feature Calculation

In order to realize the in-plane rotation invarience, 3D object recognition method is adopted to realize the in-plane rotation invarience. First, a few typcial viewpoints are chosen to be representative template library through the clustering and training. Then, the recognition is conducted according to recognition algorithm. So there is no need to find the main direction and rotate based on it. By doing so it not only satisfies the reliability requirement of the recognition algorithm, but also achieves easy implemention in hardware. Specifically, differential operation is used to calculate the amplitude and the direction of gradient at each pixel within a 17×17 window centered at the key point in the original image. Then we draw the gradient histogram cumulated in every 4×4 block, so a 64-dimention feature vector is produced at every key point.

4 The Progressive Local Recognition Algorithm

Taking into account the high efficiency and the complexity of the matching algorithm, the method based on the nearest distance in the two directions [13] is adopted. As shown in formula (4), p_i is the $i-th$ feature in the test image, q_l is the $l-th$ feature in the template image.

$$d(p_i,q_j)=\min_{q_l\in Q}(d(p_i,q_l))=\min_{p_k\in P}(d(p_k,q_j)) \quad (4)$$

$$d(p_i,q_j)\le \min_{q_l\in Q,l\neq j}(d(p_i,q_l))*\alpha \quad (5)$$

$$d(p_i,q_j)\le \min_{p_k\in P,k\neq i}(d(p_k,q_j))*\alpha \quad (6)$$

Based on the parameter α derived from training, if the conditions shown in formula (4)(5)(6) are satisfied, the target is considered to be in the image . In particular, when the number of the image rows is multiple of 16, the feature is calculated. If there is no feature, we needs to wait for another 16 rows. Otherwise the matching process is performed by comparing features of the test image to those of the templates. In particular, every feature of the test image gives a distance, and if the feature with the smallest distance meets the parameter requirements in the recognition algorithm, we take the test image as a candicate. When the number of the successful matching features meet the requirement of the threshold, the recognition is successful.

5 Results and Conclusions

In order to test peformance of the recognition algorthm, a test library is built by the simulation platform. Based on the 3D model of the target, the angle of defection、pitch and roll in the 3D space is divided respectively by the sampling interval of $(30°,20°,20°)$. By doing so a total of 3888 images are acquired, seven of which are chosen to be the model images by a clustering method in MATLAB, So the remaining 3881 images form the test library.

	Recognition rate	False recognition rate
Our recognition method	95.7%	3.6%
YU's recognition method[2]	84.5%	20.5%

Table1: Test results of the progressive local recogniton algorithm

According to the recognition algorithm, the recognition rate can reach up to 95.7%. As shown in Table 1 above, our algorithm obtains better recognition results with the same test library, and 11.2% is increased compared to the YU's recognition. Besides, the false recognition rate reduces greatly by 16.9%, that is, our algorithm can effectively prevent false recognition.

The whole system which is composed of generating the gray-scale image module, the local feature extraction module and the progressive local recognition module is implemented on the EP2S30484 chip series of the Altera StratixII. the conditions of the resource utilization in the chip is shown in Table 2 below. For the whole recognition process, the usage of resource is relatively low.

Module	ALUTs	registers	memory bits
The whole system	(21%)	(19%)	(9%)

Table 2: The conditions of the resource utilization

Acknowledgements

Thanks for the National Natural Science Foundation of China (No.60502013)

References

[1] SUN Zhihui, DENG Jiahao, Yan Xiaowei. "Progress and current state of the development of laser imaging detection system and its key techniques", *Chinese Science Abstracts,* **14**, pp.13-21, (2008) .

[2] YU Xiaoliang, MA Huimin, XIAO Jian. "Real-time target recognition system simulation based on laser near-field detection", *Optoelectronics Letters*, **5**, pp. 299-303, (2009) .

[3] KOU Xiaoming, LIU Shangqian, HONG Ming, Hong Dabao. "An adaptive infrared image enhancement technology", *Journal of Xi'an Electronics University*, **36**, pp.1070-1074,(2009).

[4] LIU Lifeng, Ma Huimin, LU Mingquan. "A FPGA and Zernike Moments Based Near-Field Laser Imaging Detector Multi-scale Real-Time Target Recognition Algorithm". *International Symposium on Information Science and Engineering,* pp.370-374, (2010) .

[5] XU Shiwen, ZHANG Zhenghui. "Analysis of test methods based on characteristics of the laser near-field target", *Infrared and Laser Engineering*, **36**, pp.64-68, (2007) .

[6] Herbert B, Tinne T, and Luc V G. "Speeded Up Robust Features". *European Conference on Computer Vision,* 404-417, (2006).

[7] Lowe D G, "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision*, **60**, pp.91-110, (2004).

[8] WEI Bin, ZHENG Lian, WANG Keyong. "A new method for laser fuze imaging detection", *Optoelectronic Engineering*, **32**, pp. 36-39, (2005).

[9] HU Chunsheng. "Investigation into the High-speed Pulsed Laser Diode 3D-imaging Ladar": [PH.D.Dissertation]. National University of Defense Technology, (2005).

[10] Koenderink J. "The structure of images". *Biological Cybernetics*, **50**, pp.363-370,(1984).

[11] Viola P, Jones M. "Rapid object detection using a boosted cascade of simple features", *Computer Vision and Pattern Recognition* , 511-518, (2001).

[12] StratixII Device Handbook, Altera Corporation, 2007.

[13] SU Juan, XU Qingsong, ZHU Jinghua, "A scene matching algorithm based on SURF Feature". *Image Analysis and Signal Processing*, 434-437, (2010).