Application research of Mean Shift Algorithm in the embedded for real-time color Target Tracking

Shuang Qiao[1], Xiufen Li[2] , Yan Zhu[3]

[1] School of physics Northeast Normal University, Changchun, 130024, qiaos810@nenu.edu.cn
[2] School of physics Northeast Normal University, Changchun, 130024
[3] School of physics Northeast Normal University, Changchun, 130024
E-mail:johnsonsun866@163.com

Abstract

In order to meet the real-time performance requirements of Mean Shift Algorithm, this paper proposes a new kernel-based quantitative method and integralization method on calculating the weights. This method makes the complex floating-point operation to be a integer computation, which can reduce the computation. We Design a closed-loop tracking network which composted of an embedded target tracking platform which bases on a high-performance TMS320C6418 DSP chip with large-scale gate array, and a servo system with two degrees of freedom. The miss distance of the object (the distance the goal from the center of the view field)is the feedback signal to control the servo system to move the camera. So that ,the goal will be near the center of view field. Experiment result shows that the system has a good stability.

Keywords: *closed-loop tracking color; characteristics; mean shift algorithm*

1. Introduction

How to effectively track the useful goal in the sequence video, is an extremely challenge topic in the computer vision, and it has the widespread application in the monitoring, video coding and the war industry. In the initiative visual system, we not only need to get the trajectory of target fast and accurately, but also need to control the tracking turntable in the tracking process to maintain the target near the center field of view to realize a continuous tracking [1].

Color feature is applied widely in the field of computer vision, such as in the target classification and indexing [2], facial recognition [3, 4], texture segmentation [5], and so on.

Mean Shift Algorithm was firstly proposed by Fukunaga and others in 1975[6]. In 2003 Comaniciu and others proposed a Algorithm which employs mean shift tracking method and integrates the color histogram of weighted, and successfully simulated it. It can combine the calculation accuracy and the amount of calculation well, because it need no parameter, but has a rapid pattern matching features. This Algorithm has achieved very good results in image retrieval and target tracking fields.

The Mean Shift Algorithm of Color feature in tracking target has a good effect, But the computing requirements a large capacity storage and high-speed. So it was seldom used in the embedded systems which requires a high real-time. With the development of electronic technology, the computing capability of embedded systems increases rapidly and it is possible to apply complex application program in embedded system.

This paper designs a system which uses TI's high-speed processing chip TMS320C6418 and programmable gate array CPLD to build a high-speed TV tracking platform. With real-time Mean Shift tracking algorithm of RGB color space features, the system can achieve a better color target tracking.

2. Real-time tracking platform design

DSP TMS320C6418 is the core of the system. The external logic control circuits which is composed of two programmable gate array CPLDs builds a tracking platform. Its schematic was showed in figure 1.

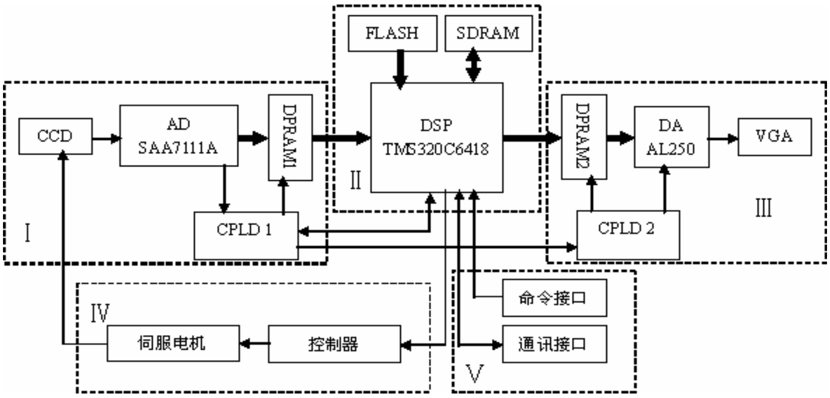

Figure1. Block diagram of the system

The system is divided into five parts, as shown in Figure I, II, III, IV and V. Module I is the forward channel of the system, and it can achieve analog video signal decodes and digital sampling. It is composed of decoder chip SAA7111A, input buffer DPRAM1 and control logic modules CPLD1. When it works, the output single of CCD is the color equalization complex analog video signal. After digitized and decoded by the video decoder chip, it puts out synchronous video control signal and digital image data. CPLD1 reads and writes DPRAM1 by the video control signal, and save the image data in DPRAM1. Module III which includes the output buffer DPRAM2, VGA encoding chip and output logic modules CPLD2, is the backward channel of the system. It realizes trans coder from digital video data to the VGA signal.

DSP will write the showed data to DPRAM2.CPLD2 generates the work timings of DPRAM2 and the encoding chip AL250, by the decoding synchronous video control signal. Finally, the video output of VGA is sent to display. Module IV which is composed with a decoding controller and servo motor is the servo and feedback control unit. Decoder controller receives PELCO-D code the output of the central processing module, and generates the level and vertical rotation voltage for the two freedom degrees servo motor to drive CCD to rotate.

Module V is the control interface for communications and the user. It realizes man-computer interactive information exchange. Module II is the core processing module, composes of the digital signal processor TMS320C6418, dynamic memory SDRAM and program memory FLASH.

The TMS320C6418 (C6418) device produced by Texas Instruments (TI) is a new DSP based on the C64x which is a code-compatible member of the C6000 DSP platform. It has very-long-instruction-word (VLIW) architecture and performance of up to 4800 million instructions per second (MIPS) at a clock rate of 600 MHz.

The C6418 uses two-level high-speed cache, and custom can use the maximum capacity of 512 k bytes. And it has a 32 bits EMIFA (external memory interface), 64 independent channels EDMA, so it can work parallel with the CPU and improve the speed of the system to access external resources. The work process of Module II is as follows:

After the system was powered, DSP load the program from the outside FLASH to internal space, and run the program if boot load is successful. Figure 2 shows the program work timing. CPLD1 sends hardware interrupts HI (Hardware Interrupt) to DSP after AD Acquaint 1 / 4 field data. DSP receives HI interruption, and generates EDMA channel 1 to read the display images (1 / 4 Field) in the external cache SDRAMsent and to write them in the external memory DPRAM2. Then it starts EMDA Channel 2 to read the dates (1 / 4 Field) from input buffer DPRAM1 and to write them in the internal memory L2. DSP will start a software interrupt SI (Software Interrupt) after the completion of HI4 interruption. In SI interrupt service routine, DSP gives an operation process to one field data according to data-processing algorithm and the operation results superimpose on the original image data in the form of text and graphics, then DSP start EDMA Channel 3 to write the new image data to external SDRAM to prepare for the next show.

3. Improved Mean Shift tracking Algorithm

3.1 color object representation

The object is expressed by the weighted color-histogram of object position accessories of ellipse image region. In order to reduce the influence of the object scale, select the different scale $h_0 = \{h_x, h_y\}$ on row and column respectively and normalize the ellipse to the unit circle. Suppose $\{x_i^*\}_{i=1\ldots n}$ are the pixel positions of normalized original object. Select one monotonic decreasing **Epanechnikov** kernel function $k(x)$ whose middle is bulged as weights between pixel and the object centre.

As follows

$$k(x) = \begin{cases} c*(1-x), if\ (x \le 1) \\ 0, otherwise \end{cases} \tag{1}$$

C is constant. Compute the color histogram of the weight of the object region. Owing to the kernel function put the little weight to the pixel which is far away from the object centre; this color histogram is robust for background pixel and partial occlusion [7]. Suppose quantification range of the color histogram is 1024, that is $BIN = \{BIN_R, BIN_G, BIN_B\} = \{8,16,8\}$, the image data format of the output of decoding chip is RGB565, as showed in figure 3 (a) . So that the image data which is **composted of the** high pass bit of R, G and B is a ten bit histogram index pointer. As showed in figure 3 (b). We express the object density function as:

$$q_u = C\sum_{i=1}^{n} k(\|x^*{}_i\|^2)\delta[b(x^2{}_i) - u] \tag{2}$$

In actual operation, in order to increase computing speed, floating-point arithmetic is replaced by fixed-point arithmetic on TMS320c6418. the weight distance is separated to N^2 grades, $k_j, j = 1\ldots N^2$, and the (1) formula is separated as:

$$k_j = (\text{int})(N^2 - (D_i * N)^2) \tag{3}$$

D_i is the normalizing-distance between pixel and the object centre 0. In the (2) formula, suppose C=1, and (2) formula will be as the following:

$$q_u = \sum_{i=1}^{n} k_j * \delta[b(x^*{}_i) - u] \tag{4}$$

$b(x^*{}_i)$ is the color quantization index pointer of the pixel x_i^*, showed in figure 3(b).

15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
R4	R3	R2	R1	R0	G5	G4	G3	G2	G1	G0	B4	B3	B2	B1	B0

(a)RGB565 Image Data Format

15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
0	0	0	0	0	0	R4	R3	R2	G5	G4	G3	G2	B4	B3	B2

(b) RGB565 10 bits quantization index data format

Figure 3. System color format

3.2 Mean Shift tracking

Mean shift is non-parametric estimation method which uses the gradient of a probability distribution to search the distribution peak value. It has good robustness. Suppose $x_i, i = 1...n$ is a finite set of d dimension space R^d. The estimation of Multivariable kernel $K(x)$ for distribution function $f(x)$ of x_i is as the following:

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} K_H(x - x_i) \tag{5}$$

H is the d×d bandwidth matrix. Its mean shift vector at x definition is as the following:

$$m(x) = \frac{\sum_{i=1}^{n} x_i w_i g(\left\| \frac{x - x_i}{h} \right\|^2)}{\sum_{i=1}^{n} w_i g(\left\| \frac{x - x_i}{h} \right\|^2)} - x \tag{6}$$

w_i is the weight, g is the gradient kernel of k, $m(x)$ always along the x direction point to the fastest-growing direction. Along the $m(x)$ direction we move the center position of the kernel function, and will find the near pattern-matching position.

In order to decrease the system inertia in the course of tracking, we calculate every filed that is to complete the tracking calculation in the services program of the software interrupt SI.

Calculation program is divided into the following steps:

(1) Calculate the quantitative histogram q_u of the object by formula (4).

(2) Predict the object position in this field by the position in the last field. And calculate the object histogram $\hat{p}(\hat{s}_0)$ of candidate object at s_0 .

(3) Calculate the weight value by the formula (7) and (8) to compost the backward projective images.

$$w_i = \sum_{u=1}^{m} \sqrt{\frac{q_u}{p_u}} \delta[b(x_i) - u] \tag{7}$$

Merge w_i to [0,255], and do integer processing with it. Formula (7) converts to formula (8)

$$W_i = (\text{int})(\frac{\sum_{u=1}^{m} \sqrt{\frac{q_u}{p_u}} \delta[b(x_i) - u]}{\max} * 255) \tag{8}$$

Max is the maximum value of w_i.

(4) Calculate the gravity center of the backward projective images.

(5) If the recurrence time is over N_{max} ,or target error $\sigma = \|s_1 - s_0\| < \Delta$, Δ is the error threshold exit the circulation, and put out the result of tracking to image and the servo-mechanism ,then exit the interrupt services program. Otherwise go to formula (2) to continue Calculating.

4. Experiment results

The computational complexity of the algorithm is decided by the tracking, window size and the maximum iteration times of mean shift .usually, mean shift need 10, 20 times recursive, and the average times is about 4 [7]. In the experiment, the quantization levels of the weight is $N^2 = 16*16$. The maximum parameter of target window is $h_0 = \{80,60\}$ (window size is 160×120). The error threshold is $\Delta = 2$, the recursive times are $N_{max} = 15$, and the system can work on-time well .finger 4 shows the tracking result in one experiment, window parameter is $h_0 = \{28,28\}$. In finger 4, fist ,target (red cups) move to the right (Frame 200), then move to the distance (Frame 300), and then part of it were obscured (400 frames). In the whole experiment course, the tracking window can track the target moment.

Frame 0 (odd field) Frame 200 (odd field)

Frame 300 (odd field) Frame 400 (odd field)

Figure4. Part result of tracking a red cup

5. Conclusion

The author discussed a embedded tracking method which is based on the color characteristics. This paper uses the color histogram with weighted distance as object feature, and the mean shift which is a rapid searching optimization algorithm to calculate.

Keeping certain precision, we use fixed-point operation to replace the floating-point operations and carry out the quantitative for the weight value to reduce the computational complexity for every field tracking. We also amplify and calibrate the weight coefficient to be integer. So that, the system can operates in real-time. The experiments show that the method can track the color target well.

Because of the heavy computation of the acquisition algorithm, it is hard to realize real-time operation. And now the system is only of using manual capture.

And the precision of servo control system is inadequate, so moving the target to the centre field of view is difficult.

How to capture the color target quickly and efficiently, improve the tracking accuracy of servo system is the problems which we will solve next stage.

References

[1] Aloimonos J, Weiss I, Bandyopadhyay. An Active vision [J]. Int J Comput Vision 1988, 1(4)：pp.333–356.

[2] Gevers T,Smeulders AWM. Color-based object recognition [J]. Pattern Recognition 1999.3 32(3) . pp.454-464.

[3] Stefano Arca, Paola Campadelli and Raffaella Lanzarotti. A face recognition system based on automatically determined facial fiducial points [J]. *Pattern Recognition,* Volume 39, Issue 3, 2006.3, 39(3), pp.432-443.

[4] Zhonglong Zheng, Jie Yang and Limin Yang. A robust method for eye features extraction on color image [J]. Pattern Recognition Letters, 2005.10, 26(14), pp.2252-2261.

[5] Yuanjie Zheng, Jie Yang, Yue Zhou and Yuzhong Wang. Color-texture based unsupervised segmentation using JSEG with fuzzy connectedness [J]. Journal of Systems Engineering and Electronics, 2006.3, 17(1), pp.213-219.

[6] K.Fukunaga and L.D.Hosterler. The estimation of the gradient of a density function with applications in pattern recognition [J]. IEEE Trans. Information Theory, 1975, vol25, 32-40

[7] Comaniciu D, Ramesh V, Meer P. Kernel-Based object tracking.[J] IEEE Trans. on Pattern Analysis and Machine Intelligence, 2003, 25(5). pp.564−575.

[8] Texas Instruments [S], TMS320C6418 Fixed-Point Digital Signal Processor Data Manual (SPRS241B), 2004.8.

[9] K.Fukunaga. Introduction to Statistical Pattern Recognition [M], 2nd Edition. Academic Press, New York, 1990.

Sun zhongshen. 1980.7, M (Han) , Zaozhuang, Shandong, the Chinese Academy of Sciences in Changchun Optical Precision Machinery and Physics Institute doctoral students, the main research direction is video image processing technology and embedded systems.

Article received: 2008-05-04