International Journal of Engineering Science Invention
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org Volume 3 Issue 4 ǁ April 2014 ǁ PP.42-48

Target Detection Based on Normalized Cross Co-relation

Niladri Shekhar Das[1], S. Arunvinthan[2], E. Giriprasad[3]
(Avionics, AISST, Amity University NOIDA)

ABSTRACT : *This paper represents a case study of target detection from an airborne camera. Such practice is very common in aviation field to confirm a waypoint or to detect a threat. Often for surveillance purpose such airborne cameras are deployed. Normalized cross correlation (NCC), where the elimination order is determined based on the gradient magnitudes of sub blocks in the current macro block, is a very common method used in detecting targets. Although very efficient, the computational cost of NCC is very high. Over the years, many researches are going on to reduce the computational cost of NCC. Here we have studied the use of a normalized cross-correlation algorithm based on multilevel Cauchy-Schwartz inequality to skip unnecessary block by block matching calculations used in general NCC algorithm. Also, additional complexity reduction is achieved re-using the normalized cross correlation values for the spatially neighboring macro block because the search areas of adjacent macro blocks are overlapped. Simulation results show that the algorithm can improve the speed-up ratio up to about 2 times in comparison with the other existing algorithms. Finally we have implemented an algorithm for reliable target detection with the help of this fast NCC calculating algorithm.*

KEYWORDS: *Target detection, Template matching, Fast normalized cross co-relation, Cauchy-Schwartz inequality, Geometric calibration of aspect angle.*

I. INTRODUCTION

Correlation is quite simple in principle. Given an image f(x, y), the correlation problem is to find all places in the image that match a sub image (also called a mask or templates) w(x, y). Typically, w(x, y) is much smaller than f(x, y). One approach for finding matches is to treat w(x, y) as a spatial filter and compute the sum of products (or a normalized version of it) for each location of w in f. Then the best match (matches) of w(x, y) in f (x, y) is (are) the location(s) of the maximum value(s) in the resulting correlation image.Normalized cross correlation (NCC) has been commonly used as a metric to evaluate the degree of similarity (or dissimilarity) between two compared images. The main advantage of the normalized cross correlation over the cross correlation is that it is less sensitive to linear changes in the amplitude of illumination in the two compared images. Furthermore, the NCC is confined in the range between –1 and 1. The setting of detection threshold value is much easier than the cross correlation. Correlation-based methods have been used extensively for many applications such as object recognition, face detection, motion analysis and industrial inspections of printed-circuit boards, surface-mounted devices, wafers, printed characters, fabrics, ceramic tiles etc.For target detection from airborne platform, normalized cross-correlation (NCC) technique is a reasonable choice for due to its capability to provide an estimate of the similarity between images. In the next few sections we will discuss about methodology of NCC implementation, a method to calculate NCC computationally faster and the target detection implementation algorithm.

Normalized Cross-Correlation: Methodology

In object recognition or pattern matching applications, one finds an instance of a small reference template in a large scene image by sliding the template window in a pixel-by-pixel basis, and computing the normalized correlation between them. The maximum values or peaks of the computed correlation values indicate the matches between a template and sub images in the scene. The normalized cross correlation used for finding matches of a reference template t (i, j) of size m×n in a scene image f (x, y) of size M×N is defined as

$$\delta(x,y)=\frac{\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} f(x+i,y+j)\cdot t(i,j)-m\cdot n\cdot \mu_f\cdot\mu_t}{\left\{\left(\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} f^2(x+i,y+j)-m\cdot n\cdot{\mu_f}^2\right)\cdot\left(\sum_{i=-m/2}^{m/2}\sum_{j=-n/2}^{n/2} t^2(i,j)-m\cdot n\cdot{\mu_t}^2\right)\right\}^{1/2}}$$

$$\text{for all } (x,y)\in M\times N \qquad (1)$$

Where

$$\mu_f(x,y) = \frac{1}{m \cdot n} \sum_{i=-m/2}^{m/2} \sum_{j=-n/2}^{n/2} f(x+i, y+j)$$

$$\mu_t(x,y) = \frac{1}{m \cdot n} \sum_{i=-m/2}^{m/2} \sum_{j=-n/2}^{n/2} t(i, j)$$

The template size m×n is smaller than the scene image size M×N. Pixel-by-pixel template matching is very time-consuming. For a scene image of size M×N, and the template of size m×n, the computational complexity is O(m× n ×M × N) . In order to alleviate the drawback of long processing time in template matching, the course-to- fine and multi-resolution search approaches have been widely used to reduce computation burden. Such algorithms first scan the image quickly and find all promising areas in the rough resolution, and then search for more accurate patterns and locations in the fine resolution.

Different Block-matching Methods Used for Object Detection in a Motion Video

Motion estimation has been employed by many video compression schemes to improve coding efficiency removing the temporal redundancy that exists in video sequences. The block-matching algorithm (BMA) is the most popular approach applied to all video coding standards, such as MPEG and H.264/AVC, due to its structural simplicity. A full search algorithm (FSA) can be the best BMA for a given block distortion criterion, as it finds the block with minimum block-matching distortion among all candidates. However, its heavy computational cost is a crucial limiting factor in terms of software implementation as well as hardware implementation.

For several decades, many fast BMAs have been developed.
These can be divided into two categories. The first category adopts pre-defined search patterns to locate candidate motion vectors (MVs) based on distortions of potential candidates [1]. The second category is entirely composed of optimal motion estimation methods, which can find the globally optimal MV within a search area [2]-[7]. Li and Salari proposed a well-known successive elimination algorithm (SEA) providing a decision boundary based on the sum norms of blocks to eliminate some checking points without the need for computationally intensive block matching [2]. Gao and others extended the SEA to a multilevel SEA (MSEA) that provides multiple levels of tighter boundaries using the sum norms of the macro block (MB) and sub blocks with reduced sizes [3]. MSEA reduces the necessary computation by detecting and rejecting unnecessary candidates from the lowest level to the highest level. Zhu and others proposed a fine granularity successive elimination (FGSE) scheme that extended the MSEA by adding greater detail levels [4]. Also, Liu and others presented an adaptive version of FGSE [5]. The FGSE is distinguished from the MSEA in that, if necessary, only a single sub block having the maximum complexity is chosen at each level and the sub block is partitioned into four smaller sub blocks at the next level. Therefore, in the case of a 16×16 block, the total number of partition levels amounts to 86. Thus, the FGSE has more potential to prune out non-optimal candidates than MSEA before wholly performing block matching. As a block distortion criterion, the sum of absolute differences (SAD) is commonly used in video compression. SAD is defined as:

$$SAD(u,v) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left| I_C(i,j) - I_R(i+u, j+v) \right|, \quad (2)$$

where (u, v) is a MV in the search area, and IC and IR denote the current and reference picture, respectively.

In addition to SAD and the sum of squared differences(SSD), the normalized cross correlation (NCC) is also a well-knownsimilarity criterion. The NCC can be defined simply as

$$NCC(u,v) = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} I_C(i,j) \cdot I_R(i+u, j+v)}{\sqrt{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} I_C(i,j)^2} \sqrt{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} I_R(i+u, j+v)^2}} \quad (3)$$

The NCC is more robust than SAD and SSD under uniform illumination changes. Accordingly, it is widely used in object recognition and industrial inspection schemes. Applying the NCC as the matching criterion to motion estimation leads to more uniform residuals. Hence, the NCC can improve subjective visual quality as well as coding efficiency in video compression [8].

Faster algorithm to calculate NCC

Song proposed BMA is a type of NCC version of FGSE that uses multilevel Cauchy-Schwartz inequality [9]-[10]. It is possible to apply the multilevel Cauchy-Schwartz inequality based on a *L*2-norm pyramid to the numerator of (3) as follows:

$$\sum_{i,j}^{N-1} C_{i,j} \cdot R_{i,j} \le \cdots \le \sum_{i,j}^{\frac{N}{2^{L-l}}-1} C_{i,j}^{l} \cdot R_{i,j}^{l} \le \cdots \le C^0 \cdot R^0. \quad (4)$$

Here,

$$X_{i,j} = X_{i,j}^{L}, X_{i,j}^{l} = \sqrt{\sum_{m=2i}^{2i+1} \sum_{n=2j}^{2j+1} \left(X_{m,n}^{l+1}\right)^2}$$

And,

$$L = \log_2 N.$$

Based on (4), it is possible to derive the following inequality:

$$NCC^0 \ge NCC^1 \ge \cdots \ge NCC^l \ge \cdots \ge NCC^L. \quad (5)$$

In (5), NCCl is defined as

$$NCC^l = \frac{\sum_{i,j}^{\frac{N}{2^{L-l}}-1} C_{i,j}^{l} \cdot R_{i,j}^{l}}{\sqrt{\sum_{i,j}^{N-1} C_{i,j}^{2}} \sqrt{\sum_{i,j}^{N-1} R_{i,j}^{2}}}. \quad (6)$$

Then a multilevel successive elimination algorithm is developed to determine the best MB with the maximal NCC value according to (5).The total number of partition levels can be extended to 86, i.e., $0 \le l \le 85$, by partitioning each of four sub blocks at a certain level one by one into another four sub blocks at the next level in the descending order of the complexities of larger-sized sub blocks as mentioned in the study [4] by employing the gradient magnitude as an image complexity measure.Furthermore, we can find that the search areas of the current MB and its leftneighboring MB are mostly overlapped. For the search range of ±16 pixels, two adjacent MBs share about2/3 of their entire search areas. Song[10] proposed a method to preventunnecessary computations of NCC using the NCC valuesobtained from motion estimation of the spatially neighboringMB.
Assume that $P_{i,j}$ denotes a pixel at (i, j) in the spatiallyadjacent 16×16 MB, that is, IC(i, j–16). For instance,we consider an example when l is equal to 1. For the samecandidate in the overlapped search area, the NCC^1s of P and Care as:

$$\frac{\sum_{i,j=0}^{1} C_{i,j}^{1} \cdot R_{i,j}^{1}}{\|C\|_2 \cdot \|R\|_2} \text{ and } \frac{\sum_{i,j=0}^{1} P_{i,j}^{1} \cdot R_{i,j}^{1}}{\|P\|_2 \cdot \|R\|_2}, \quad (7)$$

where$\|C\|_2$ stands for the L_2-norm of the current MB C.Since two adjacent MBs generally have significant spatialcorrelation, the possibility that $P_{i,j}$ and $C_{i,j}$ are equivalent is high.For example, the two NCC^1s of P and C in (7) can be thesame at level 1. Then, we can replace the NCC^1 of C with theNCC^1 of P without computation. Note that as l becomessmaller, the probability that $P^l_{i,j}$ and $C^l_{i,j}$ are equivalent becomeshigher. Thus, the NCC values of the block candidates in theoverlapped search area are stored during motion estimation ofthe spatially neighboring MB, and if $P^l_{i,j}$ and $C^l_{i,j}$ are equivalent at level l of motion estimation of the current MB, the NCC^l ofC is replaced with the stored NCC^l of P.

It is very important to choose the initialsearch point and search pattern properly in the search area. The initial search point is set to the median MV ofMVs of three spatially adjacent MBs (Fig. 1), in order tomaximize the elimination effect; a spiral search pattern with theinitial MV is presented as the starting point, as in the example shown in Fig. 1.

The algorithm is summarized as follows:
1) Offline pre-processing: Build L_2-norm pyramids for the reference frame.
2) Online processing: For each MB in the current frame, the following procedure is applied.
(a) Compute the L_2-norm pyramid of the current MB.
(b) Compute the NCC corresponding to the initial MV and set the current maximum cost (Cmax) to the computed NCC.
(c) Set l to 0. Compute NCC^l. If C0 and P0 are the same at this level and the NCC for the reference block is available, employ the stored NCC instead of the NCC^l computation.

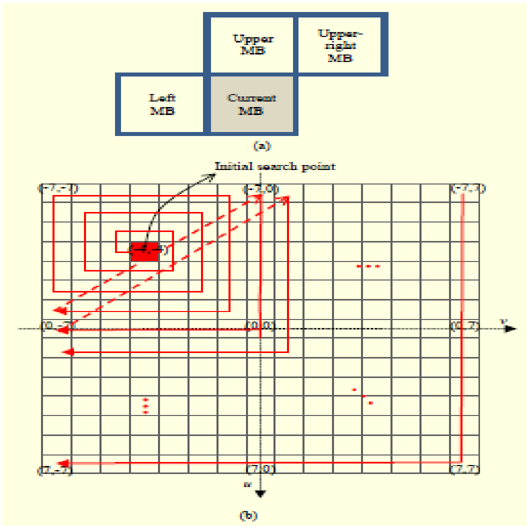


Fig 1: (a) Neighboring MBs that are used to determine initial MV; (b) initial MV-centric search pattern.

[1] (d) If $NCC^l \geq$ Cmax, replace l with l+1 and find a sub block with the largest complexity according to the given partition rule. Otherwise, reset l to 0 and go to step (c) with the next MV candidate. If no additional MV candidate exists, go to step (h).

[2] (e) Partition the sub block having the largest complexity and compute its corresponding NCC. Compute NCC^l by updating the NCC of the partitioned sub block only. If Cl and Pl are same at this level and the NCC for the reference block is available, employ the stored NCC instead of NCC^l computation. If l is equal to 85, go to step (g).

[3] (f) If $NCC^l \geq$ Cmax, replace l with l+1, find a sub block with the largest complexity according to thegiven partition rule, and go to step (e). Otherwise, reset l to 0 and go to step (c) with the next MV candidate. If no more MV candidates exist, go to step(h).

[4] (g) If $NCC^{85} \geq$ Cmax, update Cmax to the NCC85. Reset l to 0 and go to step (c) with the next MV candidate. If no additional MV candidates exist, go to step (h).

[5] (h) Select the MV corresponding to the final Cmax as the best match to the current MB. The computed NCC values are stored for the next MB.

II. TARGET DETECTION ALGORITHM

For target detection first of all we need to create a database of the targets of interest. The target of interest can be a tank, an anti-aircraft gun, a military station, a ship, etc. The database is most vital in any kind of target detection. Without a proper target profile in the database detection is impossible. For different camera angle and different sun illumination the target profile changes. NCC itself can normalize the illumination differences, but the aspect angle differences remain for different camera angels. Lynn[11] implemented a method using geometric calibration data of the airborne equipment to correct the aspect angles mismatch problem. A block diagram representation of the algorithm is given in Fig. 2.So to implement the algorithm first we prepare image database of potential targets from an inertial platform from different heights. The upper left most pixel ofthe pictures are set as (0,0) co-ordinate.The camera is installed in an inertial platform in the airborne carriers (e.g. jets, UAV's and

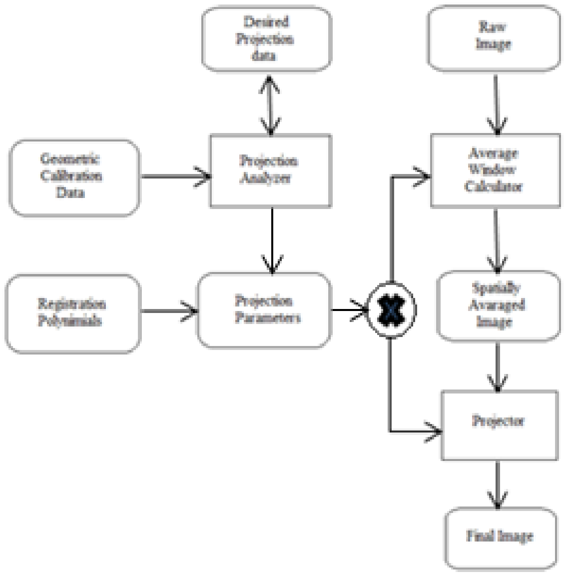


Fig 2: Aspect angle correction

LAV's). After this we can implement the previously discussed NCC method to detect the target by co-relating each frame obtained by the motion camera to the all targets in the database. As the frame size is now different, the targets will have different aspect angles other than the first window of NCC. The scenario is explained in the Fig. 3.

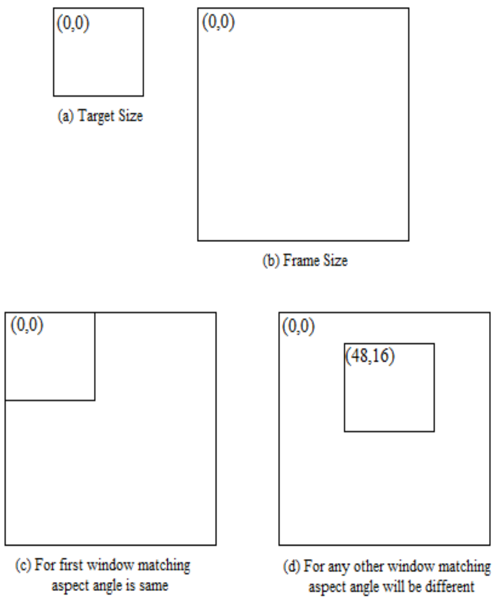
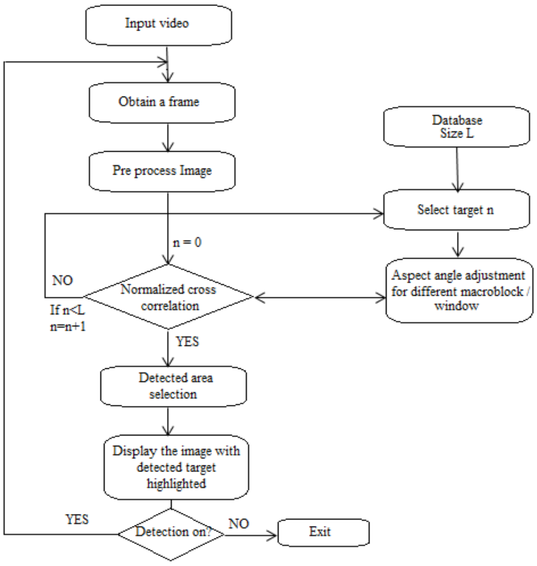


Fig 3: Aspect angle problem demonstration

In these cases we first process the frame window with Lynn's algorithm to get the correct projection and then perform the NCC algorithm on it to detect the target. The full algorithm is given below:



III. EXPERIMENTAL RESULTS

In order to evaluate the performance of the algorithm, eight CIF (352×288) video sequences were used to detect some predefined targets (Jet, Jeep, Tank, Boat). The MB sizeand the search range are fixed to 16×16 and ±15 pixels in boththe horizontal and vertical directions, respectively. Simulationwas performed on a dual core CPU at 2.66 GHz. We have calculated average no of operations per MB in different sequences to find out the reduction in computational complexity by the Cauchy-Schwartz inequality based algorithm compared to FSA and MSEA method. The result showed up to 2 times improvement in speed up ratio (refer Table 1).

	FSA	MSEA		C-S NCC	
	ANOP	ANOP	Speed up ratio	ANOP	Speed up ratio
Jet	891,065	52,043	17.1	23,964	37.2
Jeep	891,065	47,313	18.8	27,485	32.4
Tank	891,065	47,063	18.9	28,993	30.7
Boat	891,065	48,448	18.4	26,770	23.3

Table 1: Computational complexity comparison of various sequences

The following figure shows the final implementation results. Different targets are successfully detected.

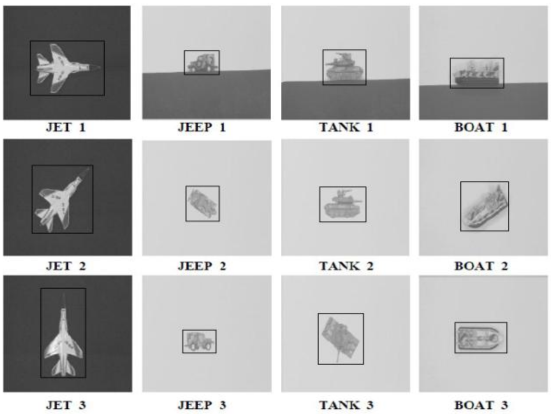


Fig 4: Simulation results

IV. CONCLUSIONS

In this case study a fast BMA based on NCC, where multilevel Cauchy-Schwartz inequality is employed to skip unnecessary block-matching calculation, is implemented and verified for target detection problems. The NCC-based algorithm considerably reduces the computational complexity of a video encoder. Finally, we have implemented an algorithm for continuous target detection from an airborne platform.

REFERENCES

[1] S. Zhu and K.K. Ma, "A New Diamond Search Algorithm forFast Block Matching Motion Estimation," IEEE Trans. ImageProcess., vol. 9, no. 2, 2000, pp. 287-290.

[2] W. Li and E. Salari, "Successive Elimination Algorithm forMotion Estimation," IEEE Trans. Image Process., vol. 4, no. 1,1995, pp. 105-107.

[3] X.Q. Gao, C.J. Duanmu, and C.R. Zou, "A Multilevel SuccessiveElimination Algorithm for Block Matching Motion Estimation,"IEEE Trans. Image Process., vol. 9, no. 3, 2000, pp. 501-504.

[4] C. Zhu, W.S. Qi, and W. Ser, "Predictive Fine GranularitySuccessive Elimination for Fast Optimal Block-Matching MotionEstimation," IEEE Trans. Image Process., vol. 14, no. 2, 2005, pp.213-221.

[5] S.W. Liu, S.D Wei, and S.H. Lai, "Fast Optimal MotionEstimation Based on Gradient-Based Adaptive MultilevelSuccessive Elimination," IEEE Trans. Circ. Syst. for VideoTechnol., vol. 18, no. 2, 2008, pp. 263-267.

[6] J.N. Kim and T.S. Choi, "Adaptive Matching Scan AlgorithmBased on Gradient Magnitude for Fast Full Search in MotionEstimation," IEEE Trans. Consum. Electron., vol. 45, no. 3, 1999, pp. 762-772.

[7] C.Y. Choi and J.C. Jeong, "New Sorting-Based Partial Distortion Elimination Algorithm for Fast Optimal Motion Estimation," IEEE Trans. Consum.Electron., vol. 55, no. 4, Nov. 2009, pp. 2335-2340.

[8] W.H. Pan et al., "A Hybrid Motion Estimation Approach Based on Normalized Cross Correlation for Video Compression," IEEE Proc. ICASSP, 2008, pp. 1037-1040.

[9] Z. Wang et al., "Image Quality Assessment: From Error Visibility to Structural Similarity," IEEE Trans. Image Process., vol. 13, no. 4, 2004, pp. 600-612.

[10] B.C. Song and J.B. Ra, "A Fast Normalized Cross Correlation-Based Block Matching Algorithm Using Multilevel Cauchy-Schwartz Inequality," ETRIJ vol. 33, no. 3, Jun. 2011, pp. 401-406.

[11] LH Quam, "Computer Comparison of Pictures," Stanford Computer Science Department Report No. STAN-CS-71-219