International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 /doi:10.14445/22315381/IJETT-V69I12P223
Volume 69 Issue 12, 197-202, December, 2021
©2021 Seventh Sense Research Group®

An Efficient Object Tracking in Thermal Imaging Using Optimal Kalman Filter

V.Teju[1], D. Bhavana[2]

[1]Assistant Professor, [2]Associate Professor & [1,2]Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram-522502, Guntur, Andhra Pradesh, India

[1]teju@kluniversity.in, [2]bhavanaece@kluniversity.in

***Abstract* — *Thermal cameras are widely used in the military. The main aim of developing new application domains in thermal imaging is due to the increase in quality and resolution of an image with lower cost and size. They are used in fields of civilian like finding out missing persons in automotive safety and in medical applications. Thermal cameras are also used to measure temperature differences. Thermal cameras are more advantageous compared to cameras of the visual spectrum as they can see them in the dark. In the present research, we will present the latest model to do object tracking to detect, monitor objects using thermal images by using an optimized Kalman filter. It has medical applications such as monitoring the mentally retorted people in detecting whether they have any weapon or not.*

Keywords — *Kalman filter, medical applications, Object tracking, Optimal features, Thermal imaging.*

I. INTRODUCTION

Tracking the people individually and their belongings is a big task in video surveillance applications. It is a challenging issue, and the literature survey on this states that tracking people and their belongings is a single object. Within applications related to vision systems of computers also, tracking an object is an important task as it is the foundation of object tracking individually. There are many algorithms on object tracking that have been proposed earlier.

Mean-shift tracking has a vital role in this domain because of its durability, ease of implement and computational efficiency. But this algorithm also has some limitations which affect the performance of tracking, and this leads to wrong tracking. Generally, mean shift tracking-based algorithms focus on the limitation of the tracking object model. But to get the best tracking, all the limitations must be considered. But most of these algorithms need manual detection of an object to find the objects that must be tracked. Target tracking has been a vast research domain since the 1960s, which is driven basically in the applications of aerospace like air traffic control, navigation system, sonars, radars, and guidance. It is used in the fields of sensor networking, robotics, econometrics, and biological systems. Generally, the target is considered as human beings, birds, air, land, vehicles related to water. Tracking is defined as the location of objects which are moved at a particular time with the help of a camera. The main theme to track is connecting the objects to a particular target in successive series of frameworks. Thus, to establish a track of the sequence of frames, an algorithm investigates a consecutive sequence of frames. Tracking of a target is explained as an issue of calculating the trajectory of a target or the thing of a picture plane that is located with a location.

The key for tracking is to monitor motion parameters like acceleration, velocity, location, and orientation which are achieved using targets. The tracking of a target is used for realizing and interpreting the behavior of a target that is exposed to pose variations, occlusion, illumination, and scale.

In tracking a target, the first step is thermal frames acquisition. The next step is to pre-process these frames. Pre-process means to remove the noise which is added during the transmission of a frame. In the next step, Background modeling or detection of the foreground of an infrared video is utilized. In the present research, background modeling is utilized as SRF-single frame differencing [1], RA- Running Averaging, and Temporal TMF-Temporal Median Filter [2]. In the next process, tracking is done by a filter called Kalman, and labeling uses a connected component. Track before detecting (TBD) is the method in which tracking is done before detecting. In the final step, the performance of this tracking model with the selected detection model is assessed by metrics of performance [3-5] such as sensitivity(s), Positive predictive value, false alarm rate, accuracy.

We will compare our proposed algorithm with other filtering models.

II. RELATED WORK

The aim of this research is to do segmentation along with tracking of one or more objects from thermal video sequences. Segmentation of a target from a sequence of frames is the most popular area of research. Segmentation is done for many purposes like retrieval of an image, summarization of video, etc. There are different ways to do the segmentation of frame objects. For the methods explained in [6]-[7] for specific frames, a symbol is needed to start the model. Later, algorithms which are using both motion and appearance are utilized to get segments of an object within a frame. The methods [8], [9] require exact target area symbol in the basic frame, and later with the use of area tracking, other frames are divided in the region of foreground and background. Usually, techniques that are semiautomatic generally give sensitive results to do segmentation. Generally, all the applications need huge information, due to which manually decreasing level formatting becomes frame prohibitor. Thus, more automatic methods are shown to do the segmentation of an object in a sequence of frames. These methods utilize motion grouping [4], [6], [7] to do object segmentation. Other methods [8]-[10] use the look of an object to do segmentation in each frame. Later, they utilize look and object movement to get the ultimate results for segmentation. Methods as [9], [11]-[12] discuss economic development layout for a spatial and temporal group of pixels to do segmentation.

At present, there are more algorithms [13]-[14] which give a brief idea of how a particular target has appeared. In the above-mentioned 3 techniques, the technique [14] isolates the targets with given images trying to make this problem easier by arranging a connection among the objects by taking beside frames. An issue of video object segmentation is discussed here [15]. Detection of an object using an optimized feature selection algorithm [16] is explained here. [17] discusses ANN established measurable estimation of underground abnormalities for quadrilateral FM imaging in thermal waves. In [18], they explain on measuring of thermic dispersion to fiber reinforce able polymeric with the help of quadrilateral FM images in thermal imaging. The CZT [19] build up increasing resolute of frequency for deepness resolving which are not stationary in images of thermal is discussed.[20] focusses on the detailed underground combining of abnormalities detecting in unbounded waves in thermal images. The videos related to selfies continual Indian's gesture mother tongue recognizing method [21] is stated here. In [22], an accent with respect to selfies recognizing system added to different elements on adaboost various labels different class classifiers is discussed. A gesture native language recognizing [23] with respect to many elements' combination and artificial neural networks classification is discussed. The classification system in Indian classic dance [24], along with Adaboost multiple class classifiers on the multi-featured combination, is stated in this. Faster and correct features extracting-related segments [25] frames for spinal cord injuries seriousness classifying are explained here. Various methods for detection, tracking, and classification are explained in [26]-[29]. The limitation of selecting a region of a particular object is shaped as an issue called maximum weighted cliques to find the original area of a thing within given sequence frameworks in a similar duration. All this issue is not deterministic, and the problem of (NPhard) is an improved method used to get the solution for this problem. There are segmentation methods that are based on the proposal of an object. The first limitation is that for every technique, the obtained proposal of the target does not depend on the proposal of a target from the next frame. The next limitation is that the segmentation results decrease as every technique does not determine the target shape in beside frames, once finding the similarities of a region. In the proposed algorithm overcomes the limitations and will perform much better on many datasets related to thermal.

There are different models that are used in the tracking of a target. Among them, many have less performance, and because of this, a misclassification rate occurs. Tracking is done by the representation of a target and localization or by using a filtering process and association of data. Representation of a target and localization is a bottom-top procedure that must be overcome with the changes in the look of a target. There are general indications of a target and methods of localization like tracking of a blob, kernel, and contour. The filtering process and association of data is generally an Up-down process that deals with the dynamics of the object to be tracked, the learning of each frame for object tracking, and the calculation of many hypotheses. For tracking, the filters used are Optimal Bayesian, Linear, Non-linear. The tracking of an object is divided into 3divisions such as point, `kernel, silhouette tracking.

III. IMPLEMENTATION

By using the fast detection technique as explained above, we can get the central positions of moving objects as detected. Because of the noise and limitation of the detection technique, a representation of central position sometimes will not show the correct location of an object which is moving. If many objects are tracked, the positions are required to indicate the exact position to track every object. By the tracking technique discussed here, we must get an optimal estimation of the path for each object to attain robust tracking of many objects which are moving with occlusion.

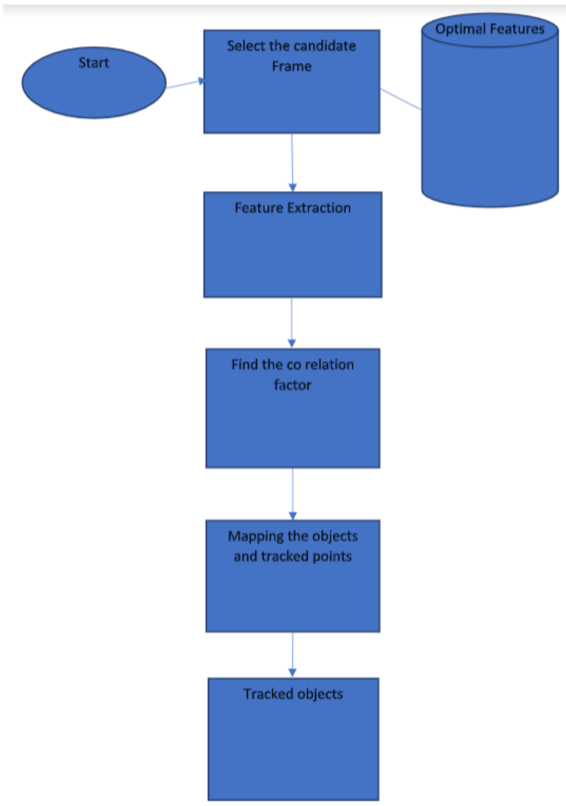


Fig.1 Architecture of Proposed Tracking System

In figure 1, it shows the proposed object tracking model. Here we get the optimal features from the candidate frame using an optimal feature selection algorithm in which a feature vector is taken as an input to the proposed tracking model. After that, we find the correlations between the optimal features and map them into a tracking point vector by applying the Optimal Kalman filter. Then, we track the objects on the frame by using the mapped features.

IV. ALGORITHM

The proposed algorithm for tracking in this research is created using the Kalman filter framework. The ability to track a moving object in thermal imaging is done by the Kalman filter.

$$A(m+1) = FA(m) + wm \quad (1)$$

$$B(m+1) = HA(m+1) + vm \quad (2)$$

where A = {p, q}, p, q denotes an object center location in both directions of parallel and perpendicular. Where **wm**, **m** indicates a noise in state and noise in measure, which are supposed to be. White noise and m indicate the time duration among both side-by-side frames. Upon the estimated position of a particular object, we get a looking zone. To change the size of the region as per the estimated exactness of Kalman, everyone can take their dimensions as a variation in reference to the middle of every frame.

Consider that every individual indicates a probable position of an object A = {p, q}, which is coded with 16 bits of a code which is gray in which lower 8 bits indicates position p, and higher 8 bits indicates position q. The main cause to choose a code that is gray is because of a reason that the value in the coding of 2 successive numerals has a differentiation of 1bit. And this outplays a code which is as it will keep away from an instance that larger the dissimilarity in code will have an influence on convergence, matching accuracy of OFSA.

The region to be estimated, which is resembled by a filter known as Kalman, is partitioned into X × Y blocks, in which X, Y is indicated as 1/4th of an object width, height, 2 singles are unevenly produced for every block their related coefficient(q) is evaluated, where the coefficient is taken as the fitness of every single frame. To all these 2MN singles, we must do sorting for them asper their fitness in decreasing manner and choose the candidate frame, then m is made permanent and utilized for the entire method of genetic evolution, and iteration number is indicated as a result.

The objective model is henceforth expected to be a square shape district loped on a point y* and comprises of n focuses demonstrated as {x∗i} i=1, ⋯, n. On apportioning the gray space of locale in m comparable parts, the wellness work histogram of target calculation is given as q^(y∗) ={q^u(y∗)} u=1, ⋯, m.

$$qu(y*) = Cn\sum i = 1k(\| y* - x*I/h1 \| 2), u = 1 \ldots\ldots m \quad (3)$$

C_1 indicates normalization constant as ∑mu=1q^u(y∗) =1; k(x) which is specified as a function of a profile of fitness function, h_1 which is indicated as the width of a window of k(x), δ which is a function of crossover which satisfies ∑mu=1δ=1, b(x∗i) is a value which is quantized of a pixel in x∗i.

Also, the target candidate cantered on a point y is indicated with function histogram, which is given by p^(y)={p^u(y)} u=1, ⋯, m:

$$p^u(y) = C2\sum i = 1sk(\| y - xi/h2 \| 2)\delta[b(xi) - u], u = 1, \ldots\ldots, m \quad (4)$$

C_2 indicates a constant of normalization as ∑mu=1p^u(y)=1, s indicates a complete numerical of points within a rectangle location of a target candidate, h_2 indicates the width of a window of k(x).

With present research, a proposed function which is indicated as K_E(x), is considered to evaluate k(x). With consideration of integral mean square error, K_E(x) is best for mostly used fitness functions which are evaluated according to

$$KE(x) = 12c - 1d(d+2)(1 - \| x \| 2), 0, \| x \| < 1 \| x \| \geq 1KE(x) = py \| x \| < 1 \| x \| \geq 1 \quad (5)$$

Step 1: Take an input frame from image sequences.
Step 2: Pre-process a frame.
Step 3: Extract the features and reduce them using OFSA.
Step 4: Construct the feature vector.
Step 5: Take the optimal feature vector as an input to the Kalman filter.
Step 6: Find the object points from the features using statistical operations.
Step 7: Perform the fusion step and combine the extracted coordinates to the optimal features.
Step 8: Draw the bounding box from the detected pixels on the frame.
Step 9: Repeat step 4 from the continuous images.
Step 10: Track the objects from the image.

The performance metrics like accuracy, sensitivity, positive predictive value, false alarm rate are calculated. Table 1 shows the proposed system experimental parameters.

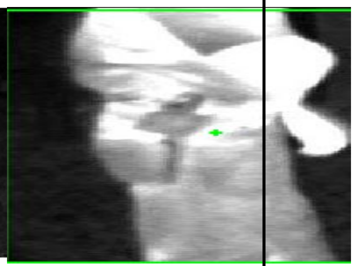
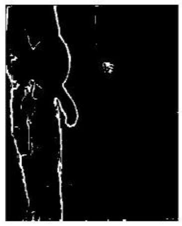
TABLEI . PROPOSED EXPERIMENTAL SYSTEM PARAMETERS

Parameters	Data
Image collected	Thermal images
Type	JPEG format
Resolution	256 *256 pixels
No of the Input images	395 images
Dataset Size	9.44 Mb
Performance Metrics	Accuracy, Sensitivity, Positive predictive value, False alarm rate

In table1, it shows the parameters used in our proposed system, and table2 explains the subjective results of our OKF, KF, CEKF, and SEKF models. We had compared our proposed optimal Kalman filter (OKF) with Kalman filter (KF), Canny edge with Kalman filter (CEKF), and SEKF models.

V. RESULTS FOR SUBJECTIVE TEST

TABLEII. SUBJECTIVE RESULTS FOR OKF, KF, CEKF, AND SEKF

Input Frames	**Pre-processed output**	**Kalman Filter**	**Optimal Kalman Filter**
			
Input Frame	**Pre-Processed output**	**Canny edge with Kalman Filter**	**Shanon entropy with Kalman filter**
			

VI. RESULTS FOR OBJECTIVE TEST

To find the efficiency of OKF method, we had evaluated parameters based on performance such as positive predicted value, sensitivity, false alarm rate, accuracy etc. Each of these parameters are estimated with TP, TN, FP, FN. The expressions to calculate these parameters are given as:

Accuracy:

$$Acc = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \quad (6)$$

Sensitivity:

$$S = \frac{T_p}{T_p + F_n} \quad (7)$$

Positive Predicted value:

$$ppv = \frac{T_p}{T_p + F_p} x100 \quad (8)$$

False Alarm Rate:

$$F_m = 1 - ppv \quad (9)$$

VII. PERFORMANCE TABLE

TABLEIII. OBJECTIVE RESULTS FOR PROPOSED RESEARCH MODEL AND EXISTING MODELS

Algorithm	**Total Frames**	**TP**	**FP**	**FN**	**Sensitivity**	**Positive Predictive value**	**False alarm Rate**	**Accuracy**
OKF	512	440	30	42	0.936	0.966	0.036	0.99
KF	512	435	25	41	0.862	0.852	0.125	0.93
CEKF	512	432	27	40	0.756	0.789	0.235	0.82
SEKF	512	432	25	40	0.723	0.763	0.256	0.79

In figure 2, it explains about the performance of proposed and existing models graphically.

In table 3, it explains about objective results of OKF, KF, CEKF, SEKF

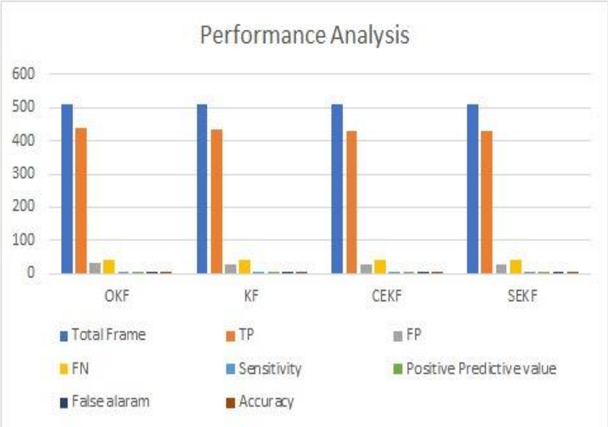


Fig.2 Performance Analysis

VIII. RESULT IN ANALYSIS

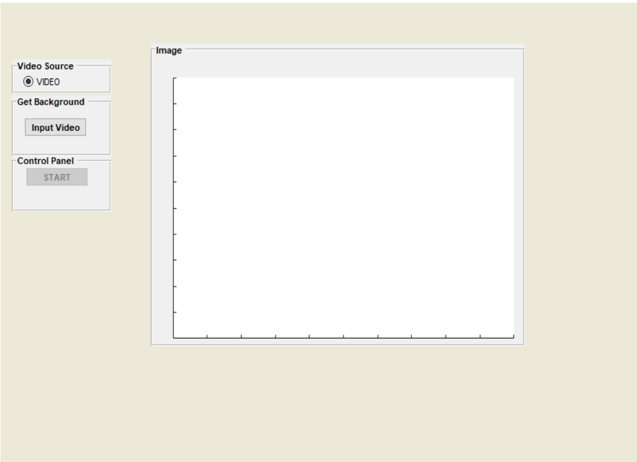


Fig. 3 Reading the continuous frames for object tracking

Fig.4 Candidate frame got the features and stored into the database

Fig. 5 Object on the frame with optimal features and coefficients of the objects are pointed

In the above, figure 3 shows reading the continuous frames for object tracking, figure 4 shows how the candidate frame gets the features and stored into the database, and figure 5 shows the object on the frame with optimal features, and coefficients of the objects are pointed.

IX. CONCLUSION

In this research, how infrared imaging is furnished is discussed. Detection of a weapon and a system for tracking are key parts of security in fields like security in the border, security in the seashore, monitoring of traffic, and rescuing operations based on robotics. The proposed Kalman filter (KF) had good performance in the tracking of many targets in thermal imaging.

REFERENCES

[1] Zhan C, Duan X, Xu S, Song Z, Luo M, An improved moving object detection algorithm based on frame difference and edge detection I, fourth international conference on image and graphics ICIG. (2007) 519-523.

[2] Hung MH, Pan JS, Hsieh CH, Speed up the temporal median filter for background subtraction, In 2010 First International Conference on Pervasive Computing, Signal Processing, and Applications. (2010) 297-300.

[3] Bashir F, Porikli F, Performance evaluation of object detection and tracking systems, InProceedings 9th IEEE International Workshop on PETS. (2006) 7-14.

[4] Nascimento J, Marques JS, New performance evaluation metrics for object detection algorithms, InIEEE Workshop on Performance Analysis of Video Surveillance and Tracking. (2004).

[5] Mariano VY, Min J, Park JH, Kasturi R, Mihalcik D, Li H, Doermann D, Drayer T, Performance evaluation of object detection algorithms, object recognition supported by user interaction for service robots .3 (2002) 561-564.

[6] Bai X, Wang J, Simons D, Sapiro, Video snap cut: robust video object cutout using localized classifiers, ACM Transactions on Graphics (ToG). 28(3) (2009) 1-11.

[7] Yuen J, Russell B, Liu C, Torralba A, Labelme video: Building a video database with human annotations, IEEE 12th International Conference on Computer Vision. (2009) 1451-1458.

[8] Sheikh Y, Javed O, Kanade T, Background subtraction for freely moving cameras, In2009 IEEE 12th International Conference on Computer Vision. (2009) 1219-1225.

[9] Tsai D, Flagg M, Nakazawa A, Rehg JM, Motion coherent tracking using multi-label MRF optimization, International Journal of computer vision. 100(2) (2012) 190-202.

[10] Vazquez-Reina A, Avidan S, Pfister H, Miller, Multiple hypothesis video segmentation from superpixel flows, In European conference on Computer vision. (2010) 268-281.

[11] Grundmann M, Kwatra V, Han M, Essa, Efficient hierarchical graph-based video segmentation, In2010 ieee computer society conference on computer vision and pattern recognition. (2010) 2141-2148.

[12] Xu C, Xiong C, Corso JJ, Streaming hierarchical video segmentation, InEuropean Conference on Computer Vision. (2012) 626-639.

[13] Endres I, Hoiem D, Category independent object proposals,InEuropean Conference on Computer Vision. (2010) 575-588.

[14] B. Alexe, T. Deselaers, and V. Ferrari, What is an object, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. (2010) 73-80.

[15] Ma T, Latecki LJ, Maximum weight cliques with mutex constraints for video object segmentation, In2012 IEEE Conference on Computer Vision and Pattern Recognition. (2012) 670-677).

[16] Teju V, Bhavana D. An efficient object detection using OFSA for thermal imaging. The International Journal of Electrical Engineering & Education. 2020.

[17] Lakshmi AV, Gopitilak V, Parvez MM, Subhani SK, Ghali VS, Artificial neural networks based quantitative evaluation of subsurface anomalies in quadratic frequency modulated thermal wave imaging, Infrared Physics & Technology .97. (2019) 108-15.

[18] Subhani S, Ghali VS, Measurement of thermal diffusivity of fiber-reinforced polymers using quadratic frequency modulated thermal wave imaging, Infrared Physics & Technology.99 (2019) 187-92.

[19] Suresh B, Subhani S, Vijayalakshmi A, Vardhan VH, Ghali VS, Chirp Z transform-based enhanced frequency resolution for depth resolvable nonstationary thermal wave imaging, Review of Scientific instruments. 88(1) (2017) 014901.

[20] Suresh B, Subhani S, Ghali VS, Mulaveesala R, Subsurface detail fusion for anomaly detection in non-stationary thermal wave imaging, Insight-Non-Destructive Testing, and Condition Monitoring. 59(10) (2017) 553-8.

[21] Rao GA, Kishore PV, Selfie video-based continuous Indian sign language recognition system, Ain Shams Engineering Journal. 9(4) (2018) 1929-39.

[22] Rao GA, Kishore PV, Selfie sign language recognition with multiple features on adaboost multilabel multiclass classifier, Journal of Engineering Science and Technology. 13(8) (2018) 2352-68.

[23] Sunitha RA, Suman M, Kishore PV, Eepuri KK, Sign language recognition with multi-feature fusion and ANN classifier, Turkish Journal of Electrical Engineering and Computer Science, 26(6) (2018) 2871-2885.

[24] Kumar KV, Kishore PV, Anil Kumar D, Indian classical dance classification with Adaboost multiclass classifier on multi-feature fusion.,Mathematical Problems in Engineering. (2017).

[25] Ahammad SH, Rajesh V, Rahman M, Fast and accurate feature extraction-based segmentation framework for spinal cord injury severity classification,.7(2019) 460 92-46103.

[26] Dheeraj D, Prasantha H S, DR-UNET: A Hybrid Model For Classification of G lioma using Transfer Learning on MR Images, International Journal of Engineering Trends and Technology, 69(10) (2021) 146-150.

[27] Suman Kumar Bhattacharyya, Sagarika Pal, Subrata Chattopadhyay, Classification and Quality Analysis of Rice Grain Based on Dimensional Measurement During Hydrothermal Treatment, International Journal of Engineering Trends and Technology.69(8) (2021)145-154.

[28] Mochammad Haldi Widianto, Ranny, Theodorus Ezra Suherman, Jefry Chiedi, Internet of Things for Detection Disaster Combined with Tracking AR Navigation, International Journal of Engineering Trends and Technology, 69(8) (2021)211-217.

[29] P. Suguna, B. Kirubagari, R. Umamaheswari, An Effective Cluster-Based Outlier Detection with Optimized Deep Neural Network for Epileptic Seizure Detection and Classification Model, International Journal of Engineering Trends and Technology, 69(3),(2021)76-84.