International Journal of Engineering Research and Technology.
ISSN 0974-3154 Volume 11, Number 11 (2018), pp. 1821-1836

http://www.irphouse.com

Fast Compressive Tracking Parameters Optimization for Enhanced Visual Tracking

Law Hooi Mee[1] and Parvathy Rajendran [2]*
[1,2] School of Aerospace Engineering, Universiti Sains Malaysia, Engineering Campus, 14300 Nibong Tebal, Pulau Pinang, Malaysia
** Corresponding author*

Abstract

Visual tracking is a hot issue in computer vision with a wide range of applications such as intelligent video surveillance. Different types of state-of-the-art methods such as self-expressive, sparse representation are proposed and several modifications are being introduced in public for respectively methods in the few decades ago. Unfortunately, due to some destabilizing factors, like occlusions and illumination changes which cause the tracking result tend to drift, until now, there are still no foolproof methods for visual tracking. This paper propose a model to enhance visual tracking by introduce the concept of regression analysis in visual tracking for optimization purpose. Instead of using random and fix values, the relationship of the dimensionality of projected space and learning parameter with the image frame size and the tracked target size has been studied. The main idea in the proposed method is to improve the overall tracking performance by increase the average overlap rate (AOR) and decrease the centre location error (CLE). The proposed method is evaluated using image sequences from various datasets; such as Babenko datasets and Kwon datasets. The AOR value has increased to 0.67 and the CLE value is decreased to 13.11. The proposed model performs favourably against state-of-the-art tracking methods on sequences in term of accuracy and robustness.

Keywords:

Nomenclature:

w : Width of image
h : Height of image
$\mu_i^1(\mu_i^0)$: Mean of the positive (negative) class
$\sigma_i^1(\sigma_i^0)$: Standard deviation of the positive (negative) class
w_1 : Tracked target width
h_1 : Tracked target height
x_t, y_t : Tracked target position
x_g, y_g : Ground truth
ROI_t : Tracking bounding box
ROI_g : Ground truth bounding box
IV : Illumination variation
SV : Scale variation

OCC : Occlusion
DEF : Deformation
MB : Motion blur
FM : Fast motion
IPR : In-plane rotation
OPR : Out-of-plane rotation
OV : Out-of-view
BC : Background clutters
LR : Low resolution

1 INTRODUCTION

Visual tracking is an important and hot research topic in computer vision with wide applications ranging from surveillance to human interactions to medical imaging such as intelligent video surveillance, automatic driving, intelligent robot, mission analysis & recognition, etc. Given the initial state; location & size of the target object, visual tracking is aim to estimate the target states in subsequent frames. Today, visual tracking is still a challenging issue to handle as there are destabilizing factors which causes drifting and overfitting problem to happen during tracking process. There are 11 attributes stated at [1] as shown in Table 1 which influence tracking performance. The most common attributes are illumination variation, background clutters, occlusion and fast motion. Until now, there is no single tracker able to deal with all these attributes at the same time.

Table 1. Annotated Sequence Attributes with the Threshold Values in the Performance Evaluation [1]

Name	Description
Illumination Variation	the illumination in the target region is significantly changed
Scale Variation	the ratio of the bounding boxes of the first frame and the current frame is out of the range ts, ts > 1 (ts=2)
Occlusion	the target is partially or fully occluded
Deformation	non-rigid object deformation
Motion Blur	the target region is blurred due to the motion of target or camera
Fast Motion	the motion of the ground truth is larger than tm pixels (tm=20)
In-Plane Rotation	the target rotates in the image plane
Out-of-Plane Rotation	the target rotates out of the image plane
Out-of-View	some portion of the target leaves the view
Background Clutters	the background near the target has the similar color or texture as the target
Low Resolution	the number of pixels inside the ground-truth bounding box is less than tr (tr =400)

In the past few decades, different type of state-of-the-art method and modifications have been proposed in order to conquer the problem. Image features model, target appearance model and motion model are examples have been introduced for visual tracking. There are three typical image features: first, colour features (e.g. Colour histogram) which requires low computational cost and are invariant to point-to-point transformations. However, they are not performed well against illumination changes. Second, texture features (Local Binary Patterns (LBP)) which well at discriminate the target from background but required high computational cost. Third, shape features (contour) which are discriminative as well but couldn't represent the target well. They need to combine with others for high level contour tracking.

Target appearance model is one of the fundamental components in visual tracking. It plays an important role in determining the success of a tracking algorithm. Target appearance model detect the variation in target appearance and also background which include moving object during tracking. Chen et al. [1] had combined the image sets and depth information into a 3D object tracking method to conquer the drift and occlusion problem. Based on their studies, a more effective appearance model can be presented with useful data variability information contained in image sets. Depth information also used by Cao et al. [2] who proposed a 3D novel object tracking method which using local patch-based appearance model to handle the deformable target and depth information to design a scheme for partial occlusions problem.

Appearance model can be categorized into either generative or discriminative based on whether the background information of the sequences is being used. Discriminative approach is classified as tracking-by-detection method which formulates tracking as a binary classification problem. Discriminative method separate target from the background and focus on the difference between the foreground and background of the image sequences. Classification score is used to determine the most likely target position, the higher the score, the higher the probability the image region is regarded as target. Superpixel method and multiple instance learning (MIL) are example of discriminative method.

Xu et al. [3] have introduced multiple instance learning (MIL) into visual tracking for dealing drifting problem. They utilized Fisher information criterion as it produced more informative for target from complex background. A tracking method through strong temporal slowness constraint and stacked convolutional autoencoders is proposed Kuen et al. [4] to learn complex-valued invariant representations from tracked sequential image patches. Cong et al. [5] presented an online metric learning (OML) tracker which incorporates adaptive metric learning and semisupervised learning into a unified framework. They design via low rank constraint to overcome overfitting problem. Sun et al. [6] proposed a novel approach to non-rigid objects contour tracking which extracts the accurate contours of the target for better description via supervised level sets model (SLSM).

Zhang et al. [7] presented a novel tracking method through multi-view learning framework which lead to more accurate and robust representation of target by using multiple support vector machines (SVM). Yin et al. [8] method directly learns and predicts the object's states. They define the objects' most confident state to combine the state-based structures support vector machine (SVM) and increment principal component analysis (PCA). They have improve the accuracy and reduce the noise interference.

Generative approach utilise template or subspace to represent the target candidates, it calculates the similarity of the tracked object and object candidates image region. Generative method use only the information of the tracked object without considering its background. It usually construct appearance model with image observations and searching the image region being generated that having the highest probability to be the true object targets. Sparse representation is one of the generative approach have been proposed.

You et al. [9] proposed a novel tracking method based on local sparse and global constraint which improves the local sparse method by adding a constraint on the distribution of the local contributions. Wang et al. [10] proposed a sparsity-based tracking method which is feature with inverse sparse representation formulation and a locally weighted distance metric. In order to adapt the appearance variations during tracking, Bai et al. [11] proposed a novel appearance method using sparse representation and online dictionary learning techniques.

Wang et al. [12] method is based on joint optimization of representation and classification and it has a well and fast performance even the target undergo several kinds of appearance variations. Wang et al. [13] proposed a tracking method by modelling targets with online-learned sparse features and classified it using Bayesian classifier to select the most likely target candidate by a binary classification process. Hu et al. [14] proposed a tracking algorithm based on a multi-feature joint sparse representation for tracking multi-objects under occlusions using a multi-feature sparse reconstruction.

A robust combination of particle filter and reverse sparse representation method is proposed by Yi et al. [15] which improve the tracking result with occlusions, background and illumination changes. A consistent low-rank sparse tracker (CLRST) which builds upon the particle filter framework has been proposed by Zhang et al. [16] for tracking. It will adaptively prunes and selects candidate particles by exploiting temporal consistency. Another which also combined these two is Cheng et al. [17] who integrate generative and discriminative approaches under particle filter framework. They encode all the tasks simultaneously in a structures multi-task learning manner. Hybrid method is used by Zhang et al. [18], a more compact spatial and temporal structure constraint ban be exploited and both generative and discriminative are utilized.

Compressive tracking is also introduced for visual tracking which extracts the features of the target object in the compressed domain by using a random measurement matrix. Compressive sensing theory suggests that if a signal is sparse or compressive, the original signal can be reconstructed by utilizing a few measured values. Chen et al. [19] proposed an online semi-supervised compressive algorithm for robust visual tracking. They introduced the weighted random projection into an adaptive compressive sensing based appearance model to obtain both local and discriminative information of the target. Chen et al. [20] suggest an algorithm to form a localized compressive sensing representation by collecting both local and global information which able to formulates the partial appearance model discriminatively. This kind of appearance model enables fast visual tracking in an intrinsic low-dimensional feature space.

Zhou et al. [21] have based on manifold ranking in designing a novel and robust tracking method. Based on CS theory, they implement non-adaptive random projections to preserve the structure of original image space and competently extracted low-dimensional compressive features using a very sparse measurement matrix for object representation. Besides, they utilised spatial context to improve the robustness to appearance variations. Wu et al. [22] proposed a multi-scale compressive tracker. They integrated an appearance model based on rectangle features extracted in the adaptive compressive domain into bootstrap filter. It increase the efficiency in feature extraction without increase the complexity.

A low-dimensional compressive features proposed by Zhang et al. [23] for the appearance modelling is utilised for doing the optimization part. Haar-like features which require high computational loads for feature extraction is compressed using a very sparse measurement matrix. The compressed features help in preserves the object structure of the original image space and can be applied efficiently in tracking. This model able to track the target object in a low computational cost, however the accuracy of the tracking result has been limited by using random and fix parameters such as learning parameter for all sequences in tracking. As there is potential for improvement, this model is chose as the core algorithm in this paper.

2 METHODOLOGY

In this section, the basic concept of the fast compressive tracking algorithm [23] will be discussed. The concept of the compressive sensing is discussed follow by the way to represent the image in this algorithm. Then, the classification method utilised to classify and determine the target location with maximum classification response is presented. A flow chart of the proposed method is showed and the proposed method will be highlighted. The concept of regression analysis and the proposed optimization method are discussed to emphasize the enhanced tracking method.

2.1 General Method

The original concept of compressed sensing was proposed by Davis L. Donoho [24]. The two important concepts stated in compressive sensing are ‘random projection’ and ‘random measurement matrix’.

In random projection, a random measurement matrix $\boldsymbol{R} \in \mathbb{R}^{nxm}$ projects a high-dimensional image feature $\boldsymbol{x} \in \mathbb{R}^{m}$ to a low-dimensional feature $\boldsymbol{v} \in \mathbb{R}^{n}$

$$v = Rx \tag{1}$$

where $n \ll m$. Baranuik et al. proved that the random matrix satisfying the Johnson-Lindenstrauss lemma holds true for the restricted isometry property in CS theory.

Therefore, the low-dimensional features can reconstruct from the original high-dimensional information with minimum error.

A typical measurement matrix satisfying the restricted isometry property is the random normal distribution $\boldsymbol{R} \in \mathbb{R}^{nxm}$ where $r_{ij} \sim \mathcal{N}(0,1)$, so a very sparse random measurement matrix is adopted with entries defined as

$$r_{ij} = \sqrt{\rho} \times \begin{cases} 1, & with\, probability\ \dfrac{1}{2\rho} \\ 0, & with\, probability\ 1 - \dfrac{1}{\rho} \\ -1, & with\, probability\ \dfrac{1}{2\rho} \end{cases} \tag{2}$$

where $\rho = \frac{m}{4} \sim \frac{m}{2.4}$ is set in this work. This can be efficiently computed for real-time tracking.

2.1.1 Image Representation

For each sample $\boldsymbol{Z} \in \mathbb{R}^{wxh}$, it is represented as the convolution of $\boldsymbol{Z}$ with multi-scale filters $\{\boldsymbol{F}_{1,1}, \dots, \boldsymbol{F}_{w,h}\}$ and defined by

$$\boldsymbol{F}_{w,h}(x,y) = \frac{1}{wh} \times \begin{cases} 1, & 1 \le x \le w, 1 \le y \le h \\ 0, & otherwise \end{cases} \tag{3}$$

Each filtered image is represented as a column vector of $\mathbb{R}^{wxh}$and concatenate these vectors to form a high-dimensional multi-scale image feature vector $\boldsymbol{x} = (x_1, \dots, x_m)^{\intercal} \in \mathbb{R}^m$ where $m = (wh)^2$. The dimensionality m is typically in the order of 10^6 to 10^{10}. Therefore, a sparse random matrix $\boldsymbol{R}$ in Eq. (2 is used to project the vector $\boldsymbol{x}$ onto a low-dimensional vector $\boldsymbol{v}$.

2.1.2 Classification

A naive Bayes classifier is used with the assumption of a uniform prior, $p(y = 1) = p(y = 0)$ and the sample label $y \in \{0,1\}$ is a binary variable.

$$\begin{aligned} H(\boldsymbol{v}) &= \log\left(\frac{\prod_{i=1}^{n} p(v_i|y=1)p(y=1)}{\prod_{i=1}^{n} p(v_i|y=0)p(y=0)}\right) \\ &= \sum_{i=1}^{n} \log\left(\frac{p(v_i|y=1)}{p(v_i|y=0)}\right) \end{aligned} \tag{4}$$

It is assumed that the conditional distributions $p(v_i|y = 1)$ and $p(v_i|y = 0)$ of the classifier $H(\boldsymbol{v})$ are Gaussian distributed with four parameters $(\mu_i^1, \sigma_i^1, \mu_i^0, \sigma_i^0)$

$$p(v_i|y = 1) \sim \mathcal{N}(\mu_i^1, \sigma_i^1), p(v_i|y = 0) \sim \mathcal{N}(\mu_i^0, \sigma_i^0) \tag{5}$$

The learning parameter λ enable the scalar parameters in Eq. (5 incrementally updated with following equation

$$\begin{aligned} \mu_i^1 &\leftarrow \lambda\mu_i^1 + (1-\lambda)\mu^1 \\ \sigma_i^1 &\leftarrow \sqrt{\lambda(\sigma_i^1)^2 + (1-\lambda)(\sigma^1)^2 + \lambda(1-\lambda)(\mu_i^1 - \mu^1)^2} \end{aligned} \tag{6}$$

where

$$\sigma^1 = \sqrt{\tfrac{1}{n}\textstyle\sum_{k=0|y=1}^{n-1}(v_i(k) - \mu^1)^2} \text{ and } \mu^1 = \tfrac{1}{n}\textstyle\sum_{k=0|y=1}^{n-1} v_i(k).$$

In the basic model, the dimensionality of projected space n is set as 100 for the random measurement matrix and the learning parameter λ is set as 0.85 to update the mean and standard deviation for classification during the entire tracing process. These two parameters have been fixed for all the sequences which border the performance of the tracking algorithm and there is still space for improvement.

2.2 Proposed Method

As there is potential in improvement, the proposed model is introduced to optimize visual tracking by improving the two parameters as mentioned in previous subsection. In this subsection, the flow chart of the proposed model will be discussed and the steps to track the target object are demonstrated. Details of the regression analysis carried out in this proposed model will be explained and the way to evaluate the performance of the tracking method is showed.

2.2.1 Flow Chart

Instead of fixing the space dimensionality and learning parameter, in our proposed method, these two parameters will change accordingly with sequence based on their frame size and target size of the image frame with equation Eq. (7) and (8). Below show the flow chart of the proposed method,

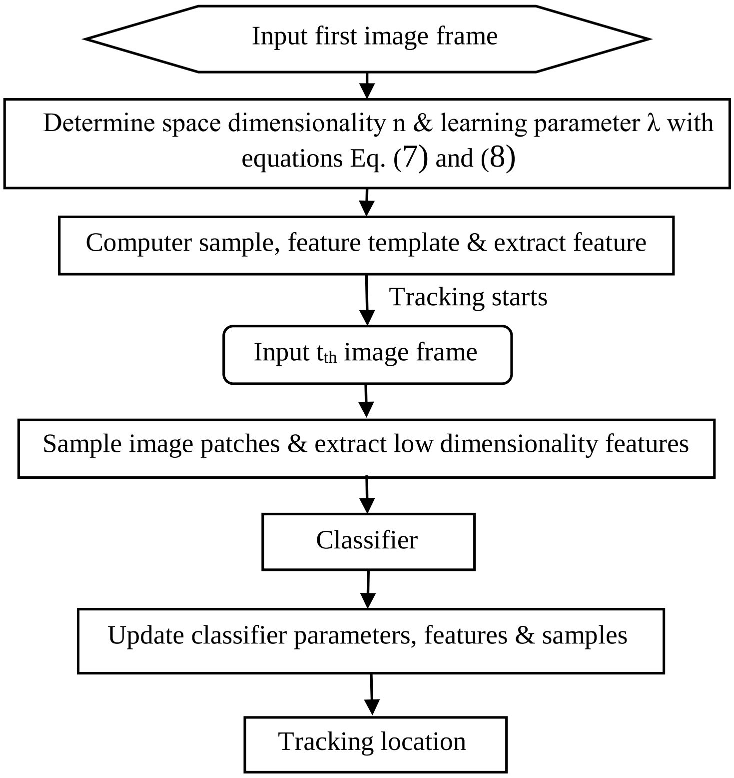

Figure 1. Flow Chart of enhanced FCT

Once the first image is read, the frame size and target size will be determined and space dimensionality Eq. (8) and learning parameter Eq. (7) will be calculated. Then, based on the initial parameter, the sample template and feature template are computed and feature is extracted for following tracking purpose. The tracking process is started after the following image is read, coarse to fine sampling is undergo in this algorithm to reduce computational complexity. Similarly, both will sample a set of image patches and extract the feature with low dimensionality, then classifier Eq. (4) is used for each feature extract and getting the tracking location with maximum classifier response. The process will continue until the last image frame. Each tracking location is recorded for performance evaluation purpose.

Below show the equations for the two parameters which obtained through regression analysis,

$$\begin{aligned}\lambda = 1.773 + 0.000195h + \frac{68.21}{h} + 0.0699\cos(h_1) \\ + 0.000331w_1h_1 + 1.021e15hw_1 - \frac{0.133h_1}{w_1} \\ - 0.00653h_1 - 0.0384w_1\end{aligned} \tag{7}$$

$$
\begin{aligned}
n = 656 + 0.143h_1^2 + h_1 \cos(1.87 + 39.69h_1) \\
- \frac{1370 \sin(3.45w_1)}{h_1 \cos(h_1)} - 16.93h_1 \\
- 228.9 \cos(h_1) \sin(3.448w_1)
\end{aligned} \tag{8}
$$

2.2.2 Regression Analysis

In order to determine the natural relations between the existing parameters, Schmidt et Lipson [25] proposed a principle for the identification of nontriviality. There are many existing data modelling methods, such as fixed-form parametric model and numerical models. Schmidt et Lipson used symbolic regression method for searching the mathematical expression with minimise error.

Symbolic regression which able to develop an analytic model which is useful for predicting response behaviours and summarize all data, is used to determine the symbolic function to describe the data effectively. Different from traditional regression method which fit parameters into a given formed equation, symbolic regression searches both the parameters and the form of equations simultaneously. They randomly combining mathematical building blocks like algebraic operators and analytical functions to form the initial expressions and continuously update the equation by recombining previous equations and varying their subexpressions. When a desired accuracy is reached, the algorithm will terminate and returning a set of equation which best describe the parameters relations.

To enhance the visual tracking, the relationship between the image frame sizes and target tracking sizes with dimensionality of projected space and learning parameters has been studied for the sequences from various datasets, such as Babenko datasets, Ross datasets and Kwon datasets as shown in Table 2. At first, the original setting of dimensionality of projected space, $n = 100$ is fixed to test with the sequences for learning parameter ranging from 0 to 1. The performance of each sequences evaluated and studied. From this study, the best performance of the sequences fall on the learning parameter ranging from 0.5 to 1.

Therefore, by setting with this range of learning parameter $\lambda = (0.5\sim 1)$, the space dimensionality ranging from 50 to 500 with interval of 50 is studied and recorded. From the result for all combination of these two parameters, the global best performance of the tracking result for each sequence with the particular space dimensionality and learning parameter are recorded. With this information, the natural relation of the parameters with image frame sizes and target tracking sizes is obtained through Eureqa [25]. Eureqa is a software tool which is used to identify the mathematical formulae that best describe a data set. It employs symbolic regression to determine the best-fitting functional equation. The result from regression analysis and the improved tracking result will be discussed in section 3.

Table 2. The information of sequences

Sequence	Frame Size	Target Size	Attributions
Bolt2	270 x 480	34 x 64	DEF, BC
CarDark	240 x 320	29 x 23	IV, BC
ClifBar	240 x 320	30 x 54	SV, OCC, MB, FM, IPR, OV, BC
Coupon	240 x 320	57 x 89	OCC, BC
Deer	400 x 704	95 x 65	MB, FM, IPR, BC, LR
Faceocc2	240 x 320	82 x 98	IV, OCC, IPR, OPR
Football	352 x 624	39 x 50	OCC, IPR, OPR, BC
Panda	233 x 312	28 x 23	SV, OCC, DEF, IPR, OPR, OV, LR
Shaking	352 x 624	61 x 71	IV, SV, IPR, OPR, BC
Skating1	360 x 640	34 x 84	IV, SV, OCC, DEF, OPR, BC
Sylvester	240 x 320	51 x 61	IV, IPR, OPR
Tiger1	480 x 640	76 x 84	IV, OCC, DEF, MB, FM, IPR, OPR
Tiger2	480 x 640	68 x 78	IV, OCC, DEF, MB, FM, IPR, OPR, OV

2.2.3 Performance Evaluation

To evaluate the tracking performance, we take centre location error (CLE) and average overlap ratio (AOR) as the parameters to evaluate qualitatively and compare the performance of the proposed algorithm with other algorithms. CLE is computed based on the ground truth of the target at the pixel level. It measured the Euclidean distance between the central locations of the tracked object and the ground truth.

$$CLE = \sqrt{\left(x_t - x_g\right)^2 + \left(y_t - y_g\right)^2} \tag{9}$$

The smaller the CLE indicates that the shorter the distance different between the tracked target and ground truth, the better the tracking performance. AOR is the indication of extent of region overlapping between tracking results and ground truths. The tracking result is considered as success if only if AOR score is more than 0.5.

$$AOR = \frac{area\left\{ROI_t \cap ROI_g\right\}}{area\left\{ROI_t \cup ROI_g\right\}} \tag{10}$$

The values of AOR range from 0 to 1. The larger the values of AOR, the higher the accuracy of the tracking results.

3 RESULT

This section will present the regression analysis result and tracking result of the proposed algorithm on the available image sequences as stated in Table 2. The best n

and λ for each sequences to obtain the global best performance is showed and the root mean square error and coefficient of determination of n and λ for each sequence are plotted. The tracking result is showed and compare with the original model and also other algorithms.

3.1 Regression Analysis

The optimization process has been carried out with the image sequences for dimensionality of projected space, $n(50, \ldots, 500)$ and learning parameter $\lambda = (0.5 \sim 1)$ for all image sequences. Through this process, the global best performance for each sequences with its own λ and n has been obtained. The equation to describe the relationship of these two parameters with image frame size and target size has been obtained through Eureqa. The result computed from the Eq. (7) and (8) which is the best fit line to describe the data set that can be determined.

The root mean square error is useful in showing how well the predicted values correspond to the actual data which attempt to model. In order to facilitates the comparison between the sequences, the normalized root mean square error (nRMSE) for the n and λ are computed. With lower nRMSE indicates less residual variance, comparatively, learning parameter having a better predicted values as most of the sequences, nRMSE shows value close to 0. While for dimensionality of projected space, nRMSE is relatively high but all the nRMSE values are less than 0.12.

The coefficient of determination is a key output of regression analysis. It is understood as the proportion of the variance in the dependent variable that is predictable from the independent variable. It ranges from 0 to 1 with 1 indicates the dependent variable can be predicted without error from the independent variable. The learning parameter in all image sequences as it shows the value close to 1. While for dimensionality of projected space, most of the sequences show a value close to 1. With nRMSE close to 0 and R^2 close to 1, it shows the learning parameter and dimensionality of projected space can be predicted with a very little error from the frame sizes and tracking target sizes.

3.2 Tracking Result

Based on the Eq. (7) and (8) obtained through Eureqa, the proposed model is evaluated and compared with the original algorithm which fix the learning parameter and dimensionality of projected space. As stated in Table 3, 10 out of 13 image sequences has showed an improvement in their tracking result based on the performance evaluation parameters, AOR and CLE. The best improvement in percentage for AOR is as high as 446.67% for CarDark sequence, at the same time, CarDark sequence also show the highest percentage improvement in CLE which is 95.71%. Besides, the average performance of these two algorithm is also computed and compared. Both AOR and CLE also show improvement based on their average performance. AOR has improved about 36.73% from 0.49 increase to 0.67 while in CLE, 45.10% has been improved as it has decrease from 23.88 to 13.11 in pixels.

Furthermore, the proposed algorithm is also evaluated with other 5 state-of-the-art tracking methods on the image sequences. The 5 evaluated trackers are distribution field (DF) [26] tracker, multi-task tracker (MTT) [27], circulant structure tracker (CST) [28], sparsity-based collaborative model (SCM) [29] tracker and adaptive structural local sparse appearance (ASLA) [30]. The tracking result of these tracking algorithms are obtained from [23]. Table 4. shows the comparison of the AOR results on the image sequences for the six methods, the five methods mentioned above and the original FCT method with the proposed algorithm. Comparing for all the image sequences, the proposed methods only performed well in 2 out of 13 sequences with the highest AOR value. However, as stated in Table 5 which shows the comparison of the CLE result, the proposed method has performed well in most of the image sequences. The proposed method having the lowest CLE in 8 out of 13 sequences. It shows the proposed model have improved the overall performance of visual tracking.

Table 3. Comparison of result between original FCT with proposed FCT

	AOR			ACLE		
	FCT	FCTop	%	FCT	FCTop	%
Bolt2	0.59	0.69	16.95	9.99	7.38	26.13
CarDark	0.15	0.82	446.67	46.34	1.99	95.71
Clifbar	0.55	0.56	1.82	11.51	10.87	5.56
Coupon	0.60	0.77	28.33	19.51	8.76	55.10
Deer	0.68	0.73	7.35	10.92	8.64	20.88
Faceocc2	0.65	0.75	15.38	20.92	13.86	33.75
Football	0.46	0.67	45.65	17.36	8.24	52.53
Panda	0.49	0.56	14.29	8.11	6.44	20.59
Shaking	0.36	0.72	100.00	36.27	8.39	76.87
Skating1	0.25	0.36	44.00	68.02	59.79	12.10
Sylvester	0.68	0.70	2.94	8.47	9.35	-10.39
Tiger1	0.58	0.73	25.86	20.08	11.05	44.97
Tiger2	0.38	0.60	57.89	32.92	16.71	49.24
Average	0.49	0.67	36.73	23.88	13.11	45.10

Table 4. Average overlap ratio (AOR) comparison

	FCT	FCTop	DF	MTT	CST	SCM	ASLA
Bolt2	0.59	0.69	0.01	0.01	**0.92**	0.01	0.01
CarDark	0.15	**0.82**	0.78	0.59	0.48	0.47	0.57
Clifbar	0.55	0.56	0.26	0.55	**0.96**	0.41	0.40
Coupon	0.60	0.77	0.34	0.39	**0.81**	0.62	0.71
Deer	0.68	0.73	0.06	0.87	**1.00**	0.98	0.96
Faceocc2	0.65	0.75	0.78	0.88	**0.99**	0.76	0.93
Football	0.46	0.67	0.56	0.67	**0.69**	0.17	0.07
Panda	0.49	0.56	0.13	0.11	0.15	0.29	**0.71**
Shaking	0.36	0.72	0.84	0.02	0.36	0.54	**0.98**
Skating1	0.25	0.36	0.19	0.10	0.09	**0.76**	0.61
Sylvester	0.68	0.70	0.32	0.67	**0.83**	0.76	0.82
Tiger1	0.58	**0.73**	0.36	0.25	0.42	0.31	0.14
Tiger2	0.38	0.60	**0.65**	0.34	0.37	0.02	0.24

Table 5. Centre location error (CLE) comparison

	FCT	FCTop	DF	MTT	CST	SCM	ASLA
Bolt2	9.99	**7.38**	277	293	12	200	210
CarDark	46.34	**1.99**	6	7	8	45	8
Clifbar	11.51	10.87	52	25	**7**	99	49
Coupon	19.51	**8.76**	23	72	21	73	23
Deer	10.92	**8.64**	252	17	15	16	13
Faceocc2	20.92	13.86	22	19	**13**	24	20
Football	17.36	**8.24**	33	9	17	200	207
Panda	8.11	**6.44**	64	47	46	156	9
Shaking	36.27	**8.39**	10	115	21	47	10
Skating1	68.02	59.79	174	78	**10**	42	72
Sylvester	8.47	9.35	56	18	**8**	10	9
Tiger1	20.08	**11.05**	30	61	25	146	49
Tiger2	32.92	16.71	**13**	24	22	230	36

Table 6 shows the average performance of the tracking result for all the seven methods as listed. The proposed method achieved the highest in both AOR and CLE average value of 0.67 and 13.11 respectively. The proposed method is only show little enhancement from CST method by 10.64% in AOR value and 4.24% in CLE value. In term of AOR value, the proposed method outperforming the DF with the highest improvement of 61.90%, follow by MTT with 40.32%, SCM with 36.36%, FCT with 26.87% and ASLA with 24.49%. For CLE average value comparison, the highest improved in performance is achieved with 273.08% if compared with MTT. The proposed method has also outperformed the DF, ASLA and SCM methods with a very high improved in percentage which is 107.22%, 175.42% and 156.31% respectively.

Table 6. Average result comparison for AOR and CLE

	AOR		CLE	
	Average	Diff in %	Average	Diff in %
DF	0.41	-61.90	77.85	107.22
MTT	0.42	-40.32	60.38	273.08
CST	0.62	-10.64	17.31	4.24
SCM	0.47	-36.36	99.08	156.31
ASLA	0.55	-24.49	55	175.42
FCT	0.49	-26.87	23.88	82.15
FCTop	**0.67**	-	**13.11**	-

4 CONCLUSION

Despite of various tracking methods have been proposed, visual tracking is still a challenging issue due to the attributes causing the tracking tend to drift. In this paper, a model including the concept of regression analysis was proposed and it enhances the overall performance of visual tracking. The relationship of the dimensionality of projected space and learning parameter with the frame size and tracked target size was studied. Instead of using random and fix values, the proposed model implemented the formulae generated through Eureqa which best describe the relationship of those parameters during tracking. The proposed model improved the AOR and CLE value in most of the sequences and the average AOR is increased to 0.67 and the average CLE is decreased to 13.11. The performance of the proposed method was compared with others tracking methods. Result shows the proposed algorithm performs favourably against several state-of-the-art tracking algorithms. The enhanced FCT model improves and shows a better performance in tracking result.

5 ACKNOWLEDGEMENTS

This research and publication were supported by Universiti Sains Malaysia Grant No. 1001/PAERO/ 814276.

REFERENCE

[1] Y. Chen, Y. Shen, X. Liu, and B. Zhong, "3D object tracking via image sets and depth-based occlusion detection," *Signal Processing*, vol. 112, pp. 146–153, 2015.

[2] L. Cao, C. Wang, and J. Li, "Robust depth-based object tracking from a moving binocular camera," *Signal Processing*, vol. 112, pp. 154–161, 2015.

[3] C. Xu, W. Tao, Z. Meng, and Z. Feng, "Robust visual tracking via online multiple instance learning with Fisher information," *Pattern Recognit.*, vol. 48, no. 12, pp. 3917–3926, 2015.

[4] J. Kuen, K. M. Lim, and C. P. Lee, "Self-taught learning of a deep invariant representation for visual tracking via temporal slowness principle," *Pattern Recognit.*, vol. 48, no. 10, pp. 2964–2982, 2015.

[5] Y. Cong, B. Fan, J. Liu, J. Luo, and H. Yu, "Speeded Up Low-Rank Online Metric Learning for Object Tracking," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 6, pp. 922–934, 2015.

[6] X. Sun, H. Yao, and S. Zhang, "Non-Rigid Object Contour Tracking via a Novel Supervised Level Set Model," vol. 24, no. 11, pp. 3386–3399, 2015.

[7] S. Zhang, X. Yu, Y. Sui, S. Zhao, and L. Zhang, "Object Tracking With Multi-View Support Vector Machines," vol. 17, no. 3, pp. 265–278, 2015.

[8] Y. Yin, D. Xu, X. Wang, and M. Bai, "Online State-Based Structured SVM Combined With Incremental PCA for Robust Visual Tracking," *IEEE Trans. Cybern.*, vol. 45, no. 9, pp. 1988–2000, 2015.

[9] X. You, X. Li, Z. He, and X. F. Zhang, "A robust local sparse tracker with global consistency constraint," *Signal Processing*, vol. 111, pp. 1–11, 2014.

[10] D. Wang, H. Lu, Z. Xiao, and M.-H. Yang, "Inverse Sparse Tracker With a Locally Weighted Distance Metric," *Appl. Soft Comput. J.*, vol. 24, no. 9, pp. 2646–2657, 2015.

[11] T. Bai, Y. Li, S. Member, and X. Zhou, "Learning Local Appearances With Sparse Representation for Robust and Fast Visual Tracking," vol. 45, no. 4, pp. 663–675, 2015.

[12] Q. Wang, F. Chen, W. Xu, M. Yang, and S. Member, "Object Tracking With Joint Optimization of Representation and Classification," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 4, pp. 638–650, 2015.

[13] Z. Wang, J. Wang, S. Zhang, and Y. Gong, "Visual tracking based on online sparse feature learning," *Image Vis. Comput.*, vol. 38, pp. 24–32, 2015.

[14] W. Hu, W. Li, X. Zhang, and S. Maybank, "Single and Multiple Object Tracking Using a Multi-Feature Joint Sparse Representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 4, pp. 1–1, 2014.

[15] S. Yi, Z. He, X. You, and Y.-M. Cheung, "Single object tracking via robust combination of particle filter and sparse representation," *Signal Processing*, vol. 110, pp. 178–187, 2015.

[16] T. Zhang, S. Liu, N. Ahuja, M.-H. Yang, and B. Ghanem, "Robust Visual

Tracking Via Consistent Low-Rank Sparse Learning," *Int. J. Comput. Vis.*, vol. 111, no. 2, pp. 171–190, 2015.

[17] X. Cheng, N. Li, T. Zhou, L. Zhou, and Z. Wu, "Object tracking via collaborative multi-task learning and appearance model updating," *Appl. Soft Comput.*, vol. 31, no. 0, pp. 81–90, 2015.

[18] S. Zhang, Y. Sui, X. Yu, S. Zhao, and L. Zhang, "Hybrid support vector machines for robust object tracking," *Pattern Recognit.*, vol. 48, no. 8, pp. 2474–2488, 2015.

[19] S. Chen, S. Li, S. Su, D. Cao, and R. Ji, "Online semi-supervised compressive coding for robust visual tracking," *J. Vis. Commun. Image Represent.*, vol. 25, no. 5, pp. 793–804, 2014.

[20] C. Chen, S. Li, H. Qin, and A. Hao, "Real-time and robust object tracking in video via low-rank coherency analysis in feature space," *Pattern Recognit.*, vol. 48, no. 9, pp. 2885–2905, 2015.

[21] T. Zhou, X. He, K. Xie, K. Fu, J. Zhang, and J. Yang, "Robust visual tracking via efficient manifold ranking with low-dimensional compressive features," *Pattern Recognit.*, vol. 48, no. 8, pp. 2459–2473, 2015.

[22] Y. Wu, N. Jia, and J. Sun, "Real-time multi-scale tracking based on compressive sensing," *Vis. Comput.*, vol. 31, no. 4, pp. 471–484, 2015.

[23] K. Zhang, L. Zhang, and M. H. Yang, "Fast Compressive Tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 10, pp. 2002–2015, 2014.

[24] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[25] M. Schmidt and H. Lipson, "Distilling Free-Form Natural Laws from Experimental Data," *Science (80-.).*, vol. 324, no. 5923, pp. 81–85, 2009.

[26] L. Sevilla-Lara and E. Learned-Miller, "Distribution fields for tracking," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1910–1917, 2012.

[27] T. Zhang, B. Ghanem, S. Liu, and N. Ahuja, "Robust visual tracking via structured multi-task sparse learning," *Int. J. Comput. Vis.*, vol. 101, no. 2, pp. 367–383, 2013.

[28] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "Exploiting the circulant structure of tracking-by-detection with kernels," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7575 LNCS, no. PART 4, pp. 702–715, 2012.

[29] W. Zhong, H. Lu, and 2012 Ieee Conference on Ming-Hsuan Yang Computer Vision and Pattern Recognition CVPR, "Robust object tracking via sparsity-based collaborative model," *Cvpr*, pp. 1838–1845, 2012.

[30] X. Jia, H. Lu, and M. H. Yang, "Visual tracking via adaptive structural local sparse appearance model," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1822–1829, 2012.