Fast Template Matching using Wavelet Transform

Srivaishnavi Sree Krishnan, Shrinidhi Venkatakrishnan [1]

Abstract— **This paper presents new algorithms for analyzing the problems of template matching in real life scenarios, specifically focusing on real time videos. Firstly we present the solution to the case where the template is translated in the image using foreground estimation. Then, we use SURF descriptors for feature matching and Hadamard transform on image and template to to perform real-time scale, translation invariant moving object area detection.**

I. INTRODUCTION

Template matching involves defining a measure or a cost to find the kinship between the reference patterns and the test pattern by performing the matching operation. It finds its application in speech recognition, in automation using robot vision, in motion estimation for video coding, and in image database retrieval systems.

The conventional methods of template matching use cross correlation algorithms. Though the method is simple to implement and understand, it is one of the slowest methods [1]. The initial approach of this paper is to perform search area reduction using Wavelet based segmentation for foreground estimation in real-time videos to improve the efficiency of classical correlation based template matching.

However, a major shortcoming of correlation based methods for template matching is that it is highly influenced by scale and rotation. Thus, as a refined approach, we also propose an algorithm for a scale-rotation invariant template matching using wavelets.

Since the input images captured are usually affected by noise during storage, transmission and acquisition, image denoising is an issue found in diverse image processing and computer vision problems. Thus, we denoise the noisy input images using wavelet based thresholding techniques to improve the accuracy of feature matching.

Most image denoising methods smoothen the input image and fail to preserve edge information. Since feature matching algorithms are largely affected by the edge information, an important property of a good image denoising model is that it should remove noise as far as possible as well as preserve edges. The feature matching between denoised image and the original template is then performed using SURF based feature description and matching. Further, we localize the moving object area in case of real time videos by employing the scaling algorithm.

[1]Srivaishnavi Sree Krishnan, Shrinidhi Venkatakrishnan
Course Instructor: Truong Nguyen
Course: ECE 251C - Fall 2019
Link to GitHub Implementation

This paper is organized as follows: Section 2 provides various recent works on template matching approaches. In Sections 3 and 4, we elaborate on the proposed approaches for template matching and their technical details. Sections 5 provides a critical analysis of the various experiments conducted and the results obtained. Section 6 gives the summary and future work of the paper.

II. RELATED WORK

Many researches have been conducted and as many number of different template matching techniques have been proposed for various applications. Among these was the classical and simple normalised cross correlation scheme for reference image based object detection. Koprinska et. al (2001) [3] proposed a cross correlation based template matching as the basic statistical approach to image registration. Template is considered as a sub-image from the reference image, and the image is considered as a sensed image. The objective was to establish the correspondence between the reference image and sensed image using the degree of similarity between an image and template. This paper describes medical image registration by template matching based on Normalized Cross-Correlation (NCC) using Cauchy-Schwartz inequality. It is observed from the results that normalised cross correlation is computationally very expensive and in order to make it faster the search area in the sensed image has to be reduced.

Instead of just using the spatial domain components the idea could be extended for frequency domain as well. Ramadhan et al. (2017) [6] and multiple other literature show that Wavelet transform on the frequencies of sub-bands split from an image is a powerful method for analysis of images. Thus application of wavelets that could possibly reduce the search area for the classical approach of NCC was analyzed in our paper.

In the paper by Parida et al(2018)[7], transition region based approaches are used as recent hybrid segmentation techniques that are well known for its simplicity and effectiveness. Here, the segmentation effectiveness depends on robust extraction of transition regions. This paper proposed a transition region method which initially decomposes the gray image in wavelet domain to obtain a feature matrix. Using this feature matrix the corresponding prominent wavelet coefficients of different bands are found. Inverse wavelet transform are then applied on the modified coefficients to get edge image with more than one pixel width and Otsu thresholding is applied on it to get transition regions. This research motivated us to apply wavelet based filtering for segmenting the area of interest for template matching algorithm.

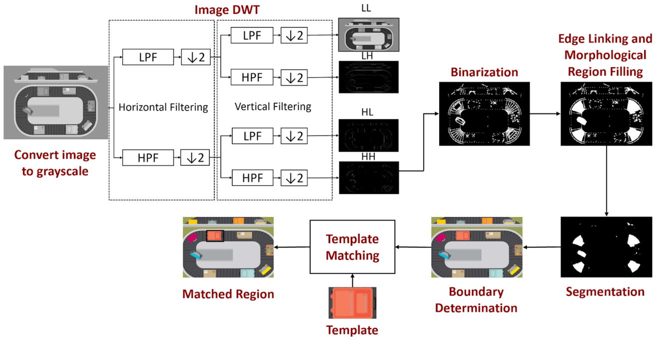

Fig. 1. Initial Approach: Segmentation based search area reduction

We also realised that template matching in real life should be scale, translation and rotation invariant which could be achieved through various wavelet based algorithms using the fine-course representation [8].

Jiang et al.(2012)[9] proposed a procedure based on wavelet based Speeded-Up Robust Features (SURF) to detect and describe the feature points of image sequence, used normalized correlation (NCC) for the initial match. Further, they eliminated mismatching points by using random sample consensus algorithm (RANSAC). Lastly, they used the least square method for precision matching. This paper inspired us to use a similar algorithm to detect and match features between template and the image.

Finally, Sauerwald (2012) [10] proposed a Hadamard transform based approach to detect the scale of a template in an image without translation. In this paper, we merge robust feature matching with this scaling algorithm to detect the moving object area of a template in an image.

III. INITIAL APPROACH

A. Segmentation based search area reduction

Figure 1 gives a summary of the various steps involved in the proposed initial approach for foreground estimation.

Discrete Wavelet Transform (DWT) is first performed on the input image to divide it into its components. The HH component is then passed to binarization step, followed by the edge linking and region filling that increase the efficiency of segmentation. The foreground boundary is chosen from the segmented region of interest. The template is finally matched with reduced boundary image using Normalized cross correlation.

B. Wavelet based filtering

A discrete wavelet transform is a wavelet transform for which the wavelets are discretely sampled. A major advantage it has over Fourier transforms is temporal resolution; it captures both frequency and location information [4] which are useful to extend to videos. The first step is to choose a wavelet type, and perform a level N decomposition. The rows of the array are processed first with only one level of decomposition. This essentially divides the array into two vertical halves, with the first half storing the average coefficients, while the second vertical half stores the detail coefficients. This process is repeated again with the columns, resulting in four sub-bands within the array defined by filter output [5]. In our approach, DWT is performed to get the high frequency components of the image in the vertical, horizontal and diagonal direction along with the low pass components.

C. Segmentation

In computer vision, image segmentation is the process of partitioning a digital image into multiple segments. The goal of segmentation is to simplify and change the representation of an image into something that is more meaningful and easier to analyze. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent sections of the segments are significantly different with respect to the same characteristic parameter[3].

In our algorithm, we try to find the boundaries and the foreground of images to decrease the search area for cross correlation. The diagonal components retain the edges in the vertical and the horizontal direction and hence this frequency component can be used to detect the needed information in terms of edges. This particular diagonal component is thus taken and binarized with the mean of the image as threshold. This is followed by edge linking and hole filling operations as pre-processing steps to get the segmented region (using global thresholding).

D. Foreground/Boundary Estimation

Taking the extremities of the filled region obtained from the above method, we can find the exact foreground. This is extracted for reduction search area for normalised cross correlation.

E. Principle of normalised cross correlation

Cross correlation is a measurement that tracks the movements of two variables or sets of data relative to each other [11]. In its simplest version, it can be described in terms of an independent variable, X, and two dependent variables, Y and Z. If independent variable X influences variable Y and the two are positively correlated, then as the value of X rises so will the value of Y.For image-processing applications in which the brightness of the image and template can vary due to lighting and exposure conditions, the images can be first normalized. This is typically done by subtracting the mean and dividing by the standard deviation. That is, the cross-correlation of a template is given by (1)

$$corr(x,y) = \sum_{n=0}^{n-1} x[n] * y[n] \tag{1}$$

IV. REFINED APPROACH

A. Feature Extraction based Template Matching: Overview

Since the initial approach for foreground estimation was not scale and rotation invariant, we propose a new approach that incorporates scale, translation invariance. The overview

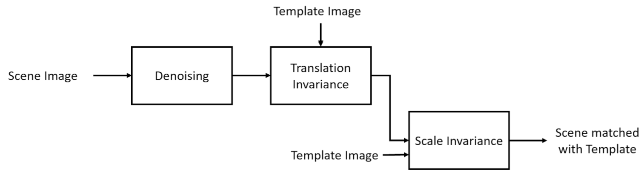

Fig. 2. Refined approach: Scale, translation invariant template matching

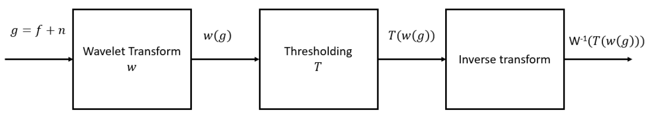

Fig. 3. Stages of denoising

of the approach is given in figure 3. The elementary step performs wavelet based denoising as a pre-processing step (Section II B). This is then followed by the determination of translation between image and template using SURF feature matching (Section II C). The template is then translated by appropriate coordinate and passed to the scaling algorithm (Section II D) that finds the corresponding bounding box for the template.

B. Denoising

Denoising is the first pre-processing step to improve input image prior to template matching. The image f was taken and gaussian white noise n of varying variance was added to it. Wavelet transform w that splits the image into the frequency channel using Haar/ db4 was implemented. This was further followed by certain thresholding methods, like the hard and soft thresholding and finally the inverse wavelet transform was employed to get back the denoised image. The whole idea of denoising is to preserve the features and at the same time remove the noise.

*1) **Haar vs db4**:* Haar wavelets are memory efficient, exactly reversible without the edge effects characteristic of other wavelets and computationally cheap [12]. It has no overlapping windows and takes only the changes between the adjacent pixel pairs. It uses two scaling and wavelet function coefficients. db4 uses overlapping windows and the result of the wavelet transform changes between pixel intensities. As it has 4 scaling functions it produces a smoother output in the image. Figure 4 gives the graphical representation of Haar and db4 wavelets.

*2) **Soft vs Hard Thresholding**:* The basic idea of thresholding lies in the concept that the features in the images are localised and the image can be reconstructed with the higher magnitude wavelet coefficient by effectively suppressing the lesser magnitude wavelet coefficient which probably represents noise. In hard thresholding, to suppress the noise we apply the following nonlinear transform to the empirical wavelet coefficients:

$$F(x) = x \cdot I(|x| > t) \tag{2}$$

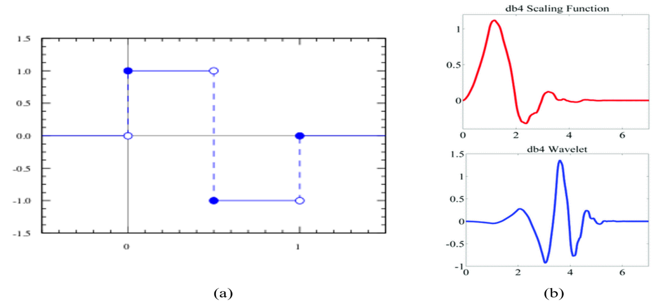

Fig. 4. (a) Haar (b) db4 wavelets

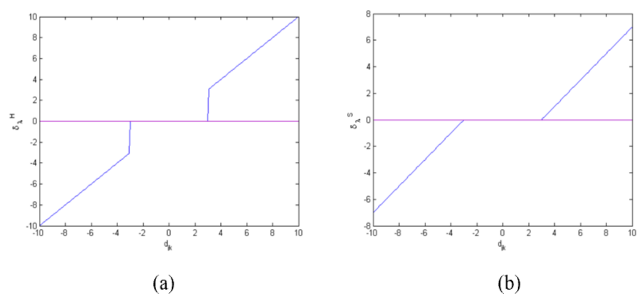

Fig. 5. (a) Soft Thresholding (b) Hard Thresholding

where t is a certain threshold. The choice of the threshold is a very delicate and important statistical problem. On the one hand, a big threshold leads to a large bias of the estimator. But on the other hand, a small threshold increases the variance of the smoother. It is given by the following formula

$$t = \sqrt{2\sigma^2 \log(n)/n}, \tag{3}$$

where n is the length of the input vector and σ^2 is the variance of the noise. The variance of the noise is estimated based on the data. We do this by averaging the squares of the empirical wavelet coefficients at the highest resolution scale [14].

The only difference between the hard and the soft thresholding procedures is in the choice of the nonlinear transform on the empirical wavelet coefficients. For soft thresholding the following nonlinear transform is used

$$S(x) = \operatorname{sign}(x)(|x| - t)I(|x| > t) \tag{4}$$

where t is a threshold. In addition you can also use a data driven procedure for choosing the threshold. This procedure is based on Stein's principle of unbiased risk estimation [14]. The main idea of this procedure is the following. Since $S(x)$ is a continuous function, given t, one can obtain an unbiased risk estimate for the soft thresholding procedure. Then the optimal threshold is obtained by minimization of the estimated risks. A smoother image is obtained when the soft thresholding is utilised in comparison to the hard threshoding but however in our applications, the hard thresholding performed better as it preserved the edges.

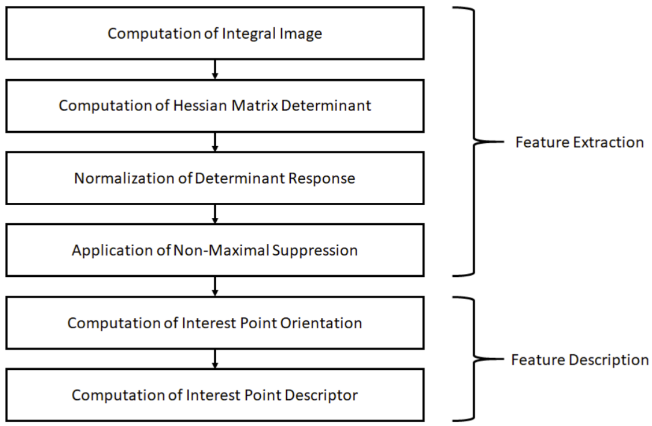

Fig. 6. Stages of SURF based Feature Matching

C. Translation Invariance using Feature Matching

In this division, we will describe the algorithm to determine the translation between the denoised image and template using Speeded Up Robust Feature (SURF) vectors. The SURF method is a fast and robust algorithm for local, similarity invariant representation and comparison of images. [9]The main interest of the SURF approach lies in its fast computation of operators that enables real-time applications such as tracking and object recognition. SURF is composed of two steps: (a) Feature Extraction (b) Feature Description and Matching.

Figure 6 gives an overview of process of feature representation using SURF.

*1) **Feature Extraction/Detection**:* The approach for interest point detection uses a very basic Hessian matrix approximation. The first stage in this is to compute the determinant of the Hessian matrix, given by the equation in [10]. For this, we need to apply convolution with Gaussian kernel, then perform second-order derivative. This process is repeated multiple times for different values of σ (scale). Similar to Laplacian of Gaussian (LoG) approximation in SIFT [15], SURF pushes the approximation(both convolution and second-order derivative) even further with box filters. One big advantage of this approximation is that, convolution with box filter can be easily calculated with the help of integral images. Fig 7 gives an overview of the approximation of LoG filter as box filters for the case $\sigma = 1.2$.

$$H(f(x,y)) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \tag{5}$$

Thus, with the box filter approximations denoted by D_{xx}, D_{yy}, and D_{xy}, the approximate determinant of the Hessian can be given by

$$det(H_{approx}(f(x,y))) = D_{xx}D_{yy} - wD_{xy}^2 \tag{6}$$

where w is chosen to be 0.9 from literature. [9] The above process is repeated for different values of σ, and doubling the filter size for each new octave(scale). This is illustrated in figure 8(a). This is followed by application of Non-maximal suppression on a 3x3x3 neighborhood.

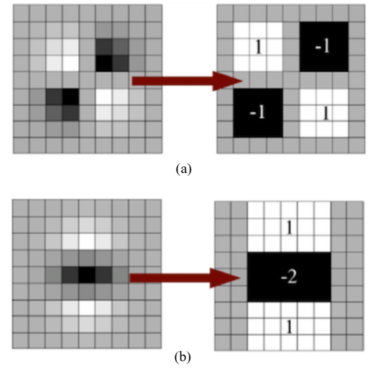

Fig. 7. Box filter approximation for (a) L_{xy} (b) L_{yy}

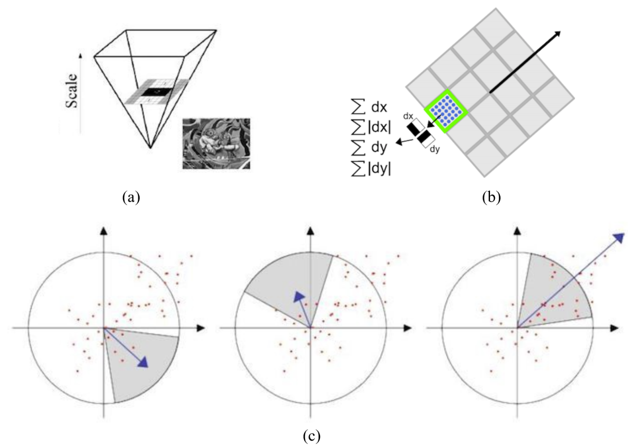

Fig. 8. (a) Scale-Space Representation (b) Interest point feature description (c) Interest point orientation computation

*2) **Feature Description and Matching**:* As seen in figure 6, the feature description stage comprises of two processes, namely, getting the orientation assignment and getting the 64 component descriptor for each feature.

For computing the orientation of a feature point, SURF first calculates the Haar-wavelet responses in x and y-direction in a circular neighborhood of radius 6s around the key point, where s is the scale at which the key point was detected. Also, the sampling step is scale dependent and chosen to be s, and the wavelet responses are computed at that current scale s. Accordingly, at high scales the size of the wavelets is big. Therefore integral images are used again for fast filtering.

Then the sum of vertical and horizontal wavelet responses in a scanning area is calculated at different orientations in increments of $\pi/3$. The direction with maximum Haar response is selected as the orientation for that feature. This process is shown in figure 8(c).

Once the orientation is determined, it is followed by the process of extracting the descriptor. This involves constructing a square region around the feature key point oriented along the orientation found in the previous step. The size of this window depends on the scale obtained during feature extraction. This region is divided into small 4x4 sub-regions and the sum of horizontal and vertical ($\sum dx$,$\sum dy$) Haar responses along with the sum of the absolute values of horizontal and vertical responses($\sum dx$, $\sum dy$) are calculated as shown in figure 6(b). Hence, each sub-region has a four-dimensional descriptor vector V for its underlying intensity structure ($V = \sum dx, \sum dy, \sum |dx|, \sum |dy|$). This results in a similar descriptor vector for all 4x4 sub-regions, giving a total of length 64 per feature.

The above algorithm is applied on both denoised image (obtained from the previous step) and the noise-free template to detect and describe feature points. Once the feature points of both images are described, the sign of LoG which is already computed in the detection process of SURF is used for underlying interest points.

The sign of the LoG distinguishes bright blobs on dark backgrounds from the reverse case. For matching, the features are compared only if they have same type of contrast (based on sign) which allows faster matching.

Once the features of the template are matched with that of the denoised image using an appropriate similarity metric, the translation is determined by finding the average difference between the coordinates of matched features for both images.

D. Moving Object Area Detection using Scale Invariant Hadamard Transform

While literature indicated that Haar transform has the properties we are looking for to detect scaling, it is a bit difficult to work with as a matrix. There is no simple formula for determining each value in the matrix, rather it follows a pattern which is clear to a viewer but harder to program into a computer and access without a relatively complicated formula.

For these reasons, and also with an eye to translating this algorithm into a quantum setting, we decided to use the Hadamard matrix instead. Each entry is determined by the simple formula

$$H_{i,j} = (-1)^{i.j} \tag{7}$$

where i, j are matrix indices.

Using the Hadamard matrix, we can construct a scaling algorithm which is fairly similar in form to the original convolution algorithm for correlation. The first step is to reduce our input images with pixel values of zero to 255 down to vectors of ± 1. This was accomplished by computing the average pixel value of the image, then setting any pixels greater than this average equal to 1, and all pixels with value less than the average were set to -1. This creates a kind of silhouette of the original image, maintaining the overall shapes. So the input template has entries of only ± 1, padded by zeros to be the same size as the image vector.

We only consider scalings by a factor of 2^n, given that the scales encoded by the Haar and Hadamard wavelets are only powers of two. One important part of the algorithm is an operation which will be referred to as S. S is a permutation matrix which takes a binary string and moves the last bit to the front, permuting the vector entries or matrix rows accordingly. By extension, S^k cyclically moves the last k bits of the string to the front [10].

The following equation explains the working of the permutation matrix when multiplied by an 8x1 column vector

$$\begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 2 \\ 4 \\ 6 \\ 1 \\ 3 \\ 5 \\ 7 \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 4 \\ 1 \\ 5 \\ 2 \\ 6 \\ 3 \\ 7 \end{pmatrix} \tag{8}$$

In (8), we see how S reorganizes the entries of a vector. The leftmost vector is the original v, the middle vector shows Sv, and the last shows S^2v. Based on the above mentioned ideas, we adapt the scaling algorithm [10] from literature to find the object area/scale difference using the following equation.

$$w_i = H(Ht \circ S^i Hv) \tag{9}$$

The result of this algorithm is a matrix where each column of index i is given by w_i, S is the permutation matrix explained in 8, t is the template column vector and v is the image column vector.

We could even reduce this to only consider the first row of this result, further minimizing computations by only multiplying by the first row of H, instead of the entire matrix as

$$w = \vec{\mathbf{1}}(Ht \circ S^i Hv) \tag{10}$$

The maximum value of this vector w will occur in the column which represents the correct scaling.

Hence, with the translation obtained in the previous step using SURF, the template can be translated to that specific coordinate and then the scaling algorithm in 10 is performed to determine the area boundary of the object of interest.

TABLE I
PERFORMANCE OF INITIAL APPROACH ON IMAGES OF FIGURE 9

Algorithm	Area Reduction	Maximum Correlation	Time taken
Image1	38%	0.985	60%
Image2	44%	0.988	51%

V. RESULTS

A. Initial Approach

The results for segmentation based search area reduction is shown in figure 9 and their corresponding performance measure in table I.

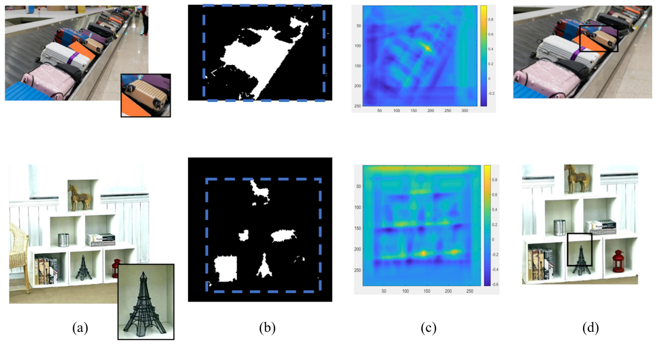

Fig. 9. (a) Input image and template (b) Segmented region with detected boundary (c) NCC result on the reduced search area (d) Object area detected on image

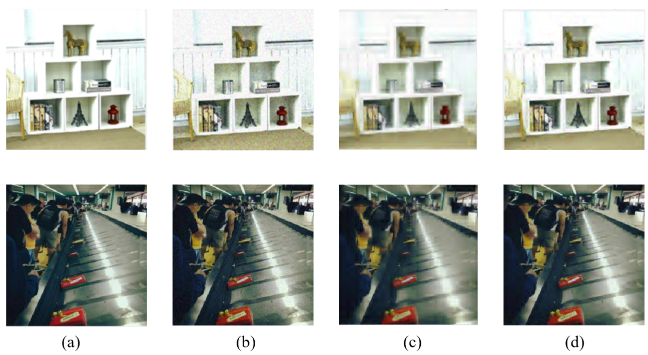

Fig. 10. db4 based denoising: (a) Noise-free image (b) Noisy image with Additive Gaussian Noise (c) Denoised image with soft thresholding (c) Denoised image with hard thresholding

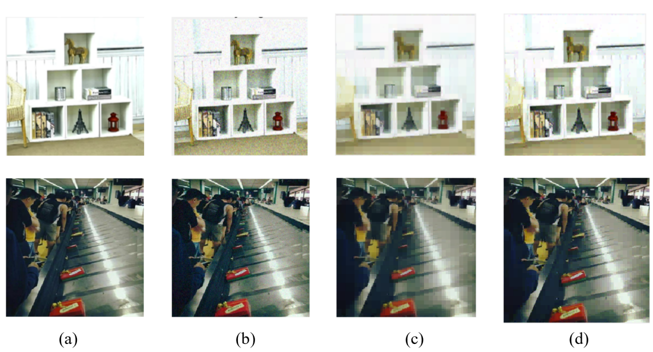

Fig. 11. Haar based denoising: (a) Noise-free image (b) Noisy image with Additive Gaussian Noise (c) Denoised image with soft thresholding (c) Denoised image with hard thresholding

B. Refined Approach

This division consists results for the denoising component and subsequent template matching. The denoising was performed for 2 different types of wavelets, namely haar and db4. This was further analyzed for different types of thresholding methods. Fig 11 gives the result of different thresholding methods on haar wavelet based denoising and fig 10 is obtained from db4 based denoising. A 4 level decomposition is performed on each of the wavelets with a noise variance of 20.

The metrics used to evaluate our denoising results are:

TABLE II

PERFORMANCE COMPARISON OF DENOISING FOR DIFFERENT WAVELET TRANSFORMS AND THRESHOLDING

Algorithm	PSNR	SSIM
db4 hard threshold	33.71	0.94
db4 soft threshold	29.74	0.90
Haar hard threshold	36.81	0.94
Haar soft threshold	31.31	0.92

*1) **PSNR**:* Peak signal-to-noise ratio (PSNR) is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale. The expression for PSNR is given by [17].

$$PSNR = 10log_{10}(R^2/MSE) \tag{11}$$

Where R represents the peak value of intensity in the image and MSE is the mean square error.

*2) **SSIM**:* The structural similarity (SSIM) index is a method for predicting the perceived quality of digital television and cinematic pictures, as well as other kinds of digital images and videos.The SSIM index is calculated on various windows of an image. The measure between two windows x and y of common size NN is:[16]

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{12}$$

where μ_x and μ_y are the averages and σ_{xy} is the co-variance and σ_x is the variance of x and σ_y is the variance of y.

Table II gives the average PSNR and SSIM over 10 different images.

We choose the Haar wavelet based hard thresholding as input for template matching algorithm as it gave a better performance. The denoised image is then passed to SURF based feature matching for translation invariance that is followed by scaling algorithm. The corresponding results are shown in fig 12

We further implemented real-time object tracking for videos using our algorithm. The results of object tracked for various frames are shown in figure 13. The template (blue box) is translated and scaled over the frames, which is detected accurately by our algorithm.

VI. SUMMARY AND FUTURE WORK

In this project, we proposed two wavelet based template matching algorithms for real-time object tracking in videos. Initially, we developed a search area reduction algorithm as an improvisation to the classical correlation based template matching. Further, we developed a scale/rotation and translation invariant algorithm using SURF based feature matching. Additionally, we proposed a Hadamard transform based scaling algorithm for object area localization.

From our experiments, we can empirically conclude that Hadamard Transform can be used as a potential complement

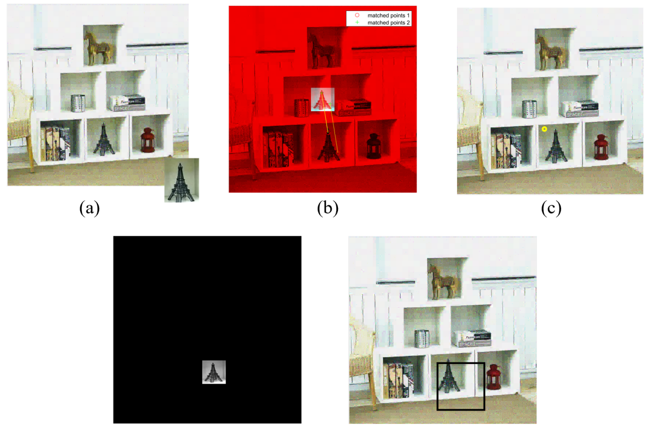

Fig. 12. (a) Input image and template (b) SURF based matched descriptors (c) Translation determination (d) Translated template (e) Boundary determination

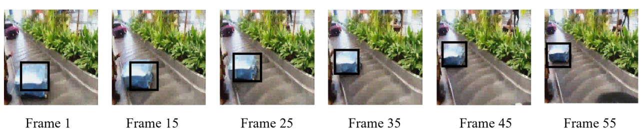

Fig. 13. Object area detection for real-time video frames

to SURF based feature description and also as an input feature space for training Neural Network models. This project can be extended further to applications such as Face Recognition, 3D Object Reconstruction and Medical Imaging. In case of low lighting conditions or poor resolutions, certain pre-processing methods like image toning can also be employed for enhancing the results.

References

[1] Pei, Liang, Zhiwei Xie, and Jiguang Dai. "Fast normalized cross-correlation image matching based on multiscale edge information." 2010 International Conference on Computer Application and System Modeling (ICCASM 2010). Vol. 10. IEEE, 2010.

[2] Koprinska, Irena, and Sergio Carrato. "Temporal video segmentation: A survey." Signal processing: Image communication 16.5 (2001): 477-500.

[3] A.N. Akansu, R.A. Haddad and H. Caglar, Perfect Reconstruction Binomial QMF-Wavelet Transform, Proc. SPIE Visual Communications and Image Processing, pp. 609–618, vol. 1360, Lausanne, Sept. 1990. Akansu, Ali N.; Haddad, Richard A. (1992), Multiresolution signal decomposition: transforms, subbands, and wavelets, Boston, MA:

[4] Discrete Wavelet Transform for Image Processing Dipalee Gupta1, Siddhartha Choubey2, International Journal of Emerging Technology and Advanced Engineering Website:www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 3, March 201.5

[5] Wang, Z.; Simoncelli, E.P.; Bovik, A.C. (2003-11-01). Multiscale structural similarity for image quality assessment. Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers, 2004. 2. pp. 1398–1402 Vol.2. CiteSeerX 10.1.1.58.1939. doi:10.1109/ACSSC.2003.1292216. ISBN 978-0-7803-8104-9.

[6] Afrah Ramadhan, Firas Mahmood and Atilla Elci,IMAGE DENOISING BY MEDIAN FILTER IN WAVELET DOMAIN,The International Journal of Multimedia Its Applications (IJMA) Vol.9, No.1, February 2017

[7] Parida, Priyadarsan, and Nilamani Bhoi. "Feature based transition region extraction for image segmentation: Application to worm separation from leaves." Future Computing and Informatics Journal 3.2 (2018): 262-274.

[8] Gilbert Strang, Truong Nguyen,Wavelets and Filter Banks,SIAM, 1996

[9] Jiang, Zetao, Qiang Wang, and Yanru Cui. "A fast method for feature matching based on SURF." International conference on intelligent science and intelligent data engineering. Springer, Berlin, Heidelberg, 2011.

[10] A Wavelet-Based Approach to Template Matching Under Scaling and Translation Natalie Sauerwald ,Advisor: David Meyer,Department of Mathematics,University of California, San Diego,June 6, 2012

[11] Kaso, Artan. "Computation of the normalized cross-correlation by fast Fourier transform." PloS one 13.9 (2018): e0203434.

[12] Radomir S. Stankovic , Bogdan J. Falkowski ,*,The Haar wavelet transform: its status and achievements

[13] Issaoui, Hadhami, Aïcha Bouzid, and Noureddine Ellouze. "Comparison between soft and hard Thresholding on selected intrinsic mode selection." 2012 6th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT). IEEE, 2012.

[14] Donoho, David L., and Iain M. Johnstone. "Adapting to unknown smoothness via wavelet shrinkage." Journal of the american statistical association 90.432 (1995): 1200-1224.

[15] Rublee, Ethan, et al. "ORB: An efficient alternative to SIFT or SURF." ICCV. Vol. 11. No. 1. 2011.

[16] Zhou Wang, Member, IEEE, Alan Conrad Bovik, Fellow, IEEE, Hamid Rahim Sheikh, Student Member, IEEE, and Eero P. Simoncelli, Senior Member, IEEE,Image Quality Assessment: From Error Visibility to Structural Similarity,IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 13, NO. 4, APRIL 2004

[17] Hore, Alain, and Djemel Ziou. "Image quality metrics: PSNR vs. SSIM." 2010 20th International Conference on Pattern Recognition. IEEE, 2010.