Improved Copy-Move Forgery Detection Using Sift With Adaptive Over Kernel Principal Component Analysis

A. Jeyalakshmi
Associate Professor, Department of Computer Science,
Sri Ramakrishna College of Art and Science,
Coimbatore, Tamil Nadu, India-06

Dr. D. Ramya Chitra
Assistant Professor, Department of Computer Science,
Bharathiar University, Coimbatore,
Tamil Nadu, India-46

***Abstract*—In the digital world, photographs served as a major source for any work. In an innovative improvement of digital device technologies, these photographs has been captured quickly at no cost, and to store them easily on many of digital supports, or share them on the Internet. At the same time, with the wide availability of advanced image editing multimedia tools (e.g. Adobe Photoshop) modifying a digital photo, with little or no obvious signs of tampering. This leads to discarding the image authenticity. These tampering are not visible to the naked eye; Copy-move and image splicing forgeries are common image tampering manipulations. In image splicing forgery, a part of another image is copy or cut and pasted on another image. Where as in copy-move, part/region of the same image has copied and moved into the same image. In this paper the copy-move forgery detection has implemented by a joint method of the SIFT with A-KPCA (Adaptive over Kernel Principal Component Analysis) framework.**

Keywords: **Digital image, Copy-Move Forgery Block based and Key point based detection, SIFT, A-KPCA.**

I. INTRODUCTION

Image Forgery is not a new thing, in ancient days the forgery had been performed in an art and fiction, but did not affect the society. Nowadays due to the increasing of digital devices and technology improvement an image can be easily manipulated and modified without any obvious signs [1]. It is very difficult for humans to identify visually whether the image is original or manipulated. Images can be used as an evidence for any case in the court of law, information broadcasting, the images broadcasted in TV news are accepted as the certificate for the truthfulness of that news. Digital images are being used in many applications ranging from military to medical diagnosis.

The key problem faced by the researchers is to categorize an image as forged or authentic and to localize the forgery. Several methods were proposed, but a proper method which can accurately detect image forgery is yet to be invented. Image Forensics is divided into two categories: active (intrusive) and passive (non-intrusive). An active forgery detection method (digital watermarking) inserts a pre-computed code as a part of the image data before it is transferred to the receiver side. At the receiving side this code is verified with the original inserted code for authenticity. The major drawback of this method is the requirement of special tools to embed the pre computed code as a part of the image before it is sent out. Image Forgery itself is of two types: cloning (copy-move) and splicing. Copy-move forgery [2] is performed by replacing a part of the given image by another portion which is taken from the same image. In image splicing, the copied region and the pasted region belong to different images. The purpose of the image forgery is to duplicate or conceal a certain object into an image or to make false propaganda. Fridrich et al. proposed forgery detection technique in which input image is segmented into overlapped rectangular blocks to find tampered regions with the help of Discrete Cosine Transform (DCT)[3].

The another efficient block based methods which gave the most precise results proposed by Christlein et al.[4] The work was aimed at detecting copy-move forgery using Zernike moments which have desirable properties like rotation invariance, robustness to noise etc. The average rate of precision achieved was 83.59%.

Ryu et al. proposed [5] detecting copy-rotate-move forgery using Zernike moments which have desirable properties like rotation invariance, robustness to noise etc. The average rate of precision achieved was 85.59%. Accurate results were obtained for rotation with 30◦. The method was found to be weak against scaling and affine transformations and it proposed the use of efficient data structures to reduce computational complexity. Another method for rotation invariant copy-move forgery detection was proposed in [6]. It was based on a method called Same Affine Transformation Selection (SATS), which had the benefit of shift vectors with some additional computational complexity. Use of Kd tree algorithm for matching reduced the number of false positives. It can detect arbitrary variations in rotation and scaling in the copied part and is a processing scheme in which any suitable feature can be used in combination with the SATS post processing.

Li Jing et. Al. [7] proposed firstly analyzes and summarizes block matching technique, then introduces a copy-move forgery detecting method based on local invariant feature matching. It locates copied and pasted regions by matching feature points. It detects feature points and extracts local feature using Scale Invariant Transform algorithm. More recently Xun yu Pan et. al[8] suggested a method to detect duplicated regions with

continuous rotation regions. As described in [4] the new method was based on the image SIFT features. First the SIFT features are collected from the image, and the image is segmented into non-overlapping examination blocks. The matches of SIFT key points in each non-overlapping pixel blocks are computed. After which the potential transform between the original and duplicated regions are estimated and the duplicated regions are identified using correlation map. Even though using SIFT key points guarantee geometric invariance and their method enables to detect rotated duplication, these methods still have a limitation on detection performance since it is only possible to extract the key points from peculiar points of the image. Cao et al. **[9]** proposes region duplication detection algorithm which is based on dividing circular blocks of image and apply with improved DCT this exhibits low computational complexity. The algorithm uses the DCT quantization with quantization table and these circled blocks are again divided into small size, in addition the Euclidian distance between these circular blocks are calculated. This algorithm is good quality in case of blurring and noising image but in poor it's robust.

Davarzani et al. **[10]** proposed algorithm that detects tempered image based LBP.It detects geometry of the forged region even if it is polluted by noise, blurring, JPEG compression, scaling or rotates in multiples of 90-degrees. Here the image transformed into gray scale then subdivided into overlapping blocks. With multi-resolution Local Binary Pattern features of each block are identified by applying different types of LBP operators. The feature vectors are put together to form feature matrices which their numeral counts which is equivalent the numeral count of LBP operators employed.

The matrices are sorted in lexicographically then k-d tree for determining the matched blocks. Hence Random Sample Consensus used elimination false matches. However, the method is still have high cost of computational steps and detect in high resolution images, and it cannot detect duplicated regions with arbitrary rotation angles either. The main goal of this paper is: to implement a cost effective less computational complicity method to detect copy-move forgery detection with the combined framework of key point and block based methods.

This paper is organized as follows: overview of different methods of detecting copy-move forgery of research method in section 2, section 3 describes proposed method for copy-move forgery detection, Experimental results and analysis for different existing and proposed method provided in section 4 and section 5 concludes this paper.

II. RESEARCH METHOD: COPY-MOVE FORGERY DETECTION

Copy-move is a simple and effective technique to make image forgeries in the digital image. A part of the image itself is copied and pasted into another part of the same image (Fig 1).In Figure 1 the image contained a single tree; the same tree has been copied and pasted in the same image with image editing tool. The purpose of this kind of forgery is usually to hide or to add some content or object in the image. Since the forged region came from the same image, it is impossible to use some statistical properties (for example, camera noise or illumination conditions) for forgery detection because they are very well matched within the image. Taking the forged region from the same image also simplifies the forgery process because it is easier to fit the forged region into the image due to the similarity of properties of the copied region and the rest of image.

Fig 1 : Example of Copy-Move Forgery Image

A number of techniques proposed to detect copy-move forgery which can be classified into two main categories such as block-based and key point-based methods. Good forgery detection method should be robust to manipulations, such as scaling, rotations, JPEG compression and Gaussian Noise addition made on the copied content. These attacks are not detected by the single method. The improved approach is proposed to detect image forgery by copy-move under above attacks by integrating block-based and feature-based method.

A. Block-based Feature detection

In block-based methods, the image will be divided into overlapping blocks of specified size and a feature vector will be computed for these blocks. Similar feature vectors are then matched to find the forged regions. For e.g. DCT, DWT, PCA, KPCA etc. In this way, the detected regions are always composed of regular blocks, which cannot represent the accurate forgery region well; therefore, the recall rate of the block-based methods is always very low. Also as the size of the host images increases, the matching computation of the overlapping blocks will be much more expensive.

Although key point based methods can reduce two problems but the recall rate is very poor. To address these issues, the proposed scheme integrates both block based and key point based methods. Figure 2 illustrates the general integration framework of copy-move forgery detection. With this combined framework the forgery

regions has to be efficiently detected with less computation time .

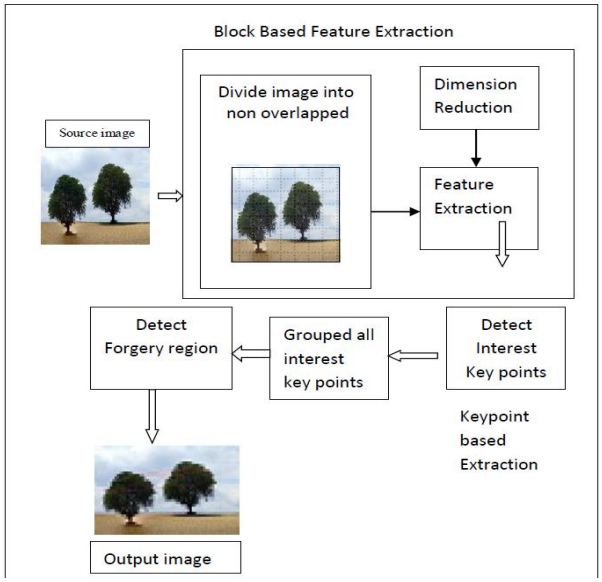

Fig 2: Integrated Block based and Key point based Framework for Copy-Move Forgery Detection

III. PROPOSED METHOD: COPY-MOVE FORGERY DETECTION

Copy-move forgery detection process incorporates a block level and key point based forgery detection taken place. Adaptive over kernel principle component analysis (A-KPCA) with SIFT based key points detection algorithm is proposed in this paper. The A-KPCA algorithm will divide the host image into non-overlapping blocks in an adaptive manner. The feature points are extracted as block level features from each block. These blocks level features are then mapped with one another to locate the labeled feature points. This leads to indicate the suspected forgery regions.

A. SIFT (Scale Invariant Feature Transform):

The S**cale-Invariant Feature Transform** (**SIFT**) is a feature detection algorithm that can extract local feature, the main idea is to find extreme points in scale space, extract invariant when location, scale, rotation, illumination changed. The algorithm is proposed by David Lowe [11] in 1999 and to be completed in 2004. The algorithm is robust with scale, rotation, brightness, affine, and increase or decrease in the target object and blocks. SIFT method includes three steps: detection of extreme point in scale space, the formation of feature point descriptor and feature point matching.

B. Extreme point in scale space detection:

Detection of key point interest in the sift algorithm framework, The image is convolved with Gaussian filters at different scales, and then the difference of successive Gaussian-blurred images are taken. Key points are then taken as maxima/minima of the Difference of Gaussians (DoG) that occur at multiple scales. Hence, a DoG image D(x, y, σ) is given by

$$D(x,y,\sigma) = L(x,y,k_i\sigma) - L(x,y,k_j\sigma) \quad \text{-(1)}$$

Where $L(x,y,k_i\sigma)$ is the convolution of the original image $I(x,y)$ with the Gaussian blur $D(x,y,\sigma)$ at scale $k\sigma$, i.e. $L(x,y,k\sigma) = D(x,y,\sigma) * I(x,y)$,Hence a DoG image between scales $k_i\sigma$ and $k_j\sigma$ is just the difference of the Gaussian-blurred images at scales $k_i\sigma$ and $k_j\sigma$.

For scale space extrema detection in the SIFT algorithm, the image is first convolved with Gaussian-blurs at different scales. In the convolved image each pixel compared with the adjacent eight pixels of the same scale and around the scale 9 × 2 pixels surround ding neighborhood to ensure that local extreme points can be detected in the scale of space and two-dimensional image space. Determine the main direction of feature points through the gradient direction distribution characteristic of neighborhood pixel, the 16*16 window of the center of feature points is divided into 4 * 4 sub-windows, each sub-window calculate gradient direction histogram of 8 directions by Gaussian-weighted, form the 4 * 4 * 8=128 dimension vector that is the feature point descriptor.

C. Feature Point matching:

Take a feature point of reference image, find out the feature points in the registration image that Euclidean distance between them is smallest and second smallest. If the value that divide the nearest distance by second nearest distance is less than a certain threshold that is the right match. In order to ensure the accuracy of matching points, two-way matching is used.

D. Kernel Principal Component Analysis(KPCA):

Principal component analysis is one of the most popular techniques for feature extraction. The principal components of input data X can be obtained by solving the eigenvalue problem of the covariance matrix of X. This conventional PCA can be generalized as a nonlinear one, the kernel PCA by $\Phi : \mathrm{R}^d \to F$,a mapping from input data space to a highly dimensional feature space F . The space F and therewith also the mapping Φ might be very complicated. To avoid this problem, the kernel PCA employs a kernel trick to perform feature space operations by explicitly using the inner product between two points in the feature space:

$$\left(\Phi(x_i), \Phi(x_j)\right) \to K(x_i, x_j) \qquad (2)$$

The covariance matrix can be written as,

$$W_{\Phi} = \frac{1}{N}\sum_{i=1}^{N}\Phi\left(x_i\right).\Phi\left(x_i\right)^T \quad (3)$$

For any eigenvalue of $W_{\Phi}, \lambda \geq 0$, and its corresponding eigenvectors $V \in F \setminus \{0\}$, the equivalent formulation of eigenvalue problem in F can be defined as:

$$N\lambda\alpha = K\alpha \quad (4)$$

Where eigenvector V spanned in space F as:

$$V = \sum_{i=1}^{N}\alpha_i\Phi\left(x_i\right) \quad (5)$$

where $K_{ij} = K\left(x_i, x_j\right) and \alpha = \left(\alpha_1, \alpha_2, ..\alpha_N\right)^T$ [12]. For the kernel component extraction, compute projection of each data sample x onto eigenvector V

$$\left(\Phi\left(x\right), V\right) = \sum_{i=1}^{N}\alpha_i K\left(x_i, x_j\right) \quad (6)$$

The Kernel PCA allows obtaining the features with high order correlation between the input data samples. Generally, the kernel projection of data samples onto the kernel principal component might undermine the nonlinear spatial structure of input data. Namely, the inherent nonlinear structure inside input data is reflected with most merit in the principal component subspace. As a powerful nonlinear feature extraction method, Kernel Principal Component Analysis (KPCA) [13] has been widely used in many applications. Kernel PCA has all the advantages of the regular PCA, as well as the implicit mapping to a feature space F where the features representing the structure in the data may be better extracted. In this paper the copy-move forgery performed like part of the image itself is copied and pasted into another part of the same image. Hence, it is not necessary explicit form of mapping in the feature space, only the inner products between two data points in the feature space are needed. Adaptive Kernel Principal Component Analysis (A-KPCA) [14], which can effectively perform.

E. Adaptive over Kernel Principal Component Analysis (A-KPCA)

Adaptive Kernel Principal Component Analysis (A-KPCA), which can effectively learn the kernels under the unsupervised learning setting. First transform original one dimensional (1D) vector into two dimensional(2D) feature matrices through a set of nonlinear mappings induced from various kernels, each corresponding to one column of the 2D feature matrix. Then, two coupled sets of projective vectors are extracted from those feature matrices using an iterative procedure. One set of projective vectors corresponds to the column direction of feature matrices and is used for nonlinear feature exaction, while the other corresponds to the row direction of feature matrices and is used for searching the optimal combination of kernels. The experiment result show that proposed detection improved method can achieve better results than the existing forgery detection methods.

$$Let \quad \Phi_i : x \in X \rightarrow \Phi_i\left(x\right) \in \mathrm{H}_i, i \in \left\{1, 2, ..N\right\} \quad (7)$$

be a set of non linear mappings from the original input space X to high dimensional feature space H_i. The inner products in H_i defined as the kernels

$$\kappa^i\left(x, y\right) = \Phi_i\left(x\right)^T\Phi_i\left(y\right) \quad (8)$$

respectively. From Φ_i ,define as

$$\Phi_i : x \in X \rightarrow \Phi_i\left(x\right) = \left(\underbrace{0^T..0^T}_{i..i-1}, \Phi_i\left(x\right)^T, \underbrace{0^T..0^T}_{i+1..f}\right)^T \in \mathrm{H}, i \in \left\{1, 2, .. f\right\} \quad (9)$$

Here H is the Hilbert space as the direct sum of H_i and the j[th] 0 vector lies in H_j .The inner product in H can be defined as: $\Phi_i\left(x\right)^T\Phi_i\left(y\right) = 0, \left(i \neq j\right)$ and $\Phi_i\left(x\right)^T\Phi_i\left(y\right) = k^i\left(x, y\right)$, kernel feature space has to be retrieved from higher dimensional original space. These feature space has to be the non overlapping block level features of the forgery image. The adaptive kernel principal component analysis is used to extract the two dimensional (2D) feature matrices through a set of nonlinear mappings induced from various kernels, each corresponding to one column of the 2D feature matrix.

F. Copy-Move Forgery Detection using SIFT with A-KPCA

An integrated block level and key point based copy-move forgery detection method is proposed. The A-KPCA algorithm will divide the host image into non-overlapping blocks in an adaptive manner. The feature points are extracted as block level features from each block. These blocks level features are then mapped with one another to locate the labeled feature points. This leads to indicate the suspected forgery regions. Figure 3 illustrates the proposed method copy-move forgery region detection. In which the image is first transformed in an adaptive manner and construct the kernel matrix K_k for each x_k . Then to Estimate the initial KPCA model i.e the eigen values and vectors and Calculate the initial control limit of the feature space .afterword's get the next testing sample and compute its principal components and kernel parameters.

With the Kernel principal components analysis detect the minima distance block features extracted and coordinate together in the form non-overlapped manner. Then extract the SIFT based interested key points. This

way of feature detection has been performed in an effective manner. First suspect the related portion of the forgery regions in an nonlinear mapping orderly. Through this the dimensionality of the feature space has been reduced and then to spot out the accurate points of forgery region. The combined framework of SIFT with A-KPCA method produce an improved results than other existing methods.

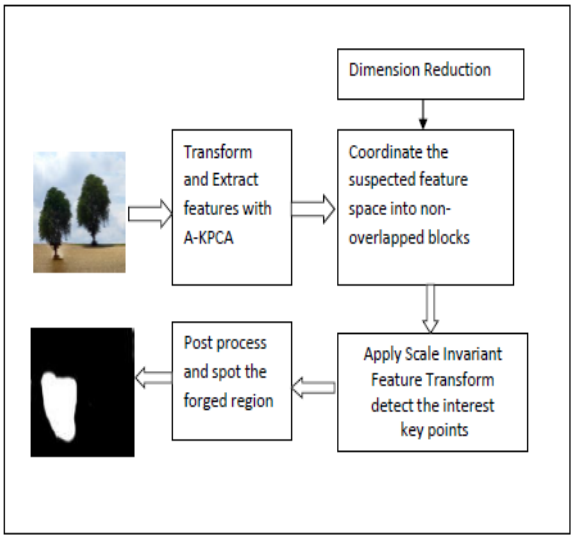

Fig 3: Copy-Move Forgery Detection using SIFT with A-KPCA

Algorithm A-KPCA

Input: Training input $X = \{x_1, x_2, ..x_n\}$

Step1: construct the kernel matrix K_k for each x_k and scale it $\overline{K_0}$

Step 2: Estimate the initial KPCA model i.e the eigen values and vectors of the $\overline{K_0}$

Step 3: Calculate the initial control limit of the feature space

Step4: Obtain the next testing sample x, calculate the principal components and kernel parameters

Step5: compute K_{new} and scale it $\overline{K}$

Step 6: project $\overline{K}$ into KPCA and obtain $\widehat{K}$

IV. EXPERIMENTS AND RESULTS

In result analysis, six cameras have been used in this experimentation. Table 1 lists the digital still cameras model. Each scene has been captured as an image at three different timings of the day at 9am, 12noon and 3pm on each camera. Thus 20 images of same scene have been captured at various timings by each of the camera. The camera specific parameters have not been altered. Totally (50x3timesx6camera) 900images were taken; from that 300 images have been tampered as copy-move with the help of Photoshop CS3 software.

Table1. Digital cameras used in this experimentation

S.No	**Camera Model**	**Max.Image Size**	**Image format**
C1	Canon PowerShot A495	3648x2048	JPEG
C2	SAMSUNG PL120	4320x2432	JPEG
C3	SONY-DSCW330	4320x3240	JPEG
C4	Canon-DSLR	5184x3456	JPEG
C5	NikonD90	4288x2848	JPEG
C6	NikonD300	4288x2848	JPEG

This paper is focused to find a given image is forged or not. if it is True; then calculate the performance of computation time. The image level fact has to be measured through precision *P* and recall *r* which are defined as:

$$p=T_P/(T_P+F_P) \quad and \quad r=T_P/(T_P+F_N) \qquad (10)$$

Table 2*:* Evaluation measures Description

Evaluation Measures	***Description***
True Positive(T_p)	*No. of Images that have been correctly detected as forged*
False Positive(F_P)	*No, of Images that have been falsely detected as forged*
False Negative(F_N)	*No.of Images that have been falsely missed but they are forged*
True Negative(T_N)	*No.of Images that have been correctly detected as not forged(without Tampering)*

Table 2 shows the evaluation measures description [12]. Recall is also called as True positive rate (TPR).The simulation results are tabulated in Table 3.

Table 3.*Performance Result of Different methods*

Methods	*No.of Images Detected out of 300*				*Precision (P)*	*Recall (r)*
	T_P	T_N	F_P	F_N		
SIFT	*236*	*244*	*64*	*56*	*78.6%*	*83.5%*
SIFT+ KPCA	*259*	*253*	*41*	*47*	*86.3%*	*84.6%*
SIFT+ A-KPCA (proposed)	*281*	*280*	*19*	*20*	*93.6%*	*93.3%*

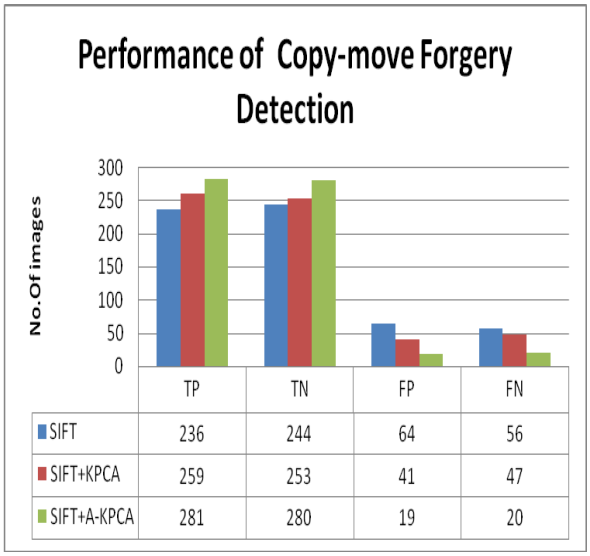

Fig.4 Performance of Copy-Movr Forgery Detection through Various Methods

In the above chart No. of Images that have been correctly detected as forged rate True positive is highly detected by the proposed method than other existing methods. In this the positive predictive value is called precision. Recall is also called as True positive Rate (TPR).The F-Score can be used as a single measure of performance of the test for the positive values. The F-score is the harmonic mean of precision and recall: This can be calculated as

$$F\text{-}Score = 2T_P/(2T_P+F_P+F_N) \qquad (10)$$

Table 4: Performance Accuracy of Various Methods

Methods	***TPR***	***FPR***	***F1Score***	***Total Elapsed Time(ms)***
SIFT	*78.6%*	14%	79.7	14395
SIFT+ KPCA	*86.3%*	11%	85.4	11060
SIFT+ A-KPCA (proposed)	*93.6%*	6%	93.5	7090

From the above comparison, the proposed method gives much better results than other existing methods.

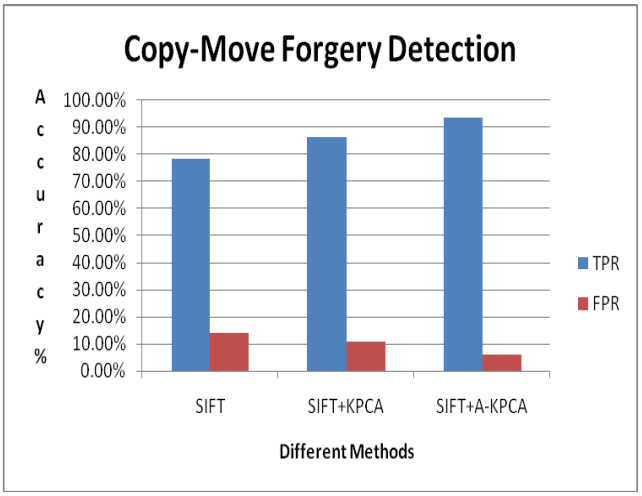

Fig 5. Performance Measure of TPR and FPR in Different methods

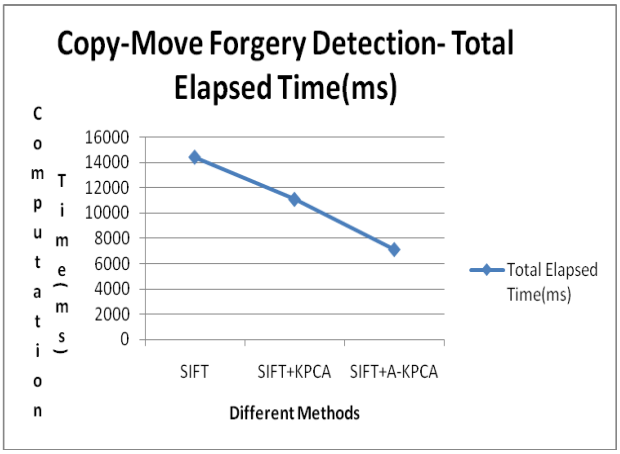

Fig.6 Computation Time taken for Copy-Move Forgery Detection by various methods

V. CONCLUSION

In this paper, Copy-Move forgery has been detected by a combined method of SIFT with A-KPCA feature detector. Which produce an improved better results than other existing methods; with combination of the SIFT with A-KPCA gives a very quick response than other existing methods. In the court premises the forged images are easily and effectively detected by this method.

REFERENCES

[1] R. E. J. Granty, T. S. Aditya, and S. Madhu, "Survey on passive methods of image tampering detection," in IEEE International Conference on Communication and Computational Intelligence (INCOCCI), pp. 431-436, 2010 .

[2] Amerini, L. Ballan, R. Caldelli, A. D. Bimbo, and G. Serra, "A SIFT-based Forensic Method for Copy-Move Attack Detection and Transformation Recovery," IEEE Transactions on Information Forensics and Security, vol. 6, no. 3, pp. 1099-1110, 2011.

[3] J. Fridrich, D. Soukal and J. Lukas, "Detection of copy–move forgery in digital images", Proceedings of Digital Forensic Research Workshop, IEEE Computer Society, Cleveland, OH, USA pp. 55–61,Aug'2003.

[4] Christlein V, C Riess , C Riess, E Angelopoulou, J Jordan, "An Evaluation of Popular Copy-Move Forgery Detection Approaches", IEEE Transactions on Information Forensica and Security, Vol 7 No 6, pp. 1099-1110, 2012.

[5] S. Ryu, M. Lee, H. Lee, "Detection of copy-rotate-move forgery using zernike moments," in Proc. Int. Workshop Information Hiding, Springer , pp. 51–65, 2010.

[6] Christlein V, C Riess, E Angelopoulou, "On rotation invariance in copy-move forgery detection ", IEEE International Workshop on Information Forensics and Security (WIFS), pp.98-102,2010.

[7] Y. Huang, et al., "Improved DCT-based detection of copy-move forgery in images", International Journal of Forensic Science Vol.206 No.(1–3) pp.178–184, 2011.

[8] X. Pan, S. Lyu," Detecting image region duplication using SIFT features", in: IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pp. 1706–1709, 2010.

[9] Cao G, Zhao Y, Ni R, Li X."Contrast enhancement-based forensics in digital images",IEEE Transactions on Inormation Forensics and Security,Vol. 9,No 3,PP.515–525,2014.

[10] Davarzani R, Yaghmaie K, Mozaffari S, Tapak M., "Copy-Move Forgery Detection Using Multiresolution Local Binary Patterns",International Journal of Forensic Science, vol. 231, issue: 1–3, pp.61– 72,2013.

[11] Lowe D G., "Distinctive image features from scale-invariant key-points", International Journal of Computer V ision, Vol.60, No.2, pp.99-110, 2004.

[12] B. Scholkopf, S. Mika, A. Smola, G. Ratsch, and K.R. Muller, "Kernel PCA pattern reconstruction via approximate pre-images,"8th International Conference on Artificial Neural Networks, Perspectives in Neural Computing, Springer Verlag, pp. 147-152, Berlin, 1998.

[13] B. Scholkopf, A. J. Smola, and K.-R. Muller, "Nonlinear component analysis as a kernel eigen-value problem", Journal of Neural Computation, Vol.10, No.5, pp.1299–1319,1998.

[14] S. Yang, S. Yan, D. Xu, X. Tang, and C. Zhang. Fisher,"Kernel criterion for discriminant analysis", International IEEE Computer Society Conference on Computer Vision and Patter Recognition, San Diego, CA, pp.197–202, 2005.

[15] Anand, Vijay, Mohammad Farukh Hashmi, and Avinash G. Keskar, "A Copy-move Image Forgery Detection Based on Speeded Up Robust Feature Transform and Wavelet Transforms", 5th International Conference on Computer and Communication Technology, Research Gate,pp:148-152, 2014.