IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)
e-ISSN: 2278-2834,p- ISSN: 2278-8735.Volume 11, Issue 1, Ver. III (Jan. - Feb .2016), PP 44-54
www.iosrjournals.org

Hybrid Technique to Identity the Faces Using Key Point Hypothesis Prediction Algorithm (Khpa)

K.RaghavendraPrasad[1], H.Girisha[2]
[1]*Department of EEE,Rao Bahadur Y.Mahabaleswarappa Engineering college, India*
[2]*Department of CSE,Rao Bahadur Y.Mahabaleswarappa Engineering college, India*

Abstract: *To achieve accurate detection of faces and extraction of criminal records through the detected faces, a feedback accuracy system is developed. This model improves the efficiency of accurate detection of criminal faces using Anisotropic Scale Invariant Feature Transform (A-SIFT) algorithm. SIFT algorithm uses anisotropic scaling on the query images for accurate corner detection. The identified true positive cases are used for detecting database image using unsupervised RESVM technique. The data is further tested using a hypothesis testing proposed with definedrule that compares the resultant RESVM image with the actual image. Depending upon the rules, the test results provide whether the image is accurate or not. Degree of accuracy determines the acceptance or rejection of images, level of accuracy above a certain threshold value resubmits the image to RESVM. Then the entire process is repeated until the image attains an accuracy level of 95% or more. This proposed method is tested with both similar and different reference and input facial images. Finally, it is found that the results are better and efficient in terms of retrieval time and keypoint reduction.*

I. Introduction

In recent years, feature detection and analysis plays a major role in image analysis and processing to obtain relevant information [1]. This feature detection is done using pattern recognition [2], annotation [3], recognition [4] and sequence analysis [5]. Out of which a high distinctive invariant, Scale Invariant Feature Transform (SIFT) proposed by David G. Lowe proved to be an effective technique to pre-process an image but robust in pattern recognition [6]. This SIFT can detect and describe the entire local feature in an image for object recognition, tracking and mapping application.

SIFT locates the extreme value in the scaling space to detect the location, scaling and rotation invariant. Disruption in recognition could be reduced using local image feature extraction than filtering that can solve hand deformation problem. In addition to this, robustness exist in SIFT due to compression, noise addition and filtering. This SIFT perform inefficient due to the presence of numerous patterns generated at each filtration stage [7]. Isotropic filter for preprocessing and matching the preprocessed images doesn't perform well in terms of matched points [8].

Apart from robustness due to filtering, accuracy is affected when SIFT is used with 2D symbolic aggregate approximation feature [9] and with Comp-Code feature extraction during recognition [10]. The main problems with SIFT technique is to discriminate the key points on an image, orientation histogram descriptors are not sufficient [11]. Another major problem is the matching of key points between the images to attain rotation and translation invariance is affected entirely due to topological relation ignorance.

SIFT is mainly used in extracting feature in a multispectral gray image, while doing this it posses certain drawbacks. The first drawback is that the conversion of multispectral to gray information reduces the level of extraction of image information, since the image formation is rapidly affected due to gray image conversion. Thus it leads to degraded SIFT extracted features and the second drawback is that it fails in detecting spectral information that is used in differentiating regional characteristics. This could avoid the number of useful descriptors or distinctive point features in a multi-spectral image leading to less accuracy [12].

There are certain considerations while pre-processing an image with respect with above drawbacks. To avoid the problem due to gray scaling, [13] used a technique that uses opponent information to retain the information of image even if the image is subjected to gray scaling with different intensity shifts. To improve the descriptor in an image after the application of SIFT, a combination of color and global information builds the color invariance components and global components using log-polar histogram [14]. This makes the descriptors to appear rich after pre-processing the image in relevance to other images.

Unsupervised recognition plays a crucial role in detecting the face of human in an image. To achieve this, Support Vector Machine (SVM) could contribute to a better part. Here to improve promising detection rates, Ensemble learning approach using SVM is used. Robust Ensemble Support Vector Machine (RESVM)plays a major in learning patterns using positive unlabelled data (PU) [ressvm.pdf] for studying the patterns given by the descriptors through Modified SIFT technique. Furthermore, the accuracy of system is increased to a vast level using this PU pattern studying that reduces the noises present in an image. This can be justified in terms of PU pattern with descriptors that represents the actual image representation rather the presence of false unlabelled data. Initially to attain this PU learning, class-weighted SVM (CWSVM) approach considers the unlabelled data to be negative cases [15]. Then the CWSVM is trained to avoid the contamination of positive instances else it will lead to accumulation of unlabelled data. Initially Bootstrap re-sampling on both positive and unlabelled instances reduces the accumulation of these attributes. In RESVM uses an extra Degree of Freedom to attain misclassification penalty between the positive and unlabelled data. The ensemble learning is done based on base model decision value that re-samples the positive and unlabelled data in SVM.

This paper aims at detection of criminal faces from a group photograph, a complex process in terms of its accuracy. The model involves two approaches; one is giving query at front end, pre-processing and classification using A-SIFT and RESVM respectively. Prior method includes the accuracy increment and retrieval process. Here, dynamic detection of human faceis done using an anomaly detection method called SIFT technique.In SIFT technique, due to unavailability of angular points large number of features are extracted leading to redundant information. Due to generation of high dimensional feature points, calculation time increases thus affecting real time performance [16]. This concern is been addressed in [17] and our concern is to improve the directionality of key points. To increase this objective we are going for anisotropicscaling over SIFT (A-SIFT) technique. Rich directional key points are obtained using A-SIFT thatimproves the key pointbehaviour of the featured transform. The obtained key points are processed for further stages to obtain better descriptors. Training is done over these descriptors with RESVM and compared with original image to eliminate the accumulation of unlabelled data. Thus, PU learning with positive labels increases more the accuracy. Thereby generating rules for accurate detection, the accuracy of detection is improved through hypothesis testing of the predicted values with generated rules. Depending on the results obtained, if the accuracy value is less than the threshold value of the model; then the system is made to re-configure all steps repeatedly over the given query image. Upon increased accuracy greater than threshold level, further process involves the retrieval ofrelated image from the database.

The contribution of the paper: section 2 includes the proposed model for accuracy improvement and section 3 includes the correctness of the proposed model and the validation is done in section 4. Finally the conclusion and future scope is included in section 5.

II. Methodologies

The proposed method uses Scale Invariant Feature Transform (SIFT), over a group photo for extracting its feature. The main feature to be extracted is the human face that is visible in the group photo.This is done to detect the criminals using facial detection technique.An automated model is mainly used for extraction of selected features' i.e. face from the group photo. Remaining attributes are eliminated;facial features are accurately selected using automated selection of threshold accurate levels for each regions in a face. The facial attribute is extracted using A-SIFT technique that uses anisotropic scaling for improving the accuracy of scaling. In the next stage, PU learning using Robust Ensemble-SVM (RESVM)removes the unlabeled data and increases the efficiency of retrieval. In the third stage, threshold accuracy value (TAV) determines the accuracy of retrieval in previous two stages. This TAV value is generated based on the rule defined in terms of hypothesis, proving whether the obtained image is significant or in-significant. Once if the accuracy value is lesser than the TAV, the pre-processing and classification is repeated until the image retrieval attains greater value than TAV.Otherwise the image is retrieved from the database whose accuracy value is greater than TAV.

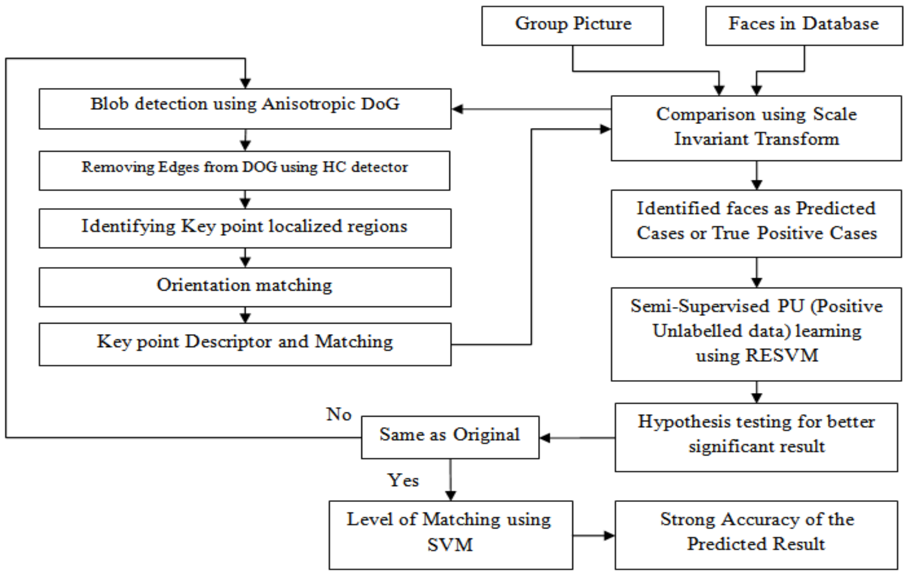

Fig. 2 Key Point Hypothesis Prediction Algorithm (KHPA)

2.1 Anisotropic-SIFT

The main steps in Scale Invariant Feature Transform involves: Scale space extreme detection that searches and detects all the scale space extreme points of an image using Difference of Gaussian (DOG). These size and location of keypoints are accurately positioned using accurate positioning keypoints. Orientation of these keypoints is assigned over a desired direction; thus converting the image data to feature points. Descriptors are assigned to feature points for counting the key points of the current scaled area [17].

2.1.1.Difficulty In SIFT Algorithm

The robustness in SIFT occurs due to the selection of key points that is been addressed in [17]. Also, corner detectors may not help when an image is scaled and [12] proposed edge extraction using canny edge detector, which fails when image is scaled. Our concern is to provide better directional scaling of the key points to achieve a proper orientation of key points. Adding anisotropy for directional scaling improves the accuracy or coherence of the directional vectors in obtaining proper feature points. Thus the main difficulty in SIFT algorithm could be removed using this coherent technique. Thus, the redundant feature vectors arising due to improper scaling in featured direction could be eliminated through this technique. The process associated with modified SIFT technique is described below:

2.1.2. Adaptive Thresholding

Initial process involves detection of key points using a windowing function. Since the image varied based on various scaling, exact detection of accurate key points is required. To achieve this, an adaptive window with similarity difference through weighted average processing between the images [18]. Objects are approximated in terms of key point collection in an image and to reduce the noise, anisotropic diffusion is used. This adaptive smoothening parameter helps in attaining smoothness images and protecting the edge features [19]. In this principle, the diffusion co-efficient K diffuses near smoothened areas than boundary regions. Texture and noise areas are set with small variations for edge preservation and for flattened areas, large variations are set. Thus using a constant K_0, semi-adaptive diffusion co-efficient are given in Eqs. (1) and (2).

$$:K' = \frac{K_0}{1 + \left\| \nabla I \right\| / 255} \tag{1}$$

$$:K' = \frac{K_0}{1 + \exp\left(-255 / \left\| \nabla I \right\|\right)} \tag{2}$$

The constant K_0 is larger than K in normal anisotropic processing and in semi-adaptive diffusion, K is large in flat region and small in edge region to diffuse more and less, respectively. Multiple diffusion process could help in achieving better denoising effects. Thus, the diffusion process is performed over several iterations follows automatically the smoothening functions. This is because K' values are dependent completely on manually specified K_0 values, named semi-adaptive threshold. Finally, improved co-efficient for smoothening function with K' is defined as:

$$: c\left(\|\nabla I\|\right) = \exp\left[-\left(\frac{\|\nabla I\|}{K_0} 1 + \|\nabla I\| / 255 \right)^2 \right] \tag{3}$$

$$: c\left(\|\nabla I\|\right) = \exp\left[-\left(\frac{\|\nabla I\|}{K_0} 1 + \exp\left(-255 / \|\nabla I\|\right) \right)^2 \right] \tag{4}$$

The smoothening function helps in removing noises present in reference image and that is represented in fig. 2. Using adaptive thresholding technique.

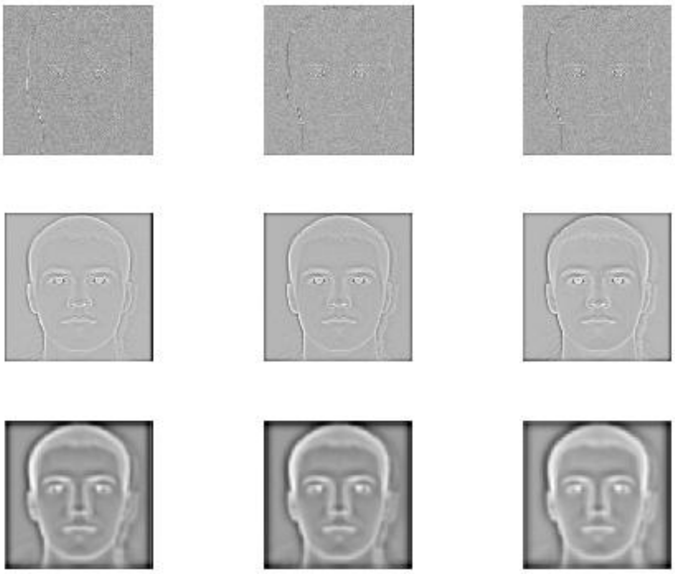

Fig. 2. Initial Pre-Processing and filtering of noises using Adaptive Thresholding

2.1.3. Scale Space Detection

This semi-adaptive threshold smoothening co-efficient finds the potential key points through local maxima solution across scale and space. This correspondingly provides set of values in *x* and *y* axis at *K'* scale.Here,instead of Difference of Gaussian (DoG)we have applied a database of keypoints taken from several facial images with various expressions of faces. Since, DoG helps in finding the feature points in the whole image. However, the facial regions alone cannot be extracted. So, the relative keypoints from the whole image keypoints set is found using database values containing facial image values as ground vectors. The database values are chosen as ground truth values from sample image of it is shown in fig. 3 (a). Likewise, 3(b) shows the crowded inefficient SIFT points over the reference image.

(a) (b)

Fig. 3. Reference image (a) normal face (b) initial SIFT points

Initially, the images of ground truth value are used for extraction of features and saving it in database. From the ground values, a new image *m* is trained with this training data. The trained image is now tested with same database after several iterations. To achieve proper keypoints in scale space, weighted Euclidean distance is used. This helps in finding the minimum distance between the keypoints over a given descriptor. The weights are based on the training data samples from the database using neural networks with inputs from ground truth value.

$$d(a,b) = \sum_{i=1}^{n} \sqrt{w_i \left[(x_i^1 - b_i^1)^2 + (x_i^2 - b_i^2)^2 \right]} \tag{5}$$

where x_i^1 and x_i^2 represents i^{th} measure of edge regionsalong *x* and *y* coordinates relative to point 1, respectively. b_i^1 and b_i^2 represents i^{th} measure of edge regions along *x* and *y* coordinates relative to point 2, respectively.w_{ij} represents a value between i^{th} measure that relates to ground truth value from the database.

2.1.4 Keypoint Localization

A threshold parameter is used to find more matching points relative to facial region. The value range is chosen as between [0,1]. Here, if the matching points are more, then the range value could be more. Greater value range directly affects the accuracy and time matching precision. This parameter is defined as:

$$d_R = w_1 \frac{d_{12}(x_i, b_i)}{d_{13}(x_i, b_i)} \tag{6}$$

Numerator in the above equation represents the weighted Euclidean distance between points 1 and 2. Denominator represents weighted Euclidean distance between points 1 and 2 based on weights w_i.Thus, using key point localization, the reference points with training data samples helped in removing redundant and reduces the keypoints to a larger extent shown in fig. 4.

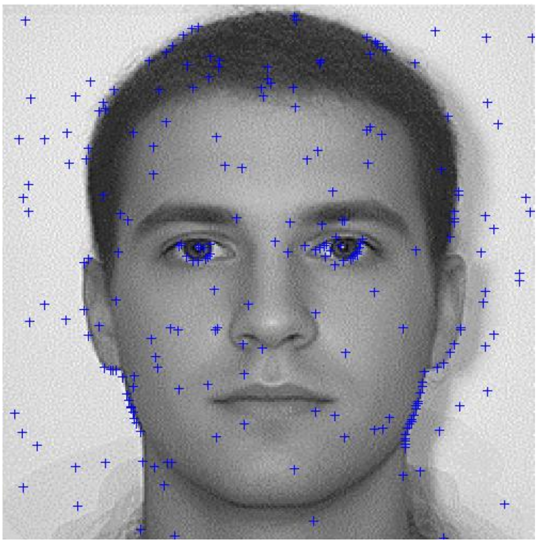

Fig. 4. Removal of Redundant Keypoints over the facial regions using Keypoint localization

2.1.5 Orientation Assignment

For each keypoint on the image contour, the orientation is assigned in terms of directional derivative, defined as:

$$D(q) \equiv \Delta\theta \arg\max_k \left\{ \left| \nabla_{\sigma,\rho} I(P_q;k) \right|, k = 1,2,\ldots,k \right\} \tag{7}$$

whereP_q represents corresponding pixel on the contour, *q* with keypoint. The use of pixel edges with keypoints correlates with its corresponding directions. $D(q)$ of the adjacent pixel is same when the edge segment is smooth. $D(q)$ of the adjacent pixel is indifferent in the corner regions. The interior point of the contour is defined as:

$$\tilde{D}(q) \equiv \min\left\{ \left| D(q-1) - D(q+1) \right|, \pi - \left| D(q-1) - D(q+1) \right| \right\} \tag{8}$$

Endpoint contour is defined as:

$$\tilde{D}(1) = \tilde{D}(q) \equiv \min\left\{\left|D(2) - D(q-1)\right|, \pi - \left|D(2) - D(q-1)\right|\right\} \quad (9)$$

Also, for the open contour, the equation is defined as:

$$D(1) \equiv \min\left\{\left|D(1) - D(2)\right|, \pi - \left|D(1) - D(2)\right|\right\} \quad (10)$$

$$D(q) \equiv \min\left\{\left|D(q) - D(q-1)\right|, \pi - \left|D(q) - D(q-1)\right|\right\}$$

The above expressions helps in reducing the robustness in measuring the directional derivatives of the edges or corners with keypoints. This robustness occurs in the edges or corners with keypoints due to noise and local neighborhood variations in the contours [20]. The orientation assignment makes a comparison between the facial access points between all the points in the fig. 2b and fig. 3. Eqs. 8, 9 and 10 helps in adding relevant keypoints from fig. 2b and thus producing fig. 5.

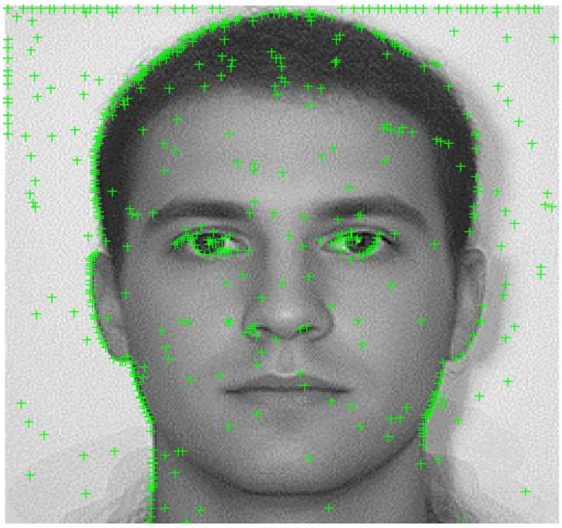

Fig. 5. Adding relevant keypoints using Orientation Assignment

2.1.6Keypoint Matching

Thus, weighted Euclidean distance with sample weights chosen from ground truth value helps in achieving nearest keypoints in the scale space. The threshold value d_R is compared between the ground image value and the actual image. When the obtained result is lesser than d_R, then the corresponding point is excluded. This helps in removing the extrema point with less contrast and eliminates poor localization. Thus, the redundant extrema point is removed using large curvature across edges and small curvature in the DoG function. When the difference is lesser than Eigen vector with larger to smaller value, Hessian matrix is formed at the keypoint location; thereby rejecting the corresponding keypoint. Thus, with the help of training data from the neural networks, accurate keypoints from ground truth image accurate the keypoints, precisely.Finally eliminating the poor and localized regions and edges of a facial image using scale space detection with adaptive thresholding. One more advantage of this technique is that use of Laplacian operators for maximum distance finding is further eliminated. With manual computation of keypoints, we could help in achieving exact keypoint over facial regions represented in fig. 5.

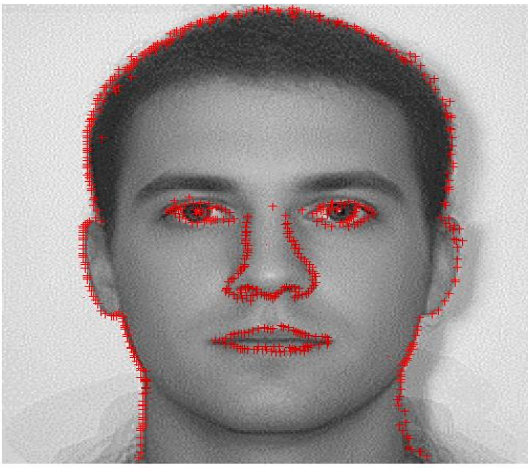

Fig. 5. Exact keypoint matching over Facial regions

2.2 RESVM

In a given image, description is given to training classes, training data and test points. The classes of interest is based on faces in an image and that is treated as the classes of interest within image regions. Training data belonging to the facial classes involves training the system with classes and learning its parameters. Here facial part in an image is chosen that represents class to obtain training data. This is done by choosing the facial access points near facial regions in an image. Facial access points are simply an irregular submatrix of the ground truth image. The pixel value in the irregular submatrix is converted into a vector. This is formed by connecting the submatrix column under previous column. This pixel vector is treated as training data of relevant class. In image classification, random pixel is selected as a test point belonging to selected classes. In RESVM classification method equal amount of test points are selected over each regions that represents facial classes. Misclassification error on test points helps to estimate the performance of the RESVM after being classified from A-SIFT.

Robust Ensemble Support Vector Machine (RESVM)bagging method uses Class-weighted Support Vector Machine (CWSVM). This is a supervised technique that uses penalty for misclassified labels in A-SIFT that differs per class. CWSVM uses PU (P – Positive, U – Negative) learning through unlabeled dataset from facial that contains negative labels i.e. noise on keypoints. Training using CWSVM is done to distinguish Positive labels (P)from Negative labels (U).During training phase, positive instances misclassification is penalized to a larger extent than unlabeled cases. This is done to emphasize higher degree certainty on P. In the context, the optimization for PU learning in training phase of CWSVM is defined as:

$$\begin{aligned} &\min_{\alpha,b,\xi} \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j \kappa(x_i,x_j) + C_P\sum_{i\in P}\xi_i + C_U\sum_{i\in U}\xi_i \\ &s.t.\, y_i\left(\sum\nolimits_{j=1}^{N}\alpha_i y_j \kappa(x_i,x_j)+b\right) \ge 1-\xi_i \qquad i=1,\dots,N \\ &\qquad\qquad \xi_i \ge 0 \qquad\qquad\qquad\qquad i=1,\dots,N \end{aligned} \tag{11}$$

with $\alpha \in \mathbb{R}^N$ is considered as a support value and $y \in \{-1,+1\}^N$ is the label vector and the kernel function is defined using $\kappa(.,.)$. Slack variables are termed as $\xi \in \mathbb{R}^N$ and b is the bias terms.C_P and C_U are the misclassification penalties that are required in order to make $C_P > C_U$. This penalties with various classes is used for tackling imbalanced datasets. RESVM resamples P along with U and uses degree of freedom for controlling misclassification penalty between P and U instances. Further this reduces the variability between base models (ground truth image).

The RESVM resamples potentially contaminated P and U that is treated as resampled sets with replacement persuadesinconsistencyacross resampled sets of U and P during training.The inconsistency between resamples increases with increased contamination on original dataset. Contamination levels lesser than 50% in a given dataset is treated as mislabeled. Large contamination in resamples increases the convergencesize of the expected contamination that equals the original dataset being resampled. Thus, with increased size of resamples the inconsistencyreduces in contamination. With varied contamination among training sets, could induceinconsistencyamongstthe facial imageand createsdiverse set of facial image samplesthough resampling both the P and U instances. Variancereduction exploits the inconsistencyin facial image model using resampling and tradeoff istaken place between increased inconsistencyby training smaller resamples and base models with improved stability with large training sets. Finally better classification of criminal faces could be obtained with P and U training in RESVM by dividing into subclasses as shown in fig. 6.

Training sets on SVM modelis quantified as dual weight and cases can be distinguished into three sets: (i) the training sets correctlyclassified lying outside the margin (α =0), (ii) the training setscorrectly classified lying on the margin ($\alpha \in [0, \square]$)and (iii) the training sets incorrectly classified lyinginside the margin ($\alpha = \square$). The training sets misclassified during trainingis treated as bounded with maximal α and considered as a controlSVM point. With bounded SVM mislabeled training instances could be eliminated to a smaller extent during training phase on learning the noise labels. There is a best case where mislabeled training sets classifies in relation with true label using training procedure of bounded SVM.

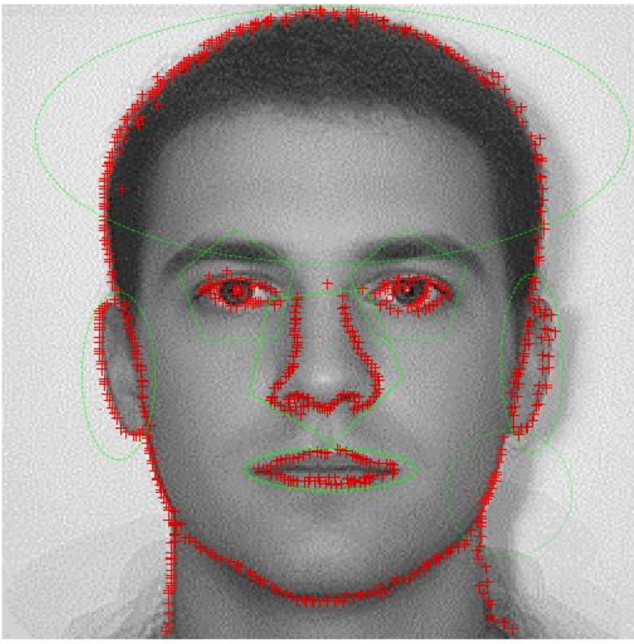

Fig. 6. Classification into subclasses using RESVM with P and U training

2.3 Hypothesis Testing

Consider pixel classes of two image (one main image and one ground truth image) with mean μ_1, μ_2. Test point for the image is chosen as x_0 with (x_{11}, x_{12}, x_{13} ,..., x_{1n_1}) and (x_{21}, x_{22}, x_{23},..., x_{1n_2}) as training data for class 1 and 2, respectively. Here, the hypothesis between the classes 1 and 2 of image 1 and 2 respectively is chosen to predict the pixel level keypoint hypothesis test. To test the image with hypothesis, null hypothesis is chosen between the image classes having equal mean with identical distribution between the classes 1 and 2. So, t-test appropriates null or equal mean hypothesis and Wilcoxon rank sum test is used for other hypothesis to find the alternate hypothesis. Here, this research considers two tests that is defined as follows:

1. Hypothesis Test 1 (H1): Place test point over training data of class 1 of image 1 for testing the null hypothesis H0.
2. Hypothesis Test 2 (H2): Place test point over training data of class 2 of image 2 for testing with class 1.

p-values for H1 and H2 is denoted through $PV_1(x_0)$ and $PV_2(x_0)$. The probability of classes with data for training is denoted by p_1 and p_2 .A small $PV_1(x_0)$ and large $PV_2(x_0)$ maintains a difference between the classes, when observation is been done on class 1.When this small and large value is applied over class 2, this will result in blurred boundary between the classes.

The relative test probability is calculated over test point x_0 is:

$$\frac{PV_1(x_0)}{PV_1(x_0) + PV_2(x_0)} \tag{12}$$

This relative test probability does not belong to class 1 and the probability of x_0 is calculated using class 1 is defined as:

$$1 - \frac{PV_1(x_0)}{PV_1(x_0) + PV_2(x_0)} \tag{13}$$

This method classifies x_0 through class 1 when $PV_2(x_0)\,p_1 > PV_1(x_0)\,p_2$ and class 2 when $PV_1(x_0)\,p_2 > PV_2(x_0)\,p_1$. Training data is chosen from classes 1 in image 1 or ground truth image with subclasses. The subclasses refers to class within facial regions that represents the eyes (subclass-1), nose (subclass-2), ear (subclass-3), chin (subclass-4), checks (subclass-4)etc. The image is fixed with standard size 512 × 512 and this training data is chosen as null hypothesis data. This is done to check the accurate relevancy of keypoints and test points in the class 1 image. When the accuracy of the resultant keypoints using RESVM and test points is greater than threshold value (TAV), image 1 is set as reference image for finding the faces in image 2. Then with the keypoints and test points, the image with multiple faces is identified using A-SIFT and RESVM. The accuracy of the resultant points with multiple subclasses in image 2 is tested with the reference points of multiple subclasses in reference image using t-test. When the accuracy is greater than the TAV, then the keypoints are extracted and matched with the image from the database.
The t-test case for class 1 with subclasses is represented with n_1 and n_2 points is denoted as:

$$T_1 = \frac{\dfrac{x_0 + x_{11} + x_{12} + \cdots + x_1 n_1}{n_1 + 1} - \dfrac{x_{21} + x_{22} + \cdots + x_2 n_2}{n_2}}{\sqrt{\dfrac{sd_1^2}{n_1 + 1} - \dfrac{\hat{\sigma}_2^2}{n_2}}} \tag{14}$$

$\hat{\sigma}_2^2$ is represented as a variance from class 2 and the standard deviation for class 1 is represented as:

$$sd_1^2 = \frac{\left[\left(x_0 - \hat{\mu}_1 \right)^2 + \sum_{i=1}^{n} \left(x_{1i} - \hat{\mu}_1 \right)^2 \right]}{n_1}$$
$$\hat{\mu}_1 = \frac{\overline{x_1} + \left(x_0 - \overline{x_1} \right)}{n_1 + 1} \tag{15}$$

The t-test 2 is represented as:

$$T_2 = \frac{\dfrac{x_{11} + x_{12} + \cdots + x_1 n_1}{n_1} - \dfrac{x_0 + x_{21} + x_{22} + \cdots + x_2 n_2}{n_2 + 1}}{\sqrt{\dfrac{sd_2^2}{n_1 + 1} - \dfrac{\hat{\sigma}_1^2}{n_2}}} \tag{16}$$

$\hat{\sigma}_2^2$ is represented as a variance from class 1 and the standard deviation for class 2 is represented as:

$$sd_2^2 = \frac{\left[\left(x_0 - \hat{\mu}_2 \right)^2 + \sum_{i=1}^{n} \left(x_{2i} - \hat{\mu}_2 \right)^2 \right]}{n_2}$$
$$\hat{\mu}_2 = \frac{\overline{x_2} + \left(x_0 - \overline{x_2} \right)}{n_2 + 1} \tag{17}$$

Thus for a given threshold value (TAV), the conditions are written as follows:

1. When maximum(PV_1, PV_2) ≥ TAV, test p-value point is greater than the TAV, then x_0 represents class 1 $PV_2(x_0)\,p_1 > PV_1(x_0)\,p_2$ and if $PV_2(x_0)\,p_2 > PV_1(x_0)\,p_1$ for class 2
2. When maximum(PV1, PV2) <TAV, test p-value point is smaller than the TAV, then x_0 represents class 1 $PV_2(x_0)\,p_1 < PV_1(x_0)\,p_2$ and if $PV_2(x_0)\,p_2 < PV_1(x_0)\,p_1$ for class 2

Results for Hypothesis Testing

Initially the test result is obtained over SIFT and compared with A-SIFT. Then, the image with A-SIFT keypoints are classified with RESVM to remove the robust keypoints and resulting in authorized keypoints or the facial access points from the input faces. This is initially done over a test or a reference image for the initial logging into the system. Then, the input photo of the criminal face is taken and the process using A-SIFT and RESVM is applied over it to find the accurate keypoints. Then, hypothesis testing is used for comparing the image with the reference image shown in fig. 8.Finally, image with higher threshold value is selected as displayed as output image.

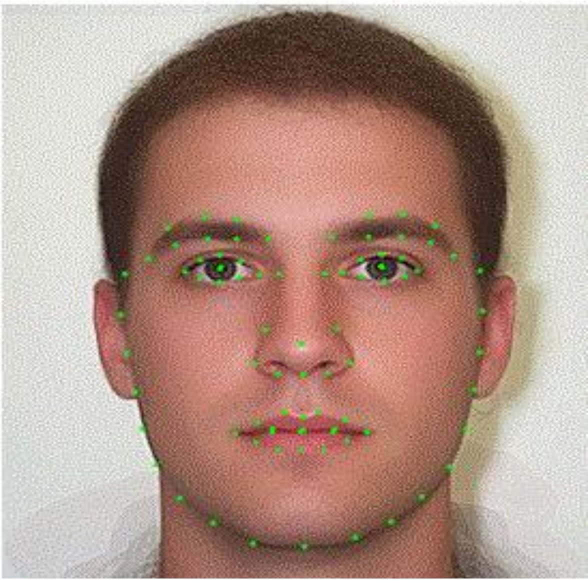

Fig. 8. Testing of Hypothesis with reference image

In the first test, same reference image is given as an input image to detect the face. The key points are refined using the cases of RESVM, using this we could further refine more keypoints and made the detection simpler as shown in fig. 8. Then the keypoints on the fig. 8 is made to match with the keypoint from fig. 7 and suitable areas were divided into subclasses. Then, each subclass is divided and measured finally retrieved the original face as shown in fig. 9.

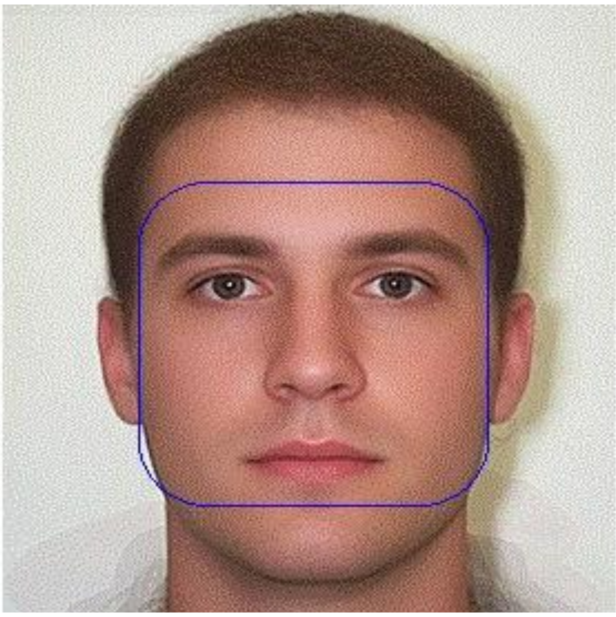

Fig. 8. Detected face with proposed technique

2.4 Accuracy Detection

The TAV is obtained by cross-validation with training data for choosing the value between the range on (0,1). This makes the error rate of cross-validation to be less for the training data giving best performance. Fixing threshold has a dominant reason that helps in finding the best face from database using hypothesis testing. Here significant levels 0.01 and 0.05 declares the significant solution for the desired input training data. When the significance is 0.05 the risk of error is more and vice versa. The value lesser than 0.01 i.e. 0.001 results in less test results and would not lead to a better significant results. Also, necessary improvement in the result is affected if the result is in 0.001 state. So, finally the result is maintained between 0.01 and 0.05 range to find the find the relative face obtained from the database. To detect the accuracy of the system, the several images are tested and we could find significance with 0.01 and 0.05 range. The image that are tested with same and different reference images and for the similar images the results were within this range. Reference images that are slightly indifferent produced a significance higher than this range. Finally, the efficacy of the proposed system is proved in terms of A-SIFT, RESVM and Hypothesis testing.

III. Conclusions

The Proposed Keypoint Hypothesis technique proved effective in terms of detecting the faces. The use of various techniques at each stage results in effective removal of redundant keypoint and making the detection simpler. Here, to achieve this at each stage we used Anisotropic-SIFT with modification in SIFT at each stage and making A-SIFT suitable for facial images. This is followed by the use of RESVM that enables the usage of P and U learning over the keypoint facial image and segregating the subclasses over facial regions using Facial Access Points grouping over a particular region. This helped the further stages to effectively compare the trained image or the reference image with suitable database of image. This comparison is done using Hypothesis testing with 2 hypothesis that enables the accuracy detection to be more effective. Depending on the results of TAV, the process might get continued or might end. Thus the effectiveness of the system is proved in terms of facial image. The proposed system could be applied over other environments for detecting the objects. Also, comparison with other facial techniques will make the system to get enhanced further.

Acknowledgement

References

[1] Deng, Houtao, et al. "A time series forest for classification and feature extraction." Information Sciences 239 (2013): 142-153.

[2] Tao, Dacheng, et al. "General tensor discriminant analysis and gabor features for gait recognition." Pattern Analysis and Machine Intelligence, IEEE Transactions on 29.10 (2007): 1700-1715.

[3] delRivero, José Antonio Sánchez, et al. "Feature selection for classification of animal feed ingredients from near infrared microscopy spectra." Information Sciences 241 (2013): 58-69.

[4] Tao, Dapeng, et al. "Similar handwritten Chinese character recognition using discriminative locality alignment manifold learning." Document Analysis and Recognition (ICDAR), 2011 International Conference on. IEEE, 2011.

[5] Huang, Qing-Hua. "Discovery of time-inconsecutive co-movement patterns of foreign currencies using an evolutionary biclustering method." Applied Mathematics and Computation 218.8 (2011): 4353-4364.

[6] Wang, Yu, et al. "License plate recognition based on SIFT feature." Optik-International Journal for Light and Electron Optics 126.21 (2015): 2895-2901.

[7] Wang, Shen, Chen Cui, and XiamuNiu. "Watermarking for DIBR 3D images based on SIFT feature points." Measurement 48 (2014): 54-62.

[8] Wu, Xiangqian, Qiushi Zhao, and Wei Bu. "A SIFT-based contactless palmprint verification approach using iterative RANSAC and local palmprint descriptors." Pattern Recognition 47.10 (2014): 3314-3326.

[9] Chen, Jiansheng, and Yiu-Sang Moon. "Using SIFT features in palmprint authentication." Pattern Recognition, 2008. ICPR 2008. 19th International Conference on. IEEE, 2008.

[10] Zhao, Qiushi, Wei Bu, and Xiangqian Wu. "SIFT-based image alignment for contactless palmprint verification." Biometrics (ICB), 2013 International Conference on. IEEE, 2013.

[11] Kong, Adams Wai-Kin, and David Zhang. "Competitive coding scheme for palmprint verification." Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. Vol. 1. IEEE, 2004.

[12] Li, Yanshan, et al. "GA-SIFT: A new scale invariant feature transform for multispectral image using geometric algebra." Information Sciences 281 (2014): 559-572.

[13] Abdel-Hakim, Alaa E., and AlyFarag. "CSIFT: A SIFT descriptor with color invariant characteristics." Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. Vol. 2. IEEE, 2006.

[14] Li, Canlin, and Lizhuang Ma. "A new framework for feature descriptor based on SIFT." Pattern Recognition Letters 30.5 (2009): 544-557.

[15] Claesen, Marc, et al. "A robust ensemble approach to learn from positive and unlabeled data using SVM base models." Neurocomputing 160 (2015): 73-84.

[16] Yang, Xiao-Min, et al. "Image feature extraction and matching technology [J]." Optics and Precision Engineering 9 (2009): 033.

[17] Gao, Xiaorong, et al. "Vehicle bottom anomaly detection algorithm based on SIFT." Optik-International Journal for Light and Electron Optics 126.23 (2015): 3562-3566.Wang, Yingmei, et al. "Adaptive filtering with self-similarity for low-dose CT imaging." Optik-International Journal for Light and Electron Optics 126.24 (2015): 4949-4953.

[18] Xu, Jiangtao, et al. "An improved anisotropic diffusion filter with semi-adaptive threshold for edge preservation." Signal Processing 119 (2016): 80-91.

[19] Zhang, Wei-Chuan, and Peng-Lang Shui. "Contour-based corner detection via angle difference of principal directions of anisotropic Gaussian directional derivatives." Pattern Recognition 48.9 (2015): 2785-2797.