International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 13, Number 12 (2020), pp. 5113-5122

Hole Filling And Image Fusion Approach For RGBD Database

Aniketh A. Gaonkar[1], Narayan T. Vetrekar[1], Rajendra S. Gad[1]
[1] Department. of Electronics, Goa University, Taleigao Plateau, Goa, India

Abstract

3D facial biometric is one of the areas with paramount attraction to the researchers. It is an easily accessible trait of biometrics as the captured information is having the additional distance component termed 'Depth' along with the RGB component. The low-cost 3D Kinect camera sensors are used in various biometric applications, however the low-resolution camera of Kinect incorporates the information loss i.e. Hole (zero pixel value) in the depth map image. Thus, to address this issue, we present in this paper, three different filters based on kernel function to fill the hole, such that the performance of face recognition system can be improved. We present the results using GU-RGBD database collected for 64 subjects in eight different variants using seven different face recognition methods. The experimental results are evaluated on depth map image, fusion of RGB and depth map and score level fusion to demonstrate our approach. We obtain the recognition rate and verification rate to present the benchmark results. Further, the recognition rate obtained with the filter outperforms the without filter results, demonstrating the significance of this approach.

Keywords: Depth Images, RGBD Images, Hole Filling, Face Recognition

I. INTRODUCTION

Facial biometrics has grabbed the unique position in research world as compared to other biometrics traits such as iris, voice, gait etc. as it is easily accessible and convenient trait of biometric. With the development of low-cost 3D Kinect camera sensor the research in the area of 3D biometrics has shown a remarkable expansion in the literature. The captured 3D images have proficiency to overcome the limitations due to illumination variation and pose variation which commonly affects the 2D imaging system [1]. This is mainly because of the inherent property associated with 3D faces i.e. 3D camera system has ability of capturing more spatial information as compared to 2D systems in form of depth map images (distance from each pixel to the sensor) along with RGB images [2].The Kinect camera has employed VGA resolution camera for capturing RGB images and the depth information is captured with the help of infrared emitter and CMOS sensor [8,9].The captured images are of low resolution and noisy [6]. The information loss in the form of holes in the depth images needs to be addressed and filled at the pre-processing stage, as it affects the overall performance of the system.

In literature one can find the contribution of Yu Mao et al. towards identification and filling of expansion holes. Based on depth histogram and the linear interpolation and the graph-based interpolation methods; the holes are identified at initial stage and are filled. [3]. Research in the line of dis-occlusion removal in Depth Image-Based Rendering (DIBR): hierarchical hole-filling (HHF) and depth adaptive hierarchical hole-filling tactics, where a pyramid approach is followed from lower resolution estimate of 3D wrapped image to estimate the hole pixels value, is contributed by Mashhour Solh et al. research [4]. Hole filling algorithm to improve image quality of DIBR has been proposed by Dan Wang et al. Here the order of hole filling is determined by the sum of priority calculation function and the depth information. Further the gradient information is used as auxiliary information to find the best matching block [5][6]. Litong Feng et al. has worked on an adaptive background biased depth map hole-filling method [7]. Amir Atapour-Abarghouei et.al., has addressed hole filling by Fourier transform and Butterworth high/low pass filtering. Here the texture synthesis method is used for high frequency details and structural inpainting is used for inpainting the low frequency information. Further high frequency depth synthesis has been performed by query expansion concept with the final output and then recombined in Fourier space. [8]Liang et al. [9] had proposed a segmentation-based approach for inpainting stereo images. Here the constraint was that the missing information in one stereo image may be filled from other image in both colour and depth images using depth-assisted texture synthesis. In absence of stereo or multi-camera views some other approaches have to be used to fill the holes in the depth images. Toby P. Breckon et al. [10] had proposed an surface completion technique based on the nonparametric propagation of existing scene information from the known /visible scene areas to the unknown/invisible 3D regions in combination with the preliminary underlying geometric surface completion.

I.I Contributions

This entire literature survey directs that, to obtain better performance the filling of holes/missing information is necessary. Here in this paper we present interpolation-based hole filling techniques/filters for the depth images. We have engaged the RGB and Depth map images from our own GU_RGBD database to demonstrate the study. Extensive study has been performed to compute identification and verification rate of the depth map images and the fused RGBD images by engaging various State of the Art feature extraction algorithms such as Histogram of Oriented Gradient (HOG) [11], Principal Component Analysis (PCA) [12], GIST [13], Local Binary Pattern (LBP) [14], LogGabor [15], Local Phase Quantization (LPQ) [16] and Binarized Statistical Image Features (BSIF) [17]. The RGB and Depth map images are fused using the Pixel level Average Image fusion technique. Further, to demonstrate the significance of our approach, we

present evaluation results on depth map images, RGB-D fused images (Fusion of depth map is performed after hole filling) and score level fusion (On two best performing algorithm). Further in due course of this work summarizes number of contributions as follows:

- A hole filling approach in the depth map images acquired from Kinect sensor to enhance the performance of 3*D* face recognition system.
- A simple and effective approach based on variable kernel size for filling the holes with the contribution from neighboring pixels.
- The study presenting the significance of employing hole filling techniques to improve the performance of the state-of-the-art face recognition methods.
- Experimented on seven different feature extraction methods such as Principal Component Analysis (PCA), Histogram of Oriented Gradient (HOG), Local Binary Pattern (LBP), Local Phase Quantization (LPQ), GIST, Binarized Statistical Image Features (BSIF) and LogGabor to demonstrate the applicability of our approach for improved performance analysis.
- Presents the evaluation results in the form of recognition rate and verification rate on depth map image alone, fusion of RGB with Depth map image, and Score level fusion extensively.

Rest of the paper is structured as follow: 'Database' section present the details related to the database employed in this work. 'Hole filling filters' Section presents the mathematical details and description of each of the three different hole filling filters (linear interpolation, exponential averaging, and weighted averaging) employed in this work. 'Experimental and Result' section introduces the exhaustive experimental evaluations results in verification and recognition rate using seven different state-of-the-art methods for RGB-D face recognition followed by our final conclusion along with future remarks based on the evaluation results.

II. DATABASE

In this section we present in detail the GU-RGBD database employed in this work. The facial images collected using RGB-D Kinect sensor (Xbox 360 Kinect depth camera from Microsoft) under controlled (session 1) and uncontrolled (session 2) environmental condition. The controlled conditions were maintained by illuminating the dark room with the two QTH light sources of 600watt, which were mounted at an angle 45˚ normal to the subject position and for the uncontrolled condition the subjects were exposed to an ambient light

The GU-RGBD database is a collection of 2048 images (comprises of 2*D* color image and corresponding depth map image) acquired from 64 subjects consists of 15 female and 49 male subjects. The database consists of facial variants such as front, pose, expression and occlusion. The pose variation consists of variations like 45◦, -45◦, 90◦, -90◦. Expression variation, includes smile and eye closed pattern and the occlusion was obtained by covering the half of the face vertically using paper. Figure 1 illustrates the sample images collected under GU-RGBD database using 3*D* Kinect sensor. In the GU_RGBD database we have tried to cover all the possible wide angles of the human faces, which can be captured by the camera along with the different poses and occlusion to have a diversified study.

Fig. 1. Sample images from GU-RGBD Database

III. HOLE FILLING FILTERS

The low resolution of the Kinect camera incorporates noise and inaccuracy in the captured images [18]. It is observed that the holes (zero value pixels) are generated in the captured images mostly due to fluctuation in distance between subject and camera, poor reflectance of surface to projected IR light. The degradation of image quality due to unavailable information affects the system performance and the recognition rate. The performance can be enhanced by engaging appropriate filtering techniques at the pre-processing stage. Here we have designed three interpolation based hole filling techniques/filters using information from populous of the neighbouring pixel. The description and the acronym for the said filters are mentioned in table 1.

Table 1. Description of three designed filters

Acronym	Description
WO	Depth map image without filtering
LI- Filter	Linear interpolation based filtered depth map image
EA- Filter	Exponential averaging based filtered depth map images
WA- Filter	Weighted Average based filtered depth map images.

The conceptual illustration of the working of filter for filling holes on face triangle is given in figure2. In our work, the N/4 dummy rows and columns pixels are added to outspread the depth images of N x N dimension to the higher proportion of aspect ratio, so as to circumvent the occurrence of computational errors at the peripheral pixels of the depth images as kernel function inflates. There is no contribution from dummy elements (pixels) in hole filling as they have

been allotted with the NaN value to avoid the false computations.

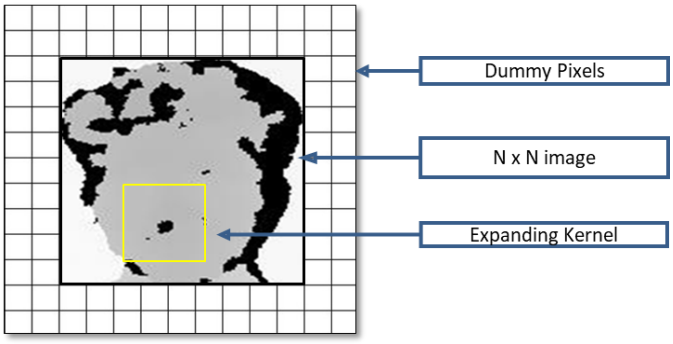


Fig. 2. Illustration of working of filters for hole filling in the depth amp images

Let v(x,y) be the depth image of dimension N x N on which the hole filling operations has to be performed. The process is initiated by defining the kernel $u_{(l,l)}$ of dimensions *l* x *l* and scanning it over the depth map image N x N so as to locate the hole (zero value pixel) $y_{(i,j)}$ to be filled. Mathematically the kernel expression is given as in equation.1

$$u_{(l,l)} = \sum_{k_1=0}^{l} \sum_{k_2=0}^{l} u_{(i\mp k_1, j\mp k_2)} \; u_{(l,l)} = \sum_{k_1=0}^{l} \sum_{k_2=0}^{l} u_{(i\mp k_1, j\mp k_2)} \; \sum_{k_1=0}^{l} \sum_{k_2=0}^{l} u_{(i\mp k_1, j\mp k_2)}$$

Where *l*= 1, 2, 3,……,N/4 and with *l*=0, kernel does not exist.

III.I LI - Filter: Linear interpolation based filtering method

The kernel given in equation 1 is the average contribution of zero and non-zero pixels and it can be written as follows.

$$u_{(l,l)} = \overline{u_{(i,j)_o}} + \overline{u_{(i,j)_{\bar{0}}}} \qquad (2)$$

Here $\overline{u_{(i,j)_o}}$ is selected as 5% and $\overline{u_{(i,j)_{\bar{0}}}}$ as 95% contribution of zero and non-zero pixels thus giving higher importance to the populous of non-zero value to interpolate the zero pixels. The expansion of kernel is restricted by the condition: n_1*nos($\overline{u_{(i,j)_o}}$) ≤ n_2*nos($\overline{u_{(i,j)_{\bar{0}}}}$), (where n_1=0.05 and n_2=0.95)**Error! Digit expected.**. On fulfilling the above condition, the average value is generated of the kernel equation and interpolates the zero-value pixel as described in equation. 3.

$$y'_{(i,j)} = y_{(i,j)} + u_{(l,l)} \qquad \forall\, y_{(i,j)} = 0 \qquad (3)$$

III.II EA- Filter: Exponential averaging based filtering method

Consider the kernel function defined in the equations 1. The set kernel expands until the condition n1*nos($\bar{u}_{(I,J)o}$) ≤ n2*nos($\bar{u}_{(I,J)\bar{0}}$) is obtained, where n_1=0.05 and n_2=0.95**Error! Digit expected.**. Further the factor α^n is introduced to give exponential weightages to the varying kernel sizes. Here α=0.9 (variable) and n=total number of windows through which the kernel propagates, the value of 'n' is assigned only to the windows with the non-zero average values in such a way that the maximum importance is given to the immediate neighbor and it exponentially decreases as window propagates. Hence, the filter equation 3 can be modified as:

$$y'_{(i,j)} = y_{(i,j)} + \sum_{l=0}^{n} \alpha^n * (u_{(l,l)}) \quad \forall y_{(i,j)} = 0 \qquad (4)$$

III.III WA- Filter: Weighted Average based filtering method

This filtering technique follows the similar conditions as that of the previous filters for selection and propagation of the kernel, i.e. n_1*nos($\bar{u}_{(I,J)o}$) ≤ n_2*nos($\bar{u}_{(I,J)\bar{0}}$), (where n_1=0.05 and n_2=0.95)**Error! Digit expected.**.**Error! Digit expected.** Here the weighted factor 1-(n•α) has been introduced, where n=number of windows and α = 1/ (M/4). The weights are distributed as per the weighted factor such that the highest weights is assigned to the immediate neighboring pixels having non-zero average value of the set kernel and gradually reduces over the number of windows. Thus, the filter equation can be written as:

$$(1)u_{(l,l)} =$$
$$y'_{(i,j)} = y_{(i,j)} + \sum_{l=0}^{n} \big(1 - (n \bullet \alpha)\big) * (u_{(l,l)}) \quad \forall y_{(i,j)} = 0 \qquad (5)$$

IV. EXPERIMENTS AND RESULTS

In this section of the paper, we present in detail the experimental evaluation protocol and related experimental results obtained in this work. Basically, we perform the hole filling on 3D depth map image using three different filters: LI-Filter: Linear Interpolation, EA-Filter: Exponential Averaging, and WA-Filter: Weighted Averaging. Further we make use of kernel window function to give proper weightage to the neighbouring elements in the kernel, while employing the filtering method. The experimental evaluation results are evaluated on GU-RGBD database consisting 64 subjects. To demonstrate the applicability of using kernel based filtering approach, we present a systematic results on seven different state-of-the-art feature extraction methods Principal Component Analysis (PCA), Histogram of Oriented Gradient (HOG), Local Binary Pattern (LBP), Local Phase Quantization (LPQ), GIST, Binarized Statistical Image Features (BSIF) and LogGabor, employed in 3*D*. We present the evaluation results in the form of recognition rate and verification rate in tabular and graphical form.

IV.I Evaluation Protocol

This section presents the evaluation protocol employed for experimental analysis. Using the GU-RGBD database, we partitioned our dataset into reference and probe. Reference set

consists of 64 subject corresponding to front face (Neutral) including their samples from session 1 data, while prob set consists corresponding 64 subject belongs to either 45°, -45°, 90°, -90°, smile, eye closed, paper occlusion from session 1 independently when operated on seven different feature extraction methods discussed above. In the similar manner, we also present the results when the reference belongs to session 1 and prob set belongs to session 2 (for simplicity, we did not include these results in the paper for now).

Using the experimental protocol, we present the three sets of experimental evaluation to demonstrate our approach of hole filling. The Evaluation 1 presents the results related to depth map, Evaluation 2 presents the results related to the fusion of RGB and Depth map (after filtering), and Evaluation 3 illustrate the score level fusion best performing algorithm separately on depth map image as well as on RGB image fused with depth map. Also, in our experimental evaluation, RGB image was first converted to gray scale image for the sake of processing.

IV.II Evaluation 1: Depth Map Image

In this section we present performance accuracy of three different filters using seven feature extraction methods. The idea is to demonstrate the performance accuracy of face recognition methods using three different filters compared to without filtering depth map image. Thus, in this section we present the benchmark results on seven different facial variants to present the significance of our approach using kernel function in implementing the filter. Table 2, 3, 4 presents the recognition rate at Rank-5, Figure 3 presents the Cumulative Match Curve (CMC) plots and Figure 4 present the Receiver Operating Curve (ROC) for this set of evaluation. Clearly, a reasonable improvement in performance accuracy of face recognition system can be seen based on our proposed filtering approach compared to without filtered depth map image. Also the major improvement with exponential average and weighted average can be observed compared to linear interpolation filtering. This further validates our idea of employing the kernel based filtering to give weighted importance to the neighboring pixel values in filling the hole. Further, based on the evaluation results obtained for facial expression, we present our major observations as follows:

- The invariants such as smile and eyes closed obtained the recognition rate at rank-5, higher as compared to other existing facial variants. The better performance in these cases is due to presence of full-face geometry compared to the reference depth map image of frontal face. Specifically, the recognition rates obtained for smile variant using PCA algorithm is 89.06% (without filter) and 90.63% (with LI-filter and WA-filter).
- Using HOG the recognition rate obtained in 93.75% (without filter) and 96.88% (with LI-filter), while the lower performance is noted for EA-filter and WA-filter. Further, with LBP the recognition rate is 48.44% (without filter) and 50% (with LI-filter) compared to highest recognition rate of 54.69% (with EA-filter and WA- filter). On the other hand the recognition rate of LPQ without filter is poor, but significant improvement can be observed using filters (42.19% (without filter) and 62.50% (with LI-filter and EA-filter), while further enhancement was seen for WA-filter obtaining 64.06% recognition rate).
- Similarly, the methods such as GIST, BSIF, LogGabor (LG) indicates the better results compared to LBP, HOG, LPQ, demonstrating the robust of these methods for face recognition. The maximum recognition rate of 89.06%, 84.38% and 92.19% has been noted using GIST algorithm (with LI-filter and EA-filter), BSIF (with EA-filter and WA-filter) and LogGabor (with LI-filter) respectively. Thus, overall it can be seen that the recognition rate for smile using all three designed filters outperforms the baseline results without filter, for face recognition algorithms used in this work.

On the other hand recognition rate for 45° and -45° pose variation is dominant over 90° and -90° pose variation. As expected this decrease in the performance due to the larger angular variation with ±90° as compared to ± 45° for all the three filters. While we note that from the observation table that, with the application of filters the performance is improved reasonably well across most of the face recognition algorithms. Although, due to angular face variability the performance of at 45° and -45° is lower than the full face, effect of the filters has shown remarkable noted improvement in the performance accuracy.

IV.III Evaluation 2: Fused (RBG + Depth) image

In this section we perform the performance analysis based on fusing RGB image and Depth map image, while the depth map images used for fusion is after employing the filters. Although depth map gives more information than 2D image, depth map presents less variability among the subjects, thus combining the RGB and depth map image can improve the performance reasonably. In this evaluation to fuse the two images, we used simple, yet an effective method by sim- ply averaging RGB and depth map image. Further, the performance analysis is obtained based on the evaluation protocol across seven different face recognition algorithm mentioned above. Table 5, 6, 7 presents the recognition rate at Rank-5; Figure 5 presents the Cumulative Match Curve (CMC) plots and Figure 6 present the Receiver Operating Curve (ROC) for this set of evaluation. Further to have fair comparison, we also compute the results on fusion of RGB and depth map, when filters are not employed. As expected the overall results based on fusing RGB and depth map shows the considerable improvements in the results across all the algorithm, while facial variants such as smile show overall highest recognition accuracy. Therefore, in this section we present the specific observation related to smile variants in comparison with the previous evaluation result obtained only with depth, as follow:

- For PCA, the highest recognition rate of 95.31% for fusion is obtained using LI, EA filter compared to

90.63%, 89.06% recognition rate noted using LI, EA filter for only depth map image.

- For HOG, the highest recognition rate of 98.44% for fusion is obtained using EA-filter compared to 90.63%, recognition rate noted using EA filter for only depth map image. For LBP, the highest recognition rate of 68.75% for fusion is obtained using WA-filter compared to 54.69%, recognition rate noted using WA filter for only depth map image. For LPQ, the highest recognition rate of 81.25% for fusion is obtained using LI, EA-filter compared to 62.50%, recognition rate noted using LI, EA-filter for only depth map image.

Compared to the algorithms employed in this work, GIST, BSIF, LG demonstrates the consistently higher performance analysis. For GIST, the highest recognition rate of 93.75% for fusion is obtained using EA-filter compared to 89.06%, recognition rate noted using EA filter for only depth map image. For BSIF, the highest recognition rate of 96.88% for fusion is obtained using EA-filter compared to 84.38%, recognition rate noted using EA filter for only depth map image. Similarly for LG, the highest recognition rate of 100% for fusion is obtained using LI, EA-filter compared to 92.19%, 90.63%, recognition rate noted using LI, EA filter for only depth map image.

IV.IV Evaluation 3: Fusion based on scores

Similar to the previous section, we present the evaluation results based on score level fusion. Precisely, we fused the scores of best performing algorithm based on the previous two evaluation results. Based on the previous results, we employed PCA and HOG score level fusion approach to demonstrate this set of results. Further, the results we repeated for depth map image alone and fusion of RGB with depth map. The scores are fused using simple sum rule to demonstrate the results. Table 8, and Table 9, presents the recognition rate at Rank-5 for depth map and the fusion of RGB with depth map.

The results obtained with the smile are found to be better in all the evaluation results performed in this experiment. Hence we present the observation related to smile in this section of experiment com- pared to the performance of depth map image and RGBD (fused) images results performed using the three designed filters. Based on the obtained results we make our observation as follow:

- Although, the score level fusion results shows the improvement in the recognition accuracy compared to the individual algorithm operated on depth map image, however, not much difference is observed when compared with individual algorithms performance operated on fusion of RGB and depth map image.
- Further, when compared with without filter results, the score level fusion obtains highest recognition rate of 98.44% for WA-filter com- pared to 93.75% obtained without filter. This indicates the improvement in the performance with the inclusion of filtering method for better performance.

Table 2. Recognition rate at Rank-5 using Depth map image after employing LI-Filter

Methods	45		90		-45		-90		Smile		Close Eyes		Paper Occlusion	
	WO	LI	WO	LI	WO	LI	WO	LI	WO	LI	WO	LI	WO	LI
PCA	21.88	21.88	15.63	14.06	17.19	**17.19**	12.50	**15.63**	89.06	**90.63**	89.06	85.94	32.81	21.88
HOG	17.19	**21.88**	14.06	10.94	18.75	10.94	14.06	12.50	93.75	90.63	89.06	**90.63**	46.88	23.44
LBP	18.75	**21.88**	14.06	14.06	17.19	**20.31**	14.06	**17.19**	48.44	**50.00**	34.38	**53.13**	7.81	**10.94**
LPQ	10.94	**15.63**	12.50	**18.75**	14.06	12.50	14.06	7.81	42.19	**62.50**	37.50	**54.69**	7.81	**9.38**
GIST	20.31	**20.31**	14.06	**14.06**	18.75	15.63	15.63	14.06	89.06	**89.06**	89.06	84.38	18.75	14.06
BSIF	18.75	**18.75**	15.63	10.94	14.06	**15.63**	12.50	10.94	82.81	79.69	78.13	**78.13**	7.81	**10.94**
LG	18.75	10.94	12.50	9.38	9.38	**14.06**	17.19	14.06	92.19	**92.19**	90.63	89.06	31.25	18.75

Table 3. Recognition rate at Rank-5 using Depth map image after employing EA-Filter

Methods	45		90		-45		-90		Smile		Close Eyes		Paper Occlusion	
	WO	EA	WO	EA	WO	EA	WO	EA	WO	EA	WO	EA	WO	EA
PCA	21.88	**21.88**	15.63	**17.19**	17.19	**18.75**	12.50	**17.19**	89.06	**89.06**	89.06	87.50	32.81	31.25
HOG	17.19	**29.69**	14.06	12.50	18.75	12.50	14.06	**14.06**	93.75	90.63	89.06	**90.63**	46.88	28.13
LBP	18.75	**25.00**	14.06	12.50	17.19	**21.88**	14.06	**15.63**	48.44	**54.69**	34.38	**53.13**	7.81	**10.94**
LPQ	10.94	**21.88**	12.50	**14.06**	14.06	**14.06**	14.06	7.81	42.19	**62.50**	37.50	**51.56**	7.81	7.81
GIST	20.31	**20.31**	14.06	**15.63**	18.75	15.63	15.63	14.06	89.06	**89.06**	89.06	85.94	18.75	**21.88**
BSIF	18.75	**25.00**	15.63	**17.19**	14.06	9.38	12.50	7.81	82.81	**84.38**	78.13	75.00	7.81	**20.31**
LG	18.75	9.38	12.50	7.81	9.38	**14.06**	17.19	12.50	92.19	90.63	90.63	89.06	31.25	21.88

Table 4 Recognition rate at Rank-5 using Depth map image after employing WA-Filter

Methods	45		90		-45		-90		Smile		Close Eyes		Paper Occlusion	
	WO	WA	WO	WA	WO	WA	WO	WA	WO	WA	WO	WA	WO	WA
PCA	21.88	20.31	15.63	**17.19**	17.19	**23.44**	12.50	**15.63**	89.06	**90.63**	89.06	85.94	32.81	26.56
HOG	17.19	**23.44**	14.06	**15.63**	18.75	**18.75**	14.06	**14.06**	93.75	**96.88**	89.06	**90.63**	46.88	25.00
LBP	18.75	**18.75**	14.06	**18.75**	17.19	**23.44**	14.06	9.38	48.44	**54.69**	34.38	**48.44**	7.81	9.38
LPQ	10.94	**14.06**	12.50	**20.31**	14.06	**18.75**	14.06	**14.06**	42.19	**64.06**	37.50	**53.13**	7.81	7.81
GIST	20.31	17.19	14.06	10.94	18.75	12.50	15.63	9.38	89.06	84.38	89.06	87.50	18.75	12.50
BSIF	18.75	**23.44**	15.63	14.06	14.06	10.94	12.50	10.94	82.81	**84.38**	78.13	**78.13**	7.81	**17.19**
LG	18.75	10.94	12.50	7.81	9.38	**10.94**	17.19	12.50	92.19	87.50	90.63	85.94	31.25	14.06

Table 5. Recognition rate at Rank-5 using Fused (RBG + Depth) image after employing LI-Filter

Methods	45		90		-45		-90		Smile		Close Eyes		Paper Occlusion	
	WO	LI	WO	LI	WO	LI	WO	LI	WO	LI	WO	LI	WO	LI
PCA	25.00	**25.00**	14.06	**14.06**	21.88	**23.44**	15.63	**18.75**	95.31	**95.31**	92.19	**93.75**	45.31	28.13
HOG	17.19	**18.75**	12.50	10.94	18.75	14.06	14.06	6.25	98.44	96.88	93.75	**95.31**	51.56	32.81
LBP	15.63	12.50	10.94	7.81	18.75	14.06	10.94	**10.94**	73.44	59.38	79.69	67.19	10.94	10.94
LPQ	25.00	23.44	7.81	**9.38**	10.94	7.81	7.81	**7.81**	81.25	**81.25**	82.81	**82.81**	28.13	14.06
GIST	31.25	**32.81**	14.06	7.81	14.06	9.38	17.19	**18.75**	96.88	92.19	90.63	**90.63**	26.56	21.88
BSIF	26.56	21.88	18.75	9.38	20.31	18.75	10.94	**15.63**	96.88	92.19	93.75	**93.75**	15.63	**21.88**
LG	15.63	**18.75**	7.81	**10.94**	18.75	**20.31**	15.63	14.06	98.44	**100.00**	96.88	**96.88**	18.75	15.63

Table 6. Recognition rate at Rank-5 using Fused (RBG + Depth) image after employing EA-Filter

Methods	45		90		-45		-90		Smile		Close Eyes		Paper Occlusion	
	WO	EA	WO	EA	WO	EA	WO	EA	WO	EA	WO	EA	WO	EA
PCA	25.00	23.44	14.06	**17.19**	21.88	20.31	15.63	**18.75**	95.31	**95.31**	92.19	**93.75**	45.31	40.63
HOG	17.19	**20.31**	12.50	10.94	18.75	15.63	14.06	7.81	98.44	**98.44**	93.75	**96.88**	51.56	**53.13**
LBP	15.63	14.06	10.94	**10.94**	18.75	17.19	10.94	**10.94**	73.44	56.25	79.69	67.19	10.94	**17.19**
LPQ	25.00	18.75	7.81	**7.81**	10.94	7.81	7.81	**7.81**	81.25	**81.25**	82.81	75.00	28.13	18.75
GIST	31.25	29.69	14.06	7.81	14.06	12.50	17.19	**21.88**	96.88	93.75	90.63	**90.63**	26.56	25.00
BSIF	26.56	15.63	18.75	14.06	20.31	17.19	10.94	**12.50**	96.88	**96.88**	93.75	**96.88**	15.63	**17.19**
LG	15.63	**17.19**	7.81	**10.94**	18.75	18.75	15.63	**15.63**	98.44	**100.00**	96.88	**96.88**	18.75	**18.75**

Table 7. Recognition rate at Rank-5 using Fused (RBG + Depth) image after employing WA-Filter

Methods	45		90		-45		-90		Smile		Close Eyes		Paper Occlusion	
	WO	WA	WO	WA	WO	WA	WO	WA	WO	WA	WO	WA	WO	WA
PCA	25.00	**26.56**	14.06	**18.75**	21.88	**21.88**	15.63	**18.75**	95.31	93.75	92.19	90.63	45.31	37.50
HOG	17.19	**25.00**	12.50	**12.50**	18.75	17.19	14.06	10.94	98.44	96.88	93.75	**95.31**	51.56	**56.25**
LBP	15.63	14.06	10.94	**10.94**	18.75	**21.88**	10.94	**12.50**	73.44	68.75	79.69	67.19	10.94	**20.31**
LPQ	25.00	20.31	7.81	**7.81**	10.94	7.81	7.81	**9.38**	81.25	76.56	82.81	78.13	28.13	23.44
GIST	31.25	20.31	14.06	4.69	14.06	12.50	17.19	12.50	96.88	90.63	90.63	**93.75**	26.56	15.63
BSIF	26.56	21.88	18.75	**18.75**	20.31	15.63	10.94	9.38	96.88	**96.88**	93.75	**98.44**	15.63	**21.88**
LG	15.63	**17.19**	7.81	**12.50**	18.75	**20.31**	15.63	9.38	98.44	96.88	96.88	95.31	18.75	12.50

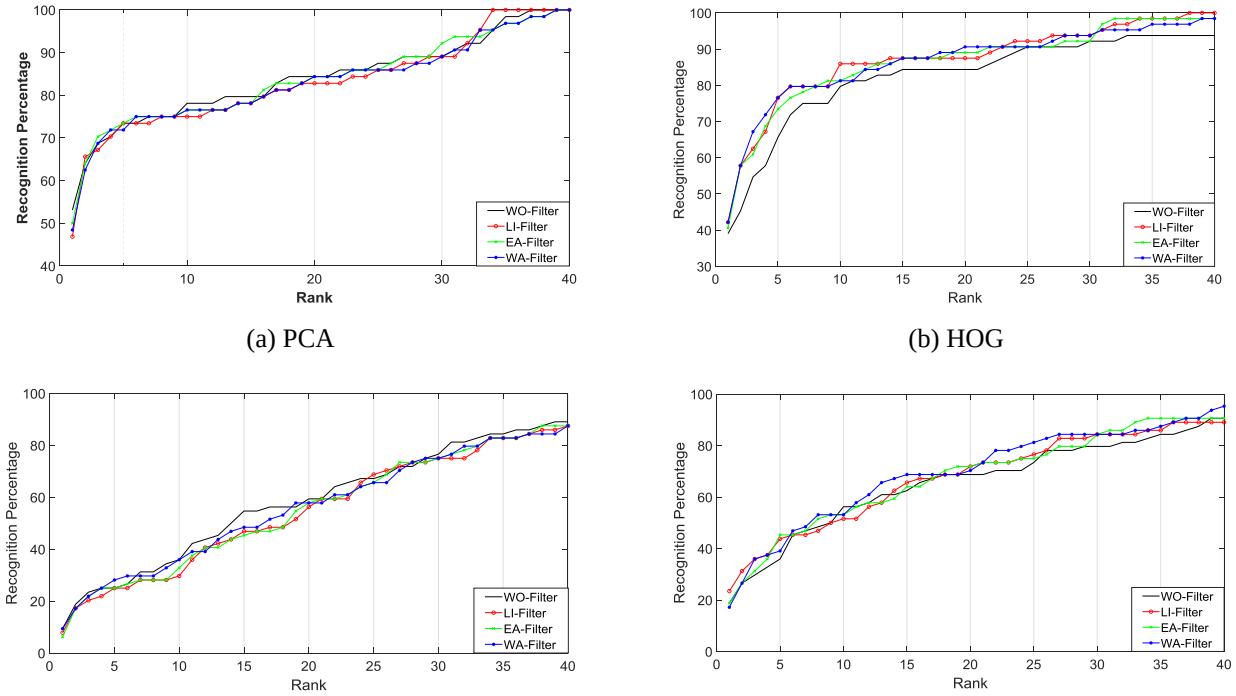
To summarize, the effect of employing filters based on kernel function indicates the improvement in the performance accuracy, demonstrating the applicability of our approach for 3D depth map images, when the number of holes present in the image significantly degrades the quality as well as the overall performance of biometric face recognition system.

Table 8. Recognition rate at Rank-5 using Fused scores (PCA + HOG)

Filters	45	90	-45	-90	Smile	Eyes Closed	Paper Occlusion
WF	26.56	17.19	15.63	12.5	93.75	92.19	35.94
LI	26.56	15.63	18.75	20.31	95.31	90.63	25
EA	25	20.31	18.75	18.75	96.88	92.19	34.38
WA	26.56	21.88	23.44	17.19	98.44	90.63	32.81

Table 9. Recognition rate at Rank-5 using Fused scores (PCA + HOG)

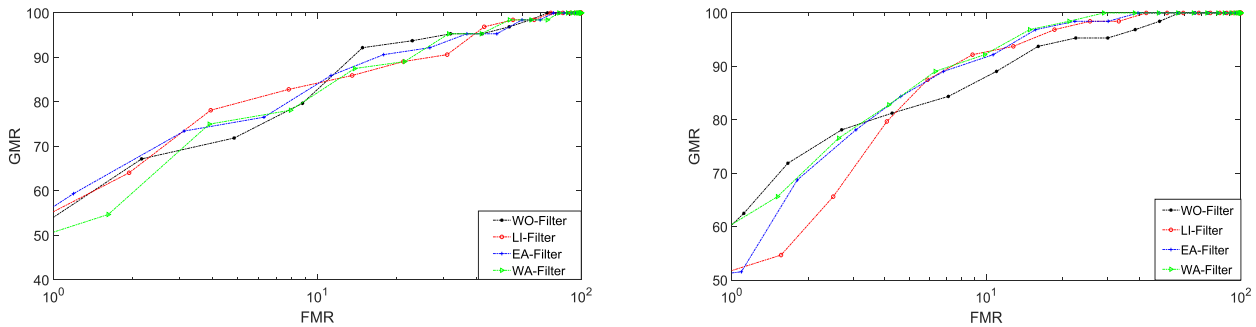
Filters	45	90	-45	-90	Smile	Eyes Closed	Paper Occlusion
WF	32.81	18.75	28.13	14.06	98.44	96.88	64.06
LI	32.81	15.63	23.13	14.06	98.44	96.88	43.75
EA	**37.50**	18.75	**32.81**	12.50	98.44	96.88	60.94
WA	**35.94**	17.19	29.69	**17.19**	96.88	**98.44**	59.38



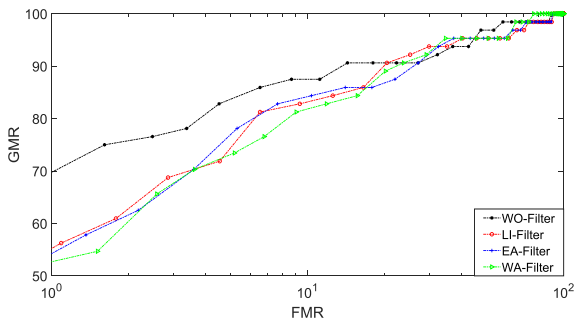
(a) PCA (b) HOG

(c) GIST (d) BSIF

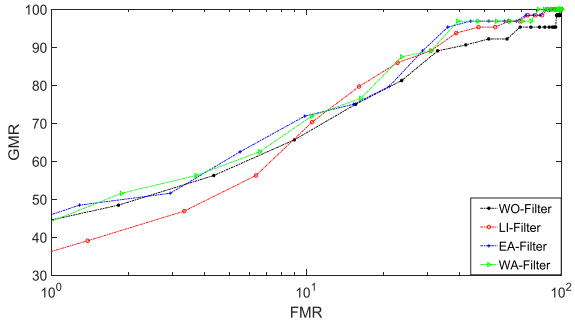
Fig. 3. Cumulative Match Curve (CMC) plots demonstrate the face recognition performance on depth map image using three different filters and without filter. The best results related to facial variant smile is presented



(a) PCA (b) HOG

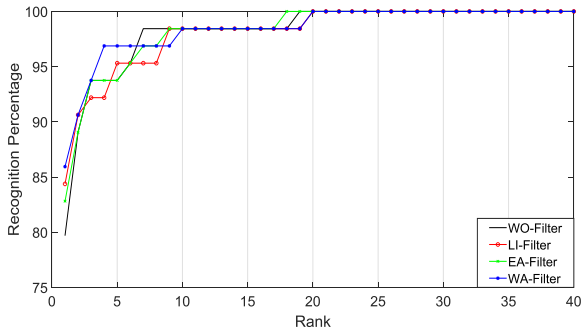


(c) GIST

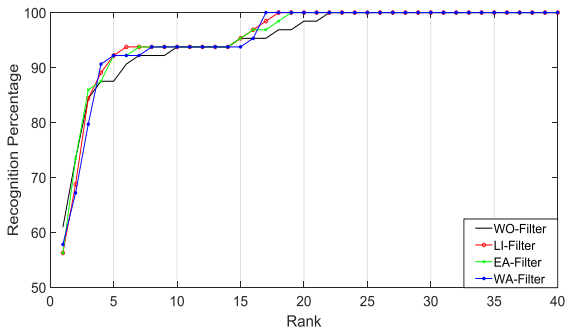


(d) BSIF

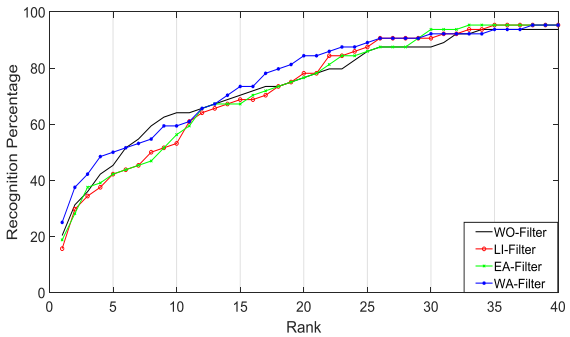
Fig. 4: Receiver Operating Curve (ROC) plots demonstrate the face recognition performance on depth map image using three different filters and without filter. The best results related to facial variant smile is presented



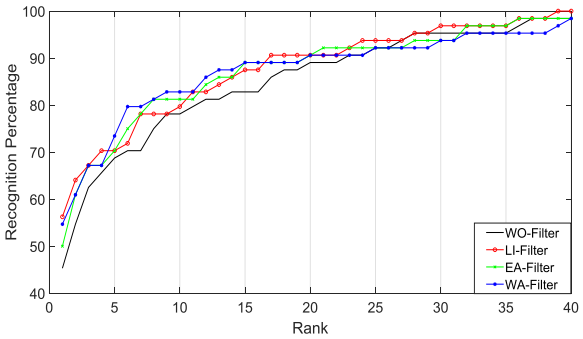
(a) PCA



(b) HOG

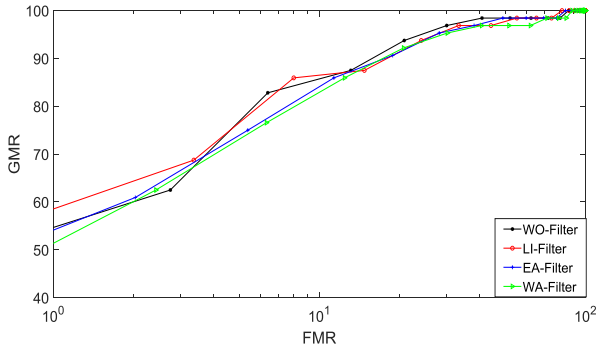


(c) GIST

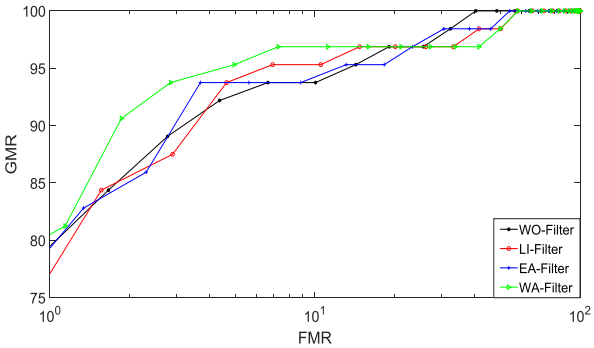


(d) BSIF

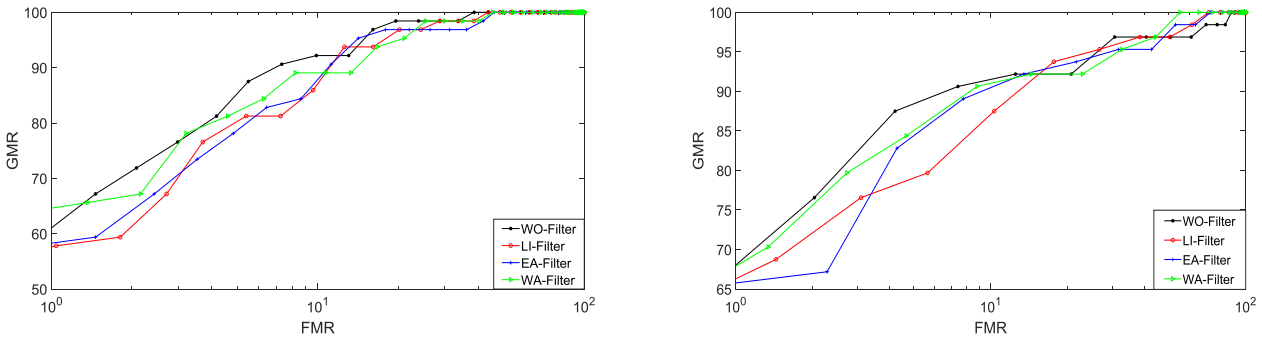
Fig. 5: Cumulative Match Curve (CMC) plots demonstrate the face recognition performance on Fused (RGB + Depth) image using three different filters and without filter. The best results related to facial variant smile is presented



(a) PCA



(b) HOG



(c) GIST (d) BSIF

Fig. 6: Receiver Operating Curve (ROC) plots demonstrate the face recognition performance on Fused (RGB + Depth) image using three different filters and without filter. The best results related to facial variant smile is presented

V. CONCLUSIONS

Authentication based on facial biometric traits has been widely used in various security applications due to its non-intrusive nature of image capture in covert manner. Although, face recognition shows great potential, the performance of the face recognition is challenged by various covariates such as pose, expression, and more recently the presentation attack. Considering this issues, 3D biometric finds alternative approach over traditional face recognition operated in 2D. While 3D depth map obtained especially using Kinect sensor introduces the holes in the depth map consisting zero pixel value, which significantly contributes the overall degradation of performance of the 3D biometric system. In this paper, we present three different filtering techniques such as linear interpolation, exponential averaging, and weighted averaging filter on depth map images. Further, we used the kernel function to demonstrate use the filtering technique efficiently. Specifically, we employed filters on variable kernel size to give appropriate weightage to nearest neighbor surrounds the hole such that contribution of required pixel can be taken to fill the missing pixel values. We experimented on our GU-RGBD database of 64 subjects collected across seven facial variants such as 45˚, -45˚, 90˚, -90˚, smile, eye closed paper occlusion. To present our results, using three different filtering approaches, we presented the results with seven different face recognition algorithms such as PCA, HOG, LBP, LPQ, GIST, BSIF and Log Gabor used in 3*D* biometrics. The extensive experimental analysis is obtained independently with depth map image, fusion of RGB and depth map, and score level fusion of two best performing algorithms. All the results related to the each of the evaluation is presented using recognition rate and verification rate and further compared the results against without filter depth map image. The results obtained after employing kernel based filtering outperforms the without filtered depth map image performance.

ACKNOWLEDGEMENTS

The Authors would like to acknowledge Goa University for all support from time to time. Also to Minister of Electronics and Information Technology (MeitY) under Visvesvaraya PhD Scheme for the financial assistance provided at the initial phase of the project.

REFERENCES

[1] Preeti.B.Sharma, Mahesh M. Goyani, 2012, “3d Face Recognition Techniques - A Review”, International Journal of Engineering Research and Applications (IJERA), Vol. 2, 1, pp.787-793

[2] G. Goswami, M. Vatsa and R. Singh, 2014, “RGB-D Face Recognition With Texture and Attribute Features”. IEEE Transactions on Information Forensics and Security, Vol. 9, no. 10, pp. 1629-1640

[3] Y. Mao, G. Cheung, A. Ortega and Y. Ji, 2013 “Expansion hole filling in depth-image-based rendering using graph-based interpolation”. IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1859-1863

[4] M. Solh and G. AlRegib. 2012, “Hierarchical Hole-Filling For Depth-Based View Synthesis in FTV and 3D Video”. IEEE Journal of Selected Topics in Signal Processing, vol. 6, no. 5, pp. 495-504

[5] D. Wang, Y. Zhao, J. Wang and Z. Wang. 2013, “A hole filling algorithm for depth image based rendering based on gradient information”. Ninth International Conference on Natural Computation (ICNC), Shenyang, pp. 1209-1213

[6] Dan Wang, Yan Zhao, Zheng Wang, Hexin Chen. 2014, “Hole-Filling for DIBR Based on Depth and Gradient Information”. International Journal of Advanced Robotic Systems

[7] Litong Feng, L. Po, X. Xu, K. Ng, Chun-Ho Cheung and K. Cheung. 2013, “An adaptive background biased depth map hole-filling method for Kinect”. IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society, Vienna, pp. 2366-2371

[8] A. Atapour-Abarghouei, G. P. de La Garanderie and T. P. Breckon. 2016, “Back to Butterworth - a Fourier

basis for 3D surface relief hole filling within RGB-D imagery". 23rd International Conference on Pattern Recognition (ICPR), Cancun, pp. 2813-2818

[9] L. Wang, H. Jin, R. Yang and M. Gong. 2008, "Stereoscopic inpainting: Joint color and depth completion from stereo images". 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, pp. 1-8

[10] T. P. Breckon and R. B. Fisher. 2008, "Three-Dimensional Surface Relief Completion Via Nonparametric Techniques". IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 12, pp. 2249-2255

[11] Dalal, N., Triggs, 2005, "B. Histograms of oriented gradients for human detection". In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. vol. 1, pp. 886–893

[12] Turk, M.A., Pentland, A.P. 1991, "Face recognition using eigenfaces". In: Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. pp. 586–591

[13] Oliva, A., Torralba, 2001, "A. Modeling the shape of the scene: A holistic representation of the spatial envelope". International Journal of Computer Vision, 42, (3), pp. 145–175

[14] Ojala, T., Pietikainen, M., Maenpaa, T. 2002, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns". IEEE Transactions on Pattern Analysis and Machine Intelligence, 24, (7), pp. 971–987

[15] Cook, J., McCool, C., Chandran, V., Sridharan, S.2006, "Combined 2d/3d face recognition using log-gabor templates'. In: 2006 IEEE International Conference on Video and Signal Based Surveillance, pp. 83–83

[16] Ojansivu V., Heikkilä J. 2008, "Blur Insensitive Texture Classification Using Local Phase Quantization". In: Elmoataz A., Lezoray O., Nouboud F., Mammass D. (eds) Image and Signal Processing. ICISP 2008. Lecture Notes in Computer Science, vol 5099

[17] Kannala, J., Rahtu, E.2012, "Bsif: Binarized statistical image features". In: Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012). pp. 1363–1

[18] R. Min, N. Kose and J. Dugelay. 2014, "KinectFaceDB: A Kinect Database for Face Recognition". IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 44, no. 11, pp. 1534-1548.