Feature and Keypoint Selection for Visible to Near-infrared Face Matching

Soumyadeep Ghosh*, Tejas I. Dhamecha*, Rohit Keshari, Richa Singh, Mayank Vatsa
IIIT-Delhi, New Delhi, India
{soumyadeepg,tejasd,rohitk,rsingh,mayank}@iiitd.ac.in

Abstract

Matching near-infrared to visible images is one of the heterogeneous face recognition challenges in which spectral variations cause changes in the appearance of face images. In this paper, we propose to utilize a keypoint selection approach in the recognition pipeline. The proposed keypoint selection approach is a fast approximation of feature selection approach, yielding two orders of magnitude improvement in computational time while maintaining the recognition performance with respect to feature selection. The keypoint selection approach also enables to visualize the keypoints that are important for recognition. The proposed matching framework yields state-of-the-art approaches results on CASIA NIR-VIS-2.0 dataset.

1. Introduction

Heterogeneous face recognition includes matching images pertaining to different imaging modalities [4, 5, 9, 16]. One of the heterogeneous face recognition problem is to efficiently match a near infrared (NIR) face image with a visible (VIS) spectrum face image. In NIR-VIS face matching scenario, the NIR face image obtained from specialized hardware [13] may be compared with legacy face datasets such as driver's license and voter's ID which contain face images in visible spectrum. Figure 1 shows some samples of face images captured in VIS and NIR spectrum. Due to the difference in the spectrum, traditional face recognition approaches may not be directly applicable for matching these cross-spectral images.

To address this heterogeneity, two types of approaches are proposed: synthesis based and direct matching based. Yi *et al*. [19] introduced the problem of visible to near infrared matching and proposed a solution which was based on learning the correlation between visible and NIR images. Lei and Li [12] proposed a subspace learning framework termed as coupled spectral regression, which learns the associated projections of VIS and NIR images to project heterogeneous data including VIS-NIR matching. Klare and Jain [9] proposed a generic heterogeneous face recognition method and introduced a prototype random subspace to address the problem and showed results on different types of heterogeneous face matching problems. Kang *et al*. [7] proposed a restoration based approach to address cross distance and cross spectral face recognition. They used locally linear embedding to restore face images captured at large standoff and then used a heterogeneous face matching algorithm [9] for recognition. Zhu *et al*. [22] proposed a transduction based framework for matching NIR and VIS face images. Jin *et al*. [5] proposed a feature learning based method to extract learned discriminative features for heterogeneous face matching. Recently, Dhamecha *et al*. [2] had studied the problem of visible to near infrared face matching to understand the effectiveness of histogram of oriented gradient features for this problem. Lu *et al*. [16] used an unsupervised feature learning algorithm with pixel difference vectors.

(a) Visible (b) Near Infrared

Figure 1: Example visible and near infrared spectrum face images from CASIA NIR-VIS-2.0 [14] dataset.

Typically, a VIS-NIR face matching pipeline involves keypoint extraction on a uniform grid (UG) [2, 8], local face patches extraction [16], or facial fiducial point (FFP) detection [18], which is followed by feature extraction, subspace and/or classifier learning. However, to the best of our knowledge not much research has been done to understand/learn and utilize keypiont selection approaches for VIS-NIR matching. Such understanding and approach could help unravel the impact of important keypoints and choose only those keypoints which are meaningful for face recognition. In order to facilitate this, we propose a keypoint selection approach using fast correlation based filter

* Equal contribution from S. Ghosh and T.I. Dhamecha.

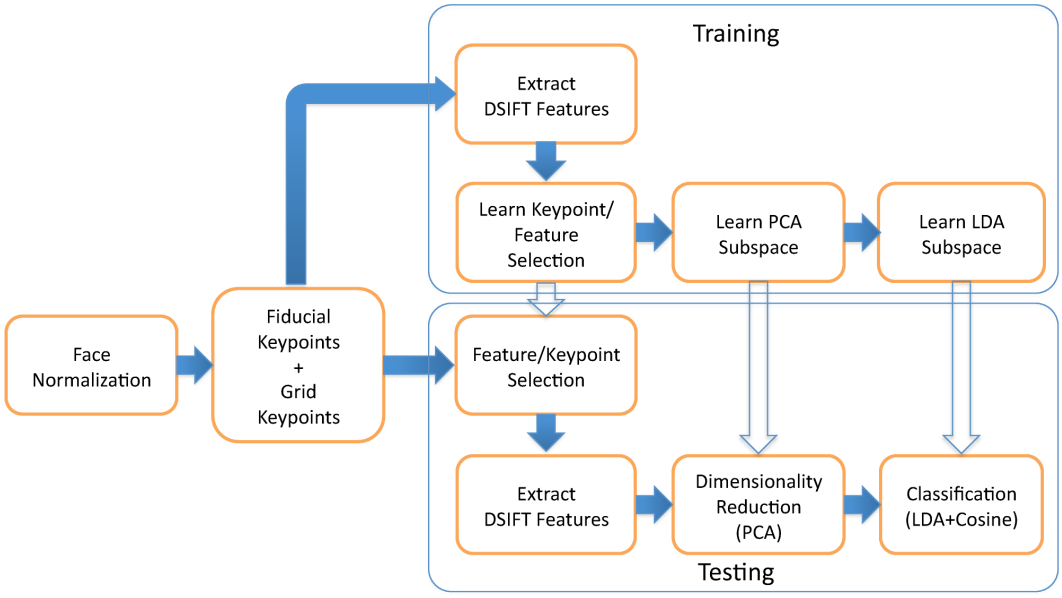


Figure 2: Block diagram of pipeline utilizing feature/keypoint selection for matching VIS to NIR face images.

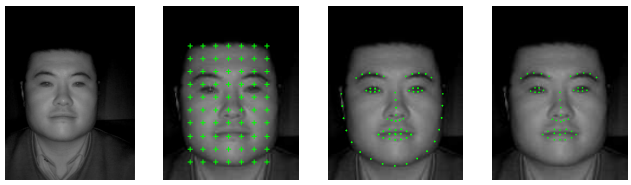
[20]. The proposed keypoint selection approach enables us to compute a fast approximation of feature selection process which, at times, is computationally expensive. Experiments are performed on the CASIA NIR-VIS-2.0 dataset [14] and it is observed that the proposed algorithm is computationally very fast and yields state-of-the-art results.

2. VIS-NIR Matching Framework

Recently, Dhamecha *et al*. [2] have studied the effectiveness of histogram of oriented gradient features on VIS-NIR face matching. Since they showcase one of the state-of-the-art results of heterogeneous face recognition on CASIA NIR-VIS-2.0, we follow a similar recognition pipeline [2]. The matching pipeline is illustrated in Figure 2. First, all the face images are geometrically normalized and all the fiducial keypoints and grid keypoints are obtained. During training stage, dense scale invariant feature transform (DSIFT) [15] features are extracted from all the keypoints. At the time of training, a model of feature or keypoint selection is learned using this feature representation. Further, PCA and LDA subspaces are learned on selected features/keypoints. During the testing phase, either only selected features or features extracted from the selected keypoints are retained. The next step is to reduce the dimensionality of retained feature set using PCA, which are further projected onto discriminative LDA space. The match scores are obtained using cosine distance measure.

2.1. Grid and FFP Keypoints

We have used two different ways of defining keypoints.



(a) (b) (c) (d)

Figure 3: (a) Original image, (b) uniform rectangular grid based keypoints shown on registered image, (c) fiducial points are extracted on the geometrically normalized face image, (d) set of fiducial points after discarding keypoints corresponding to the jawline.

1. **Uniform Grid:** We use a uniform rectangular grid (UG) to extract keypoints at specific intervals along rows and columns as shown in Figure 3(b). Uniform grid keypoints can be obtained using a grid of $p \times q$ keypoints on the face region of the registered face images.

2. **Facial Fiducial Points:** A recent approach that uses a cascaded deformable shape model [21] is employed to detect facial fiducial points. Figure 3(c) shows a sample of FFP obtained on a face image. The approach detects 66 fiducial keypoints, referred to as FFP(66), on a given face image. As shown in Figure 3(d), a variant of this FFP detection is also obtained by discarding the 17 keypoints across the jawline. This version of FFP keypoints is denoted as FFP(49). Each FFP is treated as a keypoint to extract the DSIFT features.

2.2. Feature Extraction

DSIFT features have been shown to yield good results for NIR-VIS face matching [2]. Therefore, DSIFT features [15] are extracted from the defined keypoints. The descriptors for all the keypoints are concatenated to form the feature descriptor of an image.

2.3. Feature and Keypoint Selection

For heterogeneous face matching, we represent an image with the concatenation of DSIFT features extracted from each keypoint. Thus, each element in image descriptor is a gradient pertaining to a keypoint.

2.3.1 Feature Selection

Due to the availability of large amount of high dimensional data [1], the research pertaining to feature selection has gained significant importance within the machine-learning community [17]. The presence of irrelevant or redundant features may increase the dimensionality, thereby increasing the computational time for learning and may also affect the accuracy. Feature selection [3, 17] is often utilized to reduce the dimensionality of the feature space and to retain the most discriminative features. The aim of the feature selection is to select the features which are meaningful for classification.

Feature selection algorithms may be broadly classified into two major types, namely filter and wrapper models [20]. Filter models select features by looking into the characteristics of the features and evaluating some basic statistical objective functions such as correlation and information gain. Wrapper models [11] are learning based and use a predefined model to analyze the discriminability of the features. In this research, a filter model based feature selection technique known as fast correlation based filter (FCBF) [20] is utilized which is suitable for high dimensional data. It operates on two broad objective functions: retaining the features that are highly correlated with the class and removing the features that are highly correlated with other features. The FCBF algorithm [20] uses information theoretic measure of information gain to measure the correlation between two random variables which is defined as

$$IG(A|B) = H(A) - H(A|B) \tag{1}$$

Here, $H(\cdot)$ is the entropy, defined as

$$H(A) = -\sum_i P(a_i)\log_2(P(a_i)) \tag{2}$$

Similarly, the conditional entropy is defined by

$$H(A|B) = -\sum_j P(b_j)\sum_i P(a_i|b_j)\log_2(P(a_i|b_j)) \tag{3}$$

Using Eq. 1, a normalized measure of information gain, termed as symmetrical uncertainty, is used as the measure of correlation between two random variables. The symmetrical uncertainty is defined as

$$SU(A,B) = 2\left[\frac{IG(A|B)}{H(A)+H(B)}\right] \tag{4}$$

The FCBF consists of two stages:

1. **Step 1:** Measure the correlation between features and class labels, and retain the top features with highest correlation. In this stage, A and B are the features and class labels respectively.

2. **Step 2:** Measure the correlation among features, and discard the ones which are highly correlated. In this stage, both A and B are the features.

In the first stage, a threshold θ is used as the selection criterion. The features with higher symmetrical uncertainty (SU) value than θ are retained and the rest are discarded. In the second stage, a feature is retained if its SU with all other features is less than its SU with the class labels. This procedure ensures that the selected features have high correlation with class and are less correlated with each other. For the recognition pipeline utilized in this research, the feature selection essentially chooses the *important gradient directions all over the face*. It can select few gradients belonging to a keypoint, while rejecting the other gradients.

2.3.2 Keypoint Selection

Although feature selection has its own importance, it may be computationally expensive when feature dimensionality is large. Since face recognition approaches typically deal with high dimensional features, feature selection may not be a practical solution. In this section, we propose a *keypoint selection* approach as an approximation of the feature selection. The aim of keypoint selection is to obtain keypoints that are most useful for recognition. Tracing the selected keypoints back to their spatial location is expected to provide insights into regions that are useful for recognition. Moreover, keypoint selection is computationally faster than feature selection. In this work, we extend FCBF feature selection approach to perform keypoint selection.

An image representation is obtained as concatenation of features from each keypoint. In order to perform keypoint selection, the training matrix is transformed in a way such that all the features at one keypoint are all either selected or discarded together. Let n, d and h be the number of samples, descriptor dimensionality, and the number of keypoints on an image, respectively. Training data matrix T is of size $n \times hd$ and each row of the training matrix represents one sample. The transformed training matrix T' is obtained by reshaping T into a $nd \times h$ matrix by vertically stacking

up features from the same keypoint. In the transformed matrix T' the columns corresponds to each keypoint K_i. The matrix transformation procedure is illustrated in Figure 4. By utilizing T' as the training matrix, feature selection approach can be modified into a keypoint selection approach. Operating on this training matrix, Eq. 4 is adapted for keypoint selection as follows. For step 1, symmetric uncertainty between a keypoint K_i and class labels C is obtained as

$$SU(K_i, C) = 2\left[\frac{IG(K_i|C)}{H(K_i)+H(C)}\right] \quad (5)$$

Similarly, for step 2, the symmetric uncertainty between two keypoints K_i and K_j is obtained as

$$SU(K_i, K_j) = 2\left[\frac{IG(K_i|K_j)}{H(K_i)+H(K_j)}\right] \quad (6)$$

2.3.3 Time Complexity

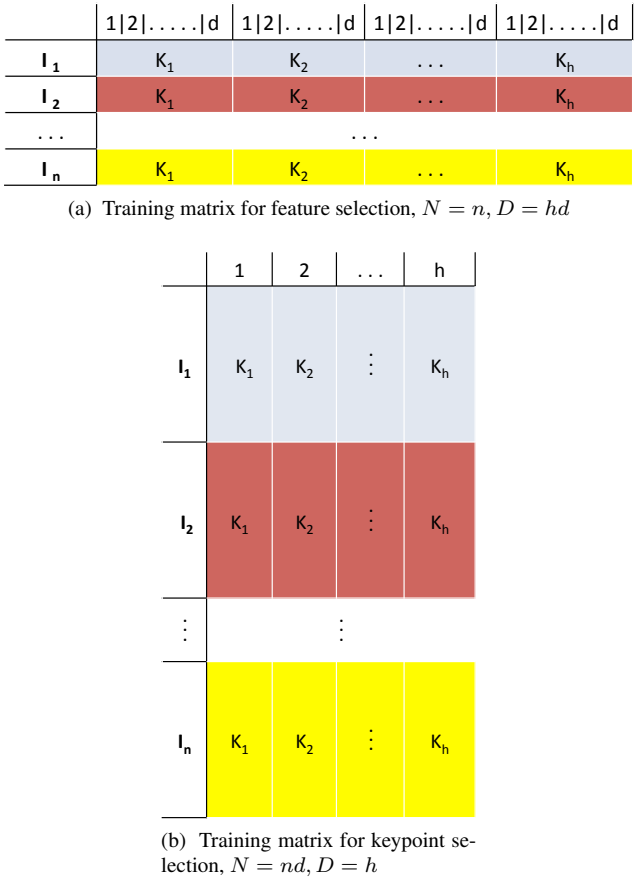
The time complexity of FCBF is $O(ND\log(D))$, where N is the number of samples and D is the feature dimensionality. If the train set contains n face images, each represented with d dimensional h keypoints, the time complexity of feature selection is $O(nhd\log(hd))$, as $N = n$ and $D = hd$. However, if the selection is performed at keypoint level, the time complexity of keypoint selection is $O(nhd\log(h))$, as $N = nd$ and $D = h$. Therefore, keypoint selection may be viewed as faster approximation of feature selection. Not only does this transform a feature selection problem into a keypoint selection problem, it also reduces the dimensionality of the feature space from hd to h. This enables feature selection algorithm to run on this transformed feature matrix with reduced time.

2.4. Dimensionality Reduction and Classification

The dimensionality of DSIFT features extracted at the selected keypoints is large and therefore, it is highly desirable to perform dimensionality reduction. Dimensionality is reduced with the help of Principal Component Analysis (PCA) [6]. We choose 98% of the eigenenergy preserving top principal components. Over the reduced dimensional representation, Linear Discriminant Analysis (LDA) [10] subspace is learned and the cosine distance is used as a distance metric for matching a pair of images. The distance between a gallery sample $\mathbf{X}$ and a probe sample $\mathbf{Y}$ is obtained as

$$distance(\mathbf{X}, \mathbf{Y}) = Cos(\mathbf{x}, \mathbf{y}) = 1 - \frac{\mathbf{x}'\mathbf{y}}{|\mathbf{x}||\mathbf{y}|}, \text{ where}$$
$$\mathbf{x} = \mathbf{X}'\mathbf{W_{pca}}\mathbf{W_{lda}}, \text{ and } \mathbf{y} = \mathbf{Y}'\mathbf{W_{pca}}\mathbf{W_{lda}}$$

where, $\mathbf{W_{pca}}$ and $\mathbf{W_{lda}}$ are the projection matrices of PCA and LDA, respectively.



(a) Training matrix for feature selection, $N = n, D = hd$

(b) Training matrix for keypoint selection, $N = nd, D = h$

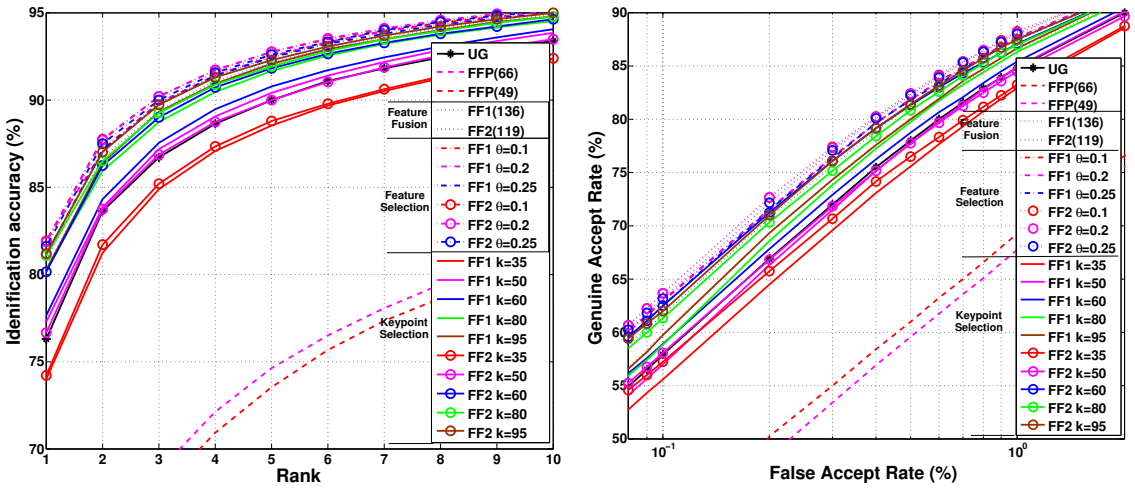
Figure 4: The illustration of training matrix transformation for keypoint selection. K_j and I_j represents the j^{th} keypoint and image respectively. The training matrix of (a) is transformed in (b) such that it represents nd samples of h dimensionality.

3. Experiments and Analysis

The experiments are performed on the CASIA NIR-VIS 2.0 [14] dataset using the predefined protocols. This section describes the dataset, protocols, and presents the results along with analysis.

3.1. Dataset and Protocol

The CASIA NIR-VIS 2.0 dataset [14] contains predefined visible and near-infrared images of 725 subjects. The dataset contains protocols pertaining to two views. View 1 is the development set which can be used for optimizing the parameters of the algorithm. Here, it is utilized to perform feature and keypoint selection. The View 2 contains training and testing splits for 10 times cross validation. Subjects in the training and testing sets are non-overlapping. The gallery set contains images acquired in VIS spectrum and the probes are NIR images. In each experiment the selected features are fixed from View 1 and results are reported on



(a) (b)

Figure 5: (a) CMC and (b) ROC curve demonstrating the effect of variations in keypoint definition style, number of fiducial points, feature fusion, feature selection, and keypoint selection.

10 times cross validation of View 2.

3.2. Experiments

The input face images are geometrically normalized using the eye coordinates provided as part of the dataset and size of the detected face image is 125×160. This process fixed the interocular distance and the eyes at a particular location across all images. For obtaining keypoints on the rectangular grid, we select keypoints using a grid of 7×10 (10 keypoints along the column and 7 along the row) on the face region. 66 FFPs are obtained by employing the fiducial point detection technique proposed by Yu *et al*. [21]. The DSIFT feature obtained at each keypoint is of length $16n_r$ where $n_r = 8$ is the number of orientations. The features are obtained with spatial bin size of 8. Thus, for each keypoint a descriptor of size 128 is generated.

Using the predefined protocol of CASIA NIR-VIS-2.0, evaluation is performed in identification and verification mode. For identification scenario, the mean and standard deviation of rank 1 identification accuracy and cumulative match score curves (CMC) are reported in Table 1 and Figure 5(a) respectively. For verification scenario, genuine accept rate at 0.1% and 1% false accept rate (FAR) and receiver operating characteristics curves (ROC) are reported in Table 1 and Figure 5(b) respectively. In order to study the effectiveness of feature and keypoint selection, following set of experiments are performed.

- **Exp. 1: Uniform Grid vs Facial Fiducial Points:** This experiment compares the recognition performance of features extracted from keypoints defined on uniform grid and facial fiducial points (FFP) [21]. Two variants of FFP, with and without keypoints on jawline, are explored; they are denoted as FFP(66) and FFP(49), respectively.

- **Exp. 2: Feature Fusion of UG and Facial Fiducial Points:** Since both the kind of keypoints, UG and FFP, do not exactly overlap in terms of spacial regions encoded by them, we explore the effectiveness of concatenating the features obtained from these two kinds of keypoints. The feature concatenation is denoted as FF1 (or UG+FFP(66)) and FF2 (or UG+FFP(49)).

- **Exp. 3: Feature Selection [20]:** This set of experiments are performed to study the effectiveness of feature selection on UG+FPP(66) and UG+FPP(49). The results are reported with varying the threshold of feature selection technique.

- **Exp. 4: Keypoint Selection:** In this experiment, the performance of the proposed keypoint selection approach is evaluated. Input to the keypoint selection approach is the concatenation of features obtained from UG and FFP. One of the motivations of this experiment is to be able to analyze selected keypoints and associated facial regions useful for heterogeneous face recognition.

During keypoint and feature selection training, DSIFT features and associated class labels are used as input. For experiments pertaining to feature selection [20], results are reported with three feature selection thresholds, $\theta = 0.1, 0.2,$ and 0.25. As the threshold increases, less number of features are selected. For keypoint selection, θ is varied such that 35, 50, 60, 80, and 95 keypoints are selected.

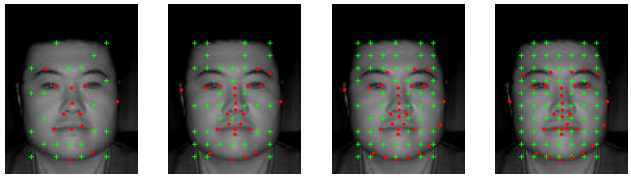
Table 1: Identification and verification results on View 2 of CASIA NIR-VIS-2.0 dataset. θ and k represent the threshold of feature selection and number of keypoints selected, respectively.

Approach		Rank-1 Accuracy	GAR @ f FAR	
			$f = 0.1\%$	$f = 1\%$
UG (70)		76.29±1.80	57.90	84.80
FFP(66) w/ jawline		54.82±1.76	41.52	69.19
FFP(49) w/o jawline		53.21±1.99	40.06	67.68
Feature fusion				
FF1(136): UG(70)+FFP(66)		80.35±1.54	59.83	86.53
FF2(119): UG(70)+FFP(49)		**80.87±1.51**	**61.61**	**87.43**
Feature selection				
FF1	$\theta = 0.1$	81.06±1.42	62.87	87.62
	$\theta = 0.2$	**81.90±1.58**	**62.97**	**87.65**
	$\theta = 0.25$	81.11±1.30	62.44	87.47
FF2	$\theta = 0.1$	81.90±1.05	63.62	88.17
	$\theta = 0.2$	81.87±1.89	63.65	88.09
	$\theta = 0.25$	81.59±1.52	63.16	88.00
Keypoint selection				
FF1	$k = 35$	74.03±1.59	55.88	82.90
	$k = 50$	77.26±1.61	56.97	84.91
	$k = 60$	77.62±1.59	58.86	85.37
	$k = 80$	80.27±1.85	58.81	86.33
	$k = 95$	80.24±1.65	59.69	86.65
FF2	$k = 35$	74.17±1.68	57.20	83.20
	$k = 50$	76.61±1.70	58.01	84.44
	$k = 60$	80.11±1.43	62.43	87.03
	$k = 80$	81.09±1.21	61.32	86.98
	$k = 95$	**81.19±1.60**	**61.98**	**87.47**
State-of-the-art				
FaceVACS [2] (2014)		58.56±1.19	-	-
Dhamecha *et al*. [2] (2014)		73.28±1.10	-	-
Lu *et al*. [16] (2015)		81.80±2.30	47.30	75.30

3.3. Observations

The results obtained from the proposed algorithm are compared with the results reported in the literature on the CASIA NIR-VIS-2.0 dataset. The analysis of the results are as follows,

- Dhamecha *et al*. [2] have used a uniform grid of size 8×8. By utilizing a different uniform grid of 7×10, the results are enhanced by 3%.
- Utilizing the features extracted only from the FFP does not yield comparable performance. This shows that although, FFP might be good at describing face shape, it may not be sufficient for improving recognition performance.
- It is observed that the FFP detection approach [21], which contains a detection module trained for visible spectrum images, efficiently works on NIR face images without any fine-tuning.



(a) 35 (b) 60 (c) 80 (d) 95

Figure 6: Visualization of keypoints selected from combined pool of uniform grid and FFP keypoints (with jawline). The selected FFP and grid keypoints are represented using red dots and green plus signs, respectively.

- **Feature fusion:** Concatenating DSIFT features extracted from UG and FFP yields better recognition performance than each of them individually. This indicates that different keypoint approaches should encode some non-overlapping information, which eventually improves the performance, in a fusion scheme.
- **Feature selection:** Not only does feature selection improve the performance over feature fusion, it is also comparable to existing state-of-the-art results [16] in terms of rank-1 identification accuracy, while outperforming it in verification scenario by 15% GAR at 0.1% FAR.
- **Keypoint selection:** The best results obtained with the proposed keypoint selection technique are comparable to that of the feature selection. The visualization of selected keypoints is shown in Figure 6.
  - Interestingly, an intuitive set of keypoints are learned and selected. For example, if we are to select a small number of keypoints ($k = 35$), the FFP points retained correspond to two eyes, two nostrils, two nose bridge points and three lips points, whereas the retained grid keypoints broadly encode the chin area. As we increase the number of selected keypoints, finer details of fiducial points such as eye corners, overall shape of eyebrows, and finer points on nose bridge can be observed in set of retained FFP points; while the retained grid keypoints become more dense (See Figure 6). Note that these keypoints are automatically selected using the proposed keypoint selection approach, which aligns with intuitive notion of important facial regions. This analysis pertaining to which keypoints are important for recognition might help build improved facial fiducial point detectors.
- **Jawline:** As shown in Table 1, rows 2 and 3, the presence of jaw-line fiducial points shows a marginal impact on recognition performance. We also observe that

when jawline keypoints are discarded, the performance is improved in both feature and keypoint selection approaches. However, with $k = 95$, 69.8% (95 out of 136) and 79.8% (95 out of 119) of keypoints are retained in FF1 and FF2, respectively. Therefore, a direct comparison of their corresponding performances may be unfair.

- **Time:** It is empirically observed that feature selection (FF2 $\theta = 0.2$) requires about 8.65 hours, whereas keypoint selection (FF2 $k = 60$) requires about 3 minutes to run on a desktop with 3.4 GHz Intel Core i7 processor and 16GB RAM. The keypoint selection achieves 173x speedup over feature selection by trading-off 1.7% in rank-1 accuracy.

4. Conclusion

This research proposes to utilize keypoint selection approach in VIS-NIR face recognition pipeline. The proposed keypoint selection approach is a fast approximation of the FCBF feature selection algorithm. When provided with a pool of keypoints, it efficiently selects a set of keypoints which are highly correlated with the class labels, and least correlated amongst each other. The evaluation on CASIA NIR-VIS-2.0 shows that utilizing the proposed keypoint selection approach yields state-of-the-art results both in terms of verification and identification accuracy. The proposed keypoint selection approach requires two orders of magnitude lesser time than that of feature selection without significantly affecting the recognition accuracy. Moreover, the visualization of keypoint selection technique helps to understand the important facial parts in context of heterogeneous face recognition.

Acknowledgement

This research is in part supported by a grant from the Department of Electronics and Information Technology, India. T. I. Dhamecha is partly supported through TCS PhD research fellowship. R. Keshari is partly supported through Visvesvaraya PhD fellowship.

References

[1] D. Chen, X. Cao, F. Wen, and J. Sun. Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification. In *IEEE CVPR*, pages 3025–3032, 2013. 3

[2] T. I. Dhamecha, P. Sharma, R. Singh, and M. Vatsa. On effectiveness of histogram of oriented gradient features for visible to near infrared face matching. In *IAPR ICPR*, pages 1788–1793, 2014. 1, 2, 3, 5, 6

[3] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *JMLR*, 3:1157–1182, 2003. 3

[4] C.-A. Hou, M.-C. Yang, and Y.-C. F. Wang. Domain adaptive self-taught learning for heterogeneous face recognition. In *IAPR ICPR*, pages 3068–3073, 2014. 1

[5] Y. Jin, J. Lu, and Q. Ruan. Coupled discriminative feature learning for heterogeneous face recognition. *IEEE TIFS*, 10(3):640–652, 2014. 1

[6] I. Jolliffe. *Principal component analysis*. Wiley Online Library, 2002. 4

[7] D. Kang, H. Han, A. K. Jain, and S.-W. Lee. Nighttime face recognition at large standoff: Cross-distance and cross-spectral matching. *PR*, 47(12):3750–3766, 2014. 1

[8] B. Klare and A. K. Jain. Heterogeneous face recognition: Matching NIR to visible light images. In *IAPR ICPR*, pages 1513–1516, 2010. 1

[9] B. F. Klare and A. K. Jain. Heterogeneous face recognition using kernel prototype similarities. *IEEE TPAMI*, 35(6):1410–1422, 2013. 1

[10] W. R. Klecka. *Discriminant analysis*, volume 19. Sage, 1980. 4

[11] R. Kohavi and G. H. John. Wrappers for feature subset selection. *AI*, 97(1):273–324, 1997. 3

[12] Z. Lei and S. Z. Li. Coupled spectral regression for matching heterogeneous faces. In *IEEE CVPR*, pages 1123–1128, 2009. 1

[13] S. Z. Li, R. Chu, S. Liao, and L. Zhang. Illumination invariant face recognition using near-infrared images. *IEEE TPAMI*, 29(4):627–639, 2007. 1

[14] S. Z. Li, D. Yi, Z. Lei, and S. Liao. The CASIA NIR-VIS 2.0 face database. In *IEEE CVPR Workshops*, pages 348–353, 2013. 1, 2, 4

[15] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004. 2, 3

[16] J. Lu, V. Liong, X. Zhou, and J. Zhou. Learning compact binary face descriptor for face recognition. *IEEE TPAMI*, PP(99):1–1, 2015. 1, 5, 6

[17] L. C. Molina, L. Belanche, and À. Nebot. Feature selection algorithms: A survey and experimental evaluation. In *IEEE ICDM*, pages 306–313, 2002. 3

[18] D. Yi, Z. Lei, S. Liao, and S. Z. Li. Shared representation learning for heterogeneous face recognition. In *IEEE FG*, 2015. 1

[19] D. Yi, R. Liu, R. Chu, Z. Lei, and S. Li. Face matching between near infrared and visible light images. In *Advances in Biometrics*, volume 4642 of *LNCS*, pages 523–530. 2007. 1

[20] L. Yu and H. Liu. Feature selection for high-dimensional data: A fast correlation-based filter solution. In *ICML*, volume 3, pages 856–863, 2003. 2, 3, 5

[21] X. Yu, J. Huang, S. Zhang, W. Yan, and D. N. Metaxas. Pose-free facial landmark fitting via optimized part mixtures and cascaded deformable shape model. In *IEEE ICCV*, pages 1944–1951, 2013. 2, 5, 6

[22] J.-Y. Zhu, W.-S. Zheng, J.-H. Lai, and S. Z. Li. Matching NIR face to VIS face using transduction. *IEEE TIFS*, 9(3):501–514, 2014. 1