# Nonsubsampled Contourlet Transform Based Descriptors for Gender Recognition

Muhammad Hussain[1], Sarah Al-Otaibi[2], Hatim Aboalsamh[2], George Bebis[4], Ghulam Muhammad[3]

[1]Department of Software Engineering, [2]Department of Computer Science, [3]Department of Computer Engineering

College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia

[4]Department of Computer Science and Engineering, University of Nevada at Reno, USA

{mhussain, ghulam, hatim}@ksu.edu.sa

## Abstract

*Gender recognition using facial images plays an important role in biometric technology. A key component of a gender recognition system is feature extraction. Motivated by the success of multiresolution techniques in various applications, we investigated four different feature extraction techniques based on Nonsubsampled Contourlet Transform (NSCT) to identify the best performing technique. We present a gender recognition system that uses SVM, two-stage feature selection and different feature descriptors based on NSCT. Among different NSCT based feature descriptors, the one based on NSCT and Weber Law Descriptor (WLD) gives the best accuracy (99.5±1.05) and it outperforms the state-of-the-art gender recognition systems on FERET database. This research reveals the best feature description technique using NSCT for gender recognition problem.*

## 1. Introduction

In category specific face recognition, facial images are first categorized based on visual cues like gender and race etc. and then face recognition is performed using category specific descriptors. This approach can improve the performance but the bottleneck is categorization [15]. In this paper, we address the problem of face cate-gorization based on gender i.e. gender recognition problem. This problem is im-portant from other aspects as well; it can enhance the performance of a wide range of applications such as search engine retrieval, demographic data collection, and surveillance etc.

Different gender recognition systems have been proposed in the literature that use different approaches for feature description and classification. One of the earlier gender recognition systems was proposed by Gutta et. al. [3]. Moghaddam and Yang [2] used SVM gender classification and obtained a 96.62% accuracy on FERET da-tabase. Baluja and Rowley [5] used an Ada-Boost based method for gender recogni-tion; they reported the best accuracy of 94.4% on FERET database. Lu and Shi [6] fused the features from left eye, upper face region and nose in their gender recogni-tion approach; their results showed that the fusion of features from different regions outperformed. Inspired by this idea, Alexandre [7] used a fusion approach based on three different types of features

from multiple scales. He achieved an accuracy of 99.07% on FERET database. Zheng and Lu [1] gave a comparison of 6 types of features using three classifiers and showed that for FERET database the best accuracy (99.1%) was obtained with features based on local Gabor binary pattern, LAD and SVM with automatic confidence. Ihsan et al. [4] used NN classifier and multiscale texture features based on dyadic wavelet trans-form and LBP and reported 99.25% accuracy on FERET.

Motivated by the success of multiscale texture features and SVM, we employed NSCT for feature extraction and SVM for classification in gender recognition sys-tem. The features can be extracted in different ways using NSCT. The question is which way of extracting features using NSCT will result in the best recognition per-formance of a gender recognition system. We investigated four different approaches for feature extraction based on NSCT: (1) NSCT, only coefficients of subbands, (2) NSCT-PCA, (3) NSCT-WLD, (4) NSCT-FOS, first order statistics (FOS) from sub-bands and found that NSCT-WLD gives the best performs and the gender recogni-tion system has improvement over the stat-of-the-art methods. For reducing the dimension of the feature space, we used two-stage feature subset selection tech-nique.

The rest of the paper is organized as follows. Section 2 gives the detail of the gender recognition system. In Section 3, the detail of experimental setup is given. Section 4 presents experimental results and their discussion. Paper is concluded in Section 5.

## 2. Methodology

In this section we describe the detail of the Gender recognition system, which is a standard pattern recognition system and is shown in Figure 1.

### 2.1. Preprocessing

Color facial images are converted into YCbCr color system, where Y is luminance component; human eyes are sensitive to this component, as such it plays key role in the recognition process. We used Y component for gender recognition.
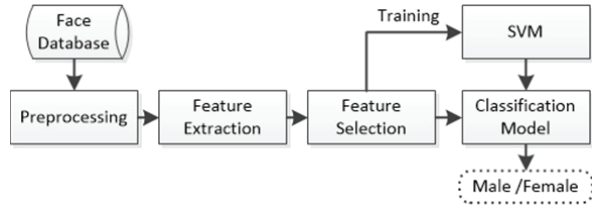
**Figure 1. System diagram for gender recognition**

## 2.2. Feature Extraction

NSCT is a multiresolution tool that decomposes an image into a hierarchy of low frequency and high frequency directional subbands, which can be used to extract textural features. Different ways are possible to extract features using NSCT. In this section, we discuss four feature extraction techniques that are based on NSCT and employ WLD, PCA, and FOS. First we give an overview of NSCT and WLD.

### 2.2.1 Nonsubsampled Contourlet Transform (NSCT)

Zhou at el. [8] proposed NSCT that decomposes an image into subbands of various scales and orientations. It involves two key operations: Nonsubsampled Pyramid (NSP) decomposition and directional decomposition using Nonsubsampled Directional Filter Bank (NDFB). First, an image is decomposed into lowpass (low) and highpass (high) subbands using NSP decomposition. Second, applying NDFB, the high subband is decomposed into directional sub-bands. This process is repeated on low subband to create a hierarchy of directional subbands of various scales.

Figure 2 shows the multiscale hierarchy of an image created with NSCT; the image has been decomposed into 13 (one low and 12 high) subbands; there are three scales and at each scale there are 4 directional subbands.
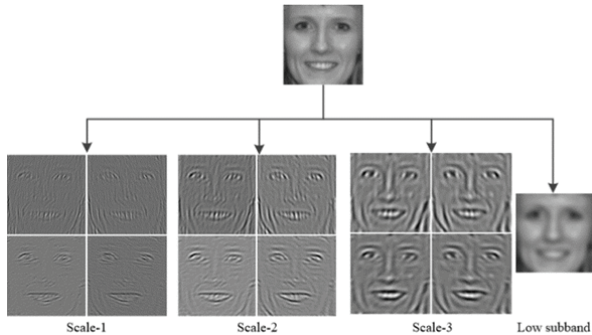


**Figure 2. Decomposition of an image using NSCT into three scales. At each scale, the image has been decomposed into 4 directional high subbands.**

### 2.2.2 WLD Descriptor

This descriptor is based on Weber's law, which states that the ratio of the increment threshold to the background intensity is constant [9]. Its construction involves the computation of Differential Excitation (DE), Gradient Orientation (GO) and WLD histogram. WLD histogram is

used as a describer. For each pixel $x_c$, DE is calculated using the following equation:

$$\zeta(x_c) = \arctan\left[\sum_{i=1}^{p}\left(\frac{x_i - x_c}{x_c}\right)\right]$$

where $x_i$, i = 1, 2, …, p, are pixels in the neighborhood of the pixel $x_c$, as shown in Figure 3(a) for p = 8. The DE $\zeta(x_c)$ can take positive as well negative values. The arctangent function is used to prevent quick increase or decrease in the output when the input becomes very small or very large.

The next important step is the computation of GO. For each pixel $x_c$, the GO $\theta(x_c)$ is calculated as follows:

$$\theta = \arctan\left(\frac{x_8 - x_4}{x_6 - x_2}\right)$$

and is mapped to the range [0, $2\pi$]. Then the $\theta$ values corresponding to all pixels are quantized to T dominant GOs $\Phi_t$, t = 1, 2, …, T; a $\theta \in \left[\Phi_t - \frac{\pi}{T}, \Phi_t + \frac{\pi}{T}\right]$ is quantized to $\Phi_t$.

The last step is building WLD histogram. Using both DE and GO for each pixel, WLD histogram is calculated, see Figure 4; the detail can be found in [9]. This descriptor involves three free parameters:

- T, the number of dominant GOs $\Phi_t$: t = 0, 1, 2, …, T-1,
- M, the number of segments $H_{m,t}$ of each sub-histogram $H_t$ corresponding to a dominant GO $\Phi_t$, and
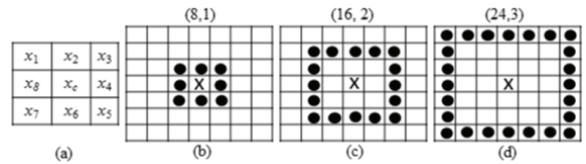- S, the number of bins in each sub-histogram $H_{m,t}$.



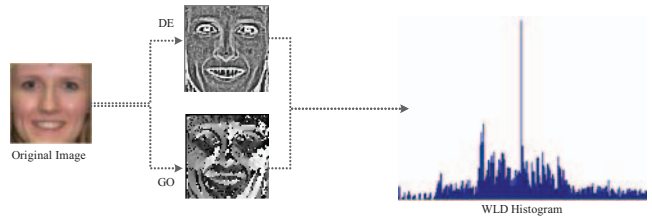**Figure 3. Neighborhoods of a pixel for computing DE.**



**Figure 4. The computation of WLD descriptor. First DE and GO are computed, then using both, WLD histogram is calculated.**

The fixed size neighbourhood 3x3 is unable to discriminate local disciminatory patterns in various granularities. To represent local salient patterns at various scales, WLD is calculated using neighbourhoods (P, R) of various sizes and shapes; we used three square neighbourhoods (P, R) with P = 8, 16, 24 and R = 1, 2, 3, as shown in Figure 3(b-d). Multi-scale WLD descriptor is built by concatenating the WLD histo-grams calculated with three neighbourhood sizes.

### 2.2.3 NSCT Based Feature Extraction Methods

Now we describe the methods for feature extraction using NSCT; we call them NSCT, NSCT-PCA, NSCT-FOS and NSCT-WLD Descriptors. In each case, a facial image is decomposed using NSCT into one low (L) and n high subbands: Hij, where i = 1, 2, …, s, j = 1, 2, …, o with s being the number of scales and o being the number of orientations at each scale, and n = s×o; Hij represents a high subband with scale i and orientation j.

**NSCT Descriptor**

In this case, the coefficients of a subband are treated as features and the subbands are converted into vectors, which are used to define feature vector describing a subject.

**NSCT-PCA Descriptor**

In this approach, a subband is projected onto an eigen space of low dimension using the eigenface technique proposed by M. Turk and A. Pentland [14]. Subbands are converted into vectors of dimension $m$ and the PCA transformation matrix $M_{ij}$ (of order $m \times n$) is computed for each of low and high frequency subbands. Using $M_{ij}$, the m-dimensional vectors corresponding to subband $H_{ij}$ are projected to n-dimensional vectors, where $n \ll m$; these vectors are used to describe the subject.

**NSCT-FOS Descriptor**

Each subband is divided into overlapping windows by first partitioning it into non-overlapping blocks, as shown in Figure 5. From each window, we compute first order statistics: mean, std, skewness, kurtosis, entropy and concatenate them to form a vector that represents the subband. We used overlapping windows to compensate for the boundary artifacts between blocks.
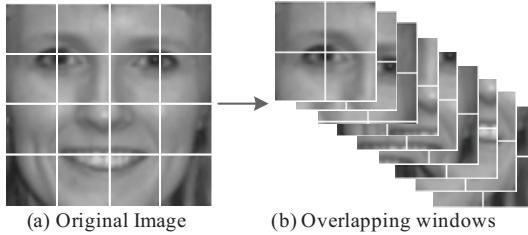


(a) Original Image    (b) Overlapping windows

**Figure 5. Partition of a facial image into 4x4 blocks and 9 overlapping windows**.

**NSCT-WLD Descriptor**

Multi-scale WLD descriptor is computed from each subband. WLD described above is not spatially enhanced; to improve its discriminatory power, we incorporate spatial locality of microstructures by dividing each subband into $m$ blocks, see Figure 5a. The detail is given below.

Step-1.  Partition each subband $H$ into $m$ non-overlapping blocks $B_1, B_2, …, B_m$.

Step-2.  For each block $B_i$, compute WLD histograms $h_{i1}$, $h_{i2}$, $h_{i3}$ using neighborhoods (8, 1), (16, 2), (24, 3), respectively, and concatenate them to build multi-scale WLD histogram $h_i = [h_{i1}, h_{i2}, h_{i3}]$.

Step-3.  Concatenate multi-scale WLD histograms $h_i$, $i = 1$, 2, …, $m$ corresponding to each block $B_i$ and build the descriptor h for H i.e. h =[ $h_1$, $h_2$, …, $h_m$].

We investigated the effect of NSCT-WLD descriptors from each subband and also using decision level fusion.

### 2.3. Feature Subset Selection

To tackle the problem of the curse of dimensionality, we applied two-stage feature subset selection approach. In the first stage, we used Fisher method [10], which ranks the features using the class information; features are selected based on their ranks. The selected features are not independent because Fisher method does not take into account the dependence of features. To remove the redundancy due to fea-ture dependence, we used Local Learning Based (LLB) method [11].

### 2.4. Classification

For classification, we used Support Vector Machines (SVM) because it is based on large margin theory and has better generalization. As RBF kernel gives good performance for many applications, we used RBF in our gender recognition system. For implementation, we used LIBSVM.

## 3.  Experimental Setup

In this section, we give the detail of the set-up used to perform the experiments.

### 3.1. Database

We used FERET [12] database to measure the performance and efficiency of the proposed system. FERET is the most popular and difficult database which was developed for research related to face recognition. The number of each subject's facial images is different, these images differ from each other based on expression, illumination, and pose. We used *fa* and *fb* sets of the database, the set *fb* contains facial images of the same subjects but with variations of expression, illumination and pose; both sets contain 1797 colored facial images of 403 male and 403 female subjects. Some sample images are shown in Figure 6.



**Figure 6. Sample female (top row) and male (bottom row) images**.

### 3.2. Parameter Tuning

NSCT, NSCT_PCA, NSCT_FOS and NSCT-WLD descriptors involve many parameters that have significant effect on performance. NSCT decomposition has two parameters: $s$ - the number of scales and $o$ - the number of

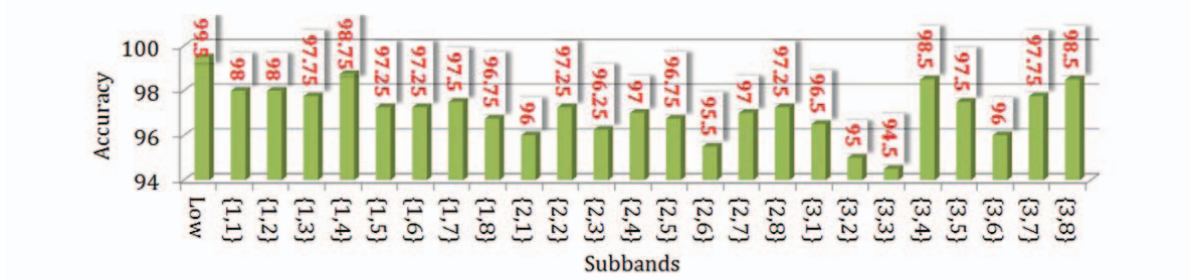orientations at each scale; for experiments, we used $s = 3$

**Figure 7. Recognition accuracies obtained with NSCT-WLD descriptor. Here low means low subband, and {i, j} stands for high subband $H_{ii}$.**
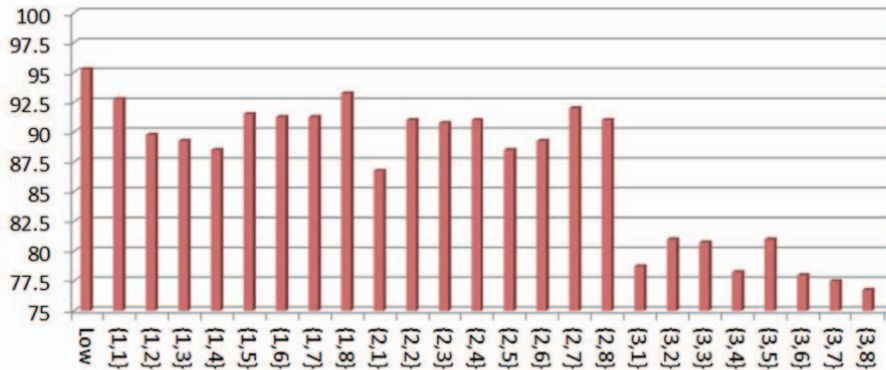


**Figure 8. Recognition accuracies obtained with NSCT-FOS descriptor. Here low means low subband, and {i, j} stands for high subband $H_{ij}$.**

**Table 1. Results with decision level fusion**

| # | Subands | Acc±std | AUC±std |
|---|---------|---------|---------|
| 3 | Low, $H_{14}$, $H_{34}$ | **99.50±1.58** | **0.993±0.02** |
| 5 | Low, $H_{11}$, $H_{14}$, $H_{34}$, $H_{38}$ | 99.00±2.10 | 0.980±0.02 |
| 7 | Low, $H_{11}$, $H_{12}$, $H_{13}$, $H_{14}$, $H_{34}$, $H_{38}$ | 98.75±2.12 | 0.987±0.02 |

and $o = 8$ in each case to be consistent.

NSCT-FOS descriptor involves one additional parameter: the number of overlapping windows; we tested 49, 225, 961 and found that 961 windows gave the best result.

For NSCT-WLD, WLD involves two parameters types: (T, M, S) and the number of non-overlapping rectangular blocks ($b$) into which an image is divided. For (T, M, S), we tested (4, 4, 5), (4, 6, 5), (12, 6, 15) and (12, 4, 5) and for $b$ we examined 4x4, 5x5, 6x6, 7x7, and 8x8 blocks. We found that (12, 4, 5) and 6x6 blocks gave the best results.

SVM with RBF kernel contain two parameters: $\gamma$ and C; for finding the optimal parameter values, we used 5-fold cross validation and grid search. First a loose grid search was conducted to find the potential optimal region in ($\gamma$, C) space. Then fine grid search was performed in this region to find the optimal parameter values.

In this section, first we preset the results of experiments performed using NSCT-WLD, NSCT-FOS, NSCT-PCA and NSCT descriptors. Then, a comparison with stat-of-the-art methods is given. For validation, we used 10-fold cross-validation and widely used performance measures: accuracy, ROC curve and AUC (area under ROC curve).

## 4.1. Results with NSCT-WLD descriptor

We decomposed Y component of each facial image into one low and 24 high subbands Hij, i = 1, 2, 3, j = 1, 2, ..., 8, and computed NSCT-WLD descriptor using T=12, M=4, S=5, and 6x6 blocks. To remove redundant features, we applied two-stage feature selection. Figure 7 shows the recognition results for individual subbands. The best accuracy of 99.5% is obtained with low subband.

In addition, we performed decision level fusion using majority vote. We fused one low and different numbers (2, 4, and 6) of high subbands, which have the highest
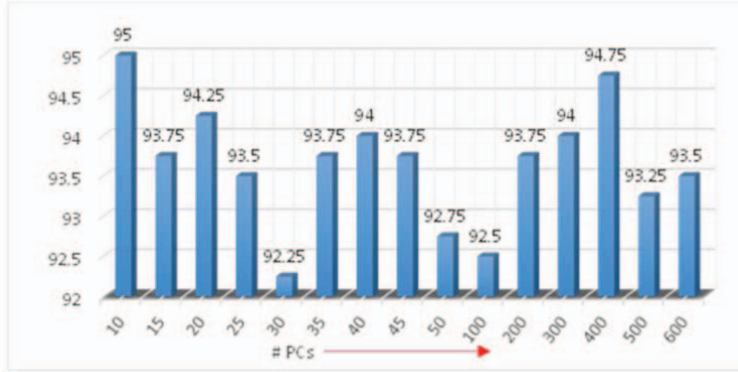
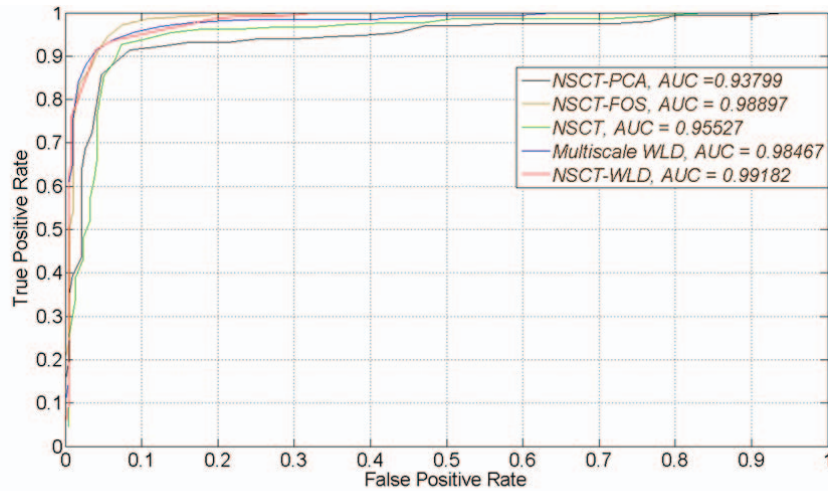**Figure 9. The effect of different number of PCs on recognition rate.**



**Figure 10. ROC curves and AUC**

**Table 2. Accuracy (%) for the four methods.**

| Method | Before FSS | | After FSS | |
|---|---|---|---|---|
| | #Feats. | Acc±std | #Feats. | Acc±std |
| NSCT-PCA (low) | --- | --- | 10 | 95.00±2.63 |
| NSCT (low) | 16384 | 0.94±0.06 | 3821 | 95.25±2.15 |
| Multiscale WLD | 25920 | 95.75±4.41 | 9647 | 98.75±1.31 |
| NSCT-FOS (low) | 4805 | 95.25±3.62 | 267 | 98.25±2.64 |
| NSCT-WLD (low) | 25920 | 98.25±2.05 | 299 | **99.5±1.05** |

recognition rates. The results shown in Table 1 indicate that fusion has no improvement; instead accuracy declines with increasing the number of high subbands. The best accuracy (**99.50±1.58**) is obtained with the fusion of low and two high subbands, which is not better than what (**99.50±1.05**) was obtained with low subband alone. This leads us to conclude that it is only low subband that plays key role in gender recognition; this is in accordance with our common observation that we can recognize a person's gender from a distance when most of high frequency features like wrinkles are not visible.

### 4.2. Results with NSCT-FOS Descriptor

We divided each subband into overlapping windows (#49, #225, #961) and found that 961 gave the best result. Like NSCT-WLD, in this case again low subband gave the best accuracy, see Figure 8.

### 4.3. Results with NSCT-PCA Descriptor

In view of the findings with NSCT-WLD and NSCT-FOS, we applied PCA only for low subband and examined the effect of different numbers of principal components (PCs), see Figure 9. The maximum accuracy is obtained with 10 PCs.

### 4.4. Results with NSCT Descriptor

In this case coefficients in each suband are treated as features and the subband is converted into a feature vector; in this case again the best accuracy was obtained with low sub-band, see Table 2.

We also tested the effect of only multiscale WLD with T=12, M=4, S=5, and 6x6 blocks. Table 2 shows the accuracy and Figure 10 shows the ROC curves and AUC for the five methods; in each case we used the parameter configurations (discussed above) that gave the best recognition rates. It is clear that feature selection not only reduces significantly, but also improves accuracy. The best accuracy (**99.5±1.05**) and AUC (**00.99**) are given by NSCT-WLD only with 299 features.

## 4.5 Comparison with State-of-the-art Methods

We perform comparison with some recent gender recognition systems: Local Gabor Binary Pattern with LDA and SVMAC (LGBP-SVMAC) [1], Local Gabor Binary Pattern with LDA and SVM (LGBP- SVM) [1], Multi-resolution Decision Fusion method (MDF) [7], and the method based on dyadic wavelet transform and LBP (DyWT-LBP) [4]. The results are shown in Figure 11.
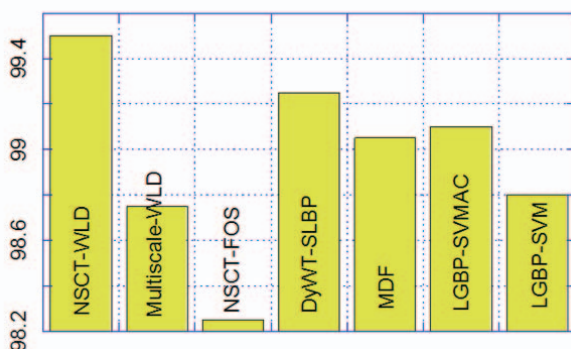


**Figure 11. Comparison with stat-of-the-art methods**

## 5. Conclusions

We investigated the feature extraction approaches based on NSCT for a gender recognition system that employs SVM, two-stage feature selection. We analyzed the impact on gender recognition of low and high subbands independently in the multiscale hierarchy of each facial image; also we examined the effect of fusing low and high subbands that gave the highest recognition rate. We found that it is only low subband that plays significant role in gender recognition; high subbands contain high frequency information that is not important for gender or face recognition; it is in accordance our common observation that we can recognize a person from a distance where high frequency information like wrinkles is not visible. The high frequency information contained in high subbands might be useful for detecting special effects like aging. We examined four feature extraction techniques based on NSCT: NSCT (coefficients only), NSCT-PCA, NSCT-FOS and NSCT-WLD. The recognition accuracy achieved using NSCT-WLD descriptor is not only better than that obtained by NSCT, NSCT-PCA, NSCT-FOS or WLD descriptors but also the state-of-the-art methods. The two-stage feature selection approach reduces the dimension of the feature space and increases the recognition rate. NSCT-WLD descriptor achieved an accuracy of 99.5% on FERET

database that is the best accuracy obtained so far. In our future work we will explore the effect of NSCT-WLD descriptor on race and face recognition, and age detection.

## Acknowledgement

## References

[1] J. Zang and B. L. Lu, "A support vector machine classifier with automatic confidence and its application to gender classification", Neurocomputing (2011) vol. 74, pp.1926-35.

[2] B. Moghaddam and M.-H.Yang, "Gender classification with support vector machines" Proc. IEEE International Conference on Automatic Face and Gesture Recognition (March 2000), pp.306-311.

[3] S. Gutta, H. Wechsler, and P. Phillips, "Gender and ethnic classification of face images", Third IEEE International Conference on Automatic Face and Gesture Recognition (FG'98), (1998) 194-199

[4] Hussain, M., Ihsan Ullah, Aboalsamh, H., Muhammad, G., Bebis, G., and Mirza, A. M., "Gender Recognition from Face Images with Dyadic Wavelet Transform and Local Binary Pattern," International Journal on Artificial Intelligence Tools vol. 22(6), 1360018 (18 pages), 2013.

[5] S. Baluja, and H. Rowley, "Boosting sex identification performance", International Journal of Computer Vision (January 2007), vol. 71(1), pp. 111-119.

[6] L. Lu and P. Shi, "Fusion of multiple facial regions for expression-invariant gender classification", IEICE Electronic Express (2009), vol. 6(10), pp. 587-593.

[7] L. A. Alexandre, "Gender recognition: A multiscale decision fusion approach" Pattern Recognition Letters (2010), vol.31, pp.1422-1427.

[8] J. Zhou, A. L. Cunha, and M. N. Do, "Nonsubsampled contourlet transform: construction and application in enhancement," Proc. ICIP 2005, pp. I 469-72.

[9] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikainen, X. Chen, and W. Gao, "WLD: A robust local image descriptor," IEEE TPAMI (2010), 32(9), pp. 1705–1720.

[10] P. E. H. R. O. Duda and D. G. Stork. Pattern Classification. Wiley-Interscience Publication, 2001.

[11] Y. Sun, S. Todorovic, and S. Goodison, "Local-learning-based feature selection for high-dimensional data analysis," IEEE TPAMI (2010), 32(9), pp. 1610–26.

[12] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," IEEE TPAMI (2000), 22(10), pp. 1090–1104.

[13] K. Veropoulos, G. Bebis, and M. Webster, "Investigating the impact of face categorization on recognition performance," Proc. First international conference on Advances in Visual Computing, Lake Tahoe, NV, 2005.

[14] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on, 1991, pp. 586–591.

[15] Taghreed Alamri, Hussain, M., Ghulam Muhammad, Hatim Aboalsamh, George Bebis and Anwar M. Mirza, "Category Specific Face Recognition Based On Gender", Proc. ICISA 2013, IEEE Computer Society Press, pp. 214-217, June 24-26, 2013, Pattaya, Thailand..