Kernel Machine Based Fourier Series

Masoumeh Abbasian, Hadi Sadoghi Yazdi, Abedin Vahedian Mazloom

Department of Communication and Computer Engineering,
Ferdowsi University of Mashhad, Iran
massomeh.abasiyan@gmail.com, h-sadoghi@um.ac.ir, vahedian@um.ac.ir

Abstract

In this paper, we present a new form of Fourier series named Kernel Fourier series (KFS) which produces time-frequency coefficients similar to wavelet. Both continuous and discrete forms of KFS are presented together with inverse KFS formulated to perform signal approximation. As KFS coefficients are dependent upon selection of kernel function, it can be used in different applications. Results from KFS test on feature extraction, signal analysis and signal estimation are presented and discussed.

Keywords: *Kernel machine, Feature extraction, Fourier series, Signal analysis*

1. Introduction

Fourier theory still remains an important tool in science and engineering with many recent applications in microwave circuit analysis [6], control theory [7], and optical measurements [8]. Combining this theory with kernel methods yields interesting results, namely time-frequency Fourier series.

Kernel Methods (KMs) are a class of algorithms for pattern recognition which deal with the issues of mapping data into a high dimensional feature space. As a generic mapping, relations found this way are consequently very general. This approach is called the kernel trick. Some recent applications of KMs include: clustering [2], time series prediction [3], recognition task [4], and tracking algorithms [5] all of which reveal ability of KMs. The Novel approach presented in this work is capable of assimilating kernel advantages to Fourier series. KFS is able to suppress noise, approximate, and extract time-frequency features from signals.

This paper is organized as follows: Section 2 explains basic concepts relevant to KFS including Fourier series and kernel methods. Section 3 is devoted to proposed KFS, its properties and applications. Some examples about the KFS navigation and applications are represented in Section 4. Finally Section 5 concludes this paper.

2. Preliminaries

KFS is the composition of kernel machine and Fourier series. Kernel machine itself is a useful tool in artificial algorithms and in pattern analysis. The relation between these fields is the kernel trick. Basic steps of KFS are presented in two subsections. The first one is a review of Fourier series and their applications and the second one is an introduction to kernel methods and concept of kernel trick.

2.1. Fourier Series

Fourier series make use of the orthogonality relationships between sine and cosine functions. The computation and study of Fourier series is known as harmonic analysis and is extremely useful as a way to break up an arbitrary periodic function into a set of simple terms which can be plugged in, individually solved, and recombined again to obtain the solution to the original problem or as an approximation of it at any desired precision.

Consider x(t) as a function of the real variable t. This function is usually taken to be periodic, of period 2π, (i.e. x(t + 2π) = x(t)) for all real numbers t. We can write x(t) as the following:

$$x(t) = \sum_{k=-\infty}^{\infty} a_k \, \varphi_k(\mathrm{t}) \tag{1}$$

According to Fourier theory $\varphi_k(t)$ are orthogonal basis functions. The inner product of $\varphi_k(t)$ and $\varphi_m(t)$ in the interval [a, b] is then:

$$\langle \varphi_k(t), \varphi_m(t) \rangle = \int_a^b \varphi_k(t)\varphi_m^*(t)dt \tag{2}$$

And

$$\langle \varphi_k(t), \varphi_m(t) \rangle = 0, \quad \forall\, k \neq m \tag{3}$$

We can rewrite (1) as:

$$\langle x(t), \varphi_m(t) \rangle = \sum_{k=-\infty}^{\infty} a_k \langle \varphi_k(t), \varphi_m(t) \rangle \tag{4}$$

We have:

$$\langle x(t), \varphi_m(t) \rangle = a_k \langle \varphi_k(t), \varphi_m(t) \rangle \; for \; k{=}m \tag{5}$$

Then:

$$a_k = \langle x(t), \varphi_k(t) \rangle / \langle \varphi_k(t), \varphi_k(t) \rangle \tag{6}$$

Now consider:

$$\varphi_k(t) = e^{jk\omega_0 t} \tag{7}$$

We can rewrite a_k as the following:

$$a_k = \frac{1}{T}\int_0^T x(t)e^{-jk\omega_0 t}\, dt \tag{8}$$

2.2. Kernel Methods

Kernel Methods (KMs) are a class of algorithms for pattern analysis with best known elements as Support Vector Machine (SVM). The general task of pattern analysis is to find and study general types of relations (for example clusters, rankings, principal components, correlations, classifications) in general types of data (such as text documents, sets of points, vectors, images, etc). KMs owe their name to the use of kernel functions which enable them to operate in the feature space without even computing the coordinates of the data in that space, but rather by computing the inner products between the images of all pairs of data in the feature space. This operation is often computationally cheaper than the explicit computation of the coordinates.

In this field, the kernel trick is a method for using a linear classifier algorithm to solve a non-linear problem by mapping the original non-linear observations into a higher-dimensional space, where the linear classifier is subsequently used; this makes a linear classification in the new space equivalent to non-linear classification in the original space. This is done by using Mercer's theorem, which states that any continuous, symmetric, positive semi-definite kernel function K(x, y) can be expressed as a dot product in a high-dimensional space.

The concept of a kernel formulated as an inner product in a feature space allows us to build interesting extensions of many well-known algorithms by making use of kernel trick, also known as kernel substitution. The general idea would be, if we have an algorithm formulated in such a way that the input vector x enters only in the form of scalar products, then we can replace that scalar product with some other choice of kernel [1].

3. Kernel Fourier Series

Kernel Fourier series, in fact, maps Fourier series into kernel domain. Since Fourier series can be expressed as an inner product we can represent it in kernel domain using kernel trick, obtaining KFS. In this section we first introduce KFS in continuous form and demonstrate some of its properties. Then discrete form of KFS is presented followed by inverse KFS and applications of KFS.

3.1. Continuous Form of KFS

To obtain KFS, the signal and its basic function should be transferred into Φ-space domain.

$$x(t) \xrightarrow{\Phi} (x(t)) \tag{9}$$

$$\varphi_k(t) \xrightarrow{\Phi} \boldsymbol{\Phi}(\varphi_k(t)) \tag{10}$$

We suppose that the basic function $\varphi_k(t)$ in the Φ-space domain is still orthogonal. Therefore we can write KFS coefficients as the following:

$$b_k = \frac{\langle \Phi(x(t)), \Phi(\varphi_k(t)) \rangle}{\langle \Phi(\varphi_k(t)), \Phi(\varphi_k(t)) \rangle} \tag{11}$$

According to $(u, v) = \Phi(u)^T \Phi(v)$, inner product of (11) can be written as:

$$b_k = \frac{k(x(t), \varphi_k(t))}{k(\varphi_k(t), \varphi_k(t))} \tag{12}$$

To make it clearer, we show that b_k will be equal to a_k through defining our kernel function in special manner.

Lemma: if we define a kernel function as:

$$k(x(t), \varphi_k(t)) = \int_0^T x(t) \varphi_k^*(t) dt \tag{13}$$

And

$$\varphi_k(t) = e^{jk w_0 t} \tag{14}$$

Then

$$k(x(t), \varphi_k(t)) = \int_0^T x(t) e^{-jk w_0 t} dt = T a_k \tag{15}$$

Hence

$$b_k = k(x(t), \varphi_k(t)) / k(\varphi_k(t), \varphi_k(t)) = a_k \tag{16}$$

In the special case of kernel, therefore, FS and KFS coefficients are equal.

3.2. Kernel Fourier Series Properties

Some of KFS coefficients properties are shown in Table I. In fact we consider the kernel functions as polynomial and Radial Basis Function (RBF) to extract these properties.

Table 1. KFS Properties

Signal	KFS coefficient
$x(t)$	$b_k(t)$
$x(at)$	$b_{k/a}(t)$
$x(-t)$	$b_{-k}(t)$
$x^*(t)$	$b^*_{-k}(t)$
$x(t)e^{JMw_0t}$ (Only for RBF kernel)	$e^{-2JMw_0t}b_{k+M}(t)$
If $x(t)$ is real signal	$b_k(t) = b^*_{-k}(t)$ i.e. $\lvert b_k(t)\rvert$ is even & $\sphericalangle(b_k(t))$ *is odd*
If $x(t)$ is even signal	$\lvert b_k(t)\rvert$ is even , $\sphericalangle(b_k(t))$ *is also even*

3.3. Discrete Form of KFS

The following equation states the discrete form of Fourier series:

$$x[n] = \sum_{k=<N>} a_k \varphi_k[n] \tag{17}$$

Because of orthogonal property of basic functions, we have:

$$a_k = \langle x[n], \varphi_k[n]\rangle / \langle \varphi_k[n], \varphi_k[n]\rangle \tag{18}$$

Consider

$$\varphi_k[\mathrm{n}] = e^{jk\frac{2\pi}{N}n} \tag{19}$$

So we can write below equation that is resulted from (18) and (19).

$$a_k = \frac{1}{N}\sum_{k=<N>} x[n]e^{-jk\frac{2\pi}{N}n} \tag{20}$$

Now we introduce discrete KFS in the same way that we defined continuous KFS. After transferring the signals into

Φ-space we can write:

$$b_k = \frac{\langle \Phi(x[n]), \Phi(\varphi_k[n])\rangle}{\langle \Phi(\varphi_k[n]), \Phi(\varphi_k[n])\rangle} \tag{21}$$

By kernel substitution we have:

$$b_k = \frac{k(x[n], \varphi_k[n])}{k(\varphi_k[n], \varphi_k[n])} \tag{22}$$

3.4. Inverse KFS for Signal Estimation

As any periodic signal can be represented in the form of:

$$x(t) = \sum_{k=-\infty}^{\infty} a_k\, \varphi_k(\mathrm{t}) \tag{23}$$

Supposing that basic functions in new domain are still orthogonal, we have:

$$\Phi(x(t)) = \sum_t \sum_k b_k(t)\Phi(\varphi_k(t)) \tag{24}$$

The ability of kernel in simplifying the inverse procedure can be easily identified. Multiplying two sides of the above equation in an arbitrary vector such as $\Phi(x_1)$ yields:

$$\Phi\big(x(t)\big)\Phi(x1) = \sum_t \sum_k b_k(t)\Phi(\varphi_k(t))\Phi(x1) \tag{25}$$

Now with kernel trick we obtain:

$$k(x(t), x1) = \sum_t \sum_k b_k(t)k\,(\varphi_k(t), x1) \tag{26}$$

To get *x(t)* it suffices to choose a proper kernel function. Hence the inverse operation depends on the selected kernel function. We obtain the inverse KFS with linear kernel function. Our experiments have shown that this procedure captures signals fairly well.

3.5. KFS Applications

We now introduce three types of KFS applications for feature extraction, signal analysis such as noise suppression, and signal estimation. We illustrate them in experimental results thorough examples.

3.5.1. KFS-Based Feature Extraction: In the pattern recognition and image processing, feature extraction is a special form of dimensionality reduction. Transforming input data into the set of features is called feature extraction. If the extracted features are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of high dimension input. Similar to Fourier series, wavelet, and curvelet transforms, we can use this method for feature extraction.

3.5.2. Signal Analysis: Once extraction of suitable features (time dependant features) is performed, we can use them for signal analysis. Some of the proposed signal analysis include: signal classification, clustering of time dependant features of signal, noise suppression or interference cancellation.

3.5.2.1. Classification and Clustering using KFS: KFS introduces time-dependant features which can be classified using Hidden Markov Model (HMM), Time Warping and Neural Network based approaches. Suitable kernel must be used in specific applications of course. Clustering techniques over time-dependant signal are another interesting field in the pattern analysis. Time-dependant features are extracted using wavelet technique and similar approaches while KFS can extract similar features.

3.5.2.2. Noise Suppression: Noisy condition is one of the main problems in signal processing and pattern recognition. Many filters were presented for this purpose as linear filters (such as Least Mean Square, Recursive Least Square, and Kalman filter) and non linear filters (like Kernel Least Mean Square, Kernel Recursive Least Square, and Extended Kalman filter). We demonstrate that a suitable kernel can suppress noise too.

3.5.2.3. Signal Estimation: Similar to Fourier series we can use KFS components to estimate signals. We show that recursive relationship of KFS could be obtained and we can approximate original signal using appropriate kernel function. KFS was applied to both standard and natural signals with results represented in the following section.

4. Experimental Results

Several examples are provided to clarify the KFS navigation for its applications mentioned earlier. These include feature extraction, an abrupt change in signal's effect on KFS, noise suppression, and signal estimation. At first we obtain KFS coefficients of a signal and test the lemma stated before. In signal estimation issue, we discuss standard signal, e.g. *sin(.)* and natural signal, e.g. speech signal.

4.1. KFS Coefficients of a Sine Function

Initialize: ***x*** is a desired signal (such as *cos(at), sin(at),*e^{-at}). Here we choose *sin(t).*

t is a time interval (i.e. $[0,2\pi]$).

k is the number of coefficients (i.e.[-5,5]).

T is the period of our signal (i.e. 2π). Consequently $\omega_0 = \frac{2\pi}{T} = 1$.

Procedure: Initial step to calculate $b_k(t)$ is the kernel function selection. We choose *tanh*$(auv^T/L + b)$ as the kernel function namely sigmoid, where a, b are constant (e.g. a=1, b=0) and L is the number of input signal components (e.g. L=2001). Given k and t, $b_k(t)$'s coefficient is resulted from (16). Figure 1 represents the KFS coefficients of *sin(t).*

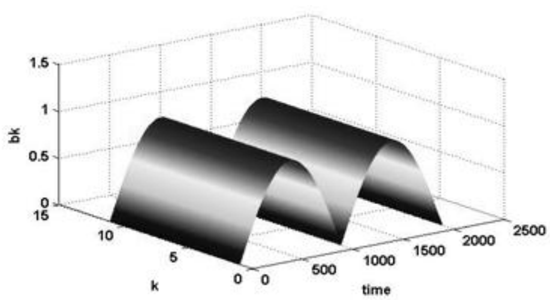

Figure 1. $b_k(t)$'s coefficients of sin(t). To calculate these coefficients we choose sigmoid kernel function.

4.2. KFS-Based Feature Extraction

Our variables in this section are similar to example 1. We calculate $b_k(t)$ and show these coefficients in 40th sample in Figure 2.

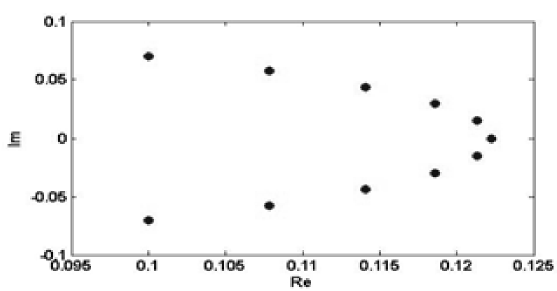

Figure 2. 11 coefficients of b_k(t) in 40th sample

These features can be used in applications like speech recognition. We can select the best features to use in classification and recognition tasks. One future work would be to apply this method in speech recognition, finding out the extent to which this method is able to improve the recognition accuracy.

4.3. Study of Influence of an Abrupt Change in Signal on KFS Coefficients

Now we make an abrupt change in the signal by setting *x(40)=0.5* as shown in Figure 3, left. Our experiment shows that this affects only 40th sample but no other samples are altered.

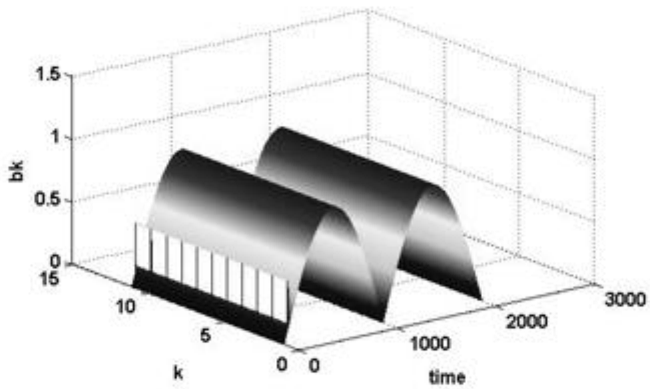

Figure 3. (a) An abrupt change in original signal (left), (b) Influence of an abrupt change in signal on $\mathbf{b_k}(\mathbf{t})$ (right).

This indicates that KFS coefficients can be used for detection of event or occurrence of an abrupt change in signals. This enables us to detect noisy data in classification applications hence obtain the best classifier.

4.4. Noise Suppression

We now demonstrate the ability of KFS to noise reduction. If we use *exp*$(auv^T/L+b)$ as essential kernel function instead of *tanh*$(auv^T/L+b)$ in previous example, the effect of noise on extracted features would be limited. This property is shown in Figure 4.

Therefore, for noise reduction in signal or feature selection in noisy environment, we can find appropriate kernel function which yields suitable and noise-free features.

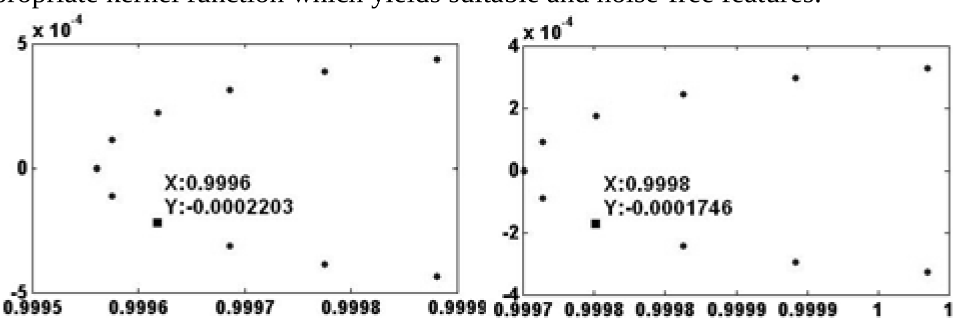

Figure 4. There is not an abrupt change in signal (left), There is an abrupt change in signal (right).

4.5. Signal estimation

This requires the following steps to be taken to estimate a desired signal:

1) Choosing proper kernel function to calculate KFS coefficients.
2) Choosing suitable kernel function to use (26) (the equation of signal reconstruction).
3) Using (26) to obtain *x(t).*

We select $tanh(\mathrm{auv}^{\mathrm{T}}/\mathrm{L}+\mathrm{b})$ for the first step and uv^{T} for the second.

Considering

$$k(x(t),x_1)=xx_1^T \tag{27}$$

And

$$k(\varphi_k(t),x_1)=\varphi_k(t)x_1^T \tag{28}$$

We can write (26) as:

$$xx_1^T=\sum\sum b_k(t)\,\varphi_k(t)x_1^T \tag{29}$$

Without loss of generality, we may assume x_1 as a unit vector. Hence:

$$x(t)=\sum\sum b_k(t)\varphi_k(t) \tag{30}$$

Therefore for each sample (e.g. t_0) in the given time interval, we can write:

$$x(t_0)=\sum_k b_k(t_0)\varphi_k(t_0) \tag{31}$$

We now use the above equation to capture some signals.

4.5.1. Signal reconstruction using KFS for standard signals: Consider $\boldsymbol{x(t)}=-\boldsymbol{t^2e^{-t}\cos(t)}$. Figure 5 shows the result of signal reconstruction using this new method which demonstrates that KFS components capture signals very well.

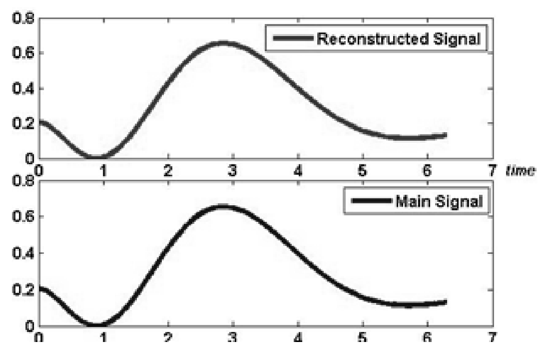

Figure. 5 The comparison between estimated signal and original signal.

4.5.2. Speech reconstruction using KFS: Based on the discussion of signal reconstruction, the method is examined for speech waveforms (e.g. /a/ phoneme) using the following structure.

- Speech in KFS domain

***x*:** The frame of original signal with a length of 20ms.

Feature*:** KFS(x***)

- Reconstruct speech signal from extracted features.

X: inverse KFS(feature)

Estimated signal**:** reconstruction from ***X*** vectors.

Figure 6 shows the result. We can see that KFS has captured this signal as good as did before for standard signal.

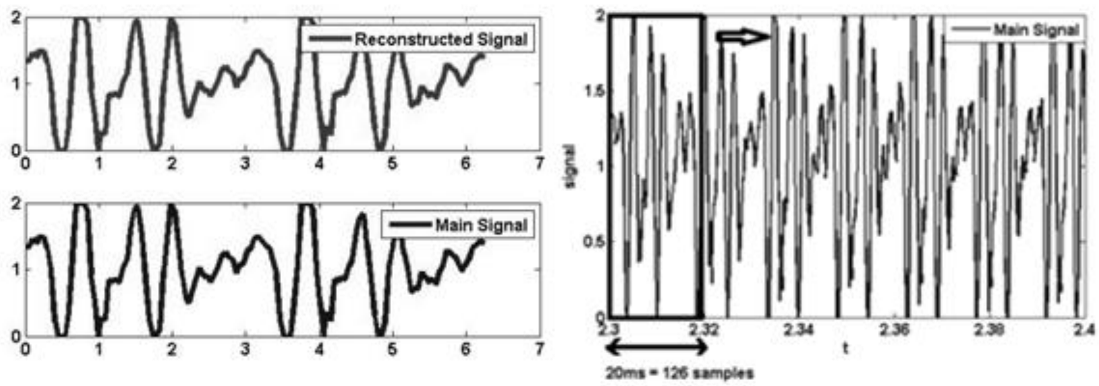

Figure. 6 reconstruction of speech waveform by a frame of 20ms length (/a/ phoneme): (a) right picture shows the framing process of signal. (b) left picture shows samples from 11340 to 11466.

5. Conclusion and Future Work

In this paper a novel idea has been proposed to define Fourier series in kernel domain. This kernel substitution is based on the concept of a kernel formulated as an inner product in a feature space. In special case with an integral kernel, we proved that KFS is similar to conventional Fourier series. KFS can be used as feature extractor and is capable of producing frequency coefficient in terms of time. We presented some applications for KFS such as pattern recognition (classification and clustering), signal analysis (noise suppression) and signal estimation. Selection of an appropriate kernel function remains a challenge however. For future work it is required to find a reliable approach to select proper kernel function. Fourier transform in kernel domain would be a research ground for image processing applications.

References

[1] A. Ralf Herbrich, "Learning Kernel Classifiers Theory and Algorithms", Massachusetts Institute of Technology, 2002.

[2] A. Heinz, C. and B. Seeger, B., "Cluster Kernels: Resource-Aware Kernel Density Estimators over Streaming Data"; Knowledge and Data Engineering , *IEEE Trans. Power Delivery*, Vol. 20, No. 7, pp. 880-893, 2008.

[3] A. Richard, C. and B. Bermudez, J. C. M. and Honeine, P., "Online Prediction of Time Series Data With Kernels"; Signal Processing, *IEEE Trans. Power Delivery*, Vol. 57, No. 3, pp. 1058 - 1067, 2009.

[4] A. Chang Huai You , B. Kong Aik Lee and C. Haizhou Li, "An SVM Kernel With GMM-Supervector Based on the Bhattacharyya Distance for Speaker Recognition"; Signal Processing Letters, *IEEE Trans. Power Delivery*, Vol. 16, No. 1, pp. 49-52, 2009.

[5] A. Leichter, I. , B. Lindenbaum, M. and C. Rivlin, E.,"Tracking by Affine Kernel Transformations Using Color and Boundary Cues", Pattern Analysis and Machine Intelligence, *IEEE Trans. Power Delivery*, Vol. 31, No. 1, pp. 164 – 171, 2009.

[6] A. Ditri, J., "Efficient Fourier Series Solutions to Nonlinear Steady-State Heat Conduction Problems in Microwave Circuits"; Components and Packing Technologies, *IEEE Trans. Power Delivery*, Vol. 32, No. 1, pp.110 – 119, 2009.

[7] A. Lasiecka, I. and B. Edward, J., "Series in Control Theory-[Book review; V. Komornik and P. Loreti]"; Automatic Control , *IEEE Trans. Power Delivery*, Vol. 52, No. 2, pp. 379 – 382, 2007.

[8] A.Vanherzeele,J. , B. Vanlanduit,S. and C. Guillaume,P. "Processing optical measurements using a regressive Fourier series"; A review Optics and Lasers in Engineering , *Power Delivery*, Vol. 47, No. 3-4, pp. 461-472, 2009.

Authors

Masoumeh Abbasian received the B.S. degree in robotic engineering from polytechnic shahrood University of Shahrood in 2006, and then she received to the M.S. degrees in communication engineering from Ferdowsi University of Mashhad, Iran, in 2010. Her research interests include pattern recognition, optimization in signal processing, image processing, machine vision, and multi resolution methods.

Hadi Sadoghi Yazdi received the B.S. degree in electrical engineering from Ferdowsi University of Mashhad in 1994, and then he received to the M.S. and PhD degrees in electrical engineering from Tarbiat Modarres University of Tehran, Iran, in 1996 and 2005 respectively. He works in Computer Department as an associate professor at Ferdowsi University of Mashhad. His research interests include pattern recognition, optimization in signal processing.

Abedin Vahedian Mazloom received the B.S. degree in electrical engineering from Ferdowsi University of Mashhad in 1987, and then he received to the M.S. and PhD degrees in electrical engineering from UNSW University of Australia, in 1995 and 1999 respectively. He works in Computer Department as an associate professor at Ferdowsi University of Mashhad. His research interests include machine vision, video conference, information technology and image coding.