The Orthogonal Decomposition Algorithm for Speech Signals in Reproducing Kernel Space

Sen Zhang, Lei Liu, and Luhong Diao
College of Applied Sciences, Beijing University of Technology, Beijing, China
Email: zhangsen@yahoo.com

Abstract—An orthogonal decomposition method and implementation algorithm for speech signal processing are proposed in this paper. In the reproducing kernel function of Hilbert space \( W_2^1 \), a set of normalized orthogonal function system \( \{ \phi_j(x) \}^n \) is generated, and speech signals can be orthogonally decomposed in \( W_2^1 \) according to the basis \( \{ \phi_j(x) \}^n \), the orthogonal decomposition coefficients can be computed by a fast algorithm based on the properties of reproducing kernel function. This approach mapped the speech signals represented by discrete samples to continuous functions which is different from the canonical form represented by series of triangle functions, and the inner product computation in Hilbert space was transformed into function evaluation problem only at some discrete points.

Index Terms—Speech Signal, Reproducing Kernel Space, Orthogonal Decomposition, Signal Analysis

I. INTRODUCTION

The orthogonal decomposition methods (ODM) have been widely applied in speech signal analysis and processing fields. By orthogonal decomposition, speech signals can be projected into a special subspace and the projections or coefficients can be used to represent the speech signals’ features which could be used in speech coding to compress the redundancy, etc. As we know, Fourier transformation and wavelet transformation are typical orthogonal decomposition methods which have been successfully utilized in speech signal processing fields and other engineering fields. However, the disadvantages of these orthogonal decomposition methods have been thoroughly investigated and were proposed in many research reports.

In recent years, Reproducing Kernel (RPK) space theory has been used in some engineering fields, such as classification methods based on Support Vector Machine (SVM), pattern recognition, machine learning, image compression and reconstruction, and these successful applications in other fields encouraged researchers to study whether and how RPKS could be used in speech signal analysis and processing fields. As a result, we know the discrete speech signals can be represented in continuous form by RPK which could reduce the computation load in speech analysis process, and the RPK operator eigenvalue theory was used to investigate the nonlinear features and the high order harmonics of speech signals. Besides, the RPK was also used in speech signal compression and reconstruction based on the property of RPK which can produce uniform approximation for speech signals by using almost optimal interpolation operator of RPK. The speech features could be also extracted in RPK space and applied in speech recognition. Some research reported that the speech features extracted in RPK space were more robust compared to MFCC which was widely used in speech recognition today.

This paper includes six sections which are arranged as follows: firstly, the up-to-date applications of RPK theory in speech signal processing fields are introduced; next, a special RPK space \( W_2^1 \) is discussed; thirdly, some problems related to orthogonal decomposition of speech signals in \( W_2^1 \) are investigated; fourthly, the orthogonal decomposition approaches and algorithms of speech signals in \( W_2^1 \) are proposed in details; the fifth section presents some experimental results and the evaluation, and some conclusions are given in the last section.

II. RPKS \( W_2^1 \)

To begin with, some fundamental results about space \( W_2^1 \) will be discussed in this section. Suppose a special functional space \( W_2^1 \) contains some real or complex functions defined in interval \( [a, b] \) as its elements. By properly defining inner product and norm, \( W_2^1 \) could become a Hilbert space, and it is also a RPK space. The RPK of \( W_2^1 \) could be represented in some simple function form.

Definition 2.1 \( W_2^1 = W_2^1 \) is an absolutely continuous function defined in \( [a, b] \), and its derivative function \( f'(x) \in L^2[a, b] \), where \( L^2[a, b] \) is a function set which contains all the functions which are integrable in squared form.

Definition 2.2 The inner product and norm in space \( W_2^1 \) were defined respectively as follows:
\[
(f, g) = \int_a^b (f(x)g(x) + f'(x)g'(x))dx,
\]

\[\|f\| = (f, f)^{1/2} \tag{2.1}\]

**Theorem 2.1**[16] The function space \(W_2^1[a,b]\) is a complete inner product space with the inner product defined by formula (2.1).

**Theorem 2.2**[16] The inner product space \(W_2^1[a,b]\) has a unique RPK function \(K(x, y)\) which can be represented as follows:

\[
K(x, y) = \frac{1}{2sh(b - a)}[ch(x + y - b - a) + ch([x - y] - b + a)]
\]

(2.2)

Thus, \(W_2^1[a,b]\) is a RPK space with RPK function \(K(x, y)\).

The RPK function \(K_i(y)\) of \(W_2^1[a,b]\) has the reproducing property, that is \(\forall f(x) \in W_2^1[a,b]\) and \(\forall y \in [a,b]\), we have

\[
(f(x), K_i(y)) = f(y) \tag{2.3}
\]

Next, a complete function system in space \(W_2^1[a,b]\) will be constructed. Suppose the set \(T = \{t_1, t_2, \cdots\}\) is dense in the interval \([a,b]\), and for each \(t_i \in T\), according to the property of the RPK function \(K_i(y)\), we have a function \(K_i(x)\), and for simplicity, we rewrite \(K_i(x)\) as \(\varphi_i(x)\), namely \(\varphi_i(x) = K_i(x)\). Then we have the following Theorem 2.3.

**Theorem 2.3** The function system \(\{\varphi_i(x)\}_{i=1}^{\infty}\) is a complete function system in space \(W_2^1[a,b]\).

Proof: Suppose function \(f(x) \in W_2^1[a,b]\). If \((f(x), \varphi_i(x)) = 0, i = 1, 2, \cdots\), note the definition of \(\varphi_i(x)\) and its reproducing property, then \(f(t_i) = (f(x), \varphi_i(x)) = 0, i = 1, 2, \cdots\). Since the set \(T = \{t_1, t_2, \cdots\}\) is dense in interval \([a,b]\), and \(f(x)\) is continuous in \([a,b]\), it is easy to know that \(f(x) \equiv 0\), which means the function system \(\{\varphi_i(x)\}_{i=1}^{\infty}\) is a complete function system in space \(W_2^1[a,b]\). □

Furthermore, take \(\{\varphi_i(x)\}_{i=1}^{\infty}\) to be orthogonalized through Schmidt approach, a standard orthogonal base \(\{\varphi_i(x)\}_{i=1}^{\infty}\) in \(W_2^1[a,b]\) is obtained, that is

\[
\varphi_i^*(x) = \sum_{j=1}^{i} \beta_{ji} \varphi_j(x) \tag{2.4}
\]

where \(\beta_{ji}\) is the orthogonal coefficient.

**Theorem 2.4** Suppose \(f(x)\) is a function in \(W_2^1[a,b]\), then the Fourier series of \(f(x)\) regarding to \(\{\varphi_i^*(x)\}_{i=1}^{\infty}\) is convergent, and we have the following formula:

\[
f(x) = \sum_{i=1}^{\infty} \sum_{j=1}^{i} \beta_{ji} f(t_j) \varphi_i^*(x) \tag{2.5}
\]

Proof: Since the standard orthogonal base \(\{\varphi_i^*(x)\}_{i=1}^{\infty}\) in \(W_2^1[a,b]\) is complete, function \(f(x) \in W_2^1[a,b]\) can be decomposed as follows:

\[
f(x) = \sum_{i=1}^{\infty} (f(x), \varphi_i^*(x)) \varphi_i^*(x)
\]

based on formula (2.4), we obtained \(f(x)\) as follows:

\[
f(x) = \sum_{i=1}^{\infty} \sum_{j=1}^{i} \beta_{ji} f(t_j) \varphi_i^*(x)
\]

\[
= \sum_{i=1}^{\infty} \sum_{j=1}^{i} \beta_{ji} \varphi_j(x) \varphi_i^*(x)
\]

\[
= \sum_{i=1}^{\infty} \sum_{j=1}^{i} \beta_{ji} f(t_j) \varphi_i^*(x)
\]

We write \(\alpha_{ji} = \sum_{j=1}^{\infty} \beta_{ji} f(t_j)\) and call \(\alpha_{ji}\) generalized Fourier coefficients, where \(\beta_{ji}\) is the Schmidt orthogonal coefficient. Obviously, the computation of \(\alpha_{ji}\) only requires to evaluate \(f(x)\) at some fixed points (e.g., \(t_j\)). In general, the computation of Fourier coefficients requires to calculate integral which usually is a heavy computation load. Comparatively, the computation of generalized Fourier coefficients in \(W_2^1[a,b]\) requires much less computation cost.

**Theorem 2.5** The projection of function \(f(x) \in W_2^1[a,b]\) in subspace \(S = \text{span}\{\varphi_1, \varphi_2, \cdots, \varphi_n\}\) is represented as follows:

\[
\overline{f}_n(x) = (P_n f)(x) = \sum_{i=1}^{n} \sum_{j=1}^{i} \beta_{ji} f(t_j) \varphi_i^*(x) \tag{2.6}
\]

344
Moreover, we have \( f(t_i) = \overline{f}_n(t_i) \), and \( \overline{f}_n(x) \) is uniformly convergent to \( f(x) \) while \( n \to \infty \), where \( P_n : W^1_2[a, b] \to S \) is projection operator.

Proof: Since \( P_n \) is projection operator, it is easy to obtain the following formula:

\[
(P_n f)(x) = \sum_{i=1}^n (f(x), \varphi_i^* (x)) \varphi_i^* (x)
\]

Note that \( \varphi_i^* (x) = \sum_{j=1}^n \beta_{ji} \varphi_j (x) \) and the property of RPK, then we can obtain:

\[
(P_n f)(x) = \sum_{i=1}^n \sum_{j=1}^n \beta_{ji} f(t_j) \varphi_i^* (x)
\]

On the other hand, since \( P_n \) is a conjugated operator, it satisfies the following conditions:

\[
\left( f(t_i), (P_n f)(x), \varphi_j (x) \right) = \left( f(x), (P_n \varphi_i) (x), \varphi_j (x) \right) = f(t_i) \varphi_j (x)
\]

Next, we will prove that \( \overline{f}_n(x) \) is uniformly convergent to \( f(x) \).

From formula (2.3), we know:

\[
\left\| \overline{f}_n(x) - f(x) \right\| \leq \sum_{j=1}^n \sum_{k=1}^n \beta_{ji} \beta_{kj} f(t_k) \varphi_i (x) - f(t_j)
\]

and based on Th.2.4, we can obtain that \( \left\| \overline{f}_n(x) - f(x) \right\| \to 0 \) thus, \( \overline{f}_n(x) \) is uniformly convergent to \( f(x) \).

\[\square\]

Theorem 2.5 indicates that the projection of function \( f(x) \) in space \( W^1_2[a, b] \) is one kind of interpolation operation of \( f(x) \). Furthermore, we know that projection approximation is optimal in some sense, and the interpolation operation of \( f(x) \) ensures that interpolating function (the projection of function \( f(x) \)) equals to the original function \( f(x) \) at all interpolation points. So, the computation of the projection of function \( f(x) \) can be realized in space \( W^1_2[a, b] \) with much less calculation load.

Based on theorem 2.4 and 2.5, we can obtain that the orthogonal decomposition of \( f(x) \) in space \( W^1_2[a, b] \) regarding to finite orthogonal function set \( \{ \varphi_i(x) \}^n_{i=1} \) is as follows:

\[
f(x) = \sum_{i=1}^n (f(x), \varphi_i^* (x)) \varphi_i^* (x) = \sum_{i=1}^n \sum_{j=1}^n \alpha_{ji} \beta_{ij} \varphi_j (x)
\]

The above formula (2.7) indicates that function \( f(x) \) can be denoted by finite samples \( \{ f(x_j) \}^n_{j=1} \) in interpolation form. In practical application, we do not require to compute all the coefficients \( \beta_{ji} \). In fact, only about \( 3^n \beta_{ji} \) is enough.

The 3-dimensional and 2-dimensional images of the RPK function \( K(x, y) \) of \( W^1_2[a, b] \) were shown in the following Fig. 2.1 and Fig. 2.2, where \( x, y \in [0, 1] \), and \( y \) was sampled at 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.9, etc.

![Fig. 2.1 The 3-dimensional image of \( K(x, y) \)](image)

![Fig. 2.2 The 2-dimensional image of \( K(x, y) \)](image)

**III. DECOMPOSITION IN \( W^1_2[a, b] \)**

Suppose speech signal samples were represented as \( f_1, f_2, \ldots, f_n \). We know from above formula (3.1), the speech signal can be decomposed in space \( W^1_2[a, b] \), the orthogonal decomposition coefficients \( \{ \alpha_{ji} \}^n_{i=1} \) can be computed by following formula:

\[
\alpha_{ji} = \sum_{k=1}^n \beta_{kj} f_k
\]

where \( \beta_{kj} \) is the orthogonal coefficient of \( \{ \varphi_j (x) \}^n_{i=1} \) through Schmidt approach. In general, the computation of \( \beta_{kj} \) requires inner product computation, which usually needs to calculate integral. It is obvious that the key problem of speech signal decomposition in space
$W_2^1[a,b]$ is how to calculate $\beta_{kj}$ quickly by using the properties of RPK space $W_2^1[a,b]$.

From the definition of $\{\varphi_j(x)\}_1^n$, it is a linearly independent function set and has the reproducing property. The Schmidt orthogonalizing process of basis functions $\{\varphi_j(x)\}_1^n$ is as follows:

$$\begin{align*}
\psi_1(x) &= \varphi_1(x), \\
\varphi_1^*(x) &= \psi_1(x)/\|\psi_1\| \\
\psi_k(x) &= \varphi_k(x) - \sum_{j=1}^{k-1} (\varphi_j(x), \varphi_j^*(x))\varphi_j^*(x) \\
\varphi_k^*(x) &= \psi_k(x)/\|\psi_k\|, \quad k = 2,\ldots,n
\end{align*}$$

(4.1)

From above process we know $\varphi_k^*(x)$ is linear combination of $\{\varphi_j(x)\}_1^n$, and the computation of the inner product $(\varphi_k(x), \varphi_j^*(x))$ can be transformed into the computation of $(\varphi_k(x), \varphi_j(x))$, $j = 1,\ldots,k-1$.

Besides, to compute $\|\psi_k\|$ requires to compute $(\varphi_k(x), \varphi_k(x))$ first. Hence, the computation of $(\varphi_k(x), \varphi_j(x))$, $j = 1,\ldots,k$ is necessary.

From the properties of $W_2^1[a,b]$ and the definition of $\{\varphi_j(x)\}_1^n$, it is easy to obtain the following:

$$(\varphi_k(x), \varphi_j(x)) = \varphi_j(x), \quad j = 1,\ldots,k$$

Therefore, we can compute the orthogonal coefficient $\beta_{kj}$ according to the following algorithm:

(a) for $k=1$ and $j=1$, compute $\|\psi_1\|, \beta_{11}$ as:

$$\begin{align*}
\|\psi_1\| &= [\varphi_1(x_1)]^{1/2} \\
\beta_{11} &= 1/\|\psi_1\| = 1/[\varphi_1(x_1)]^{1/2}, \varphi_1^*(x) = \beta_{11}\varphi_1(x)
\end{align*}$$

(b) for $k=2$, based on the results of step (a), compute:

$$\|\psi_2\|^2 = \varphi_2(x_2) - [\beta_{11}\varphi_2(x_1)]^2$$

thus, $\beta_{12}$ and $\beta_{22}$ can be computed as:

$$\begin{align*}
\beta_{22} &= \frac{1}{\sqrt{\varphi_2(x_2) - [\beta_{11}\varphi_2(x_1)]^2}} \\
\beta_{12} &= -\beta_{22}\beta_{11}\varphi_2(x_1) \\
\varphi_2^*(x) &= \beta_{12}\varphi_2(x) + \beta_{22}\varphi_2(x)
\end{align*}$$

(c) for $1<k-1$, suppose all $\beta_{ji}$ ($j = 1,\ldots,i, \quad i \leq k-1$) were computed and available, the left work is to compute $\beta_{jk}, j = 1,\ldots,k$.

For $j < k$, we have

$$\varphi_j^*(x) = \sum_{i=1}^{j} \beta_{ij}\varphi_i(x)$$

$$\begin{align*}
(\varphi_k(x), \varphi_j^*(x)) &= \sum_{i=1}^{j} \beta_{ij}(\varphi_k(x), \varphi_i(x)) \\
&= \sum_{i=1}^{j} \beta_{ij}\varphi_j(x), \quad (j<k)
\end{align*}$$

$$\|\psi_k\|^2 = \varphi_k(x_2) - \sum_{j=1}^{k-1} \left[\sum_{i=1}^{j} \beta_{ij}\varphi_j(x_i)\right]^2$$

$$\beta_{kk} = 1/\|\psi_k\|$$

for $j = 1,\ldots,k-1$, we can obtain $\beta_{jk}$ as follow:

$$\beta_{jk} = -\beta_{kk}\sum_{i=1}^{k-1} \left[\beta_{ji}\sum_{m=1}^{i} \beta_{mi}\varphi_k(x_m)\right], \quad j = 1,\ldots,k-1$$

(d) Computing the decomposition coefficients $\{\alpha_j\}_1^n$.

For $j=1,\ldots, n$, do

$$\alpha_j = \sum_{k=1}^{n} \beta_{kj} f_k$$

and

$$f(t) = \sum_{i=1}^{n} \alpha_i \varphi_i^*(t)$$

Let us make a brief analysis about the above algorithm. Suppose $N$ samples, the computation of $\beta_{kk} (k > 1)$ requires $O(N^2)$ product operations, and $\beta_{kj}, j = 1,\ldots,k$ requires $O(N^2)$ product operations too. Hence, the total product operations requires $O(N^3)$. In the above analysis, the computation load to evaluate $\varphi_k(x_j) (k = 1,\ldots,N, j = 1,\ldots,k)$ is not taken account. Besides, the total add operations requires $O(N^3)$ too.

IV. EXPERIMENTAL RESULTS

According to the decomposition algorithm in above section, we will decompose some speech signals in space $W_2^1[a,b]$. First, the coefficients $\beta_{jk}, j = 1,\ldots,k$ will be calculated. Let’s take the interval $[0,1]$ and equally divide it into $n$ parts, thus a set $\{x_i\}_1^n$ was obtained, where $\{x_i\}_1^n \subset [0,1]$, and the value of $n$ was set to 16,32,320,512,1000, respectively. Now, for each value of $n$, a group of the coefficients $\beta_{jk}, j = 1,\ldots,k$ were calculated as shown in Table 5.1. By observing Table 5.1, we find the except $\beta_{11} = 0.8$, other $\beta_{kk} (k > 1)$ were within the range 4.12~31.63, and the values of $\beta_{k-1,k} (k > 1)$ were almost equal to $\beta_{kk} (k > 1)$, but the
sign was opposite. The values of other $\beta_{jk}$, $j \leq k - 2$ were very small, less than $10^{-10}$, which could be ignored without any significant influence.

<table>
<thead>
<tr>
<th>$N$ node</th>
<th>$\beta_{11}$</th>
<th>$\beta_{kk}$ ($k &gt; 1$)</th>
<th>$\beta_{k-1,k}$</th>
<th>$\beta_{jk}$ ($j &lt; k - 1$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0.91</td>
<td>2.95</td>
<td>-2.72</td>
<td>$10^{-15}$</td>
</tr>
<tr>
<td>16</td>
<td>0.89</td>
<td>4.12</td>
<td>-3.91</td>
<td>$10^{-14}$</td>
</tr>
<tr>
<td>32</td>
<td>0.88</td>
<td>5.71</td>
<td>-5.62</td>
<td>$10^{-13}$</td>
</tr>
<tr>
<td>320</td>
<td>0.87</td>
<td>17.91</td>
<td>-17.87</td>
<td>$10^{-12}$</td>
</tr>
<tr>
<td>512</td>
<td>0.87</td>
<td>22.64</td>
<td>-22.61</td>
<td>$10^{-11}$</td>
</tr>
<tr>
<td>1000</td>
<td>0.87</td>
<td>31.63</td>
<td>-31.61</td>
<td>$10^{-11}$</td>
</tr>
</tbody>
</table>

From above Table 5.1, we know that only $\beta_{kk}$ and $\beta_{k-1,k}$ are important in practical computation. The following figure Fig. 5.1 showed the distribution of the values of $\beta_{jk}$, $j = 1, \ldots, k$, where the node number $n=32$. The peaks of the comb shape are the values of non-zero $\beta_{kk}$ and $\beta_{k-1,k}$. The program for computing the coefficients $\beta$ and plotting the following figure Fig. 5.1 could be found in Appendix B and run with Mathematica 5.1.

![Fig. 5.1 Distribution of $\beta_{jk}$](image)

In the computation process of $\beta_{jk}$, $j = 1, \ldots, k$, only the node set $\{x_j\}_{j=1}^n$ and the RPK function set $\{\phi_j(x)\}_{j=1}^n$ were applied, without any information of speech signal $f(x)$. So, these values can be applied to the processing of other speech frames and needn’t to be re-computed.

Next, we will compute the decomposition coefficients $\alpha_i$ which used the sample values of speech signal $f(x)$. Since $\beta_{ji}$ has localization property, the computation of $\alpha_i$ only requires $\beta_{ji}$, $\beta_{ji-1}$, and $f(x_j)$, $f(x_{j-1})$. We set node number or frame length $n=512$, and computed the decomposition coefficients $\alpha_i$ of phoneme “i” and “sh” of Pinyin, respectively. The following Fig.5.2(a)–(b) showed the original waveform and the corresponding decomposition coefficients $\alpha_i$ of phoneme “i”. The program for computing the coefficients $\alpha$ and plotting the following figure Fig. 5.2(a)–(b) could be found in Appendix A and run with Mathematica 5.1.

![Fig. 5.2 (a) The original waveform of “i”](image)

![Fig. 5.2 (b) The coefficients $\alpha_i$ of “i”](image)

From above Fig.5.2(a)–(b), we can find that the images of the original waveform and the coefficients $\alpha_i$ are very similar, and the image of $\alpha_i$ keeps the pitch features of the original waveform unchanged.

![Fig. 5.3 (a) The original waveform of “sh”](image)

![Fig. 5.3 (b) The coefficients $\alpha_i$ of “sh”](image)

From above Fig.5.3(a)–(b), we can find that the images of the original waveform and the coefficients $\alpha_i$ are very similar, and the image of $\alpha_i$ keeps the noise-like features of the original waveform unchanged. Besides, the reconstructed speech waveform by Th.2.5 was identical to the original speech signal.

V. CONCLUSIONS

Speech signal orthogonal decomposition approach in the RPK space $W_2^1[a,b]$ was discussed and the realization algorithm was presented in this paper. In the RPK space $W_2^1[a,b]$, the speech signal based on some discrete samples was mapped into a continuous function, which is an exact analytical representation of the original speech signal. Such analytical representation is different from the canonical representation of speech signals by the series of triangle functions. Besides, the inner product computation in $W_2^1[a,b]$ was transformed into the
evaluation of some functions at some fixed points, which significantly improved the computation process.

ACKNOWLEDGMENT

The authors wish to thank Professor Cui Minggen. This work was supported in part by a grant from NSFC (No.60572125).

REFERENCES