High Range Resolution Profile Automatic Target Recognition Using Sparse Representation

Zhou Nuo, Chen Wei*
School of Electronics and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China

Received 29 November 2009; accepted 8 April 2010

Abstract

Sparse representation is a new signal analysis method which is receiving increasing attention in recent years. In this article, a novel scheme solving high range resolution profile automatic target recognition for ground moving targets is proposed. The sparse representation theory is applied to analyzing the components of high range resolution profiles and sparse coefficients are used to describe their features. Numerous experiments with the target type number ranging from 2 to 6 have been implemented. Results show that the proposed scheme not only provides higher recognition preciseness in real time, but also achieves more robust performance as the target type number increases.

Keywords: automatic target recognition; high range resolution profile; sparse representation; feature extraction; dictionary generation

1. Introduction

High range resolution profile (HRRP) is a one-dimensional measurement of target radar reflectivity along radar line-of-sight (LOS). It provides detailed structure information of ground moving targets for identification. Furthermore, compared with two-dimensional synthetic aperture radar (SAR) images, it also has the advantage of easy acquisition, fast processing and less storage space taking. Over the recent years, high range resolution (HRR) mode has been used as a complementary mode to ground moving target indicator (GMTI) and SAR imagery for ground moving target recognition[1-3].

Feature analysis and extraction are significant parts of HRR automatic target recognition (ATR). Their missions are to discover and capture the primary signatures of HRRP as well as to delete plenty of unnecessary information. The obtained features benefit the improvement of recognition preciseness and running speed. Several methods such as HRRP amplitudes[3-5], FFT spectrum analysis[6], S transformation[7], wavelet coefficient fitting[8,9] and nonlinear kernel transformation[10-11] have been used extensively for HRR ATR applications. The method based on HRRP amplitudes is more popular due to its clear physical interpretation, convenient acquisition and high recognition preciseness.

Recently, as a new method of signal analysis, sparse representation has received a great deal of attention. In sparse representation theory, it is considered that signals are mixtures of a wide range of waveforms with different frequency properties, time-scale properties or time-frequency properties. Instead of conventional methods which decompose signals by a single type of waveforms, sparse representation analyzes signals over a redundant dictionary consisting of multiple types of waveforms. The waveforms in the dictionary are called atoms in sparse representation theory. Compared with the conventional methods, local structures within signals could be individually adapted by the atoms which take the most similar properties, and the signal decomposition is more concise. Consequently, sparse representation provides a more flexible and exact approach for signal analysis. To date, sparse representation has been successfully applied in signal analysis, images or videos compression and reconstruction, radar signal formation, and so on[12]. In addition, because of its excellent capability in signal analysis and signal representation, sparse representation has been attempted to be applied to experiments including face recognition[13-14], human movement recognition[15] and cancer diagnosis[16].
In this article, we explore the use of sparse representation into HRRP feature analysis and expect its performance in HRRP recognition.

2. Sparse Representation

2.1. Theory

Sparse representation is originally introduced by S. G. Mallat and Z. F. Zhang in 1993[17], and developed by D. L. Donoho[18-20] later. It is a method to analyze signals under a redundant dictionary.

\[ y = AX = [A_1 \quad A_2 \quad \cdots \quad A_N] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \sum_{i=1}^{N} A_i x_i \]  

(1)

where the measured signal \( y = [y_1 \quad y_2 \quad \cdots \quad y_d]^T \in \mathbb{R}^d \) is a d-dimensional vector, the redundant dictionary \( A = [A_1 \quad A_2 \quad \cdots \quad A_N] \in \mathbb{R}^{d \times N} \) is a matrix with size of \( d \times N, d < N \), the atom \( A_i \in A \) is a d-dimensional vector, \( i = 1,2,\ldots,N \), the coefficient vector \( X = [x_1 \quad x_2 \quad \cdots \quad x_N]^T \in \mathbb{R}^N \) is a N-dimensional vector obtained by \( y \) expanding over \( A \). As \( d < N \), equations \( y = AX \) are under determined. There are infinite \( X \) satisfying \( y = AX \). The freedom of selection of \( X \) facilitates the adaption to various practical applications.

In all the possible solutions, \( X \) with fewest nonzero coefficients, \( X_0 \), is the sparse representation[21].

\[ X_0 = \min_X \|X\|_0 \quad \text{subject to} \quad y = AX \]  

(2)

In fact, \( X_0 \) is greatly attractive in actual demands since it always means fewer sensors in a radar system or less prerequisites for a machine.

2.2. Dictionary

Dictionary design is crucial for signal component analysis in sparse representation. Generally, atoms in a dictionary should meet two basic requirements: containing a wide range of types of atoms to better adapt to different local structures of signals, and having sufficient number of atoms to make the dictionary redundant.

The available methods for dictionary establishment can be classified into two categories. One category is that the dictionary consists of a collection of \( g' \):

\[ g' = \frac{1}{\sqrt{s}} g(t-u) e^{-\pi \xi^2} \]  

(3)

where the collection of \( g' \) are transformations of a base function \( g(t) \)[17]. \( s, u \) and \( \xi \) are dilation, translation and frequency modulation parameters, respectively. \( g' \) produced by different \( s, u \) and \( \xi \) takes various frequency properties, time-scale properties or time-frequency properties. Moreover, because the freedom of the selection of \( s, u \) and \( \xi \) is unlimited, the number of \( g' \) is infinite. That is, the created dictionary can be as redundant as it is needed. However, in practical application, it is sometimes hard to decide the suitable number of \( g' \) for a dictionary. Furthermore, as all atoms in the dictionary derive from the same base function, the adaptability in signal component representation is relatively limited.

Another category for dictionary generation is a training sample-based method[19]. The dictionary is combined with training samples from several target types. These training samples are collected under a wide scope of real scenes or their simulations where the measured signals, or also called test samples related to training samples, probably take place. Due to the similar collecting conditions, training samples have similar characteristics with their associated test samples. The dictionary generated by these training samples is able to represent the test samples more efficiently and precisely. Additionally, since the real scenes where signals take place are known or predictable, the atom number in the dictionary can be decided easily.

2.3. Approaches

Finding the sparse representation by the minimum \( l_0 \) norm in Eq.(2) is complex and extremely hard[18]. Several algorithms have been designed to find suboptimal or approximate solutions instead. There are three approaches widely used in sparse representation over redundant dictionaries. They are matching pursuit (MP), basis pursuit (BP) and least absolute shrinkage and selection operation (LASSO)[17-23].

MP[17-21], proposed by S. G. Mallat, is a greedy pursuit algorithm. It finds the sparse solution by an iterative matching process. Compared with BP and LASSO, MP has the advantage of easy implementing and short running time. However, since the solutions produced by MP are just local optimal, they are sometimes not precise enough.

D. L. Donoho[19] simplified problems minimum \( l_0 \) norm into problems minimum \( l_1 \) norm:

\[ X_1 = \min_X \|X\|_1 \quad \text{subject to} \quad y = AX \]  

(4)

Furthermore, he gave evidence that \( X_1 \) in Eq.(4) is equal to \( X_0 \) in Eq.(2) when the sparse representation of \( y \) exists. Recently, \( l_1 \) minimization has been considered as an effective and general method for underdetermined equations. Based on the idea in Eq.(4), D. L. Donoho proposed BP[18-21] whose solutions are much sparser and more precise than the solutions MP produces. Nevertheless, computation of BP is still too complicated to be used in applications, especially in the situation that the dimension of \( y \) is large.

LASSO is also a method dealing with sparse problem based on Eq.(4). But it tackles \( l_1 \) minimization
problem by a regression process subject to constraint such as \( \| X \|_1 \leq p \).

\[
\min_X \| y - AX \|_2^2 \quad \text{subject to} \quad \| X \|_1 \leq p \quad (5)
\]

In the regress process, sparse solution is obtained by optimizing the parameter \( p \) value and the coefficients at each chosen value of \( p \). The available literature has demonstrated that solutions obtained by LASSO are sensible approximation to \( y \), or even equal to \( X_0 \), and the performance of LASSO is fast, efficient and stable no matter the scale of dimension of \( y \) is large or small\[^{[19-23]}\]. For these merits, LASSO is selected to solve the sparse representation in this article.

3. HRRP ATR

3.1. Data description

Due to a lack of real HRRP database obtained from moving ground vehicles, the raw HRRP used in our experiments are formed from a SAR image database collected by an X-band HH-polarization SAR sensor\[^{[24]}\]. This SAR image database contains several classes of stationary ground vehicles which are no more than 3 m in height, 4 m in width and 8 m in length. Each SAR image in the database is about 128 pixel×128 pixel with 1 foot by 1 foot resolution. Each vehicle corresponds to two groups of SAR images. These two groups of images are obtained under comparable conditions except that one group of SAR images are collected at 15° depression angles over the 360° coverage in aspect while 17° depression angles for the other. At the same depression angle for each vehicle, the interval between two neighboring SAR images is less than 3° aspect angles, and the number of SAR images is more than 120. According to the instruction of the database\[^{[24]}\], SAR images from 17° depression angles are used as training samples in a recognition experiment while SAR images from 15° depression angles are used as test samples. The detailed transformation process about the raw HRRP generation from SAR images is explicitly described in Ref. [25].

Due to the formed raw HRRP containing amount of background information which mainly expands in several first and last range bins and is unnecessary for an identification task, only the information in the central 60 range bins of raw HRRP is kept as the HRRP data used in the following experiment.

3.2. Dictionary generation

Considering the superiorities in decomposition accuracy and predictability of the dictionary size, dictionary generation based on training samples is attempted to be used in this experimental study. In the remainder of this section, its feasibility will be discussed from two aspects: the variability of atom types and the dictionary redundancy.

As radar bandwidth increases over the years, radar resolution in range has been greatly improved. Radar is no longer a blob detector, but a sophisticated system with the abilities to acquire detailed information for target type recognition\[^{[26]}\]. Point scattering models are universally adopted to describe target scattering properties which are produced by wideband radars\[^{[27]}\].

Assume there are \( L \) scattering centers for a given target. Distance of each scattering center to the radar receiver along radar LOS is \( r_i, i=1,2,\ldots,L \). The corresponding intensity and initial phase are \( \sigma_i \) and \( \psi_i \) individually. \( L \) scattering centers are divided into \( N \) contiguous range bins according to \( r_i \).

Assume further that there are \( L_n \) scattering centers in \( n \)th range bin, \( n=1,2,\ldots,N, L=L_1+L_2+\cdots+L_N \). Their distances to the radar receiver, intensities and initial phases are \( r_{ni}, \sigma_i \) and \( \psi_i \) in sequence, \( i=1,2,\ldots,L_n \). HRRP in \( n \)th, \( h_n \), is the sum of echoes from \( L_n \) scattering centers:

\[
h_n = \sum_{i=1}^{L_n} \sigma_i \exp\left[-j\left(\frac{4\pi}{\lambda} r_{ni} - \psi_i \right)\right] \quad (6)
\]

where \( \lambda \) is radar wavelength.

It is seen that there is a close relationship between point scattering models and HRRP. However, point scattering models are strongly dependent on target types and the aspect and depression angles where targets stay ahead. For targets with different types, since the sizes, shapes, materials and scattering properties are distinct, their scattering properties are highly different from each other\[^{[28]}\]. Various point scattering models are required to represent them individually. For a given target, a relatively large rotation in aspect or depression, about 10°\[^{[29]}\], also results in the changes of target scattering properties. Similarly, a series of point scattering models thus are also needed to describe the scattering properties in continuous angle ranges. In fact, although the point scattering model is constant for a given target, a small rotation in aspect or depression equally produces the changes of scattering center distribution in range bins. It is called migration through resolution cells (MTRC). Assume the range resolution is \( R \), the maximum size of the given target in height and the azimuth plane is no more than \( Z \). If the target rotation \( \delta \phi \) satisfies

\[
\delta \phi = \frac{R}{Z} \times \frac{360}{2\pi} \quad (7)
\]

MTRC can be avoided\[^{[29-30]}\]. Therefore, it is observed that all of the changes caused by target types, large rotation in aspect or depression and MTRC can cause the changes of \( L_n, r_{ni}, \sigma_i, \psi_i, L_n, r_{ni}, \sigma_i, \psi_i \), and finally result in the variability of HRRP local structures.

According to the discussion above, it is known that if a dictionary is combined with training samples from multiple target types, the diversity of atom types in the dictionary can be satisfied easily. Furthermore, since the scale of experimental HRRP is 60 and the number
of training samples for each target is more than 120 as described in Section 3.1, the dictionary generated by the concatenation of training samples of various types of targets is sufficiently redundant. Meanwhile, the dictionary size can be correctly predicted by summing up the number of training samples.

In addition, according to Section 3.1, it can also be seen that the maximum size of targets in height and the azimuth plane is no more than 8 m, and thus the rotation $\delta \phi$ avoiding MTRC should be no more than 2.18$\degree$. That is, if a rotation in aspect or depression is less than 2.18$\degree$, $L$, $r$, $\sigma$, $\psi$, and $L_{\text{at}}, r_{\text{at}}, \sigma_{\text{at}}, \psi_{\text{at}}$ are relatively stable. The components of HRRP can be viewed as constant. But if the rotation is larger than 2.18$\degree$, $L$, $r$, $\sigma$, $\psi$, and $L_{\text{at}}, r_{\text{at}}, \sigma_{\text{at}}, \psi_{\text{at}}$ will vary. For the training or test samples of a given target collected at a fixed depression angle, since the angle increment in aspect between contiguous samples is between 1$\degree$ and 3$\degree$, the local structures of samples vary following the change of aspect angles over the 360$\degree$ aspect coverage. However, the training samples and the test samples gained at the same aspect angle consist of similar local structures in that the difference between them is just 2$\degree$ in depression. Therefore, atoms combined with training samples of multiple target types not only meet the two basic acquirements of dictionary generation, but also take superiorities to adapt the local structures of individual test samples more flexibly and exactly.

3.3. Feature sparsity

According to Section 3.2, it is observed that training samples are highly similar to local structures with their corresponding test samples. If a test sample is given from one of the target types, the ideal signal component analysis in theory should approximately locate on the atoms related to this target type, especially the ones obtained at the same aspect angle or around it. Therefore, the decomposition result is sparse. It will be utilized as features in recognition experiments.

3.4. Recognition

Refer to the definition in Eq.(1) and assume that there are $c$ types of targets $T_1, T_2, \ldots, T_c$. The dictionary is $A = [B_1, B_2, \ldots, B_c]$, $B_i = [A'_{i,1}, A'_{i,2}, A'_{i,m_i}]$, $i = 1, 2, \ldots, c$, $m_i$ is the number of training samples associated with $i$th target type. The resolved sparse coefficient vector is $X'_i = [X'_{i,1}, X'_{i,2}, \ldots, X'_{i,m_i}]^T \in \mathbb{R}^{M}$, $X'_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,m_i}]^T$, $M = m_1 + m_2 + \ldots + m_c$. $y$ is a test sample from $k$th target type, $k \in \{1, c\}$.

According to the analysis in Section 3.2 and Section 3.3, the nonzero elements should mainly locate on $X'_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,m_i}]^T$, and few locate on $X'_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,m_i}]^T$, $i \in \{1, c\}$, $i \neq k$. Therefore, there are

$$y = AX = \sum_{i=1}^{c} B_i X'_i = B_i X'_i = \sum_{j=1}^{m_i} A'_{i,j} x_{i,j} \tag{8}$$

$$\sum_{j=1}^{m_i} A'_{i,j} x_{i,j} \ll y_i, \quad i \neq k \tag{9}$$

The recognition process can be divided into three steps:

1) Sum up the arithmetic product of $B_i$ and $X'_i$

$$\text{sum}_i = B_i X'_i = \sum_{j=1}^{m_i} A'_{i,j} x_{i,j}, \quad i \in \{1, c\} \tag{10}$$

2) Find the minimum distance $d$ between $y$ and $\text{sum}_i$

$$\min d = \min \|y - \text{sum}_i\|, \quad i \in \{1, c\} \tag{11}$$

3) Recognize the test sample to the type of targets which the minimum $d$ corresponding to.

4. Results and Discussion

4.1. Sparse decomposition

Numerous test samples have been experiment to validate deduction in Section 3.2 about the feasibility of dictionary generation based on training samples and the assumption in Section 3.3 that the decomposition results are sparse. The results show that the obtained decomposition results are sparse indeed and primary sparse coefficients assemble within the range of the corresponded target types.

Ranges at [1, 135], [136, 283], [284, 421], [422, 658], [659, 795], and [796, 1032] are related to atoms belonging to 1st, 2nd, 3rd, 4th, 5th, and 6th target types, respectively. As examples, Fig.1 shows LASSO decomposition results of test samples which belong to 1st, 2nd and 3rd type of targets around 0$\degree$, 90$\degree$, 180$\degree$.
and 270° aspect angles. It is seen that there are few sparse coefficients appearing in Fig.1. Significant sparse coefficients concentrate on the atoms related to the type of test samples while sparse coefficients of atoms corresponding to other target types are relatively much smaller, or even equal to zero.

4.2. Recognition comparison

Two approaches are considered in recognition comparison. One is the currently widely used template based recognition (TBR) using HRRP magnitudes [3-5]. The other is the sparse representation based recognition (SRBR) employing sparse coefficients proposed in this article. Both approaches are implemented in 5 sorts of recognition experiments. The target type number in these 5 sorts of recognition experiments ranges from 2 to 6 sequentially.

In order to make the results of SRBR and TBR comparable, test samples used in both approaches are produced as described in Section 3.1; templates employed in TBR and the dictionary for sparse representation are combined by training samples formed in Section 3.1 as well. Once the number of target types in a sort of experiments is fixed, numerous recognition experiments are implemented, and the target types in each experiment are selected randomly from the database; the probability of correct classification (PCC) of each sort of experiments is measured by the statistical mean of the results obtained in the corresponding experiments.

Fig.2 depicts the performance of TBR and SRBR during 2 to 6 recognition experiments. It is seen that, when the target type number is fixed, SRBR keeps higher PCC than TBR does. The improvement in 2 to 6 recognition experiments is 2.3718%, 3.6828%, 4.6801%, 5.3684% and 5.9817%, respectively. It can also be observed that as the number of target types increases, both of PCC in SRBR and TBR drop. The rate of PCC declining in SRBR is about 14.8096%, from 93.5082% to 78.6986% during 2 to 6 recognition experiments, while that in TBR is more than 18%, from 91.1364% to 72.7169%.

4.3. Discussion

From Fig.2, it is known that both of TBR and SRBR provide satisfied recognition results. However, SRBR is superior to TBR throughout 2 to 6 recognition experiments. The improvement in SRBR primarily benefits from the insight of sparse representation into signal decomposition. For example, Fig.3 contains a template matching result that a test sample of 2nd target type at 14° aspect angle. The maximum correlation coefficient appears at the range between 659th template and 795th template which corresponds to 5th target type. Hence, the test sample is judged as the 5th target type accord-

Fig.1 Sparse coefficients obtained by sparse analysis.

Fig.2 Performance comparison of SRBR and TBR in 2 to 6 recognition experiments.
ingly. Obviously, this recognition result is wrong. The main reason of the mistake is that the method based on template matching emphasizes on the rough similarity in general between two samples, but ignores the difference in local structures in detail. The judgment lacks accuracy.

In contrast, sparse representation, like a microscope, provides an insight into the internal local structure composition. During the signal approximation process, sparse analysis keeps adding atoms adapted to signal local structures, evaluating their significance and removing atoms with less similarity. Fig.4 shows the approximation process of the test sample of 2nd target type at 14° aspect angle in sparse representation. Similarly with template matching, at the beginning, sparse analysis firstly fixes the atom related to 5th target type as the original and rough analysis result, as shown in Fig.4(a). During the successive process, atoms are added or deleted according to their significance at each step of approximation, shown as Figs.4(b)-(g). The final result is a perfect approximation of the test sample with atoms as few as possible (see Fig.4(h)).

5. Conclusions

(1) For the experimental data with less dimension but large number of training samples, dictionary generation based on training samples is a simple but efficient means.

(2) Sparse representation provides a more accurate way for HRRP analysis. The decomposition results are sparse and exact.

(3) Compared with TBR, SRBR produces an obvious improvement in recognition precision when the target types range from 2 to 6. Furthermore, performance obtained by SRBR is more robust and stable.

(4) Compared to TBR, the running time of SRBR is little longer than TBR. But the difference is within several seconds, so it is still able to meet the demands in real applications.

References


[20] Donoho D L. For most large underdetermined systems of linear equations the minimal 1-norm solution is also the sparsest solution. Communications on Pure and Applied Mathematics 2006; 59(6): 797-829.


Biographies:

Zhou Nuo Born in 1981, she is currently a Ph.D. candidate in Beijing University of Aeronautics and Astronautics. Her research interests include signal processing and automatic target recognition.
E-mail: joycesmiling@163.com

Chen Wei Born in 1963. He is an associate professor in Beijing University of Aeronautics and Astronautics. His research interests include signal and information processing, information fusion, digital image processing and target recognition.
E-mail: chenwei@buaa.edu.cn