High Resolution Range Profile Sequence Recognition
Based on ARTRBM

Yi-fan ZHANG*, Xuan PENG, Jia-qi YE, Dan QIN,
Xun-zhang GAO and Xiang LI

College of Electronic Science, National University of Defense Technology, Changsha, China

*Corresponding author

Keywords: HRRP, RATR, RTRBM, Attention mechanism.

Abstract. In this paper, a novel stochastic neural network model named Attention based Recurrent
Temporal Restricted Boltzmann Machine (ARTRBM) is proposed for the poor performance of the
traditional HRRP recognition methods on high dimensional sequential data and noisy data. RTRBM
is efficient to model high dimensional HRRP sequence because it can extract the information of
temporal and spatial correlation between adjacent HRRPs. Attention mechanism is used in sequential
data recognition tasks, making the model pay more attention to the major features of recognition.
Therefore, the combination of RTRBM and attention mechanism makes our model extract more
internal related features effectively and choose the important parts of the extracted features. Experiment results show that our proposed model outperforms other traditional methods, indicating
that ARTRBM extracts, selects and utilizes the correlation information between adjacent HRRPs
effectively, and is suitable for high dimensional data or noise corrupted data.

Introduction

The HRRP recognition has been studied for decades in the field of Radar Automatic Target
Recognition (RATR) because it contains important structural information such as the target size and
the distribution of scattering points [1-2]. The existing reported methods can be summarized as
extracting features of HRRPs after dividing the full target radar aspect angles into several frames and
performing the target detection to select the region of interest in an HRRP. However, most traditional
recognition methods utilize the single HRRP rather than HRRP sequences, ignoring the temporal and
spatial correlation within the samples.

Noting strong relativity is contained between the adjacent HRRP, sequential HRRP is of potential
usage for recognition. HMM is often utilized for sequential problems, which utilizes the sequence
information of HRRP, and considers the structure information inside the target [3]. However, the
model is not efficient to deal with high dimensional sequential data because of the high computational
complexity. Recently, the Recurrent Temporal Restricted Boltzmann Machine (RTRBM) was
proposed as a generative model for high dimensional sequences [4-6]. RTRBM model introduces the
correlation matrix between the hidden layers of adjacent RBMs to tack the correlation inside the data
into consideration [4]. The model has achieved great success in extracting internal correlations
between adjacent HRRPs and capturing spatial and temporal patterns in high dimensional sequential
data. However, considering the contribution of each feature vector to the recognition is different,
which has been ignored in the discriminative model of RTRBM [7], it is essential for the recognition
method to gain the ability to pay more attention to the important parts of features.

To further improve the performance of RTRBM, a novel method that combines the RTRBM model
with the attention mechanism [8] for sequential radar HRRP recognition is proposed in this paper.
The attention mechanism was first proposed in the field of visual image in [9], and shown good
performance on a range of tasks including machine translation, machine comprehension and Relation
classification [10-12]. The combination of RTRBM and attention mechanism makes the model focus
its attention on specific features which are important to the classification task. More specifically, this
model encodes the HRRPs sequence through the RTRBM model, and calculates the weight
coefficient for each hidden unit according to their contributions to the recognition performance, then
the features are utilized to construct the attention layer for the recognition task. This combination brings performance improvements on high recognition accuracy achievement and strong robustness to noise. To demonstrate the effectiveness of the proposed model, two experiments are performed, which utilize the HRRP data converted from the SAR data of MSTAR. Experiment results indicate the superior performance of the proposed model against HMM, Class RBM and Principle Component Analysis (PCA), and show its strong robustness to noise.

**Recurrent Temporal Restricted Boltzmann Machine**

The Recurrent Temporal Restricted Boltzmann Machine (RTRBM) is a generative model for modeling high-dimensional sequences, which constructed by rolling multiple RBMs over time step. The RBM is an undirected graphical model that uses a layer of hidden variables \( h = [h_1, h_2, \ldots, h_m] \) to model a joint distribution over the visible variables \( v = [v_1, v_2, \ldots, v_n] \) [13]. The graphical model for the RTRBM is illustrated in Figure 1.

![Figure 1. Graphical structure of the RTRBM.](image)

In Figure 2, RTRBM gives five parameters \( \{W, W_{hh}, \hat{h}^{(t)}, b, c\} \). Here \( W \) is the weight matrix between visible and hidden layer of the RBM at each time instant. \( W_{hh} \) stands for the directed weights which connect the hidden layer at time instant \( t-1 \) and \( t \), and \( \hat{h}^{(t)} \) is a vector of initial mean-field values of the hidden units, the motivation for the choice of \( \hat{h}^{(t)} \) is that using the RBM associated with time instant \( t \), we have that \( \mathbb{E}(h(l) | v(t)) = \hat{h}^{(t)} \); i.e., it is the expected value of the hidden units vector and the major difference compared to the RBM. Additionally, \( b^{(t)} \) and \( c^{(t)} \) is the biases of visible and hidden layers respectively, which depend on \( \hat{h}^{(t-1)} \):

\[
\begin{align*}
    b^{(t)} &= W_{hh} \hat{h}^{(t-1)} + b \\
    c^{(t)} &= W_{hh} \hat{h}^{(t-1)} + c
\end{align*}
\]

where \( \hat{h}^{(t)} \) is the mean-field value of \( h^{(t)} \), which can be represented detailed as:

\[
\hat{h}^{(t)} = \sigma(W v^{(t)} + c^{(t)}) = \begin{cases} 
    \sigma(W v^{(t)} + c_0) & \text{if } t = 1; \\
    \sigma(W v^{(t)} + W_{hh} \hat{h}^{(t-1)} + c) & \text{if } t > 1.
\end{cases}
\]

Given hidden inputs \( \hat{h}^{(t-1)}(t>1) \), the conditional distributions are factorized and takes the form:

\[
\begin{align*}
    p(h_{t,j} = 1 | v, \hat{h}^{(t-1)}) &= \sigma(\sum_i \omega_{ji} v_{t,i} + b_j + \sum_i W_{hh,j,m} \hat{h}^{(t-1,m)}) \\
    p(v_{t,i} = 1 | h_{t}, \hat{h}^{(t-1)}) &= \sigma(\sum_j \omega_{ji} h_{t,j} + c_i)
\end{align*}
\]

Therefore, the joint probability distribution of the visible and hidden units of the RTRBM with length \( T \) takes the form as [6]:

\[
p(v^{(1:T)}, h^{(1:T)}; \hat{h}^{(1:T-1)}) = \prod_{t=1}^{T} p(v^{(t)}, h^{(t)}; \hat{h}^{(t-1)}) = \prod_{t=1}^{T} \frac{\exp[-E(v^{(t)}, \hat{h}^{(t)}; \hat{h}^{(t-1)})]}{Z_{\hat{h}^{(t-1)}}}
\]
where $Z_{\hat{r}(t-1)}$ denotes the normalization factors for the RBM at $T=t$, and $E(v^{(t)}, h^{(t)}; \hat{r}^{(t-1)})$ is the energy function at the time step $t$, which is defined as:

$$E(v^{(t)}, h^{(t)}; \hat{r}^{(t-1)}) = -h^{(t)T} W v - c^{(t)T} v - b^{(t)T} h^{(t)}$$

(5)

It has been demonstrated in [15-16] that how to learn the model using CD approximation.

**Attention Based Recurrent Temporal Restricted Boltzmann Machine**

Based on the original RTRBM, the newly proposed model brings the idea of attention mechanism, which named Attention based RTRBM, and the graphical structure of the proposed model is demonstrated in Figure 2. In the proposed model, RTRBM is utilized to extract features from the input data and store the extracted features in hidden vector. A new hidden layer $s$ is introduced to RTRBM by weighted sum all hidden layers to measure the role of each hidden vector in recognition tasks, and then the new hidden layer is used for classification.

In Figure 3, $\alpha_t$ stands for the weight coefficient for the hidden layer at time step $t$, the layer $s$ is determined by the hidden layer of each time step, and $y$ is a vector representing the class label, in which all values are set to 0 except at the position corresponding to a label $y$, which is set to 1.

As we can see from the figure above, layer $s$ is obtained by the weighted summation of the hidden layers of each time instant, which can be expressed as:

$$s_t = \sum_{j=1}^{T} \alpha_j h_{j}$$

(6)

where the weight coefficient $\alpha_j$ can be defined as:

$$\alpha_j = \frac{\exp(e_j)}{\sum_{j=1}^{T} \exp(e_j)}$$

(7)

where $e_j$ corresponds to the hidden layer energy at time instant $j$ given below:

$$e_j = V a \cdot \tanh(W a \cdot h_j)$$

(8)

The weight coefficient $\alpha_j$ represents the effect of the hidden layer feature $h_j$ in recognition. The attention mechanism [10, 16] is also determined by the parameter $\alpha_j$. By training the parameters $Va$ and $Wa$, the model can assign the hidden layer $h_j$ with different weights at different moment, which makes the model more focused on the parts who play a major role in the recognition tasks.

We define the training objective of the model as cross entropy, which can be shown as follows:
To learn the model, we need to obtain the partial derivatives of $f_{\text{cross}}(\theta, D_{\text{train}})$ with respect to the parameters. The gradients can be given as:

$$\frac{\partial}{\partial \theta} f_{\text{cross}}(\theta, D_{\text{train}}) = \frac{1}{|D_{\text{train}}|} \sum_{n=1}^{|D_{\text{train}}|} \frac{\partial F(y^n|s^n)}{\partial \theta}$$

The process of our proposed model can be shown as follows:

The pseudo code of model parameter update for the proposed model is summarized in Algorithm 1, which is shown as follows:

**Algorithm 1** Pseudo code for the learning steps of Attention based RTRBM model

**Input:** training pair: \{v_train; y_train\}, hidden layer size: dim_h; learning rate: $\lambda_1, \lambda_2$; momentum: $\beta$; and weightcost: $\eta_1, \eta_2$.

**Output:** label vector $y$.

**# Section 1: Extract features using RTRBM**

1. Calculate $\hat{h}_{t}^{(t)}$ according to Equation (2);
2. Calculate $P(h_{t,j} = 1|v, \hat{h}_{t}\{(t-1)\})$ and $P(v_{t,i} = 1|h_t, \hat{h}_{t}\{t-1\})$ respectively according to Equation (3);
3. Calculate the L2 reconstruction error: $\text{Loss} \leftarrow \|v_t - v_{t,k}\|_2$;
4. Update parameters of this section: $\Theta \leftarrow \Theta - \Delta \Theta$, $\Delta \Theta \leftarrow \beta \Delta \Theta - \lambda_1 (\nabla \Theta - \eta_1 \Theta)$;
5. Repeat step (1) to (4) for 1000 epochs, and save the trained $\Theta$ for test phase.

**# Section 2: Classification with Attention mechanism**

1. Calculate $\alpha_j, j \in \{1,2,\cdots, T\}$ according to Equation (7);
2. Calculate $s_{t,i} \in \{1,2,\cdots, \text{dim}_h\}$ according to Equation (6);
3. Calculate the cross entropy according to Equation (9);
4. Update parameters of this section: $\Theta \leftarrow \Theta - \lambda_2 (\nabla \Theta - \eta_2 \Theta)$;
5. Repeat step (1) to (4) for 1000 epochs, and save the trained $\Theta$ for test phase.

**Experiments**

**The Dataset**

The sequential HRRP dataset we used in the experiments were constructed by SAR images published by MSTAR. To make it suitable for our model, we first transformed the two-dimensional SAR into one-dimensional HRRP vector to train our proposed model. We chose three of the most similar targets in MSTAR for the experiment. All the targets cover 0 to 360 degrees of aspect angles, and their distance and azimuth resolutions are 0.3 meters [17]. In the dataset, each target is obtained under the depression angle of 15 and 17 degrees, the HRRPs of 17 degree of depression angle were used as the training data, while the HRRPs of 15 degree were used as the test data. The size of the training and test dataset is briefly illustrated in Table 1.
Table 1. Training and testing set of HRRPs for three targets.

<table>
<thead>
<tr>
<th>Number</th>
<th>Training Set</th>
<th>Size</th>
<th>Testing Set</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BMP2 (Sn_C9563)</td>
<td>2330</td>
<td>BMP2 (Sn_C9566)</td>
<td>1960</td>
</tr>
<tr>
<td></td>
<td>BMP2 (Sn_C21)</td>
<td>1960</td>
<td>BMP2 (Sn_C9563)</td>
<td>1950</td>
</tr>
<tr>
<td>2</td>
<td>T72 (Sn_132)</td>
<td>2320</td>
<td>T72 (Sn_812)</td>
<td>1950</td>
</tr>
<tr>
<td></td>
<td>T72 (Sn_S7)</td>
<td>1910</td>
<td>T72 (Sn_132)</td>
<td>1960</td>
</tr>
<tr>
<td>3</td>
<td>BTR70 (Sn_C71)</td>
<td>2330</td>
<td>BTR70 (Sn_C71)</td>
<td>1960</td>
</tr>
<tr>
<td>Sum</td>
<td>Training Set</td>
<td>6980</td>
<td>Testing Set</td>
<td>13650</td>
</tr>
</tbody>
</table>

As is shown in Table 1, the training set and testing set contain 6980 HRRPs and 13,650 HRRPs respectively. Noting that the target BMP2 contains only one type of training set while three different types of the testing set, and so is T72. And target BTR70 uses type Sn_C71 to train the model and tests it with corresponding type.

Experiments

**Experiment 1: Investigating the Influence of Hidden Layer Size on Recognition Performance.**

In this sub-section, Class RBM (CRBM) with different hidden layer size (size = 16, 32, 64, 128, 256, 384, 512) were trained as comparisons to the proposed method. We carry out the contrast experiments with two different data input methods, constructing an average HRRP with 15 adjacent HRRPs and connecting 15 HRRPs end to end. The recognition performance of each model is shown in Figure 5, where the test accuracy is computed by averaging the test results of three targets.

Figure 4 indicates the superior recognition performance of ARTRBM against other two models. And our proposed model gets optimal recognition accuracy on each size of the hidden layer, showing the strong ability to deal with high dimensional sequences. We design another baseline which uses PCA to reducing the dimension of input data to 15, and the classifier is Support Vector Machine (SVM). We repeat the PCA+SVM five times and the average test accuracy is 91.22%. And the test performance of HMM model is lower than 80% when the sequence length is 15, which is provided by [3]. The explanation for the best performance of our model may be that the correlation matrix between the adjacent hidden layers helps RTRBM to extract more discriminatory features and the weight coefficients make attention mechanism select more separable features, meaning that ARTRBM is more suitable for radar HRRP sequence recognition task.

To look insight into the performance of three methods on different targets, we list the confusion matrix for the three targets in Table 2. The number of hidden units for all the methods is 384.
As is shown in Table 3, our model outperforms the other two, but the misclassification of BMP2 lowers the average accuracy. One possible reason may be summarized as we train the models only on BMP2 (Sn_C9563), but test models on three types, unfortunately, these three types of BMP2 has a low similarity. However, our proposed model still achieves higher accuracy than two contrast models on BMP2, indicating that ARTRBM is a better choice when there is a great difference between the training and testing dataset.

**Experiment 2: Investigating the Influence of SNR on Recognition Performance.** For applications in real scenarios, white Gaussian noise of different Signal-to-Noise (SNR) increases from -10dB to 30dB were added to the testing data to investigate the robustness of the proposed model. Here, we trained the ARTRBM using the HRRP sequence with T=15, 384 hidden units. We choose the Class RBM with 384 hidden units as the contrast experiment, and the data input method is connecting 15 HRRPs end to end. Another contrast experiment uses PCA to reduce the dimension to 15 of input data, and the classifier is Support Vector Machine (SVM).

![Figure 5. Recognition performance on models tested with different SNR.](image)

As is shown in figure 5, our proposed model achieves better performance than other two models at all SNR levels, and it gets more than 10% advantage over other two models at -10dB. In addition, the testing accuracy keeps stable at a high level which is near the average accuracy in Table2 (0.9488) when the SNR is higher than 15dB, which inflects that our proposed model has a certain anti-noise ability. The accuracy of proposed model decreases to about 65% with the decrease of SNR, however, this number is less than 55% for CRBM. This result shows the strong anti-noise power of ARTRBM. Considering the working environment of radar system, the training samples are often corrupted by noise, the model we proposed is a better choice to perform the HRRP sequence recognition task.

**Conclusions**

In this paper, Attention based RTRBM is proposed for target recognition based on HRRP sequence. Compared with the reported methods, the proposed method has some compelling advantages. Firstly, it introduces the correlation matrix between the hidden layers to extract more correlation information, which makes the extracted features contain the previous and current information. Then, it is efficient to deal with high dimensional sequential data, which performs better than Class RBM using two different data input methods. Additionally, it can be effective to choose and utilize the important parts of the extracted features, which outperforms the RTRBM+SLP model using different input features.
Also, the proposed model performs well in case of strong noise, which indicates a strong robustness to the noise. In the near future, to better solve the problem of sequential HRRP recognition, we plan to combine other deeper models with attention mechanism as a classifier for RTRBM or other sequential feature extraction models. And in order to make the model more applicable to the real scene, we will operate related experiments with train phase and test phase at different angular sampling rate. Besides, we attempt to develop a model which can set the length of the attention mechanism adaptively. In this case, the number of $T$ will not need to be set by experience, which may achieve better performances.

Acknowledgments
This work is funded by National Natural Science Foundation of China under contract No. 6157450.

References
