Application of Wavelet Support Vector Regression on SAR data Denoising

Yi Lin\textsuperscript{1,2}, Shaoming Zhang\textsuperscript{1}, Jianqing Cai\textsuperscript{2}, Nico Sneeuw\textsuperscript{2}

\textsuperscript{1}Dept. of Surveying and Geo-Informatics, Tongji University
\textsuperscript{2}Institute of Geodesy, Universität Stuttgart

Session 6. Theoretical Geodesy
Geodetic Week 2009, Karlsruhe
Thursday, 24 September 2009

Outline

1. Introduction
2. Support Vector Regression and Wavelet Kernel Function
3. SAR data Denoising based on WSVR
4. The experimental results and conclusion
1. Introduction

SVM & SVR

- **Support Vector Machine (SVM)**
  - A kind of machine learning theory based on mathematical statistics (Vapnik, 1996)
  - Advantages:
    - Solving multiclass problem
    - Solving high-dimensional problem
    - Improving generalization performance
    - Solving the machine learning problem of small sample

- **Support Vector Regression (SVR)**
  - SVM method is introduced into SAR image filtering, by using it's restraining action for signal noise.
  - A kind of nonlinear regression
  - SVR is used for classification and regression
  - SVM is used for classification and regression
  - SVM & SVR

2. Support Vector Regression and Wavelet Kernel Function

Mathematical representations of Support Vector Regression

- Given training data set

\[
\begin{align*}
N & \in \mathbb{N}, \quad \epsilon & \in \mathbb{R}, \quad \kappa & \in \mathbb{R}, \\
\{ (\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_ N) \} & \subseteq \mathcal{F}
\end{align*}
\]

- To solve this regression problem, we minimize the objective function:

\[
\begin{align*}
\min_{\mathbf{w}, \mathbf{b}} & \quad \frac{1}{2} \mathbf{w}^T \mathbf{w} \\
\text{subject to} & \quad \mathbf{y} = f(x) + \mathbf{w}^T \phi(x) + \mathbf{b}
\end{align*}
\]

where \( \mathbf{w} \in \mathbb{R}^m \) is the weight vector, \( \mathbf{b} \in \mathbb{R} \) is the intercept of the regression function, \( \phi(x) \) is the feature mapping function of \( x \)

- Considering the regression function

\[
f(x) = \sum_{i=1}^{N} \alpha_i \kappa(x, x_i) + b
\]

\( \{ (\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_ N) \} = \mathcal{F} \)

- Wavelet Kernel Function

- Insensitive interval \([ -\epsilon, \epsilon ] \)
- \( \epsilon \) is insensitive threshold
- When controlling the training error in an insensitive interval, we think that there is no loss of training error.
- \( \epsilon \) are slack variables
- The second part of above formula is training loss.
- \( C \) is a positive constant, which is used to keep the balance of smoothing extent between training loss and fitting result.
- \( \epsilon \) and \( \xi \) are insensitive intervals

To solve this regression problem, we minimize the objective function:

\[
\begin{align*}
N & \in \mathbb{N}, \quad \epsilon & \in \mathbb{R}, \quad \kappa & \in \mathbb{R}, \\
\{ (\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_ N) \} & \subseteq \mathcal{F}
\end{align*}
\]

\[
\begin{align*}
\min_{\mathbf{w}, \mathbf{b}} & \quad \frac{1}{2} \mathbf{w}^T \mathbf{w} \\
\text{subject to} & \quad \mathbf{y} = f(x) + \mathbf{w}^T \phi(x) + \mathbf{b}
\end{align*}
\]

where \( \mathbf{w} \in \mathbb{R}^m \) is the weight vector, \( \mathbf{b} \in \mathbb{R} \) is the intercept of the regression function, \( \phi(x) \) is the feature mapping function of \( x \)

- Considering the regression function

\[
f(x) = \sum_{i=1}^{N} \alpha_i \kappa(x, x_i) + b
\]

\( \{ (\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_ N) \} = \mathcal{F} \)

- Wavelet Kernel Function

- Insensitive interval \([ -\epsilon, \epsilon ] \)
- \( \epsilon \) is insensitive threshold
- When controlling the training error in an insensitive interval, we think that there is no loss of training error.
- \( \epsilon \) are slack variables
- The second part of above formula is training loss.
- \( C \) is a positive constant, which is used to keep the balance of smoothing extent between training loss and fitting result.
- \( \epsilon \) and \( \xi \) are insensitive intervals

To solve this regression problem, we minimize the objective function:

\[
\begin{align*}
N & \in \mathbb{N}, \quad \epsilon & \in \mathbb{R}, \quad \kappa & \in \mathbb{R}, \\
\{ (\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_ N) \} & \subseteq \mathcal{F}
\end{align*}
\]

\[
\begin{align*}
\min_{\mathbf{w}, \mathbf{b}} & \quad \frac{1}{2} \mathbf{w}^T \mathbf{w} \\
\text{subject to} & \quad \mathbf{y} = f(x) + \mathbf{w}^T \phi(x) + \mathbf{b}
\end{align*}
\]

where \( \mathbf{w} \in \mathbb{R}^m \) is the weight vector, \( \mathbf{b} \in \mathbb{R} \) is the intercept of the regression function, \( \phi(x) \) is the feature mapping function of \( x \)

- Considering the regression function

\[
f(x) = \sum_{i=1}^{N} \alpha_i \kappa(x, x_i) + b
\]

\( \{ (\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_ N) \} = \mathcal{F} \)
Mathematical representations of Support Vector Regression (2)

• From this we can establish corresponding Lagrange functional and optimize it. The solution is given by:

\[
\sum_{i=1}^{N} \alpha_i (x_i - x)^T \phi(x_i) = (x - x)^T \phi
\]

are Lagrange multipliers. The result of function fitting could be represented as:

\[
q + (x^T x) \phi (\mathcal{L} - \mathcal{L}) \sum_{i=1}^{N} = (x) f
\]

and then the Wavelet Support Vector Regression can be obtained. Here we consider using Morlet mother wavelet to construct the wavelet function.

Wavelet Support Vector Regression

The advantage of wavelet kernel function lies that wavelet has better expressive ability for the signal detail, and for arbitrary complex function, the advantage of wavelet kernel function lies that wavelet has better expressiveness ability for signal detail, and for arbitrary complex function, it can obtain better fitting result than any traditional kernel.

The advantage of wavelet kernel function lies that wavelet has better expressiveness ability for signal detail, and for arbitrary complex function, it can obtain better fitting result than any traditional kernel.

\[
\mathbb{E} \left[ \frac{pZ}{(x' - x')^T} \right] d\alpha \mathbb{E} \left[ \frac{p}{(x' - x')^T} \cos 1.75 \right] = (x - x)^T \phi
\]
The feasibility of WSVR using in the signal filtering

The experiment of 1-dimension signal regression with WSVR, where x denotes scalar input. In this case, the SVR parameter $\epsilon = 0.3$, $c = 0.5$

WSVR Results (black asterisks) are sample noisy

The feasibility of WSVR using in the signal filtering

Gauss, SNR = 38.2397
Lev, SNR = 36.2937
Wavelet Soft, SNR = 44.1924
Wavelet, SNR = 51.7361
Median, SNR = 51.7361
WSVR, SNR = 51.7361

$\text{SNR} = \frac{\sum_{i=1}^{N} \left[ \frac{(x_i f_i - (x_i f_i)^2)_{\text{WSVR}}}{\varepsilon(x_i f_i)} \right]}{10 \times \text{log}_{10}}$

Filtering Result evaluation for one-dimensional signal
The noise model of intensity imagery for SAR

- The model of SAR -
  - The coherent speckle noises in SAR is multiplicative, the model is:
    \[ u = I \]
  - The model of intensity imagery for SAR -
    \[ n = \sigma \]
  - Where \( \sigma \) is the observed value of intensity imagery, \( u \) is multiplicative random noise.

The filtering based on Markov random field (Geman et al., 1994; Donoho, 1995)

- Wavelet method (Guo et al., 1994; Donoho, 1995)
  - Wavelet's soft-threshold value de-noising method
  - Assume the image prior probability using Markov random field, then solve the issue in the frame of maximum posterior probability.

Adaptive filtering based on local statistical property (Lee, 1986)

- Adaptive filtering based on local statistical property
  - Adaptive filtering based on local statistical property
    - Assume a type of noise firstly, and then to do filtering operation based on the distribution characters. The representative filter are Frost filter, Lee filter, Kuan filter and GMAP filter.
  - Wavelet method
    - Wavelet's soft-threshold value de-noising method
    - Assume a type of noise firstly, and then to do filtering operation based on the distribution characters. The representative filter are Frost filter, Lee filter, Kuan filter and GMAP filter.

Classical denoising methods for SAR

- Adaptive filtering based on local statistical property
  - Adaptive filtering based on local statistical property
    - Wavelet method
      - Wavelet's soft-threshold value de-noising method
      - Assume a type of noise firstly, and then to do filtering operation based on the distribution characters. The representative filter are Frost filter, Lee filter, Kuan filter and GMAP filter.
SAR imagery filtering based on WSVR

1. SAR imagery is a kind of two-dimensional signal, the training data set of the image is:

2. The result of function fitting could be represented as:

3. The algorithm flow chart of WSVR filtering for SAR:

   - Logarithmical transformation changes the multiplicative noise of image into additive noise
   - Regression distance is used as judgment index of the noise type:

   \[
   \begin{cases}
   \sum_{i=1}^{N} \sum_{j=1}^{M} (|f - x_i| - \alpha_0) \geq \frac{\beta_0}{2} \quad \text{loss} \\
   \sum_{i=1}^{N} \sum_{j=1}^{M} (|f - x_i| - \alpha_0) < \frac{\beta_0}{2} \quad \text{median}
   \end{cases}
   \]

   \[|f - x_i| = \frac{\beta_0}{2} \cdot a\]

   - The regression distance is used as judgment index of the noise type

4. The experimental results and conclusions:

   - We have chosen two typical areas in SAR image to proceed the experiment, which are the areas of water and land common boundary and airport. The size of two image are all of 256 × 256.

   - The results of SAR image after filtering are shown below.
In this paper, we choose four quantized indicators to evaluate the filtering result, which including the gain of $\text{ENL}$, the gain of $\text{STD}$, edge enhancement index $\text{EEI}$ and the mean of ratio $\text{ER}$. Evaluation index statistics of SAR image filtering results in the area of Fig (a) and (b). The results show that the WSVR filtering method can well reduce the additive random noise and multiplicative salt-and-pepper noise for SAR image filtering. \(\text{WSVR} \) could keep edge features better and enhanced $\text{EEI}$, while the wavelet soft threshold filtering method, such as $\text{KUAN}$ filtering, could not maintain the edge feature as well.

<table>
<thead>
<tr>
<th>Filtering Method</th>
<th>$\text{ENL}$</th>
<th>$\text{STD}$</th>
<th>$\text{EEI}$</th>
<th>$\text{ER}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelet soft threshold filtering</td>
<td>0.2680</td>
<td>0.2324</td>
<td>0.3385</td>
<td>9.4924</td>
</tr>
<tr>
<td>Kuan filtering</td>
<td>0.2670</td>
<td>0.2887</td>
<td>0.2670</td>
<td>13.6881</td>
</tr>
<tr>
<td>Lee filtering</td>
<td>0.2880</td>
<td>0.2880</td>
<td>0.2880</td>
<td>11.6879</td>
</tr>
<tr>
<td>WSVR filtering</td>
<td>0.2880</td>
<td>0.2880</td>
<td>0.2880</td>
<td>11.6879</td>
</tr>
</tbody>
</table>

Conclusion

- WSVR could well reduce the additive random noise and multiplicative salt-and-pepper noise for SAR image filtering.
- Compared to other image denoising methods, such as $\text{KUAN}$ filtering, $\text{Lee}$ filtering and the wavelet soft threshold filtering, WSVR filtering could keep edge feature better.
- But WSVR took a longer time. For real-time demanding applications, we need to research the method of high-speed realization.
Thanks for your attention!