# The Study on Predicting Respiratory Motion with Support Vector Regression

Lei Tong[1], Chaomin Chen[2,*], Kailian Kang[2] and Zihai Xu[3]
[1]Guangdong Vocational College of Mechanical and Electrical Technology, Guangzhou 510515, China
[2]Institute of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China
[3]Radiotherapy Center of PLA 303 Hospital, Nanning 530021, China
*Corresponding author

*Abstract*—**Objective: The target is usually tracked in real time at thoracic and abdominal radiotherapy due to the effect of respiratory motion, the prediction is necessary to compensate the system latency. Method: This paper proposed a prediction method based on support vector regression, it dynamically updates the training set and achieves the accurate online support vector regression. Result: The experiment selected seven respiratory motion data, using online model trained and predicted. The mean absolute error was 0.30mm. Conclusion: The online accurate support vector regression described respiratory motion accurately, and the results with high precision can be satisfied in practical application.**

*Keywords—radiotherapy; support vector regression; respiratory motion prediction; kernel function*

## I. Introduction

Most tumor patients have received radiotherapy; there are many errors in the radiotherapy process, so precise radiotherapy is very important. Precise radiotherapy is aim to achieve high accuracy, high dose, high efficacy and low damage. Due to the respiratory motion, target position will have a certain displacement in the thoracic and abdominal radiotherapy, and the tumor target escape into the normal tissue, so tracking of the target area is essential[1]. With the development of the traditional radiotherapy, respiratory control technology, such as the edge expansion of field, gating technology, breath-hold, four-dimensional radiotherapy etc. has a certain effect, but it still does not meet the need for accurate radiotherapy [2]. In which that real-time tumor tracking technique is in the free breathing state of the patient, the device follows the tumor position, and the tumor movement is tracked by adjusting the irradiation field, so that the center of the field and the center of the target area are kept relatively still. Moreover, the method is a better solution for solving the problem of tumor movement at present. Current research shows there is a significant correlation between the external respiratory signal and the tumor motion. Therefore, it is a trend to study the location of tumor in vivo by using in vitro respiratory signals [3]. However, there is a few hundred milliseconds delay from obtaining tumor location information to adjusting the field. Therefore, a model is needed to predict the in vitro signal to compensate for the delay of system.

Researchers have proposed different forecasting methods, such as neural network, local regression method, memory learning and kernel density estimation [4, 5, 6]. All of these methods than no prediction error is small, but has its defects, which affect the neural network structure and sample complexity, prone to learning or low generalization ability; local regression will be affected by the baseline drift in the regression analysis, often-think "overfitting" method of learning and memory. The pathological nuclear the density estimation method in large quantity, it is difficult to meet the real-time demand. This paper proposes a prediction algorithm of respiratory motion based on support vector regression. This method first selects respiratory motion data of a certain length for training, and then the regression model is obtained. When there is a new data, the regression model calculated the corresponding data in vitro[7,8]. Based on this, it can dynamically update that train set, update the model online, and realize the regression of the precise online support vector.

## II. Methods

Human respiratory signal can be regarded as a time series. The breath motion signal Y sampled at equal frequency is assumed as the input, firstly. The current sampling point is n, and $\lambda$ is the predicted step size. The objectives should be calculated $y_{n+\lambda}$ through time series $y_1, \dots y_n$. The essential thought of support vector regression applied to prediction of time series is, through the training set T=$\{(x_i, y_i), i = 1 \cdots k\}$, when the training set meet $x_i \epsilon R^N, y_i \epsilon R$, the regression equation $f(x)$ could be obtained, usually the expression of $f(x)$ in the feature space F is the formula (1):

$$F(x) = W^T \phi(x) + b \tag{1}$$

Among them, W is a vector in the feature space F, $\phi(x)$ is the mapping of input signal x in the feature space. The W and B in the formula could be obtained through solving the optimization problem below [9]:

$$\min_{W,b} P = \frac{1}{2} W^T W + C \sum_{i=1}^{k} (\xi_i + \xi_i^*)$$

$$\text{s.t.} \quad y_i - (W^T \phi(x) + b) \leq \varepsilon + \xi_i$$

$$(W^T \phi(x) + b) - y_i \leq \varepsilon + \xi_i^*$$
$$\xi_i, \xi_i^* \geq 0, i = 1 \cdots k \tag{2}$$

$\xi_i$ and $\xi_i^*$ are slack variable, that is, the positive and negative values of the deviation. The introduction of slack variables is used to handle training samples that do not meet the conditions: $|f(u_i) - y_{i+\lambda}| \leq \varepsilon$. C controls the degree of punishment for deviations. The optimization criterion of the above problem is to punish the data points of the y true value and the fitting value f (x) exceeding the error of $\varepsilon$. The $\alpha, \alpha^\wedge *, \eta$ and $\eta^\wedge *$ are lagrange multiplier, it used to programming the Lagrange equation of formula (2):

$$L_P = \frac{1}{2}W^T W + C\sum_{i=1}^{k}(\xi_i + \xi_i^*) - \sum_{i=1}^{k}(\eta_i\xi_i + \eta_i^*\xi_i^*)$$

$$-\sum_{i=1}^{k}\alpha_i\left(\varepsilon + \xi_i + y_i - W^T\Phi(x_i) - b\right)$$

$$-\sum_{i=1}^{k}\alpha_i^*\left(\varepsilon + \xi_i^* - y_i + W^T\Phi(x_i) + b\right)$$

$$\text{s.t.} \quad \alpha_i, \alpha_i^*, \eta_i, \eta_i^* \geq 0 \quad (3)$$

The saddle point of Lagrange equation $L_P$ matched condition; hence, the partial derivative of Initial variable b, w, $\xi, \xi^*$ is equal to zero in the optimization problem, as follows:

$$\frac{\partial L_p}{\partial b} = \sum_{i=1}^{k}(\alpha_i^* - \alpha_i) = 0 \quad (4)$$

$$\frac{\partial L_p}{\partial W} = W - \sum_{i=1}^{k}(\alpha_i^* - \alpha_i)\Phi(u_i) = 0 \quad (5)$$

$$\frac{\partial L_p}{\partial \xi} = C - \alpha_i - \eta_i = 0 \quad (6)$$

$$\frac{\partial L_p}{\partial \xi^*} = C - \alpha_i^* - \eta_i^* = 0 \quad (7)$$

Therefore, the above optimization problem can be transformed into its dual optimization problem.

$$\min_{\alpha_i, \alpha^*} D = \frac{1}{2}\sum_{i=1}^{k}\sum_{i=1}^{k}Q_{ij}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) + \varepsilon\sum_{i=1}^{k}(\alpha_i + \alpha_i^*)$$

$$-\sum_{i=1}^{k}y_i(\alpha_i - \alpha_i^*) = 0$$

$$\text{s.t.} \quad 0 \leq \alpha_i, \alpha_i^* \leq C, \sum_{i=1}^{k}(\alpha_i - \alpha_i^*) = 0, i = 1, \cdot k \quad (8)$$

$Q_{ij} = \Phi(x_i)^T\Phi(x_j) = K(x_i, x_j), K(x_i, x_j)$ is the kernel function.The regression equations required can be written as:

$$F(x) = \sum_{i=1}^{k}(\alpha_i - \alpha_i^*)K(x_i, x_j) + b \quad (9)$$

### A. The Kernel Function of Support Vector Machine

The kernel function $\Phi(x)$ denotes the feature of high dimensional space derived from the transformation of vector x

in the original space. Then the classification hyperplane is constructed in the new high dimensional feature space. This is equivalent to constructing the classification surface in the input space. Commutation is a very complex and time-consuming operation. In support vector machines, this nonlinear transformation is realized by defining appropriate kernel functions. We made:

$$K(x_i, x_j) = H(x_i) \cdot H(x_j)$$

$K(x_i, x_j)$ Instead of the point product of the optimal classification plane $x_i^T \cdot x_j$, H is called the spatial transformation function. The function is equivalent to transforming the original feature space into a new feature space. The kernel function realizes the transformation from the low-dimensional space to the high-dimensional space. At the same time, the computation of support vector regression is simplified, the concrete form of transformation function is not needed in the whole process, and the subsequent computation is realized by kernel function.

Kernel functions are mainly of the following types.

(1) Linear kernel function $K(x, y) = \langle x, y\rangle$

(2)Polynomial kernel function

$$K(x, y) = (\langle x, y\rangle + 1)^\sigma$$

(3)Gaussian Radial basis function

$$K(x, y) = \exp(-\frac{\|x - y\|_2^2}{2\sigma^2})$$

(4)Exponential Radial basis function

$$K(x, y) = \exp(-\frac{\|x - y\|_1^2}{2\sigma^2})$$

(5)Sigmoid kernel function

$$K(x, y) = \tanh(\sigma\langle x, y\rangle + \tau)$$

(6)The kernel function constructed by linear combination of kernel basis functions. This kernel function preserves the characteristics of each original kernel function and expresses the influence of different basis kernel functions through different weights.

$\sigma$ and $\tau$ are kernel function parameters. The selection of kernel function and its parameters is very important in the implementation of support vector machine (SVM) algorithm.

### B. Model Parameter Selection

Epsilon insensitive loss coefficient, penalty factor C have a great influence on learning ability and generalization ability of support vector regression , so determining the model parameters is also an important part of the SVR, and its value is related to the training sample. The penalty function C determines the complexity and deviation of the model, which is greater than the acceptance of $\varepsilon$ in the optimization process.

That is, if the penalty function C is too large (infinity), the objective in the optimization process is to minimize the empirical risk. The insensitive loss coefficient ε controls the width of the insensitive region, which is used to fit the training data, and affects the number of support vectors to construct the regression equation. Large insensitive loss coefficient ε will lead to fewer support vectors, and lower complexity regression estimation can be obtained. Therefore, we can see that insensitive loss coefficient ε and penalty coefficient C have influence on the complexity of regression model. It just changes in the opposite direction.

There are several ways to select the parameters of support vector regression: 1) The user chooses according to experience or prior knowledge. 2) The optimal insensitive loss coefficient ε is proportional to the noise level of the sample. 3) The penalty function C has the same range as the output variable. 4) The loss function based on empirical risk in support vector regression is related to the special type of additive noise in the regression equation.

In this study, the insensitive loss coefficient ε is directly proportional to the noise level of the input variable. The determination of ε depends on the noise level of the training sample and the number of the training sample, and the noise level of the training data is known or can be estimated.

## III. RESULT

Due to laboratory constraints, the experimental data used in this paper are free data made public by the Institute of Robotics and Cognitive Systems, University of Lubeck, Germany. The support vector regression is evaluated by seven sample respiratory motion data. The sampling frequency of the data is 20Hz, when the respiratory motion series is $F=[\ F_1\ \ F_2\ \ F_3\ \cdots F_{end}\ ]$, the data collected 10 seconds(about 200) conducted as training sets. Experimental results showed that the training set is decomposed into 4-sample label as a set of input S, the precision of the model is high and the time cost is not increase obviously. The model is trained with expected respiratory movement series T as an output corresponding to S, and the mapping relationship is as follows:

$$S = \begin{bmatrix} F_1 & F_2 & F_3 & F_4 \\ F_2 & F_3 & F_4 & F_5 \\ \vdots & \vdots & \vdots & \vdots \\ F_{end-\lambda-3} & F_{end-\lambda-2} & F_{end-\lambda-1} & F_{end-\lambda} \end{bmatrix}$$

$$\rightarrow T = [F_{4+\lambda}\ F_{5+\lambda}\ \cdots\ F_{end}]$$

At the same time, model parameter selection is referenced in published literature. [10]

The impact of the kernel function on the model performance is shown in Figure 1;when the penalty function C, the insensitive loss coefficient ε is invariant, the effect of kernel function parameter on RMSE is shown.

Different kernel functions show different changes with the change of parameters. Linear kernel functions are also commonly used. Because it is not affected by kernel function parameters, it is not to be compared here. By comparing and analyzing, we can see that when the parameters of

GAUSSIAN RBF kernel function are chosen in a large range, the RMSE can all have ideal values; the comprehensive stability and the algorithm complexity are all good, so the GAUSSIANRBF kernel is chosen in this paper. The results showed that the Gauss kernel function is relatively stable.
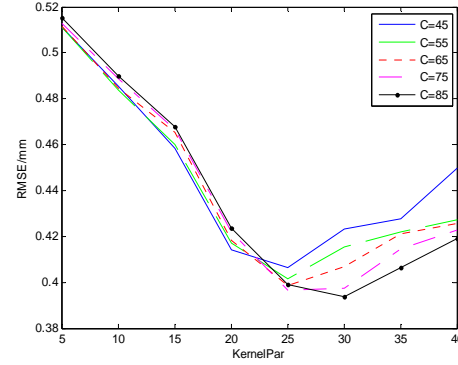


FIGURE I. THE INFLUENCE OF KERNEL FUNCTION

Figure 2 shows the influence of the change of the parameters of the kernel function on the model performance in support vector regression. From figure 2, we can see that when the penalty function C remains unchanged, the kernel function parameter changes, the change trend of the algorithm evaluation index RMSE is first decreasing and then rising, but the change of penalty function C is not very obvious when the kernel function parameter is invariant. From the above analysis, we can get that: as the AOSVR regression model, when the insensitive loss function ε is invariant, changing the penalty function C has little effect on the RMSE, but changing the parameter σ of the kernel function has a certain effect on the value of RMSE.
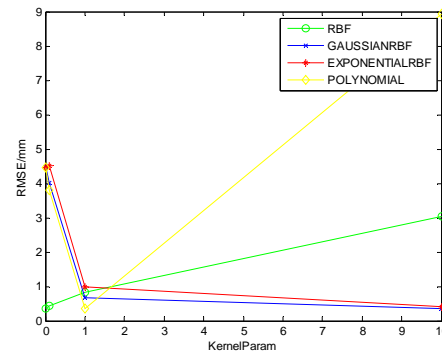


FIGURE II. THE INFLUENCE OF KERNEL FUNCTION PARAMETERS ON PERFORMANCE OF ALGORITHM

## IV. CONCLUSION

Support vector regression (SVM) is prominent in solving nonlinear regression. Respiratory motion is a typical nonlinear time series, in order to accurately fit its motion characteristics accurately, to choose the proper mapping relationship is necessary. Support vector regression (SVM) is trained by selecting a certain historical data, to get the relation between input and output. When new input is obtained, the corresponding output is calculated by model. The respiratory

motion series is divided into several vectors as the feature vectors, and the vectors are used as the input of support vector regression. The mapping combined historical information fit the current output accurately. However, time-consuming still is the defect of support vector regression, and then the acceleration method to make the algorithm more suitable for practical applications will be a topic of the future work.

## ACKNOWLEDGEMENTS

## REFERENCES

[1] WU JH, XU ZH, CHEN CHM, et al. Review of Correspondence Models and Prediction Algorithms in Real-time Tracking Radiotherapy[J]. ChineseJournal of Medical Physics.2015, 32(02):248-251

[2] JEREMY D P, HOISAK, M SC, et al. Correlation of Lung Tumor Motion with External Surrogate Indicators of Respiration[J]. Radiation Oncology Biol. Phys, 2004, 60(4): 1298–1306.

[3] Schweikard A, Glosser G, Bodduluri M, et al. Robotic Motion Compensation For Respiratory Motion During Radiosurgery[J]. Comput Aided Surg, 2000, 5(4):263–277

[4] HUANG ZHY, CHEN WF, ZHOU LH,et al. Prediction of respiratory movement based on GA_BP neural network [J]. Chinese Journal of Biomedical Engineering. 2010, 29(6):812-817

[5] Ruan D, Fessler J A , Balter J M. Real-time Prediction of Respiratory Motion Based on Local Regression Methods[J]. Phys. Med. Biol. 52 (2007) 7137–7152

[6] WAN WQ, ZHANG HL, XU ZH, et al. Prediction of respiratory motion during radiotherapy based on memory learning [J]. Chinese Journal of Biomedical Engineering. 2014,33(2):65-71

[7] Ma J, Theiler J, Perkins S. Accurate On-line Support Vector Regression. Neural Comput, 2003, 15(11):2683–2703.

[8] ALEX J. SMOLA and BERNHARD SCHO. A Tutorial on Support Vector Regression. LKOPF Statistics and Computing, 2004,14: 199–222

[9] Vapnik V. Statistical Learning Theory [M].New York: Wiley, 1998

[10] Cherkassky V, Ma Yunqian. Practical Selection of SVM Parameters and Noise Estimation for SVM Regression[J]. Neural Networks, 2004,17 :113–126