Proc. Int. Conf. on Art. Neural Networks ICANN-96, Lecture Notes in Computer Science LNCS 1112, Springer Verlag 1996, pp. 605-610

Using RBF-Nets in Rubber Industry Process Control

U. Pietruschka, R. Brause

J.W. Goethe-Universität, Frankfurt a.M., Germany
brause@informatik.uni-frankfurt.de

Abstract

This paper describes the use of a radial basis function (RBF) neural network. It approximates the process parameters for the extrusion of a rubber profile used in tyre production.

After introducing the problem, we describe the RBF net algorithm and the modeling of the industrial problem. The algorithm shows good results even using only a few training samples. It turns out that the „curse of dimensions" plays an important role in the model.

The paper concludes by a discussion of possible systematic error influences and improvements.

1 Introduction

Process control in rubber industry has the smell of a „dirty" industrial branch. This comes not only from the often very dull and dusty rubber and tyre production rooms where the products are „baked" by heat and steam, but also from the fact that the macromolecular proportions of rubber are hard to predict due to their nonlinear character. In the extruder (the melting and form-giving machine) the rubber mixture is heated up to 110°-140°C, compressed with 70-140 bar by a screw conveyor and pressed through a metal mask. On leaving the extruder, the rubber relaxes, that is it expands or shrinks, depending on the mixture, changing therefore its shape in a nonlinear manner by 10%-20% up to 50%. Figure 1 shows the basic layout for our example of tyre profile production.

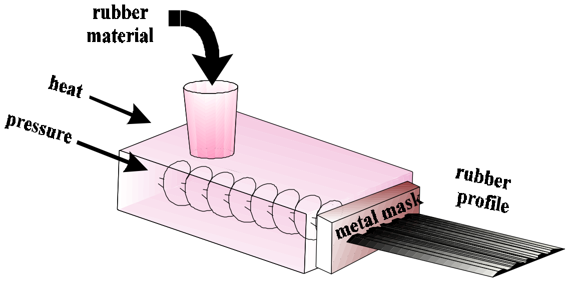

Fig. 1 The tire rubber profile extrusion

The task of process control consists of estimating the necessary extrusion parameters (i.e. the shape of the extrusion metal mask) for an acceptable rubber product after relaxation. The modeling has to reflect the following facts:

- The rubber expansion pressure and flow within the profile heavily depend whether the neighbor parts of the profile have a high level, or if the neighbor parts are low-leveled. This causes the rubber profile to be also a function of the profile height of the neighbored points.

- Additionally, the extruded rubber profile heights depend nonlinearly on the rubber mixture G, the pressure P by the screw conveyor, on the temperature T, on the extruder type E and on the weight w per meter of the band.
- By the nonlinear form of the screw conveyor the pressure along the profile mask decreases nonlinearly. Therefore, the rubber profile does also depend on the absolute position along the metal mask.

Nevertheless, the whole system is deterministic: the same rubber mixture G with the same mask g(x), temperature T and pressure P result in the same rubber profile r(x), even on a different extruder machine of the same type E.

Up to now, due to the nonlinear nature of the macromolecular mixture this task can not be solved analytically. Instead, specialized people estimate the profile of the original metal mask by their experience with the subject and correct their estimates after experience. This gives a trial-and-error production cycle that causes severe disadvantages for the production business:

- The start for a new product is delayed by the time for 2-3 cycles. Each one takes 4-5 days to make a new mask, install it on the extruder, make an extrusion try, measure the obtained rubber profile and estimate a new metal mask. This delay does not only waste time, money and natural resources, but also increase the production overhead and impedes therefore the production flexibility severely.
- The experienced employees are tied to this job (which they judge as „boring“) without the possibility of a change within the enterprise. Additionally, in the case of illness of an employee or a change to another enterprise, the knowledge is no longer accessible. This causes mayor obstacles for the production.

Now, in this paper we will show that adaptive process control methods can overcome this kind of problems. They will update the parameters purely based on the final, measured outcome data.

2 An RBF approximation network

In this section we describe the methods for approximating the exact metal mask profile f(x) at location x of the extruder by a two-layer network function F(x). This produces the desired rubber profile r(x), see figure 2. It is well known that such a two-layer neural network can approximate any continuos function to any degree, provided that we have enough neurons in the first layer, e.g. [Xu, Krzyzak, Yuille 1994].

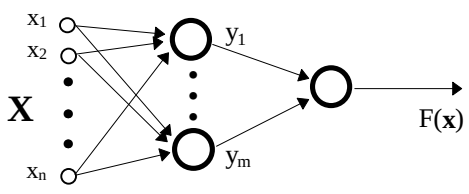

Fig. 2 The activity approximation network

The activity $\mathbf{y} = (y_1,...,y_m)$ of the first layer is defined by

$$y_i = S_i(\mathbf{x},\mathbf{c}_i) = \exp(-d^2) \qquad i=1..m$$

$$d^2 = |\mathbf{M}\,(\mathbf{x}-\mathbf{c})|^2 = (\mathbf{x}-\mathbf{c})^T\mathbf{M}^T\mathbf{M}\,(\mathbf{x}-\mathbf{c})$$

and the second layer with

$y_0 \equiv 1$, $w_0 \cong$ bias by

$$F(\mathbf{x}) = F(\mathbf{y}(\mathbf{x})) = \sum_{j=1}^{m} w_j y_j = \mathbf{w}^T\mathbf{y}$$

using *m* nonlinear basis functions S_i that depend only on the Mahalanobis distance *d* between the input **x** and a neuronal center $\mathbf{c}_i$. For adapting and scaling the ellipsoidal input field, we used the scaling equation

$$\mathbf{M}^{NEW} = (\mathbf{I} - \gamma(1-\alpha)(\mathbf{a}\mathbf{a}^T))\,\mathbf{M}^{OLD} \qquad \mathbf{a} = (\mathbf{x}-\mathbf{c})/\,|\mathbf{x}-\mathbf{c}|$$

with the scaling factor α and the learning rate γ, see [Pietruschka, Kinder 1995].

There are principally two approaches to train the network parameters: either we train the two layers separately or as a whole.

The approach of treating the two layers separately, clustering the input space first and then optimizing the weights of the second layer, is fast, but it has some flaws. This gives us a high sample density of output values where we have clusters of input samples, not where the output error is high.

Therefore, we optimize both layers at the same time. To avoid the computational problems of the backpropagation approach we choose a different strategy. We start with the lowest possible complexity of the network and gradually increase the number of neurons in the first layer until the error is sufficiently reduced. This was already proposed for RBF nets, for example by [Schiøler, Hartmann 1992]. We insert the neuron at location $\mathbf{x}_k$, the k-th sample with the maximal error, that has to be compensated by the new m-th neuron. We have to design the width $\mathbf{M}_m$ such that it fits the new basis function in the context of all neighbored neuron basis functions. In contrast to the approach of [Platt, 1995] we do not use gradient descend technique to rearrange all other neurons and adapt all their receptive fields: this is computationally intensive and is the source of new errors. Instead, we stop the adaptation process of the new neuron by the criterion of non-significant activation on a data point. Additionally, we reduce the long distance neighborhood influence by a learning rate $\gamma(d)$ which drops with increasing distance from $\mathbf{c}_m$, that is with decreasing activity level, see [Pietruschka, Kinder 1995].

3 Approximating the extrusion process parameters

To apply the approximation algorithm that we developed in the previous sections we have to model the industrial process for the example of tire production. The main task consists of estimating the profile of a metal mask that extrudes the profile of a rubber band. This band is then cut into a stripe of the perimeter length of a tire and then glued to the casing. The raw tire is then „baked“ in a metal tire form for 20 minutes, giving the preliminary profile the ultimate form.

3.1 Modeling the process

Although the extruded rubber profile is a temporary form its desired accuracy is 0.1 mm. This settles the upper limit for our approximation error. In figure 3 a sample profile is shown.

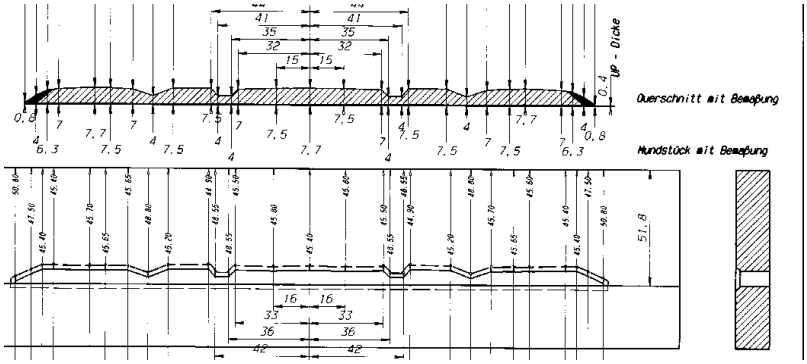

Fig. 3 A rubber profile and the corresponding metal mask

The upper profile is the desired rubber profile; the lower one shows the corresponding rectangular metal mask. On the right hand side a cut through the metal (shaded area) shows the form of the opening (not shaded). Where the rubber flows in, the profile has a wider opening. This corresponds to the dotted line that encircles the profile opening in the metal mask.

The analytical treatment of the nonlinear dependencies is very difficult. Conventional assumptions about energy (i.e. enthalpy) conservation are not valid here. Also the direct measurement of the process parameters like temperature and pressure in the profile are practically limited. In contrast to this, our approach models the system as a whole, avoiding all difference equations and constants which are hard to devise and to measure.

We devided the whole centered profile, depending on the tyre width, into 170-270 points that are placed in the regular distance of *d* mm. Each point x_i has a desired rubber profile height r(i). The intermediate points are interpolative generated, see figure 4.

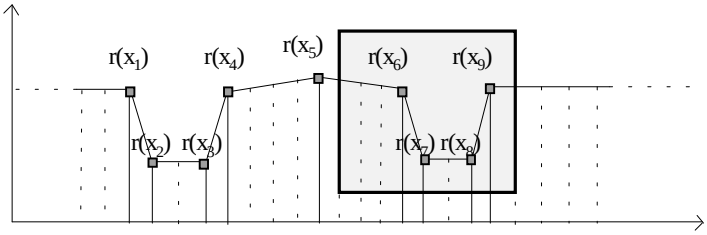

Fig. 4 The intermediate interpolation of the profile and a neighborhood window

Since the influence of the sample points is limited to the neighborhood for a certain rubber profile height r(i) we have only to consider *n=2s+1* neighbor points

$$g(i) = F(r_{i-s},...,r_i,...,r_{i+s}, i, G, E, P, w, ..)$$

By this model we implement a neighborhood window that uses n=2s+1 sampling points around location i. All values r_k for the sampling points outside the profile limits are set to zero.

3.2 Simulation results

For the determination of the two parameters, the number *n* of neighborhood sampling points, and the distance *d* between the sampling points, we decided to simulate different configurations in order to get an acceptable choice.

We generated the training set by shifting a window (determined by *d* and *n*) by an increment of 1 mm over the profile data of 5 profiles with the same values of G, E and P. This generated 1346 training patterns. The sixth profile was used for the generation of test set of 271 test points. The simulation results generally showed only a very small influence of the position i. So, let us regard other dependencies.

For the expected absolute error for 100 neurons we got different results, depending on the type of network we used. The nets with growing, radially symmetric input regions have in the average 10-90% more error than the growing elipsoidal nets. The best performance of the two types converged by training to the following expected absolute error, depending on the number of sampling points *n* and the interpoint distance *d*.

n \ d	3 mm	4mm	5mm
7	0,178	0,159	0,187
9	0,167	0,162	0,197
11	0,165	0,206	0,226

It is interesting to see that the error does not automatically decrease or increase when we increase the number of sampling points. There is a configuration of the parameters where we roughly meet the balance and the error becomes quite small. The best results are observed by n=9 and d=4 mm which corresponds to a window size of 32 mm with the expected absolute error of 0.16 mm and the maximal absolute error of 0.56 mm. In figure 5 the test profile, the result of the network and the resulting

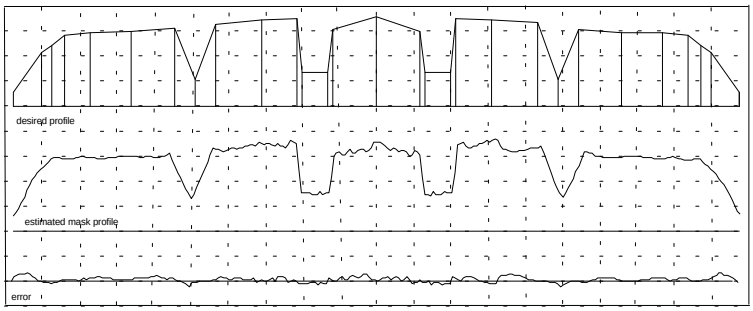

Fig. 5 The wanted profile and the profile produced by the net for n=9, d=4

error is shown for this configuration. The y-axis is scaled up by the factor of three to enhance the visibility of the errors.

4 Discussion and outlook

In the previous sections we have presented an adaptive solution for the problem of unknown process parameters in tyre production. The learning algorithm uses no internal process variables or other intrinsic knowledge but only the measurable external process parameters as the weight per meter and the resulting rubber profile.
This approach avoids many technical and economic disadvantages that were given in the introduction.

Nevertheless, our work also shows that there are still several problems to be solved:

- The current modeling uses the data provided by the production as tuples (wanted rubber profile, successful metal profile). The successful metal profile used for training was obtained after several trials and corrections, that is the training is based on an artificial and not on a real sample of the input/output mapping.

  This problem can be solved by measuring and using only the directly obtained rubber profile data.

An important key for the simulation performance turned out to be the two parameters, the number n of neighborhood sampling points, and the distance d between the sampling points. We can determine the proper choice by balancing the counteracting influences:

- If we choose d too small, we increase the number of necessary sampling points for a certain neighborhood and increase therefore the dimension of the input space. Since we have only a small limited number of training samples, the training becomes very difficult since the input space becomes very sparse. This is known as the „curse of dimensions“ [Huber 85]. On the other hand, if we choose d too big, we can lose important information due to undersampling the dependency function.
- If we choose n too big, we meet the same problem of sparseness of the training samples in the input space. Additionally, by increasing to much context information, the generalization ability of the network will be limited. On the other hand, if we limit the window too much, necessary context information which helps to distinguish between different situations is lost.

From the theoretical point of view, this is an interesting situation. By the nature of the problem, we have not hundreds of sample profiles but just a few ones. Nevertheless, we are not aware of an applicable method of determining the optimal d and n to solve the problem of optimal training. Here, the work of theorists is welcome.

5 References

P. Huber: Projection Pursuit; The Annals of Statistics, Vol.13, No.2, pp. 435-475, 1985

U. Pietruschka, M. Kinder: Elipsoidal Basis Functions for Higher-Dimensional Approximation Problems; Proc. ICANN-95, Vol II, Paris 1995, pp.81-85

J. C. Platt; Learning by Combining Memorization and Gradient Descent; NIPS; pp. 714-720;1992.

H. Shiøler, U. Hartmann: Mapping Neural Network Derived from Parzen Window Estimator; Neural Networks, Vol.5, pp.903-909 (1992)

Lei Xu, Adam Krzyzak, Alan Yuille: On Radial Basis Function Nets and Kernel Regression: Statistical Consistency, Convergence Rates, and Receptive Field Size; Neural Networks, Vol. 7, No.4, pp.609-628, (1994)