Steel Sheet Incremental Cold Shaping improvements using Hybridized Genetic Algorithms with Support Vector Machines and Neural Networks

Laura Puigpinós, José R. Villar, Javier Sedano, Emilio Corchado, Joaquim de Ciurana

Abstract The complexity and difficulties in modelling the most of nowadays real world problems increase as the computational capacity does, specially in those processes where relatively new technology arises. One of such processes is the steel sheet incremental cold shaping. The steel sheet incremental cold shaping process is a new technique for shaping metal sheets when a reduced amount of pieces per lots should be manufactured. As it is a relatively new technique, there is a lack of knowledge in defining the operating conditions, so in order to fit them, before manufacturing a lot a trial and error stage is carried out. A decision support system to reduce the cost of processing and to assist in defining the operating conditions should be studied. This study focus on the analysis and design of the decision support system, and as it is going to be shown, the most suitable features have been found using a wrapper feature selection method, in which genetic algorithms support vector machines and neural networks are hybridized. Some facts concerning the enhanced experimentation needed and the improvements in the algorithm are drawn.

Laura Puigpinós
Fundación Privada Ascamm, Avda. Universitat Autònoma, 23 08290 Cerdanyola del Vallés (SPAIN) e-mail: lpuigpinos@ascamm.com

José R. Villar
University of Oviedo, Campus de Viesques s/n 33204 Gijón (SPAIN) e-mail: villarjose@uniovi.es

Javier Sedano
Instituto Tecnológico de Castilla y León, Poligono Industrial de Villalonquejar, Burgos (SPAIN), e-mail: javier.sedano@itcl.es

Emilio Corchado
University of Salamanca, Salamanca (SPAIN) e-mail: escorchado@usal.es

Joaquim de Ciurana
University of Girona, Escola Politècnica Superior Edifici PII, Avd. Lluís Santaló, s/n 17071 Girona (SPAIN) e-mail: quim.ciurana@udg.es

1 Introduction

Over recent years there have been a significant increase in the use of artificial intelligence and Soft Computing (SOCO) methods to solve real world problems. Many different SOCO applications have been reported: the use of Exploratory Projection Pursuit (EPS) and ARMAX for modelling the manufacture of steel components [2], EPS and Neural Networks (NN) for determining the operating conditions in face milling operations [10] and in pneumatic drilling process [11], genetic algorithms and programming for trading rule extraction [4] and low quality data in lighting control systems [15], feature selection and association rule discovery in high dimensional spaces [14] or NNs and principal component analysis and EPS in building energy efficiency [12, 13].

It is known that the complexity inherited in most of new real world problems, the steel cold shaping industrial process among them, increases as the computer capacity does. Higher performance requirements with a lower amount of data examples is needed due to the costs of generating new instances, specially in those processes where new technology arises.

In this sense, the steel cold shaping is a relatively new technology in the production of lots with a low quantity of pieces, which represents an effervescent area. NNs have been used to find relationships between the mechanical properties of the cold-rolled sheets of interstitial free and the chemical composition of the steel and the rolling and the batch annealing parameters [9]. NNs have been also applied for identification of the parameters for operating conditions [20, 21]. To the best of our knowledge, no specific study has been published in steel iterative cold shaping.

This study focuses on determining the main parameters in an steel sheet incremental cold shaping. The main objective is to find the most relevant feature subset; and the second objective is to obtain a decision support system in the operating conditions design, so the costs of producing such low amount of pieces in the lots is reduced. The next Section is concerned with the problem description. In Section 3 the algorithm used is detailed, while Sect. 4 deals with the experiments carried out and the results obtained. Finally, conclusions are drawn and future work goals are set.

2 Steel Sheet Incremental Cold Shaping

The metal incremental cold shaping is based on the concept of incremental deformation. This technology allows the manufacturing of pieces of metal sheet through the iteration of small sequential deformation stages until the desired shape is achieved and avoiding the axis-symmetric restrictions due to incremental rotatory deformation.

Comparing the incremental cold shaping with traditional deformation technologies it can be said that the former reduces the cost of specific machine tools and the manufacturing costs dramatically.

This type of technology has evolved from the well-known Rapid Manufacturing, allowing to generate pieces with complex geometries in a wide spread of materials without the need of frameworks or specific tools.

The main part of cold shaping has been controlled using numerical controlled tools in order to reduce as most as possible the fast, reliable, ans low-cost manufacturing of lots with an small amount of metal pieces and prototypes. The scheme of metal sheet incremental cold shaping process is shown in Fig. 1.

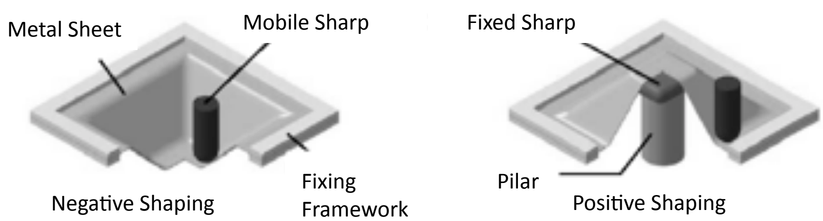

Fig. 1 The incremental cold shaping of a steel sheet. A sharpening tool is iteratively applied onto the metal sheet at a different depth. In the negative shaping only the sharp tool is moved, while in the positive shaping both the metal sheet and the sharp tool are moved.

The process of cold shaping starts with the design of a geometric shape in a 3D CAD file. This file should include as many layers as desired, each layer represents the bounds to be reached in each deforming step and are piled vertically. Consequently, the piece should be generated using the sequential and incremental layers, each one at a different depth and constraint within the defined bounds.

Plenty of parameters have to be fixed for the manufacture a metal piece, among them the force developed by the deforming head in each of the three dimensions, the speed change, the trajectory of the head, the surface roughness, the sheet pressure stress, the incremental step between layers, the number of steps or stages, the attack angle, the angle variation, the depth variation, etc.

2.1 The problem definition

The first aim of this study is to evaluate if it is possible to model the operating conditions so the suitability of the experiment could be established, in other words, to analyse whether the operating conditions would generate a faulty piece or not while the most relevant features involved are to be selected.

The second aim is to model the maximum suitable depth that can be reached with the given operating conditions. As in the former problem, the best feature subset is also required.

Therefore, there are two problems to solve, both including a feature selection process and a modelling process. The former is a two-class problem, while the second is a regression problem.

3 Feature Selection, Support Vector Machines and Neural Networks

In order to obtain a suitable feature subset some requirements are needed. As there are integer features, nominal features and real valued features, the algorithm should deal with any kind of data. Therefore, the same approach should be valid for the both subproblems, the two-class problem and the maximum depth. Besides, not only the best feature subset for each problem but also the best model are desired, a classifier in the former case and a regression model in he latter.

It is known that for this kind of problems the wrapper approach for feature selection performs better than filter solutions [3, 18]. These studies proposed wrapper feature selection methods using genetic algorithms (GA) for dealing with the feature subset selection, that is, each individual is a feature subset. To evaluate individuals a modeling technique has been applied: the former proposed a lazy learning model as the K-Nearest Neighbour, the latter made use of a NN model that iteratively fix the number of hidden neurons.

Different approaches as to how the NN is learnt have been studied. In [1] a GA approach to fingerprint feature selection is proposed and selected features are supplied as input to NN for fingerprint recognition, while in [17] a similar approach has been applied to automatic digital modulation recognition. Moreover, this type of approach has been reported to perform better than using statistical models [19]. Despite, Support Vector Machines (SVM) have been also used in conjunction with evolutionary feature selection to reduce the input space dimensionality [5, 6].

In this study, two different approaches are analysed. The first one is an specific GA wrapper feature selection method for each problem, the second approach is an evolution of the former, allowing different NN models to be found.

3.1 GA+SVM+NN feature selection

In this study we adopt two different solutions depending whether we are dealing with the two-class or the maximum depth problem. An hybridized method of GA evolving the feature subsets and a SVM classifier is chosen in the former case, while in the latter an hybridized method of GA evolving the feature subsets and a NN for modeling the desired output is used. In both modelling and feature selection problems the GA is an steady state approach with the percentage of elite individuals to be defined as a parameter. The algorithm has been implemented in MatLab [8], using both the NN and the SVM toolboxes.

The algorithm is outlined in Algorithms [1,2]. Algorithm [1] evaluates an individual (which is to say, a feature subset), while the latter, which is the main algorithm, shows the GA that evolves the feature subset and that calls Algorithm 1 whenever required.

For the sake of simplicity we have neither reproduce the algorithm for the SVM nor for the NN cases. Instead, we present the general case in the algorithms, and when it is said that a model is trained, the reader should consider which problem (the two-class or the regression problem) is related to the use of NN or SVM.

The typical steady state GA parameters, like the crossover and mutation probabilities, the number of generations, the population size and the elite population size, are all of them given for each experiment. The individual representation is the string of indexes of the chosen feature subset. The tournament selection is implemented and one point crossover is used. After each genetic operation the validity of the off-prints is analysis: repeated features indexes are erased and random indexes are introduced to fill the individual feature subset.

Third order polynomials are used as kernel functions for the SVM. The number of hidden nodes in the NN is set as a parameter. The NN models are generated randomly and trained using the corresponding Matlab library functions. In all cases, 10-fold cross validation is used, and the mean value of the mean squared error in each fold is the fitness of an individual.

Algorithm 1 IND_EVALUATION: Evaluates an individual

Require: I the input variables data set
Require: O the output variable data set
Require: ind the individual to evaluate, with its feature subset
  $model$ {the best model learned for ind}
  $mse = 0$ {the associated mean of Mean Square Error for ind}
  $indMSE = 0$ {best MSE found in the cross validation}
  for $k = 1$ to 10 **do**
    {run the k fold in the cross validation scheme}
    generate the train and test reduced feature data set
    initialize the model $indModel$
    train $indModel$ with the train data set
    $indKMSE \leftarrow$ calculate the MSE for $indModel$ with the test data set
    $mse+ = indKMSE$
    if $k == 1$ **or** $indMSE > indKMSE$ **then**
      $indMSE = indKMSE$
      $model = indModel$
    end if
  end for
  $mse = mse/10$
  return $[model, mse]$

Algorithm 2 GA$^+$ Feature Selection

Require: I the input variables data set
Require: O the output variable data set
Require: N the feature subset size
  $FS \leftarrow \{\emptyset\}$ {the best feature subset}
  $model$ {the model learned for FS}
  $mse = 0$ {the associated mean of Mean Square Error for FS}
  Generate the initial population, Pop
  for all individual ind in Pop **do**
    $[ind.model, ind.mse] = IND_E VALUATION(I, O, ind)$
  end for
  $g \leftarrow 0$
  while $g < G$ **do**
    while $size(Pop') < (popSize - |E|)$ **do**
      Generate new individuals through selection, crossover and mutation
      add valid individuals to Pop'
    end while
    extract the elite subpopulation $E \in Pop$
    for all individual ind in Pop' **do**
      $[ind.model, ind.mse] = IND_E VALUATION(I, O, ind)$
    end for
    $Pop = \{E \cup Pop'\}$
    sort Pop
    g++
  end while
  $FS \leftarrow Pop[0]$
  $[model, mse] \leftarrow$ corresponding model and MSE
  return $[FS, model, mse]$

3.2 Evaluating the effect of some trainnig and test issues

Some variations were accomplished in order to evaluate the performance of the feature selection modelling. The main question is if the features where good enough to be chosen if we change the NN and its parameters.

For this purposes, a second collection of 10 fold cross validation data sets was generated independently to those prepared by Matlab. Moreover, a different NN was used for the second problem. In this case, a feedfoward NN with three hidden layers and one linear output unit, with 10 units in the first hidden layer, 15 units in the second hidden layer and 4 units in the third hidden layer is used; all hidden unit are hyperbolic tangent units. The NN was trained with the Lenverberg-Marquart method.

3.3 A discussion on cross validation schemes

The fitness of each individual is calculated, as outlined before, using a k-fold cross validation (cv) scheme. The main aim of this evaluation is to estimate the performance of a feature subset in operating conditions, when the sample given has not been presented for training. This has been found relevant in those problems for which the data set includes few samples [16].

Nevertheless, many different schemes can be used, the leave-one-out, the k-fold cv or the 5x2 cv among them. The selection of the cv relies on the data set dimensions.

The mean error among the whole set of models learned is proposed as the fitness function; for all cases in this study, the validation has been calculated with the test data set. However, it is worth noting that it is possible to use a validation data set for evaluating the mean error and so the fitness of the individuals.

Consequently, the cv should extract a three data sets from the original one: the train, the test and the validation data sets. Clearly, choosing the kind of cv will depend on the data set dimensionality; when there are not enough samples are available the solution will tend to do use leave one out cv without a specific validation data set. Automatic selection and developing of the cv scheme is also left for future work.

It is worth mentioning which model is kept for an individual. In this study we kept the best model found, that is, the one with lower mse, but if multi objective were used then a different criteria should be selected.

4 Experiments and results

Generating data set samples is costly as the each one of the samples needs a real case to be carried out, that is, a sheet of steel has to be cold shaped; consequently, the smaller the number of experiments, the lower the cost and the smaller the data set size.

The data set comprises 19 samples, each one with the whole set of parameters values. Once the piece is processed as the corresponding sample establishes, then it is manually classified as $\{$GOOD, BAD$\}$ according to the deformation or the quality faults that could appear in the piece. Besides, the maximum depth in each case is also measured. These two latter values are appended to each sample as the output variables.

As SVM and NN are to be used in the modelling part of the feature selection GA method, then the data set is normalized with means 0 and deviations 1.

In the first experimentation, the GA parameters has been fixed as follows: 50 individuals in the population, 100 of generations, the probability of crossover equals to 0.75, while the mutation probability is 0.25. An steady state GA evolutionary scheme is used, with a number of 5 elite individuals that will be kept in the next generation.

The size of the feature subset has been fixed to three. The SVM kernel function is fixed as third order polynomials and the feed forward back-propagation NNs includes 6 neurons in the hidden layer, with $\mu = 0.001$, $\mu_{dec} = 10^{-10}$, and $\mu_{inc} = 10^{-6}$.

The parameters of the SVM and the NN have been kept constant during the feature selection and the model learning.

As stated in the previous section, the 10-fold cross validation schema is carried out. Only the validation results are used to compare individuals. For the two-class problem the mean of the classification errors among the folds is used to evaluate the models and the feature subsets. For the maximum depth estimation, each feature subset and its corresponding model are evaluated with the mean of the mean squared error on each fold.

In the case of the two-class problem, the best feature subset found includes the *step increment*, the *angle variation* and the *variation in depth*, with a mean classification error of 0.1%. For the second problem, the best feature subset found includes the variables *step increment*, the *number of stages* and *variation in depth*, with a mean error of 0.0096.

For the second approach experimentation a bayesian regularization was applied to the data sets [7]. All the parameters remain the same except for the NN learning parameters, which where fixed at $\mu = 1.0$, $\mu_{dec} = 0.8$, and $\mu_{inc} = 0.0001$. In this case, the obtained errors were meaningfully worse, and consequently, the feature subset chosen in this case was not the same as in the former experimentation. Only in the case of reducing the feature space to two variables both approaches proposed the same feature subset. Clearly, in order to obtain the best feature subset an automatic model parameters identification is needed so the obtained feature subset does not depend on the experiment decisions.

It is worth mentioning that the reduced number of data samples induces relatively high error values as the test data set includes only one or two samples. More accuracy should be obtained if a bigger number of samples is given. However, the cost of the data gathering increases considerably; this dilemma should be evaluated.

Moreover, the algorithms do not include local optimization of the models parameters. So it is possible that better performance of the models and the feature selection process should be achieved if such local optimization were implemented.

Finally, the maximum depth have been found regardless of the two-class problem, which was not the expected result in the expert opinion. It is though that the above mentioned local optimization of the models parameters should improve the performance and the experts confidence in the results.

5 Conclusions and future work

This study presents a feature selection method for choosing the best feature subset in a steel sheets cold shaping process divided in a two-class problem and a maximum depth estimation problem. Moreover, a genetic algorithm is hybridized, on the one hand, for the first case, with a support vector machine model to choose the best feature subset and on the other hand, for the second case, with a feed foward back-propagation neural network.

From the experimentation the best feature subset has been found for both problems, and some relevant facts have arisen. Firstly, the data set size should be increased in order to obtain better models fitness values. Secondly, local optimization for the models parameters should improve the obtained results. Besides, it could be desirable that the optimum number of features should be dynamically fixed, which represents an improvement in the individual representation. Finally, the analysis of the different cv schemes and the decision of which of them to use is left for future work.

Acknowledgements This research has been funded by the Spanish Ministry of Science and Innovation, under project TIN2008-06681-C06-04, the Spanish Ministry of Science and Innovation [PID 560300-2009-11], the Junta de Castilla y León [CCTT/10/BU/0002] and by the ITCL project CONSOCO.

References

1. Altun, A. A., Allahverdi, N. (2007) Neural Network Based Recognition by Using Genetic Algorithm for Feature Selection of Enhanced Fingerprints. ICANNGA '07 Proceedings of the 8th international conference on Adaptive and Natural Computing Algorithms, Part II. Springer-Verlag Berlin, Heidelberg
2. Bustillo, A., Sedano, J., Curiel, L., Villar, J. R., Corchado, E. (2008) A Soft Computing System for Modelling the Manufacture of Steel Components. Advances in Soft Computing **57:3**, 601–609
3. Casillas, J., Cordón, O., del Jesus, M. J., Herrera, F. (2001) Genetic Feature Selection in a Fuzzy Rule-Based Classification System Learning Process. Information Sciences **136:1-4**, 135–157
4. de la Cal, E., Fernández, E. M., Quiroga, R., Villar, J., Sedano, J. (2010) Scalability of a Methodology for Generating Technical Trading Rules with GAPs Based on Risk-Return Adjustment and Incremental Training. LNCS **6077**, 143–150
5. Fung, G. M., Mangasarian, O.L. (2004) A Feature Selection Newton Method for Support Vector Machine Classification. Computational Optimization and Applications **28:2**, 185–202
6. Huanga, C.-L., Wang, C.-J. (2006) A GA-based feature selection and parameters optimizationfor support vector machines. Expert Systems with Applications **31:2**, 231–240
7. Mackay, D. J. C. (1992) Bayesian interpolation. Neural Computation **4:3**, 415–447
8. The MathWorks (2011) MATLAB - The Language Of Technical Computing http://www.mathworks.com/products/matlab/
9. Mohanty, I., Datta, S., Bhattacharjeeb, D. (2009) Composition-Processing-Property Correlation of Cold-Rolled IF Steel Sheets Using Neural Network. Materials and Manufacturing Processes **24:1**, 100–105
10. Redondo, R., Santos, P., Bustillo, A., Sedano, J., Villar, J. R., Correa, M., Alique, J. R., Corchado, E. (2008) A Soft Computing System to Perform Face milling Operations. 4th International Workshop on Soft Computing Models in Industrial Applications, LNCS **5518**, 1282–1291
11. Sedano, J., Corchado, E., Curiel, L., Villar, J., Bravo, P. (2008) The Application of a Two-Step AI Model to an Automated Pneumatic Drilling Process. Int. J. of Comp. Mat. **86:10-11**, 1769–1777
12. Sedano, J., Curiel, L., Corchado, E., de la Cal, E., Villar, J. R. (2009) A Soft Computing Based Method for Detecting Lifetime Building Thermal Insulation Failures. Int. Comp.-Aided Eng. **17:2**, 103–115

13. Sedano, J., Villar, J. R., Curiel, L., de la Cal, E., Corchado, E. (2009) Improving Energy Efficiency in Buildings Using Machine Intelligence. 10th Int. Conf. on Intelligent Data Engineering and Automated Learning (IDEAL 2009), LNCS **5788**, 773–782
14. Villar, J. R., Suárez, M. R., Sedano, J., Mateos, F. (2009) Unsupervised Feature Selection in High Dimensional Spaces and Uncertainty. 4th International Workshop on Hybrid Artificial Intelligence Systems, LNCS **5572**, 565–572
15. Villar, J. R., Berzosa, A., de la Cal, E., Sedano, J., Garcıa-Tamargo, M. (2011) Multi-objective Simulated Annealing in Genetic Algorithm and Programming learning with low quality data. In publication for Neural Computing
16. Villar, J. R., Sedano, J., Corchado, E. and Puigpinós, L. (2011) Soft Computing decision support for a steel sheet incremental cold shaping process, accepted for the Proc. of the 12th Int. Conf. on Intel. Data Eng. and A. L. (IDEAL'2011)
17. Wong, M. L. D., Nandi, A. K. (2004) Automatic digital modulation recognition using artificial neural network and genetic algorithm. Signal Proc. **84:2** 351–365
18. Yang, J., Honavar, V. (1998) Feature Subset Selection Using a Genetic Algorithm. IEEE Intelligent Systems **13:2**, 44–49.
19. Zhang, P., Verma, B., Kumar, K. (2005) Neural vs. statistical classifier in conjunction with genetic algorithm based feature selection. Pat. Recog. Letters **28:7** 909-919
20. Zhao, J., Cao, H. Q., Ma, L.X., Wang, F. Q., Li, S. B. (2005) Study on intelligent control technology for the deep drawing of an axi-symmetric shell part. J. of Materials Processing Tech. **151:1-3**, 98–104
21. Zhao, J., Wang, F. (2005) Parameter identification by neural network for intelligent deep drawing of axisymmetric workpieces. J. of Materials Processing Tech. **166:3**, 387–391