2nd International Conference on System-Integrated Intelligence: Challenges for Product and Production Engineering

Data mining approach for knowledge-based process planning

Berend Denkena, Justin Schmidt*, Max Krüger

Institute of Production Engineering and Machine Tools (IFW), Leibniz Universität Hannover, Hannover Centre for Production Technology (PZH), An der Universität 2, 30823 Garbsen, Germany

Abstract

Concepts like gentelligent products, smart objects or cyber-physical systems have already proven a high potential especially for decentralized production planning and control. In this context, decentralized communication, new sensor technologies and the increased application of simulation and monitoring systems lead to an enormous increase of manufacturing data. Additionally, a new approach for the assessment of manufacturing quality based on process signals from the machine tool is proposed, which provides current tool state and surface roughness information for every manufacturing process. In order to reuse and evaluate this data for knowledge-based process planning, an approach to manufacturing data collection and evaluation using data mining methods was developed. The advantages of the proposed data mining approach for process planning is demonstrated by an exemplary testing case.

1. Introduction

   Significant advances in system-integrated intelligence, novel communication technologies and the application of autonomous objects and systems in production have led to new paradigms in production. Concepts like gentelligent components, smart objects or cyber-physical systems have already proven a high potential especially for decentralized manufacturing planning and control. First concepts have been evaluated in testing cases and lead the way to an individualized product-driven production [1, 2].

---

  * Corresponding author. Tel.: +49-511-762-18312; fax: +49-511-762-5115.
  E-mail address: schmidtj@ifw.uni-hannover.de

In this context, decentralized communication, new sensor technologies and the increased application of simulation and monitoring systems lead to an enormous increase of available manufacturing data like machine data, product quality data or information on failures and disturbances. The availability of manufacturing data offers new possibilities for a knowledge-based manufacturing planning and control, for instance finding suitable resources and optimum manufacturing process parameters. Especially process planning is a task that is strongly dependent on expert's knowledge. In the recent decades of research, subjects like feature technology, expert systems and data format standardization (e. g. STEP-NC) have been investigated and form the basis for knowledge-based computer-aided decision support systems [3].

In terms of knowledge acquisition from large volumes of data, knowledge discovery in databases has become a separate but interdisciplinary research trend over the years. It includes methods from statistics, machine learning and data modelling. Fayyad et al. proposed a procedure model that contains five steps of data processing: selection, preprocessing, transformation, data mining and interpretation [4]. The relevance of the first steps is strongly influenced by the kind and the structure of input data. The main task of this procedure is the data mining. The purpose of data mining is the recognition of unknown patterns within data. Different algorithms are used and some of them have already proved their applicability in terms of analyzing manufacturing data [5].

Up to now, the major part of the available manufacturing data is not yet exploited for manufacturing planning and control. The main reasons are the huge amount of highly detailed information and the adherence to established structures of information systems in manufacturing, which leads to an inconsistent data basis. A further obstacle is the absence of concepts for the evaluation of manufacturing data that integrate data sources from smart objects [6]. This paper presents an approach to the systematic evaluation of manufacturing data for knowledge-based process planning related to the concept of gentelligent production. First, manufacturing data is analyzed and classified and a new approach for data acquisition by online process assessment is presented. Based on the manufacturing data and new data sources, an evaluation procedure using data mining methods and its first application results are described.

2. Concept of knowledge-based process planning in gentelligent production

The idea of gentelligent production developed within the Collaborative Research Center (CRC) 653 aims at the merging of physical components with its related information. Intelligent workpieces store their manufacturing data inherently, search their way through the factory autonomously, trigger their own production processes and store their manufacturing history permanently. This enables maximum adaptability and the ability for single workpiece tracking and tracing over their whole life cycle [7]. In this context, a concept for knowledge-based process planning by reusing individualized manufacturing information was developed (see Fig. 1).
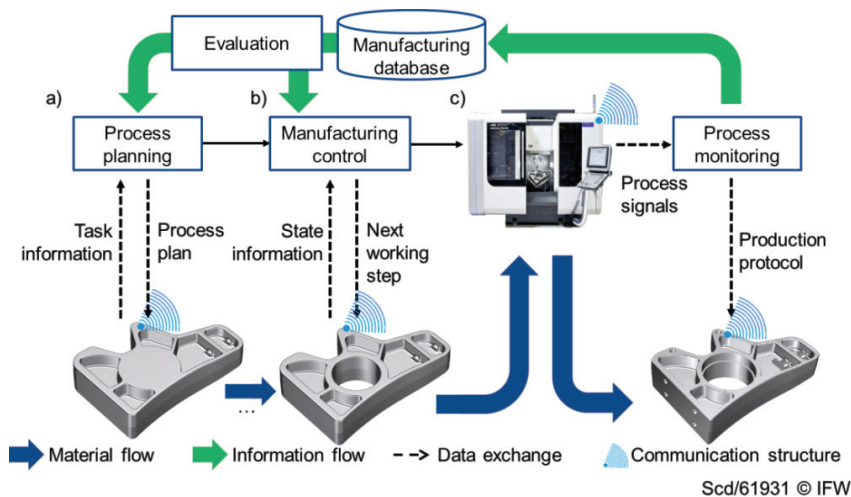

Fig. 1. Knowledge-based planning in gentelligent production.

In the first step, process plans are generated based on the workpiece's construction data (Fig. 1a). Then, considering current production state information, the next working step is selected and – if necessary – adapted within a manufacturing control system (Fig. 1b). During machining, process signals are evaluated by a process monitoring system (Fig. 1c). Finally, a production protocol can be stored inherently in the workpiece, which will be available over the whole product life cycle. In order to reproduce optimum process results, workpiece-specific process information from process monitoring will be combined with additional manufacturing data (data enrichment) and fed back to process planning (data enhancement). Therefore, an approach to manufacturing data collection and evaluation using data mining has been developed. In a first step, an analysis and classification of manufacturing data has been carried out.

3. Analysis and classification of manufacturing data

As the gentelligent production concept aims at a maximum adaptability in process planning and manufacturing control, manufacturing data contains both static and dynamic information. According to literature [8, 9], the input data for process planning consists of product, process and resource information. For metal cutting applications, the information classes order, workpiece, stock, process, machine tool, cutting tool and personnel can be identified exemplarily. Fig. 2 shows information sources for static and dynamic manufacturing data, which can serve as a data input for process planning.

Fig. 2. Manufacturing information analysis.

Existing information systems in manufacturing environments offer a wide range of static manufacturing data. Also dynamic data is acquired by systems like production or machine data acquisition (PDA/MDA). In the industrial practice, most of this information is neither collected consistently nor fed back to planning according to a lack of reliable information flow infrastructure and evaluation methods.

These deficits are addressed by the recent developments in smart systems like gentelligent components. Novel approaches in this field such as intelligent workpieces [7], intelligent machine tools [10] and new online process assessment methods deliver additional information with a higher level of detail, a direct reference to a single workpiece and almost real-time information availability (e. g. current process quality information or machine error information). In the next chapter, online process assessment is introduced as a main input source for knowledge-based process planning.

4. Online process assessment

In order to provide workpiece-specific process information for a knowledge based process planning, an online process assessment method has been developed. Beside conventional process monitoring or PDA strategies, a process assessment is necessary for compensating the lack of product specific information during the production process. In the following sections a new method for machine internal data acquisition is proposed and demonstrated by the reconstruction of tool runout and surface roughness for an exemplary milling process.

4.1. Data acquisition

In order to acquire process state information in real time from the manufacturing process, a modern machine tool has been equipped by sensory components. The so called Gentelligent Machine Tool is able to acquire and monitor internal signals from the manufacturing process [10]. This new kind of machine tool is able to sense forces between tool and workpiece during a machining process. Based on the provided force signals, the machining process can be monitored online (Fig. 3).

The determination of the process forces is achieved by measuring the mechanical strain in the integrated components. The sensory z-axis-slide and clamping system of the machine tool were completely new designed under the aspect of sensor integration (e.g. strain sensitive sensors) without affecting the stiffness of the components (Fig. 3).

The applied force on the tool can be captured with an actual resolution of 10 N by a sampling rate of 500 Hz. The available cutting force signals provide process information for a process or machine monitoring system and thereby can be used for the detection of cutter breakage or process instability [11]. Additionally, the cutting forces can be used for the reconstruction of process and workpiece conditions, which is demonstrated in the following section.
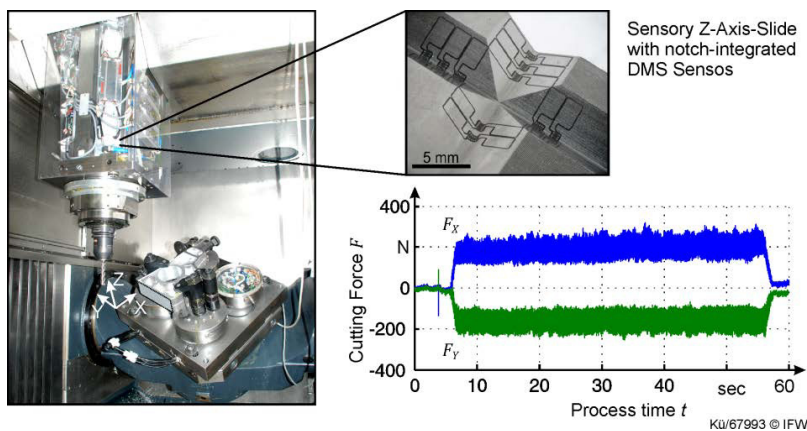

Fig. 3. Developed sensory machine components, integrated strain gauges in notches and recorded cutting forces.

4.2. Reconstruction of tool runout and surface roughness based on cutting forces

The surface topography of a milled surface in feed direction is characterized by the kinematic roughness profile. Beside the influence of process parameters of a milling process, the kinematic surface roughness is drastically influenced by vibrations and tool runout [12].

In order to assess the influence of tool runout on surface roughness in milling, an online determination of tool runout is required. Furthermore, the cutting forces from the gentelligent machine tool (Fig. 3) are used to identify the current runout parameters and to estimate the influence of surface roughness on the workpiece. Based on the extraction of characteristic signal features from cutting forces, a reference value was derived, which describes the effect of runout on the resulting cutting forces. In order to identify the tool runout parameters (amount ρ and phase

angle λ) accurately, a numerical optimization problem has been set up. The optimization problem is given by the identification of unknown runout parameters based on given process parameters and given signal features from cutting forces. For the identification of runout parameters, the estimated parameters have to be varied until the difference between the characteristic signal feature and the numerical computed reference feature is minimized. For a fast determination of runout parameters, a Nelder-Mead Simplex-algorithm was applied on the optimization problem. The developed runout identification algorithm has been tested extensively for various tools and process conditions. Thereby, the precision of runout identification has been determined statistically. The developed algorithm provides a precision of $\pm3\mu m$ for the amount of runout ρ and $\pm3°$ for the phase angle of runout λ [13].

In order to assess the influence of tool runout on surface roughness of the machined workpiece, the simplified equation for the calculation of kinematic surface roughness is not applicable. Hence, a geometrical process model has been developed. Therefore, the cutting edge movement is calculated analytically. The trajectory of cutting edge movement is influenced by the tool runout, which is calculated by the cutting edge radius R_i:

$$S_i = \begin{bmatrix} -R_i \cdot cos(\varphi - \phi_i) \\ R_i \cdot sin(\varphi - \phi_i) \end{bmatrix} \tag{1}$$

$$R_i = \sqrt{R_0{}^2 \cdot 2R_0\rho \cdot cos(\lambda - \phi_i) \cdot \rho^2}, \text{ with } \phi_i = \frac{2 \cdot \pi \cdot (i-1)}{n}, \text{ for } i = 1 \dots n \tag{2}$$

where R_0 is the nominal tool radius, ρ is the amount of tool runout, λ is the phase angle of runout and ϕ_i is the cutting edge position angle. The cutting edge radii influence the trajectory of the cutting edge S_i. The surface profile of the machined workpiece is generated by the Boolean intersection of the cutting edge trajectory S and the workpiece polygon WP:

$$WP_i = WP_i \cap \overline{S_i} \text{ , for } i = 1 \dots n \tag{3}$$

$WP(z)$ describes the shape of the workpiece at a specific height level z as a polygon. In order to assess the current surface roughness based on the identified tool runout, the functions have been implemented inside a process model, shown in the following Fig. 4.
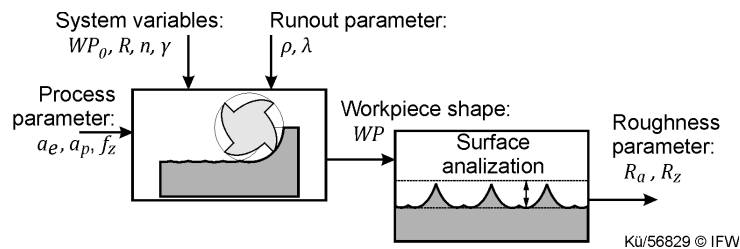

Fig. 4. Process model for the assessment of surface roughness.

The process model generates the workpiece polygon *WP*, based on given process and system variables and based on the identified runout parameters from the process. The workpiece polygon incorporates the characteristic kinematic surface roughness of a milled surface. In order to calculate the commonly used roughness parameters of the surface profile, a surface analysis function has been developed. The developed model-based approach delivers a detailed surface shape of the machined workpiece, which can be evaluated for roughness parameters over the engagement depths. An exemplary result for a milled surface is given in the following Fig. 5.
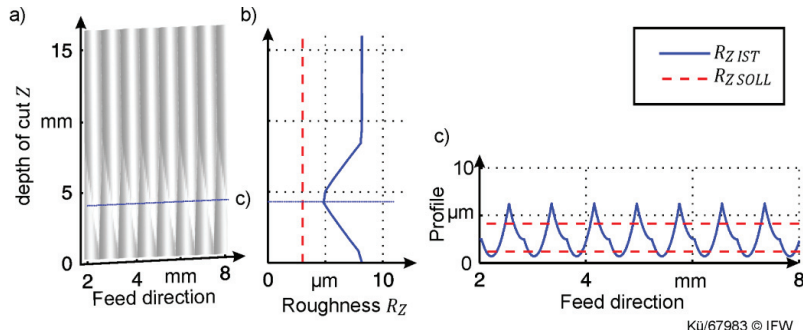
Fig. 5. (a) reconstructed surface shape; (b) reconstructed roughness over depth of cut; (c) reconstructed surface profile.

The process model delivers a three dimensional shape model of the machined surface, which incorporates characteristic feed marks from the milling operation (Fig. 5 a). Due to the tool runout error, the kinematic surface roughness can vary over the depth of cut (Fig. 5 b). Fig. 5 c) shows the surface roughness profile of a specific contour line of the surface, which is used to compute the kinematic surface roughness by roughness parameters R_z, R_a, R_{max} (cf. DIN 4768). Thereby, the roughness parameters of the machined workpiece can be reconstructed in parallel to the machining process and are provided to the knowledge base during the manufacturing process.

5. Approach to manufacturing data evaluation for knowledge-based process planning

5.1. Data structure

In order to build up a knowledge base for process planning, manufacturing data from the online process assessment (e.g. surface roughness, process time) has to be documented systematically considering related information from other information systems (see section 3). As a basis for systematic evaluation, an information model has been developed (Fig. 6).
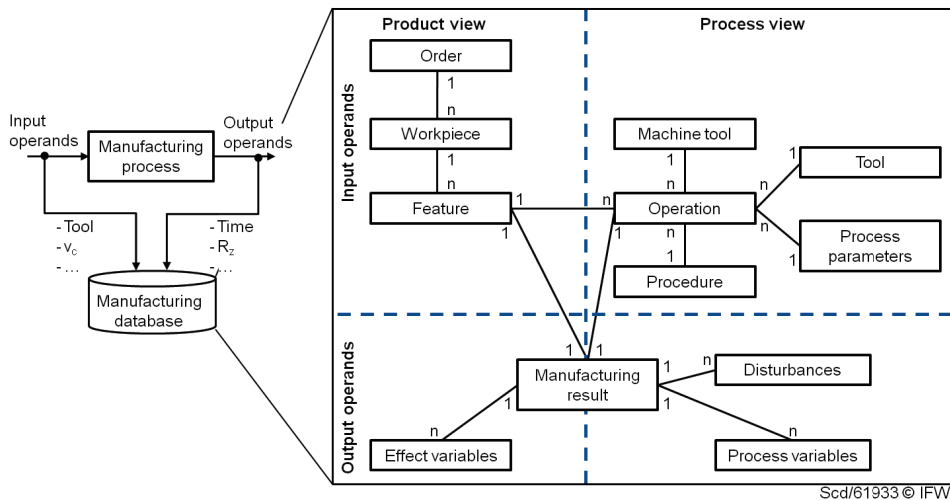

Fig. 6. Information model for the knowledge base.

The model determines the structure of the manufacturing database, which stores input and output operands for every manufacturing process. Taking into consideration that feature technology has become a standard for product modelling, the object "feature" was defined as the elementary object from the product view on the information

model. From the process view, the corresponding object is called "operation" which strictly refers to one feature. Hence, manufacturing results can be stored in the database connected to specific manufacturing operations. This allows rediscovering specific manufacturing information in the database.

5.2. Evaluation procedure

For the evaluation of the knowledge that is contained in the manufacturing database and its reuse in process planning, an evaluation procedure using data mining methods was developed (Fig. 7). Input data for the evaluation procedure is the product's feature sequence. This can also include alternative feature chains (e. g. for non-linear process planning, see [14]). Current specifications like material, tolerances, etc. are taken into consideration, too. In a first step, the remaining degrees of freedom (e. g. cutting depth or width) which have to be determined within process planning, have to be defined. Other attributes (e. g. tool radius limitations) are set as boundary conditions for the further evaluation (Fig. 7, step 1).

Then, all relevant datasets that fulfill the specifications are searched in the database. Result of this step is the identification of possible combinations of machine tools and cutting tools that have coped with the specification set in the past (Fig. 7, step 2). In order to predict the capability of a specific process with one specific set of parameters succeeding in a manufacturing operation to be planned, datasets have to be analyzed concerning similarity (e. g. in terms of cutting parameters, process forces etc.). Therefore, a cluster analysis using k-means algorithm is applied. As a quality criterion for clustering the Silhouette Coefficient is used [15]. The result of this step is finding groups of similar datasets, which represent a specific process (same or similar input operands, different output operands, Fig. 7, step 3).
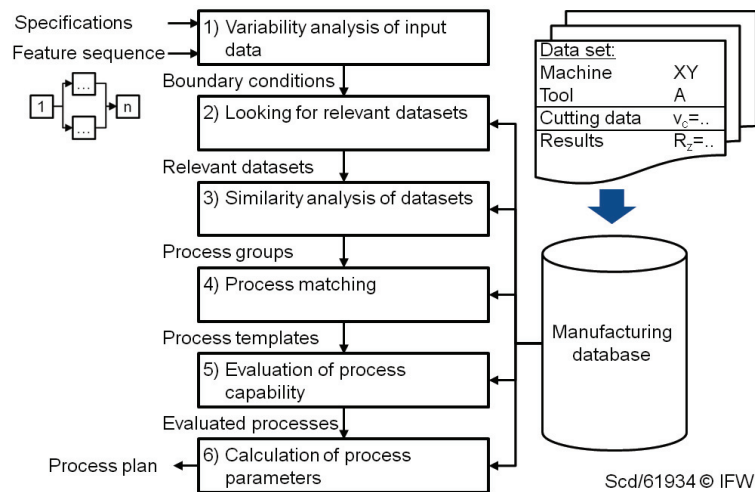

Fig. 7. Evaluation procedure.

Afterwards, those process clusters with the closest match to the specifications are selected. As the specifications depend on the planning task and can be seen as a dataset of characteristics, a k-nearest neighbor algorithm is used to identify the best process setting. This is done for every combination of machine tool and cutting tool identified in step 2. This step delivers process templates each based on a cluster of datasets (Fig. 7, step 4). As there could be more than one alternative in resources and processes, possible operations have to be evaluated concerning process capability in terms of the planned manufacturing task. Therefore, a capability index referring to critical process capability C_{pK} that is known from serial production was defined (4).

$$C_{pCK} = \frac{min(UTL - \tilde{x}_{Cluster}; \tilde{x}_{Cluster} - LTL)}{3 \cdot s_{Cluster}} \tag{4}$$

The cluster capability index C_{pCK} is calculated for a tolerated characteristic x considering the upper tolerance limit (*UTL*), the lower tolerance limit (*LTL*) and the spread of the characteristic value within the cluster. On the basis of the C_{pCK} index it is possible to identify the optimum process setting for a specific manufacturing task (Fig. 7, step 5). In the last step, appropriate process parameter sets for every possible operation are calculated from the average cluster values (Fig. 7, step 6).

5.3. First application

In order to test the evaluation procedure, a first application based on an exemplary planning task has been carried out. In need of a broad basis of documented manufacturing process data, a testing case containing 10,000 datasets has been generated by means of random numbers. In order to ensure that the datasets correspond with realistic manufacturing conditions, experimental machining data of about 100 datasets have been analyzed. The testing case was elaborated by the variation of parameters and results of the experimental datasets. Thus, datasets representing machining operations performed on two machine tools, using end-milling tools with diameters from 2 to 30 mm for aluminum and steel material and different cutting conditions were generated as testing data.

Table 1 shows the planning results for an exemplary milling operation. As manufacturing task the machining of a flat face of an aluminum workpiece was defined. Further specifications are the material (G-AlSi7Mg wa), a tolerated surface roughness <3 µm and a machining allowance of <1 mm, which determines the cutting width a_e for side milling.

Table 1: Proposed alternative operations for the process plan

Operation ID	30	20	3	37	4
Process	Milling	Milling	Milling	Milling	Milling
Machine tool	No. 2	No. 1	No. 1	No. 2	No. 1
Tool radius [mm]	4	2	15	5	6
Cutting speed v_c [m/min]	210.00	190.00	185.00	210.00	200.00
Feed rate v_f [m/min]	3.44	5.26	6.00	2.89	1.79
Cutting depth a_p [mm]	4.30	2.20	0.66	3.50	2.70
Material removal rate Q_W [mm³/s]	246.37	193.02	65.89	168.51	80.75

As a planning result, the five optimum operations identified by means of the evaluation procedure are issued as alternative operations for the process plans. In order to select an operation for the process plan, the operations have to be compared concerning the expected results in tolerated characteristics (e. g. surface roughness) and efficiency. Fig. 8 shows the evaluation of the expected surface roughness based on the related data cluster and the calculated material removal rate.
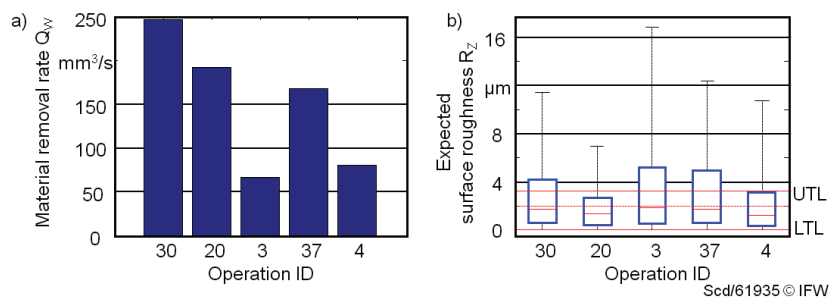

Fig. 8. Evaluation of proposed operations for the process plan: (a) calculated material removal rate; (b) expected surface roughness.

The evaluation procedure delivers a decision support for the process planner to select an optimum operation regarding productivity and quality factors for the process plan. In this case, operation no. 30 offers the most efficient machining process (Fig. 8, left). On the other hand, operations no. 20 and no. 4 lead to the optimum surface roughness. In this case, operation no. 20 would be the preferable operation for the process plan.

6. Conclusion and outlook

Smart systems in manufacturing deliver a high density of detailed manufacturing data. As manufacturing knowledge acquisition is one of the key factors to increase flexibility and efficiency in process planning, an approach to manufacturing data collection and evaluation using data mining has been presented in this contribution. An analysis and classification of manufacturing data has been carried out to identify input data for knowledge-based process planning. It was displayed, how smart objects' information is used for online assessment of the manufacturing quality. The proposed method enables the online assessment of the influence of tool runout on surface roughness based on observed cutting forces of a milling process.

Finally, the evaluation procedure was explained and applied with a testing case. Thus, the function of the approach to identify manufacturing operations for process planning by means of a knowledge base could be approved. The presented approach serves as a decision support for process planning and also offers the possibility to identify alternative processes e. g. in case of missing resources or for rescheduling in case of disturbances. As the input data for the manufacturing database has a major impact on the planning results, the range, the consistency and the level of detail of the manufacturing data should be examined in the next step. In this context, the performance and the limits of the approach have to be investigated.

Acknowledgements

References

[1] Trentesaux D, Thomas A. Product-Driven Control: Concept, Literature Review and Future Trends. In: Borangiu T, Thomas A, Trentesaux D. Service Orientation in Holonic and Multi Agent Manufacturing and Robotics. In: Studies in Computational Intelligence 472. Heidelberg: Springer; 2013.

[2] Uhlmann E, Hohwieler E, Kraft, M. Selbstorganisierende Produktion mit verteilter Intelligenz. Intelligente Werkstücke steuern ihren Weg durch die Fertigung. wt Werkstattstechnik online 2013;103(2):114-117.

[3] Xu X, Wang L, Newman ST. Computer-aided process planning – A critical review of recent developments and future trends. International Journal of Computer Integrated Manufacturing 2011;24(1):1-31.

[4] Fayyad U, Piatetsky-Shapiro G, Smyth P. From Data Mining to Knowledge Discovery in Databases. AI Magazine 1996;17(3):37-54.

[5] Köksal G, Batmaz I, Testik MC. A review of data mining applications for quality improvement in manufacturing industry. Expert Systems with Application 2011;38:13448-13467.

[6] Sánchez López T, Ranasinghe DC, Patkai B, McFarlane D. Taxonomy, technology and applications of smart objects. Information Systems Frontiers 2009;13(2):281-300.

[7] Denkena B, Mörke T, Krüger M, Schmidt J, Boujnah H, Meyer J, Gottwald P, Spitschan B, Winkens M. Development and first applications of gentelligent components over their lifecycle. CIRP Journal of Manufacturing Science and Technology 2014;7:139-150.

[8] Schneewind J. Entwicklung eines Systems zur integrierten Arbeitsplanerstellung und Fertigungsfeinplanung und -steuerung für die spanende Fertigung. PhD thesis, RWTH Aachen. Aachen: Shaker Verlag; 1994.

[9] Werner M.. Simulationsgestützte Reorganisation von Produktions- und Logistikprozessen. PhD thesis, Techn. Univ. Munich. Munich: Herbert Utz Verlag; 2001.

[10] Denkena B, Litwinski KM, Brouwer D, Boujnah H. Design and analysis of a prototypical sensory z-slide for machine tools, Production Engineering - Research and Development 2013;7:9-14.

[11] Litwinski KM. Sensorisches Spannsystem zur Überwachung von Zerspanprozessen in der Einzelteilfertigung. PhD thesis, Leibniz Universität Hanover. Hannover: PZH Verlag; 2011.

[12] Benardos PG, Vosniakos GC. Predicting surface roughness in machining: a review. International Journal of Machine Tools and Manufacture 2003;43:833-844.

[13] Krüger M, Denkena B. Model-based identification of tool runout in end milling and estimation of surface roughness from measured cutting forces. International Journal of Advanced Manufacturing Technology 2013;65:1067-1080.

[14] Lorenzen LE. Entwicklung einer Methode zur gentelligenten Arbeitsplanung. PhD thesis, Leibniz Universität Hanover. Hanover: PZH Verlag; 2012.

[15] Kaufman L, Rousseeuw PJ. Finding Groups in Data: An Introduction to Cluster Analysis. Hoboken (New Jersey): Wiley; 1990.