Automatic Generation of Integrated Process Data Visualizations Using Human Knowledge

Felix Mayer(✉), Ulrich Bührer, Dorothea Pantförder, Denise Gramß, and Birgit Vogel-Heuser

Faculty of Mechanical Engineering, Institute of Automation and Information Systems, Technische Universität München, München, Germany
{mayer, buehrer, pantfoerder, gramss, vogel-heuser}@ais.mw.tum.de

Abstract. The increasing complexity of industrial processes leads to complex process visualizations. Amongst other things, this is often due to the fact that a visualization engineer does not have deep knowledge of all physical and logical relations inside a plant. Additionally, different operators have to work with the same visualization despite the fact, that their personal preferences, abilities and needs differ. Having to work with unclear and confusing visualizations leads to an increased workload for plant operators and thus to higher error rates. Due to cost and time constraints, creating better or user-specific visualizations manually is not possible, especially because taking the operators' specific knowledge and experience into account is difficult.

This paper presents a concept to automatically generate process visualizations and support systems by the usage of a knowledge base and an influence model. This allows for operator-specific visualizations, considering preferences, abilities and needs. It also eases the visualization engineer's work by automatically choosing suitable diagrams and their properties. Additionally, by providing a system to acquire the operator's knowledge, complex relations inside a plant can be made accessible and utilized for optimizing the production process and visualization.

Keywords: Knowledge base · Visualization · 3D · Operator support

1 Introduction

Nowadays, visualizations of industrial processes need to display more and more data, as processes get more complex and devices offer more data. These visualizations are used by different personnel, e.g. technicians and process operators, for monitoring tasks. While working with such complex visualizations, the operator needs to closely monitor all relevant data, basing decisions on his observations, while performing additional tasks, such as administrative tasks. Consequently, the operator's workload and mental demand is increased. Operator's tasks are challenging and can only be

S. Yamamoto (Ed.): HIMI 2015, Part II, LNCS 9173, pp. 488–498, 2015.
DOI: 10.1007/978-3-319-20618-9_49

accomplished by skilled and trained personnel. An appropriate and elaborated mental model, which is developed through experience, is required to handle these complex tasks. This is the base for the operator's assessment of the current situation in process monitoring and leads to a selection of operator actions. Additionally, in case of critical situations, the operator's actions are often highly time-critical and errors while handling these situations may have extensive consequences, such as environmental pollution. Taking these points into account, the human-machine interface between the plant and the operator is therefore an important part for any industrial process that has to be designed thoroughly.

Designing visualization for a process is a complex task, done by engineers that preferably possess deep knowledge of the plant and process. Additionally, the visualization engineer needs deep knowledge of human-machine interaction and programming in general. While designing the visualization for a given industrial process, the engineer needs to decide on how and where the process and all relevant process values should be displayed within the visualization. In process industry, this is often solved by simply displaying an abstract 2D representation of the plant's geometry and by placing the process values close to the sensors that generate them. This often leads to cluttered and unclear visualizations that further increase the workload of the operator, who has to select and integrate the given information according to their mental model of the process.

However, the operators' mental models are not only based on monitoring of the process. Their knowledge of the process is often gathered throughout many years, enabling the operator to take appropriate actions in critical situations. This allows them to provide information beneficial to the production process, especially in recurring or critical situations. This is due to the fact, that operators possess great deals of experience and system knowledge of their plant, which has not been taken into account during the engineering of the plant and process. This experience may contain knowledge regarding parameters and their relations, detection of critical plant states, and general process optimizations. Hence, collecting this knowledge promises great potential to optimize the production process.

This paper introduces different interactive methods of knowledge acquisition and provides results of their first evaluation, according to their applicability to the production environment. These methods include text based approaches as well as visual approaches. Combining this collected information and big data from modern intelligent sensors, existing models of the plant can be optimized. These optimized plant models provide the foundation for further optimizations, which enhance the production process.

Additionally, this paper proposes a concept for the architecture of a knowledge-based approach for automatically choosing and configuring diagrams for process visualizations in order to help the visualization engineer. The provided algorithm, which supports choosing adequate diagrams and diagram configurations, can simplify the process of designing visualizations.

Both concepts can be brought together to form a comprehensive approach to improve the operator's integration into the production process and reduce the operator's workload. The concept also closes the knowledge-loop between visualization engineer, process and operator.

2 State of the Art

As stated above, operators of industrial processes combine information from different sources in order to understand the current process and the plant's state.

The assessment of the current process and plant state requires appropriate situation awareness. Situation awareness is a model of the dynamic environment, which is the activated part of the operator's mental model. It involves the perception of environmental elements, the comprehension of their meaning and the projection of future states [End95]. The operator's mental model is a prequisite for achieving situation awareness [Joh83]. The mental model of operators contains components, relations and functions of the monitored plant and process. The operator's goals determine which mental model is selected for the current situation and consequently, leads the perception of elements at a given point in time and thus selection for further processing. The mental model enables the interpretation and comprehension of the perceived element of the dynamic environment and allows the projection of future states.

Depending on the task, different visualizations are chosen [VDI3699], this includes the utilization of different colors and shapes. Multiple guidelines and norms for the design of visualizations exist, e.g. VDI 3850-1 [VDI3850], VDI 3699 and ISO 9241 [ISO9241], all of them deal with aspects of representing real-world systems. None of those guidelines and norms deal with aspects of integrated process visualizations that offer significant benefits according to Wickens' Proximity Compatibility Principle (PCP) [WiAn90]].

With the PCP, Wickens defined influences on the generation of mental models and the effort associated by doing so. According to Wickens, the operator's workload can be reduced by following the principles of mental and spatial proximity. The fewer information sources there are and the better they are mentally and spatially integrated, the lower the information access costs are. Highly integrated visualizations thus minimize the effort needed to create a comprehensive mental model. Creating high spatial proximity is comparatively easy, as values that should be considered coherently should also be displayed close to each other on the screen. Integrating multiple values into as few diagrams as possible maximizes the effect of spatial proximity, because relations are already clearly displayed and thus do not have to be created by the operator. Mental proximity means that all relevant data is displayed in such a way that an operator is relieved by the data prepared for easy access and understanding. This leads to less frustration and less workload.

The PCP states that integrative tasks, meaning the gathering of information by combining data, e.g. set different data in relation to each other, benefits from a 3D representation. Tasks requiring a high level of attention, e.g. reading a value over a longer period of time, benefit from a 2D representation. Wickens states that even the integration of undetailed information already produces better results.

According to John et al. [JCSO01], the advantage of a 3D visualization depends on the given task. In accordance to Wickens, they showed in several experiments that the identification of specific data is difficult in 3D. In contrast, if integration of various data is required for specific tasks, 3D facilitates performance and reduces cognitive demand.

For the task of human information processing using mental models, Rasmussen [Ras86], Hacker [Hac80], Wickens [WiHo00] and Card et al. [CMN86] created different sequential, capacitive and quantitative models in order to describe the complex process by using different views. Those findings, as well as findings concerning three-dimensional visualizations, have to be considered while generating process data visualizations. Depending on the complexity and integration of the data, a three-dimensional presentation could be beneficial to the operator's workload. Wickens et al. [WML94] defined tasks for three levels of information integration (low, medium, high) and the suitability of 3D representations for those three levels. The benefits of 3D visualizations were also shown by Tory et al. [TKAM06]. An example of integrated 3D process data visualization was given by Mayer et al. [MPVH13].

Optimizing operator workload while gathering knowledge at the same time is a difficult task to accomplish, as operators should not be distracted from their work, especially during the process control task in a plant. With post-tests, important information gets lost, as misperceptions can be forgotten. Therefore, data is less reliable as there is a tendency to overgeneralize and over rationalize [NiWi77]. Consequently, it is most promising to measure knowledge about causes, actions and relationships right after a problem is solved [SkAu99].

3 Goals /Requirements

Derived from to the description of the current processes for designing process visualizations as well as the state of the art, the overall goal for a new concept is to simplify the initial creation, and its successive adjustment and adaptation. During the lifetime of industrial plants – typically up to 30 years – the electrical equipment and especially the software are appended and changed on multiple occasions. The current, purely manual processes of (re-)designing and (re-)implementing a visualization can thus be improved and accelerated by automating them based on specialists' knowledge, e.g. of engineers and operators. This also includes considering all relevant legal requirements such as colors of buttons and font sizes as well as safety regulations. Additionally, evolving research findings, e.g. Wickens' PCP, have to be taken into account, in order to reduce the operators' workload, which in turn reduces failure rates.

In order to be able to also include the knowledge of stakeholders that do not have engineering knowledge, the new approach should not depend on any programming language but should instead be model-based to reach a greater number of users.

Regarding interactive operator knowledge acquisition, the main goal is to maximize production potentials by optimizing the existing models of the plant, based on existing operator knowledge on causes and relations of process components. In order to not distract the operator from controlling the plant, the knowledge collection has to be unobtrusive. To automate the collection, seamless integration into the existing modeling environment is required. Since human inputs might be erroneous or imprecise, the knowledge acquisition must be able to handle uncertainties and consist of an inherent validation process for error detection.

4 Concept

The proposed architecture consists of two systems: the visualization generator and the knowledge acquisition system, designed to help the operator as well as the engineer and technical personnel. Both systems can be viewed separately but unfold their true potential only in conjunction with each other and aim at significantly reducing the overall workload of all stakeholders.

In Fig. 1, the comprehensive approach of the visualization generator and the knowledge elicitation system are displayed. In short, both subsystems share a central knowledge base (1). The visualization generator thereof generates a customizable influence model (2) that is afterwards processed by an algorithm (3), taking into account the operator's profile (4) and generating the final visualization code (6) to be displayed (7) by using inherent criteria (5). The knowledge elicitation system, surrounded by the dotted line, extracts part of the operator's (8) mental model (9), checks for validity (11) and stores the information in the central knowledge base.

Based on this concept, plants are able to support all human operators, which are in turn able to provide beneficial knowledge about their plant. Both systems are described in detail in the following.

4.1 Acquisition of Operator Knowledge

Operators (8) often possess valuable knowledge and experience of 'their' machine or plant, which can be utilized to improve the quality of existing knowledge bases (1) and optimize production processes (12). Therefore, it is beneficial to collect existing operator knowledge (9) with the goal to gain knowledge about causes and relations between different process components as well as corrective information in specific

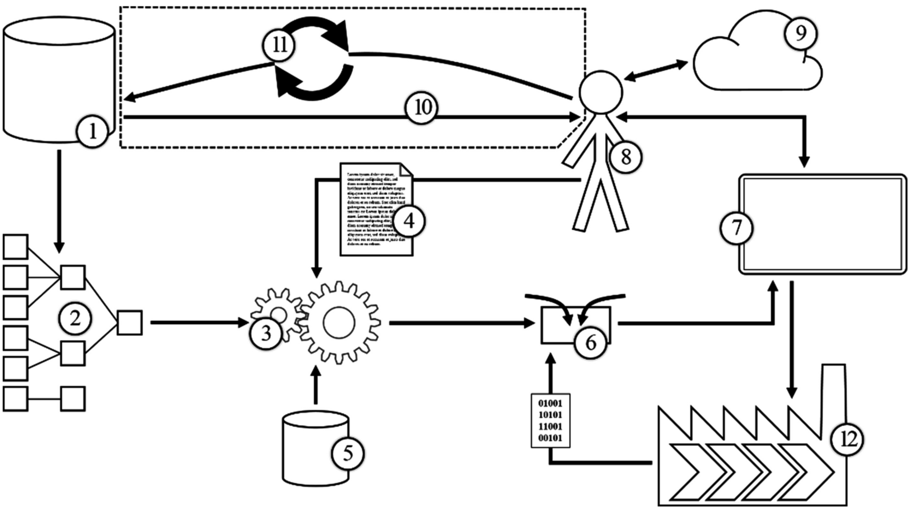

Fig. 1. Overview of the proposed architecture

situations in addition to the system model. In order to avoid unnecessary distractions from plant operation and to avoid operator's frustration, the knowledge collection must be unobtrusive and intuitive to use while also considering time efficiency.

For this purpose, different concepts for the acquisition of knowledge during plant operation are introduced in the following. In Fig. 2, an overview of the proposed concepts is given. Therein, problem trees, cause and effect graphs and input based on text blocks are illustrated exemplarily.

Problem Trees (PT) visualize causes and problem chains. Consequently, PTs aid the operator in detecting problem causes and, thereby, in performing maintenance tasks. To enable knowledge collection, the operator is able to enter additional problem causes within the existing problem tree, based on his experience and mental model of the plant.

Cause-Effect graphs (CEG) illustrate relations between different parameters of the production. Such a CEG is based on the existing knowledge base (1). In this context, the knowledge acquisition is enabled by allowing the operator to increase or decrease the influence one parameter has on another, thereby, the quality of the knowledge base is improved. In order to extend the existing knowledge base, the operator is also allowed to add additional parameters and further influences.

Text Blocks (TB), similarly to CEGs, depict dependencies between causes and effects. Furthermore, they also define the correlation between problems and their causes. For this purpose, text blocks of different characters are combined. The combination of a cause (e.g. Pump X clogged) with a specific value or intensity (e.g. high, low) with a problem (e.g. product viscosity) and its specific value through a correlation (e.g. causes, increases) enables a description of a problem and its cause. The combination of TBs regarding different parameters, their correlation and intensity, enables input and visualization of cause and effect correlations. As a result of the predefined TB-structure, the application of speech recognition technologies is

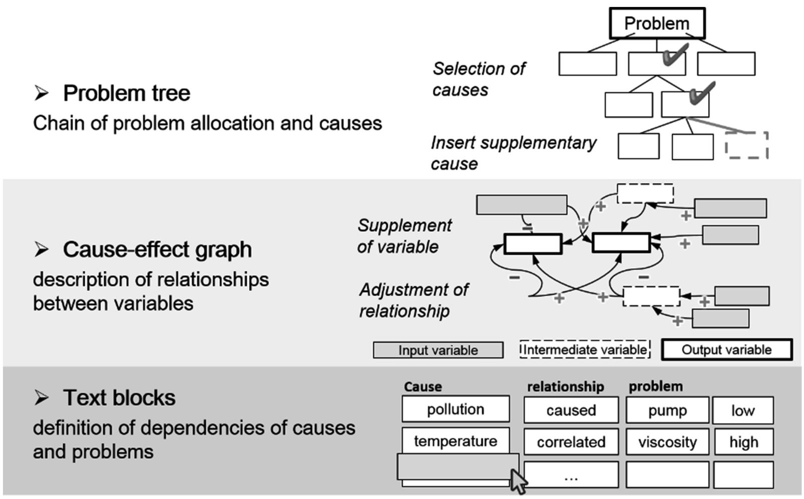

Fig. 2. Knowledge acquisition concepts

simplified and can be utilized to further reduce distractions for the operator during knowledge acquisition.

Since user interactions might be erroneous, the collected data must be verified and validated (11) before resulting process enhancements may be implemented into the production process. In order to allow verification of the collected knowledge and to continuously optimize the knowledge base and production processes, the individual operator interactions with the physical system are recorded and correlated to the collected knowledge. To avoid data privacy issues, the data collection is anonymized.

This approach to knowledge acquisition is aimed at closing the knowledge-support loop (cp. Fig. 1). Consequently, through adequate visualizations, plants are able to support human operators, which are in turn able to provide beneficial knowledge about their plant, further improving plant models and, thereby, recommended actions provided by the visualization. As plant safety is a major concern, the changes proposed by- and collected from human operators need to be verified for accuracy first, before comprehensive changes to the model of a plant or the production process can be propagated.

4.2 Visualization Generator

The visualization generator consists of different components, as displayed in Fig. 1. The overall state of a system, i.e. the current state of all actuators, sensors and products, can be extracted by combining different process values. This is normally done by the operator.

In this concept (Fig. 1), a knowledge base (1) is used to store all known interrelations between the technical system and its process, represented by process data of sensors and actuators. Thus, it can be used to deduce, which process values should be combined in order to extract additional information from a given set of process data. This state of the system can then be presented with different degrees of abstraction for different user roles. An operator e.g. needs more specific and detailed insight, than the management.

The knowledge base is directly linked to an influence model (2) that makes the knowledge base's contents accessible to humans. It also serves as one way to alter those contents, e.g. by adding new relations. Each influence model's nodes are either a data node (lowest layer), or an aggregation node (upper layers), connected by transitions without properties. Data nodes represent raw process data, whereas aggregation nodes are used to combine data from other nodes, i.e. aggregate and integrate data in order to provide more in depth information regarding the automation system. Clearly, also the amount of abstraction increases with each aggregation step. Aggregation of data can be performed in different ways, e.g. by logical functions in case of Boolean values. Each degree of abstraction calls for different means of visualization. One example of an influence model with implications is displayed in Fig. 3. The example consists of three abstraction layers, with raw process data at the bottom layer. When viewing this detailed first layer, the visualization could e.g. be done using a helix diagram (bottom right). The second layer, consisting of already aggregated data, has a different type of visualization, e.g. a radar chart. The third layer is even more abstract,

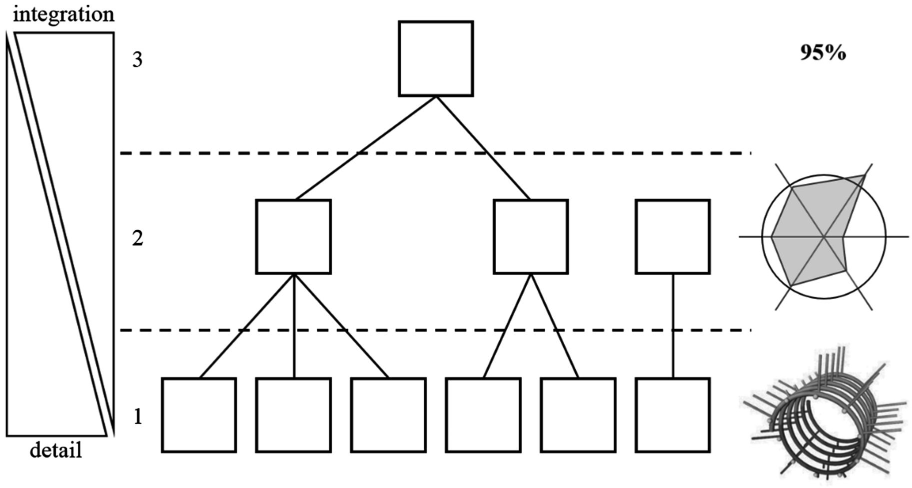

Fig. 3. The influence model

e.g. consisting solely of a simple number. That way different information for different roles is created and visualized.

The decision of what type of diagram is used to display certain nodes is not performed by the influence model, but an algorithm (3) that also considers the user's characteristics (4). The user's characteristics include its profile, e.g. color blindness, age, personal preferences, e.g. for specific diagrams and his role, e.g. operator or technical personnel. With the help of this information as well as the influence model, the algorithm chooses suitable visualization diagrams and their configuration, i.e. which colors to use, which axis to use for what value, etc. The algorithm uses a sets of rules (5) for this task, e.g. normally a temperature is displayed from blue (cold) to red (hot) in contrast to the green-to-red color gradient for other process data. Therefore, whenever a user starts using the visualization, the algorithm can create diagrams suitable for this user, as long as its characteristics are known. Afterwards it is possible to switch between the generated visualizations, based on the user.

The last step, before finally displaying the visualization on a screen (7), is to fill the diagrams with actual data from the industrial process (6). This last step differs from other visualization systems as it also applies the configured aggregations to the raw data.

Because visualizations for specific users can automatically be created after the initial model has been created, each user can thus have a customized visualization that fits its needs and preferences, thus lessening the workload.

In summary (cp. Fig. 1), all interrelations of a plant and its process are stored inside a knowledge base (1). Parts of the knowledge can automatically be elicitated and processed in order to form an influence model (2) between the process, the process values and the plant geometry. This influence model can be customized, e.g. by the visualization engineer, in order to be able to handle special situations and requirements. The influence model is processed by an algorithm (3) that chooses, in accordance to the operator's profile (4) and role, diagrams and diagram configurations based on inherent criteria (5). The operator's profile includes all personal preferences and characteristics and is used to customize the visualization. The visualization's diagrams are filled by raw or aggregated process data (6) and are displayed on a screen (7). With the help of these diagrams, the operator (8) creates a mental model (9) of the process by using the visual information provided. Aside from obtaining process information, he can also influence the process (12) through interaction with the visualization system.

The proposed approach allows for easy creation of process visualizations by providing a model-based approach for visualization engineers. It also simplifies the alteration of process visualizations during the lifetime of an industrial plant. At the same time, providing optimized diagrams, the operator's workload can be reduced. The approach also enables the integration of the knowledge elicitation system in a neat way. The extracted knowledge is utilized to extend the knowledge base and thus closes the knowledge loop.

5 Evaluation Study

As a first evaluation of the proposed concepts and their applicability to plant operation, a small study with experts from the manufacturing industry was carried out. Within this study, the application of PTs was seen critical, as the effort to enter new problems and their derived causes was seen as potentially tedious and, therefore, too distracting for the operative implementation. CEGs on the other hand, were seen as potentially applicable, due to their focus on point and click interactions. Regarding the input mechanisms, TBs were given the best rating, but their visualization was seen lacking. Therefore, a combination of the CEGs as a visualization tool and input mechanisms based on TBs were identified as the optimal knowledge collection tools. Therefore, the authors will focus future works on developing such a combined operator support system.

6 Benefits

By combining the two presented approaches, a closed knowledge loop is achieved. On the one hand, the knowledge gathered from the operator by the knowledge elicitation system is utilized to improve the diagrams and visualizations by altering the underlying influence model. Interrelations unknown at the time of engineering can be detected, acquired and implemented during the whole lifetime of a plant. Likewise, erroneous assumptions made during the engineering-phase can be corrected in later stages of a plant's life cycle.

On the other hand, the way in which information is presented by the visualization aids in extracting the operator's knowledge and storing it. This knowledge elicitation supports the identification of hidden plant interrelations, while minimizing interfering with the operator's work. The extracted knowledge of plant, process and product can e.g. be used to generate visualizations that show the required, relevant data, so that work is easier for new operators, which do not have comprehensive experience and knowledge. Providing optimized diagrams in respect to data integration and operator support reduces, according to Wickens PCP, the operator's workload and thus leads to fewer errors while operating a plant. Alternatively, operators are able to handle greater amounts of data while maintaining the same workload-level. Especially, regarding big-data aspects, the possibility to automatically aggregate and integrate data for visualization purposes, provides new possibilities.

All of these benefits can be achieved without having to write code, as the two approaches are model-based. This does not only influence the complexity to create an adapted visualization, but also the time required to do so.

7 Summary and Outlook

The work presented in this paper outlines the concept of an improved integration of the operator and the visualization engineer in day-to-day process control. By providing a system to automatically generate process visualizations from a knowledge base, containing physical and logical interrelations within an industrial plant, the individual needs and preferences of operators, e.g. concerning color-blindness, can be addressed. Likewise, physical and logical changes of the plant and new findings concerning the display of data can be taken into account, automatically adjusting the global knowledge base. In order to support plant evolution during its life cycle, the operator's knowledge about the process and the plant, which includes information about relations, and critical as well as recurring situations, is continuously acquired. Consequently, the plants models are constantly adapted, thereby optimizing the plant's models. This leads to an increased plant efficiency and improved operator support. To elicitate an operator's knowledge, multiple approaches were introduced.

Based on the first performed evaluation by industry experts, these concepts will be further evaluated in future work. Additionally, future work will consist of creating the individual modules of the presented approach and implementing the system as a whole. This also includes the evaluation of the presented approach against current and traditional approaches.

References

1. Card, S., Moran, T.P., Newell, A.: The Psychology of Human Computer Interaction. Lawrence Erlbaum Associates, Hillsdale (1986)
2. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Hum. Factors J. Hum. Factors Ergon. Soc. **37**(1), 32–64 (1995)
3. DIN EN ISO 9241-110: Ergonomics of human-system interaction (2006)
4. Hacker, W.: Allgemeine Arbeits- und Ingenieurpsychologie: Psychische Struktur und Regulation von Arbeitstätigkeiten. Huver, Bern (1980)
5. John, M.S., Cowen, M.B., Smallman, H.S., Oonk, H.M.: The use of 2D and 3D display for shape-understanding versus relative-position tasks. Hum. Factors J. Hum. Factors Ergon. Soc. **42**(1), 79–98 (2001)

6. Johnson-Laird, P.N.: Mental Models: Towards a Cognitive Science of Language. Inference and Consciousness. Harvard University Press, Cambridge (1983)
7. Mayer, F., Pantförder, D., Vogel-Heuser, B.: Design and implementation of an integrated, platform independent 3D visualization of complex process data. In: 12th IFAC IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems (HMS 2013), Las Vegas (2013)
8. Nisbett, R.E., Wilson, T.D.: Telling more than we can know: Verbal reports on mental processes. Psychological Review (1977)
9. Rassmussen, J.: Information Processing and Human-Machine Interaction: An Approach To Cognitive Engineering. North-Holland, New York (1986)
10. Skourup, C., Aune, A.: Decision support in process plants: visualisation of experiences. In: International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres, Bath (1999)
11. Tory, M., Kirkpatrick, A.E., Atkins, M.S., Möller, T.: Visualization task performance with 2D, 3D and combination displays. IEEE Trans. Visual Comput. Graphics **12**(1), 2–13 (2006)
12. VDI/VDE 3699 Blatt 2: Prozessführung mit Bildschirmen – Grundlagen (2005)
13. VDI/VDE 3850 Blatt 1: Gebrauchstaugliche Gestaltung von benutzerschnittstellen für technische Anlagen – Konzepte, Prinzipien und grundsätzliche Empfehlungen (2000)
14. Wickens, C.D., Merwin, D.H., Lin, E.L.: Implications of Graphics Enhancements for the Visualization of Scientific Data: Dimensional Integrality, Stereopsis, Motion and Mesh. Hum. Factors J. Hum. Factors Ergon. Soc. **36**(1), 44–61 (1994)
15. Wickens, C.D., Andre, A.D.: Proximity Compatibiliy and Information Display: Effect of Color, Space and Objectness on Information Integration. Hum. Factors J. Hum. Factors Ergon. Soc. Santa Monica **32**(1), 61–77 (1990)
16. Wickens, C.D., Hollands, J.G.: Engineering Psychology and Human Performance. Prentice Hall, New Jersey (2000)