International Journal of Advanced Robotic Systems

Cognitive Human-Machine Interface Applied in Remote Support for Industrial Robot Systems

Regular Paper

Tomasz Kosicki[1,2,*] and Trygve Thomessen[2]

1 Norwegian University of Science and Technology (NTNU) in Trondheim, Norway
2 PPM AS in Trondheim, Norway
* Corresponding author E-mail: tomasz.kosicki@ntnu.no

Received 24 May 2012 ; Accepted 04 Mar 2013

DOI: 10.5772/56296

Abstract An attempt is currently being made to widely introduce industrial robots to Small-Medium Enterprises (SMEs). Since the enterprises usually employ too small number of robot units to afford specialized departments for robot maintenance, they must be provided with inexpensive and immediate support remotely. This paper evaluates whether the support can be provided by means of Cognitive Info-communication – communication in which human cognitive capabilities are extended irrespectively of geographical distances. The evaluations are given with an aid of experimental system that consists of local and remote rooms, which are physically separated – a six-degree-of-freedom NACHI SH133-03 industrial robot is situated in the local room, while the operator, who supervises the robot by means of audio-visual Cognitive Human-Machine Interface, is situated in the remote room. The results of simple experiments show that Cognitive Info-communication is not only efficient mean to provide the support remotely, but is probably also a powerful tool to enhance interaction with any data-rich environment that require good conceptual understanding of system's state and careful attention management. Furthermore, the paper discusses data presentation and reduction methods for data-rich environments, as well as introduces the concepts of Naturally Acquired Data and Cognitive Human-Machine Interfaces.

Keywords Industrial Robotics, Small-Medium Enterprises, Cognitive Info-communication, Multimodal HMI, Attentive User Interface, Auditory Display, Data-Rich Environments, Bilateral Teleoperation, Internet

1. Introduction

Presently, the sector of non-financial Small-Medium Enterprises (SMEs) in Europe and other developed regions is facing a new challenge – high local labor costs force the enterprises either to use inappropriate automation solutions developed for capital-intensive large-batch manufacturing, or to move their production to under-developed countries and compete on the basis of low wages [1]. According to the report of Forge and

Blackman from 2010 [2], the number of SMEs in EU27 is estimated on 20.2 million, which is 99.8% of all non-financial enterprises. Among these, only 33% use industrial robots [3] in limited range of applications, and only 11% claim that they would introduce them within the next two years if suitable solutions were developed [3]. One might wonder why the robot density (robot/10,000 workers) remains several times higher in the sector of Large Enterprises (LEs) [4], and why industrial robots still encounter considerable difficulties in entering SMEs on a wide scale.

Traditionally, industrial robot systems were developed for automotive industry [5], where the number of industrial robot units is sufficiently large to provide them with constant professional supervision [6]. However, the supervision cannot be provided to SMEs due to economic reasons, and hence must be external. Since, in the majority of cases, robotic problems can be solved by professionals within minutes (e.g. the system is stopped due to pressed robot emergency stop which can be omitted by non-professionals), the professionals spend much more time on traveling to and from the site than on actual work. Obviously this causes significant increase in the costs of the support, and makes SMEs decide to not use industrial robots for fear of unjustified expenses [2]. One of the most promising approaches to overcome the reluctance is to provide SMEs with remote support over the Internet, either in a form of consultancy services [7], [8], or as complete support provided by periodical remote operation of industrial robots [9], [10], [11]. There have also been a large number of investigations concerning teleoperation of industrial robots [12], [13], [14], [15]. Though the purposes of these investigations were different, they demonstrated successful remote operation of industrial robots and provided relevant results.

The goal of this paper is to qualitatively investigate whether audio-visual Cognitive Info-communication (CogInfoCom), a new multidisciplinary field concerning communication among artificial and natural cognitive systems [16], [17], is an adequate method to provide remote support for industrial robots. When successful, such implementation can both vitally reduce service expanses for the systems by excluding travel costs, and provide knowledge on how Human-Machine Interfaces (HMIs) can be enhanced by means of Cognitive Info-communication.

The paper is structured as follows. First, Naturally Acquired Data (NAD), Cognitive HMI, and non-engineering concepts relevant for the investigation are introduced in the preliminaries. Then, the assumptions and structure of the experimental system are explained, and the results of quantitative observations are presented. Subsequently, data presentation and reduction methods are discussed, and the structure of Cognitive HMI is proposed. Finally, a conclusion is given.

2. Preliminaries

2.1 Naturally Acquired Data

When controlling an industrial robot system locally, one acquires data both by means of sensory devices, which present measured data via HMI, and by one's human senses, simply by direct system observation. In the remote control, however, one is physically separated from the system, and hence the data acquired by direct system observation, called Naturally Acquired Data (NAD), must be acquired and presented by an additional system (Fig. 1) or will be missed. The data are referred to as *Natural*, since similarly to Natural Language Processing [18], NAD refers to natural human capabilities:

Definition 1.
Naturally Acquired Data (NAD) *are the data that are acquired by human senses without use of artificial aids.*

Definition 2.
Natural Acquisition of Data *is the process in which human acquires data by human senses without use of artificial aids.*

Remark 1.
The concept of NAD is neither restricted to industrial robots, nor to remote operation. It can be applied for any data that are sensed by a human without use of artificial aids. For example, a driver (Fig. 2) acquires data through both direct system observation (NAD) and indirect system observation aided by, e.g., speed and fuel indicators (non-NAD).

Remark 2.
NAD cannot be easily expressed by means of numbers, words, or any other arrangements of ordered abstractive symbols, due to its complexity, strong context-dependency, and the frequent uniqueness of data entities. For example, one is neither able to easily express a complex traffic situation on a street only by means of numbers and words (Fig. 2), nor do thus with a state of an industrial robot cell (Fig. 1).

Since NAD is inherently symbolic, it shall be considered as symbolic data, which require different treatment than commonly encountered classical data. The difference between symbolic and classical data, the conversion between the types, and the principles of Symbolic Data Analysis (SDA), have all been explained by Billard and Diday in [19]. While appropriate tools to process NAD exist, the architecture of additional system acquiring and presenting NAD for remote operation of industrial robots still remains a challenge. In order to effectively comprehend system state and operate it, the remote operator shall be delivered with NAD providing sufficient degree of telepresence [20].

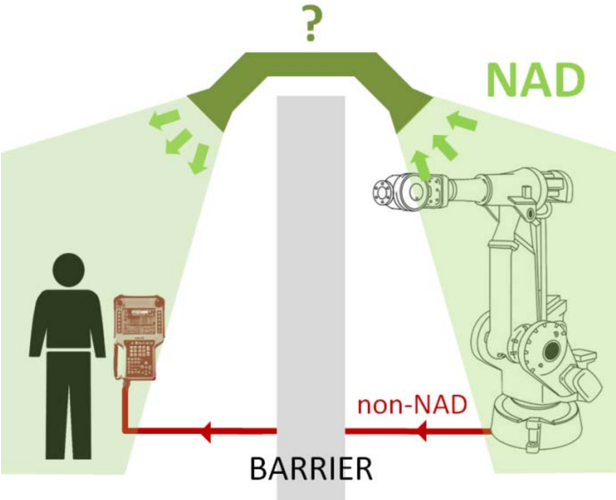

Figure 1. NAD in remote control of an industrial robot. The camera-like symbol expresses additional system used to provide NAD.

Figure 2. NAD in car driving.

2.2 Cognitive Human-Machine Interface

The majority of present HMI systems communicate with human through miscellaneous displays expressing data in a form directly inherited from books – by a mixture of structured text, numbers, and supporting illustrations, which has been adopted rather as a proven solution, than a result of multidisciplinary investigation involving engineers and psychologists. Though the mixture is accurate and unambiguous, it requires human brain to perform such mental operations as reading and counting, which have been developed in relatively recent human history due to the needs of civilization [21], [22]. Before civilization, survival in the natural environment caused other abilities to evolve, such as the ability to adaptively adjust sensing range [23], categorize perceived phenomena and objects [24], ignore insignificant and focus on significant ones [25], learn new categories [26], and solve encountered problems [27]. For instance when directly observing an industrial robot system, one can easily track several moving objects, estimate their motion parameters, and moreover envisage possible collisions and their consequences, by utilizing the abilities. On the contrary, the same data cannot be as efficiently provided by means of standard HMI, which is excellent for providing classical and fairly useless for symbolic data. There is a need, therefore, to develop a new type of HMI.

The most promising avenue appears to be development of multimodal HMI involving Cognitive Info-communication, since data acquisition and presentation in the communication are based upon the mentioned human abilities [16] – multiple senses are redundantly involved in parallel, and symbolic data is provided in a form of stimuli/action patterns corresponding to known psychological concepts:

Definition 3.
Cognitive Human-Machine Interface (CogHMI or Cognitive HMI) *is the HMI that allows humans to communicate with a machine by means of symbolic data.*

Remark 3.
In the fact, every HMI is to some extent a Cognitive HMI, since after a few uses even "a numerical displayer" becomes a unique concept of "the numerical displayer" that provides an unknown value of a known parameter. Cognitive HMI is used, therefore, to specify the HMI that at least partially is design to provide symbolic data. A trivial example of an everyday-use device which has Cognitive HMI is a TV set. Though it does not allow for real interaction as it is for Teleoperation [28] *and Virtual Reality* [29], *it provides symbolic data in the most primitive way, by direct reproduction of realty. For instance, when watching a speaker on the television news, one easily recognizes his or her emotions by facial expression and voice intonation.*

One can expect that the application of Cognitive HMI in the system will raise the efficiency of remote operation, since it is capable of providing NAD in an easily comprehended form. The use of Cognitive HMI will also allow for parallel data presentation, enhance human attention management [30], and make interaction less tiring, perhaps even enjoyable.

2.3 Non-Engineering Concepts in the Investigation

Orienting reflex is an involuntary body movement that orients an organism's receptors towards the part of environment in which sudden changes of stimuli occur [31]. For instance, when one suddenly hears the noise of a truck approaching from behind, one immediately turns one's head around to evaluate the danger by one's eyes.

Sensory adaptation occurs when a stimuli does not change and is at a safe range for the sensing organism [23]. For instance, when one is concentrated on reading an article, one does not feel the pressure exerted on one's foot in a slipper, since it is at a safe range. One will become, however, immediately aware of the pressure, when a child playing nearby steps on one's foot.

A *Concept* represents an entire class of objects (e.g., an apple), states (e.g., being warm), phenomenon (e.g., shining), or abstractions (e.g., truth, and the number four), that have common properties. One has an implicit knowledge of how concepts relate to each other, e.g., "dogs" are members of the larger concept "mammal", which includes also "humans" [32].

Categorization is a process of assigning an object to a concept. When categorizing an object, one treats it according to its properties associated with the concept, including the properties that were not perceived directly [24].

Localization and recognition are major functions of human perceptual system. The function of localization is used to spatially localize perceived objects, whereas by means of recognition, perceived objects are categorized into concepts [33].

Selective attention (e.g., selective listening) is the cognitive process in which one remains selectively focused on one aspect of the environment, while filtering or ignoring others. For example, when selectively listening, one is focused on only one sound [25].

Sensory substitution is the phenomenon in which perceptions characteristic of one sensory modality are made when another sensory modality is stimulated. For example, blind people can not only accurately recognize and localize objects by means of audition, but also perceive images comparable with visual. Sensory substitution is a result of a certain combination of needs and training [34].

3. Experimental System

3.1 Structure of the System

The system consists of local and remote rooms that are physically separated (Fig. 3). The industrial robot system is situated in the local room, whereas the operator, who exercises control over the system, is situated in the remote room. Communication between the rooms is executed by local and remote PCs implementing LabView software. The local PC is interfaced with the industrial robot system and additional sensory devices installed for the purposes of the system, whereas the remote PC is interfaced with the remote operator via HMI devices.

3.2 Sensory System in the Local Room

The robot controller (NACHI AX10CE-5111) is connected with local PC via NACHI Communication Software (RS232) and Real-Time Communication Unit (Ethernet); thus system data and movement of the robot manipulator

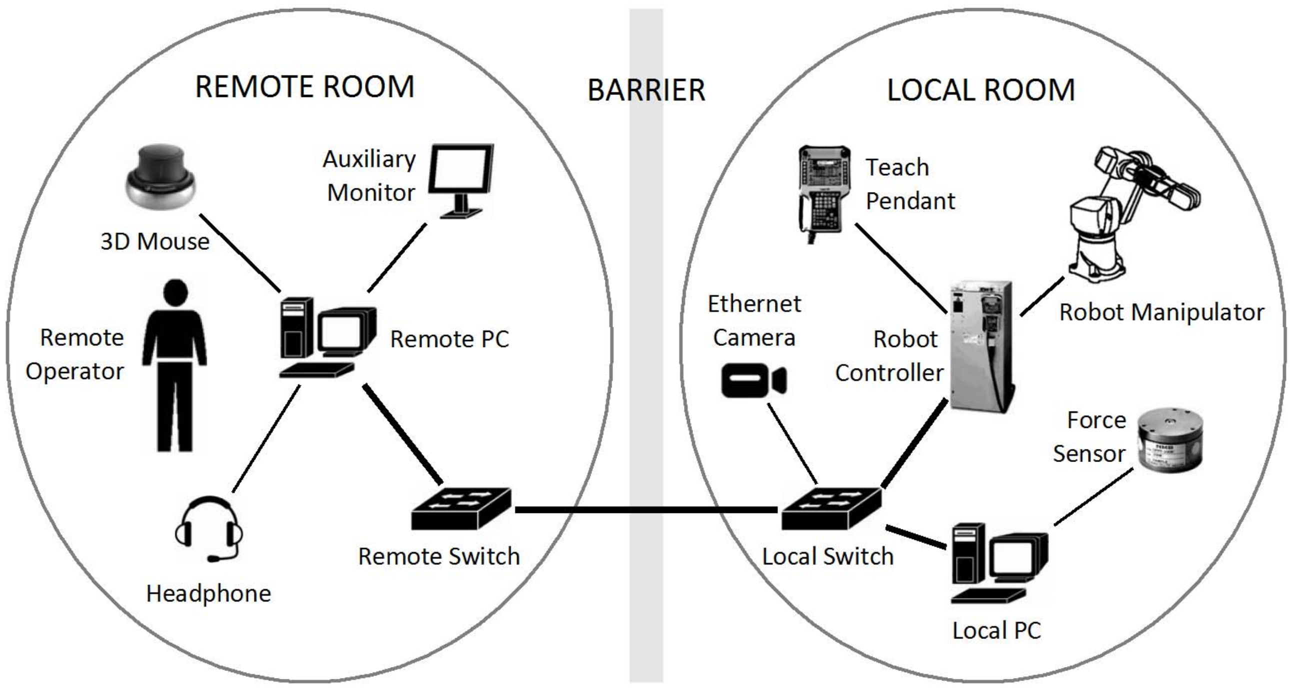

Figure 3. Structure of the experimental system.

No	Local Room			Remote Room		
	Source of Data	Acquired Data	Human Senses Involved in Sensing NAD	HMI Devices Providing the NAD	Human Senses Involved by the HMI	Data Expression
1	Robot Controller	the robot's axis movement	Vision/ Audition	Headphones	Audition	rising/falling cello melody for positive/ negative axis movement; spoken messages
2		the robot in singularity area	Vision	Headphones	Audition	disturbing sound with volume rising when immerging into the area; spoken messages
3		perpendicularity of the robot's tool to a surface	Vision	Headphones	Audition	consonance sound with volume rising when approaching the perpendicularity; spoken messages
4		close proximity to the robot's software endstop	Vision	Headphones	Audition	rhythmical beats with tempo rising when approaching the endstop; spoken messages
5		the robot's motors turned on/off	Vision/ Audition	Headphones	Audition	for motors turned on the data no. 1-4 can be perceived; spoken messages
6		the robot's emergency stop pressed	Vision/ Audition	Headphones	Audition	sound of alert, when pressed; spoken message
7	Force Sensor	occurrence of collision between the robot and environment	Vision/ Audition	Headphones	Audition	when occur, sound of crashing window; spoken message
8a		magnitude of contact forces developed on the robot's end effector	usually cannot be directly sensed	Headphones	Audition	triangle waveforms whose amplitude, distortion and panning are varying with regard to input data
8v				LCD Monitor	Vision	2D/3D graph of the forces (F_X,F_Y,F_z) that can be widely adjusted with regard to color and shape
9	Ethernet Camera	camera's vision	Vision	LCD Monitor	Vision	camera's vision

Table 1. Data acquisition and presentation.

(NACHI SH133-03) can be controlled in a real time. In addition, load cell force sensor (LORD 125/600 F/T), which is also interfaced with the PC (RS232) and is positioned on the end of the manipulator arm, allows the measurement of contact forces. The whole system can be observed remotely by means of moveable Ethernet camera (Panasonic BB-HCM515), whose pitch, yaw, and zoom are adjustable.

3.3 Human-Machine Interface in the Remote Room

When designing the experimental system, it was decided to incorporate only audio-visual HMI devices, since the devices are commonly available on the market and can be inexpensively applied. Another reason is that one can expect the dependences observed for human vision and auditory sense to be widely universal for other senses as well.

The HMI consists of two LCD monitors, headphone set, 3D mouse (3D Connexion Space Navigator) which is used to control the manipulator's movement remotely, and a standard PC mouse and keyboard.

3.4 Data Acquisition and Presentation

In Table 1, the data dealt with by means of Cognitive Info-communication methods is presented. In the part entitled "Local Room", source and type of acquired data are determined, as well as human senses involved in sensing the data locally. In the part entitled "Remote Room", one can see how the data are presented to the remote operator, and which of the human senses are utilized in the presentation. What is particularly worthy to notice in the table is that some of the data provided remotely are presented to different senses than they would be for direct system observation in local control (sensory substitution).

The data enumerated as 1-8 in Table 1 are provided also in a form of conventional numbers. They are, however, not included in the table, since they are insignificant for the following considerations.

4. Observations

4.1 Visual Presentation

The qualitative difference between the lack and presence of forces (Table 1, no. 8v) is immediately perceived when one is focused on the projection (Fig. 4). Conversely, when one is not focused, the difference cannot be perceived. Without selective (visual) attention, such projections therefore appear to be useless.

By means of 2D projections, one can easily evaluate quantitative differences between the forces, since the difference in length is clearly perceived (Fig. 4a). The

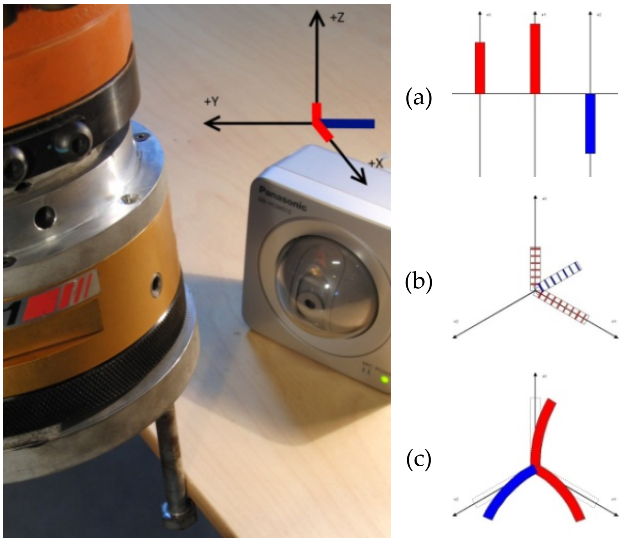

Figure 4. Camera's vision with visualization of the contact forces.

projections, however, do not provide any information about spatial relations between the forces. On the contrary, 3D projections (Fig. 4b-c), when spatially adjusted to the orientation of the camera view, allow one to comprehend and associate the relations with observed manipulator's movements by means of localization function.

Shape behavior of bars, which expresses contact forces, can be beneficially used to indicate their magnitude, only when it is associated with known concepts. For instance, bowed bar (Fig. 4c) can be easily associated with bowed ruler – one immediately associates its deformed shape with tactile sensations developed when bending the object. Similarly, a bar with moving strips (Fig. 4b) can be associated with liquid flowing into an elastic container – changes in a force are similar to changes in container's volume, or pressure developed in the container.

In contrast to shapes, one's abilities to recognize changes in colors, and to associate them with corresponding contact forces, appears to be strongly limited, since colors play minor role in recognition function [35]. Such clearly distinguishable colors as red and blue, however, can be effectively used to indicate opposite states. For example, negative and positive direction of forces (Fig. 4) can be associated with concepts "cold" and "hot".

4.2 Sounds Corresponding to Concepts

It was observed that the most efficient means of symbolic data presentation are sounds directly corresponding to known concepts (Table 1, no. 6 and 7). For instance, the sound of crashing window (no. 6) is immediately categorized into a concept of "collision", since in everyday life the sound is likely associated with a potential danger of injury. Likewise, the sound of alert (no. 7) is categorized into a concept of "emergency state", since it is similar to the sound of a fire fighters' or ambulance's siren, which is usually heard in emergency situations.

Though some of the sounds do not correspond directly to known concepts (Table 1, no. 1-3), when the sounds are clearly distinguishable and provide symbolic data, one can easily develop new concepts for them. For instance, one learns to associate a cello melody with the concept of "the robot's axis movement" (no. 1), since it is unlikely to hear any other cello when controlling the system. Likewise, the disturbing sound (no. 2) is associated with the concept of "singularity area", as it is equally undesired to hearing the sound. Deeply focusing consonance sound of organs (no. 3) is associated with the concept of "perpendicularity", which is equally desired to hearing the sound.

New concepts can be also developed by analogy to existing ones (Table 1, no. 4 and 8a). The density of rhythmical beats (no. 4), which rises when approaching the endstop, seems to be analogical to tactically perceived pressure – similarly to fast beating, high tactile pressure cannot be ignored; thus the closer is the endstop, the more one is forced to be focused on the sound (selective listening). Another such example is the sound of triangle waveforms (no. 8a). Above a certain threshold, the higher a contact force, the more corresponding waveform is distorted – the level of waveform distortion appears to be analogical to the level of roughness perceived tactically [36].

In general, it was observed that the sounds corresponding to concepts tend to be able to be used only to provide symbolic data.

4.3 Musical Sounds

The cognition of musical sounds (Table 1, no. 1-4 and 8a) is strongly dependent on individual musical skills. It has observed that well trained individuals can successfully acquire classical data with a limited resolution, while average individuals are merely able to recognize qualitative differences such as the difference between rising and falling melody (no. 1). Since the sounds are perceived rather as sounds corresponding to different concepts than musical sounds, one can develop new concepts for them to provide symbolic data. It was also observed that one quickly accustomed to consonant sounds (no. 3), and without conscious intension to follow them, one forgets about their presence. On the contrary dissonance sounds (no. 2) cannot be neglected; hence they are suitable to indicate undesired situations, e.g., robot approaches singularity. The overall drawback of musical sounds is that they require selective attention, and thus cannot be effectively comprehended in parallel.

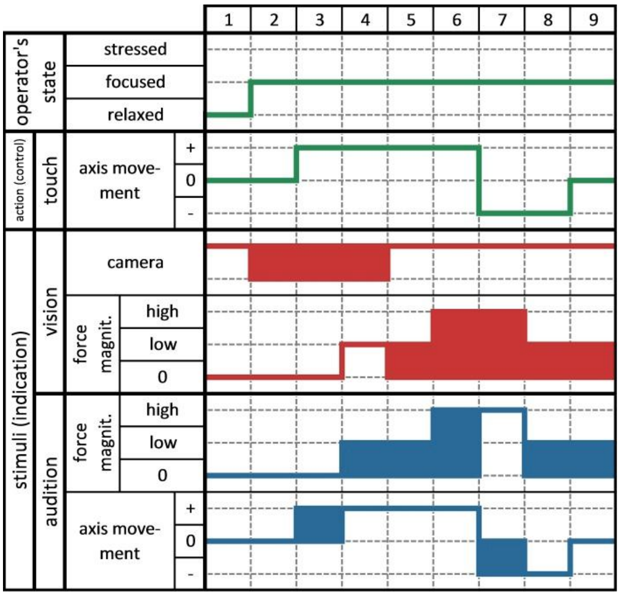

Figure 5. A typical robot operation performed remotely - adjustment of contact forces. The plot demonstrates quantitatively changes in human focus indicated by the filled parts, and the relationships between provided to human data and action.

4.4 Spoken Messages

The spoken messages appear (Table 1, no. 1-7) to be a beneficial addition to the sounds, since they can provide symbolic data, similarly as it is for two individuals exchanging the data in a conversation. One also rapidly learns to recognize the voice of speaker; thus a new concept is developed – “a spoken message in the system”. Since the content of the messages is commonly relevant for the control, the occurrence of the voice results in selective listening. The listening is, however, also a drawback of the method, since one neglects other sounds in a great range.

4.5 Mixed Audio-Visual Presentation

When sounds corresponding to known concepts are played in parallel, one does not encounter any significant difficulties in data comprehension. On the contrary, spoken messages and sounds utilizing the musical ear cannot be comprehended in parallel, since they require selective listening.

It is a remarkable fact that, without the sounds, when not observing the displays, one is not aware of occurring contact forces (Table 1, no. 8). However, when the sounds corresponding to contact forces (no. 8a) occur, one involuntarily turns one’s head/eyes into the display due to orienting reflex. In general, one usually at first realizes the occurrence of the forces by means of sound (no. 8a), and then verifies their magnitude via visualization (no. 8v), which for an average human provides more accurate data than the sound.

On Fig. 5 one can see an orienting reflex occurring in a typical robot operation – contact force adjustment. The appearance of contact force sounds (step 4), causes one to move focus from camera vision to visualization of contact forces (step 5). The operator’s state and the changes in auditory focus were subjectively reported by the operator, while the changes in visual focus and human actions were registered objectively by the camera recording gaze movement and the LabView software.

5. Discussion

The investigation elucidated in this paper has demonstrated the potential benefits of HMI based on unconventional communication channels, as well as allowed the reader to distinguish and evaluate data presentation methods. It also indicated how to reduce data in data-rich environments, and proposed a structure of Cognitive HMI.

5.1 Data Presentation Methods

Basing on the observations, three data presentation methods can be distinguished (q-DPM, Q-DPM and S-DPM: Table 2) in which, depending on the method and circumstances, data is comprehended in 1-3 steps (Table 3) - obviously, the fewer the steps in data comprehension, the shorter the comprehension time.

	Data Presentation Methods				
	Quantitative	Qualitative			Symbolic
Abbreviation	q-DPM	Q-DPM			S-DPM
Data Wrapping	Numbers	Text	Speech	Musical Sounds	Symbols
Conceptuality	no	yes	yes	yes	yes
Accuracy	unlimited	limited	limited	limited	strongly limited
Comprehension Time	long	moderate	moderate	moderate	negligible
Involuntary Comprehension	no	usually no	usually no	usually no	yes
Parallelization	no	no	no	no	yes
Multimodality	no	no	no	no	yes

Table 2. Qualitative comparison of data presentation methods. The cells with the desired properties were filled to clarify the table.

Step	Data Comprehension Process	Data Presentation Methods q-DPM	Q-DPM	S-DPM
1	Pattern recognition	yes	depends	No
2	Concept recognition	yes	yes	depends
3	Context recognition	yes	yes	Yes

Table 3. Data Comprehension for different data presentation methods. The cells are filled to clarify the table.

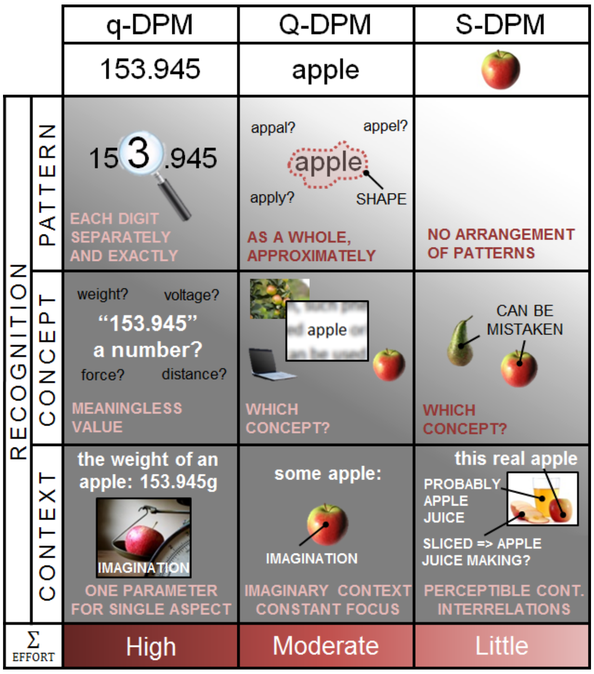

Figure 6. DPMs and typical features for exemplary data.

Since the mechanisms of human cognition are still not very well understood, the division is in a certain range contractual, and some exceptions can be found for it. For instance, while numbers undoubtedly convey quantitative data, and symbols undoubtedly convey symbolic data, one can argue that text, speech and musical sounds should not be classified as qualitative, since they can also convey symbolic data. In fact, at the final stage of each DPM, data are transformed into symbolical form as concepts, since humans usually use concepts for thinking.

5.1.1 Quantitative Data Presentation Method

Undoubtedly numbers are q-DPM, since without context they do not correspond to any non-quantitative meaning (Fig. 6). The remaining DPMs, on the contrary, provide data corresponding either to ambiguous concepts (Q-DPM and some S-DPM) or unambiguous concepts (S-DPM). For instance, the word "apple" found in a text, depending on the context, can refer to the concepts of "fruit", "apple tree", or the company "Apple" written in lower case. A real apple found on a tree, however, refers only to the unambiguous concept of "apple fruit", and in particular, to the concept of "the apple fruit", which has only one existing realization.

5.1.2 Qualitative Data Presentation Method

Text, speech, and musical sounds are Q-DPMs, since they are not restricted to quantitative data as q-DPM, and opposite to S-DPM usually provide data that convey meaning highly dependent on context (Fig. 6).

In Q-DPM, data are encoded into arrangements of contractual mono-modal patterns, such as graphemes, including Latin letters (visual), phonemes and tones (auditory), and many other less conventional patterns, such as Morse code letters (auditory) and Braille letters (tactical). Though the patterns are meaningless on their own, e.g., {"a","e","l","p","p"}, in an ordered sequence they can correspond to certain concept, e.g., "apple". Opposite to numbers, which are also ordered arrangements of patterns, in Q-DPM data can be correctly conceived even if some of the patterns are misunderstood or lost [37], and after certain training the sequence can be comprehended simultaneously as a whole [38]. For instance, healthy people generally have automated comprehension of phonemes (auditory), the majority of them have automated comprehension of graphemes (visual), and a minority, such as professional musicians, of tones (auditory). For the majority, comprehension of spoken words is superior, since it is trained from the earliest years of childhood.

Q-DPMs utilize artificial abilities which, even if brilliantly mastered, never become completely automated. To a certain degree they always require selective attention and usually cannot be parallelized, hence, for example, one cannot effectively follow written and spoken messages. Q-DPMs can be, nevertheless, used to accurately express all aspects of reality and human imagination, including symbolic aspects.

5.1.3 Symbolic Data Presentation Method

In S-DPM, one is exposed to stimuli patterns directly corresponding to concepts (Fig. 6). Opposite to Q-DPM, recognition of the patterns is accomplished in a negligibly short period of time, since it is based on innate human cognitive capabilities, and can be multimodal and parallel, e.g., the recognition of auditory and visual patterns corresponding to an event involving several objects. These properties make S-DPM the most suitable method to provide NAD, which primarily concern non-quantitative interdependencies among data. In the most straightforward form, S-DPM can be a mixed reality [29] in which the interdependencies are observed similarly to reality. However, one can also imagine S-DPM in the form of dynamic multimodal pictograms, or even as a mixed reality with physics different than real world physics, hence relevant properties of objects and events can be comprehended more easily.

5.2 Data Reduction in Data-Rich Environments

Nowadays, data acquisition is no longer a challenge [39]. The great number of easily available, inexpensive sensors and the high bandwidth of current communication standards enable the development of complex

Distributed Sensor Networks (DSN) which automatically acquire large amounts of data. The present challenges are, instead, efficient information and knowledge extraction, and efficient data presentation for human.

In biological systems, such as with humans, a small number of senses provides large amounts of data, which at the early stage of processing are reduced into the data necessary at that point in order to operate. The data is subsequently generalized into concepts among which again only a few reach consciousness. Despite that reduction, one is still capable of effectively perceiving complex states and dependencies of the surrounding environment. If data in data-rich environments, such as the remote operation of an industrial robot system, is not treated analogically, data comprehension becomes ineffective. Therefore, the two main reasons why the environments are usually difficult to operate in are information overloading (too much data at once without generalization) and no information filtering (presenting all the data, including irrelevant data).

Though one is able to measure the everyday environment with almost unlimited accuracy, usually a general understanding of the environment is sufficient to effectively operate in it. First of all, the HMI shall provide only generalized data that are easy to comprehend and sufficient to understand the system state. Then, it can provide more details, either if the need for them is inferred from the current system state and user's performance, or if it is actively requested by the user through the HMI. The investigation has shown that there are three DPMs that can be synergically used in the approach – system state shall be presented mainly on the level of S-DPM, whereas the levels of Q-DPM and q-DPM shall be entered only if details are necessary (Fig. 7). In addition, such phenomena as the observed orienting reflex can be used to support attention management and switching between DPMs.

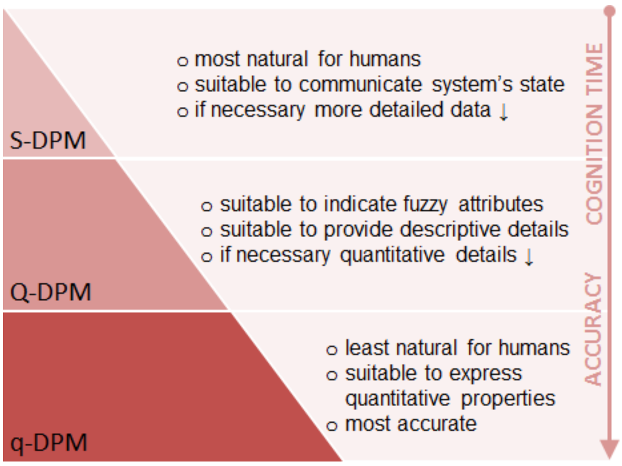

Figure 7. DPMs for data-reduction in data-rich environments.

5.3 Structure of Cognitive Human-Machine Interface

The investigation has exposed the limitations of the existing HMI, which cannot be overcome by simple improvements. Basing on the obtained results and the multidisciplinary discussion presented previously, we have proposed a structure of Cognitive Human-Machine Interface (CogHMI, Fig. 8) which allows reducing or even completely removing these limitations.

Opposite to traditional HMI, CogHMI adjusts its performance to current human needs, acting rather like an active mediator between the human and the data system, than a passive data displayer operated by a control panel. CogHMI is capable of guessing human needs, adaptively selecting data (Information Filtering and Importance Evaluation), and providing the data in the most convenient form for the human (Cognitive Info Encoding). Similarly, as it is for two natural cognitive systems, e.g., two humans, the data between CogHMI and the human flow in two feedback loops – conscious and unconscious – and the performance of the CogHMI subcomponents is adjusted with regard to estimation of the human and system states. While the system state is evaluated by means of a distributed sensor network and a system model, e.g., a model of an industrial robot cell, the human state is evaluated by means of a customizable human cognitive model and multimodal input devices. The devices provide both conscious input (Human Requests), such as keyboard typing, and unconscious input (Human Needs), such as eye movements.

Though named differently, certain CogHMI functionalities have already been developed, and some of them can even be found in commercially available systems. CogHMI's functionalities can be found, for instance, in the systems which monitor driver fatigue, lack of attention, and sleepiness [40], and in the systems detecting drunk driver such as the commercially available Driver Alcohol Detection System for Safety (DADSS) [41]. Both of them evaluate the human state without conscious

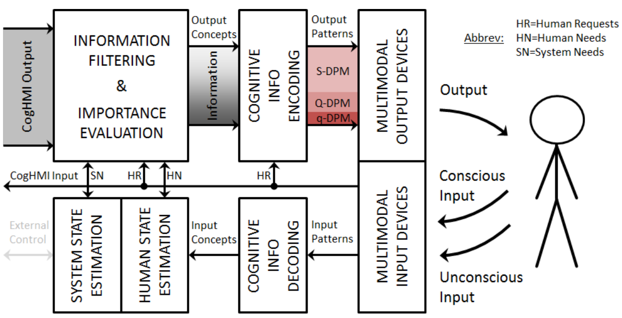

Figure 8. Structure of CogHMI - the HMI with conscious and unconscious feedback loops. The stages of data processing, which were discussed in detail previously, are emphasized using colours.

use of the systems to support human performance. Another more sophisticated example is Intelligent Space (iSpace) [42], which is an environment that attempts to evaluate human needs by means of a distributed sensor network, and to satisfy them with the aid of service robots.

6. Conclusion

The constantly rising complexity of data systems and the amount of data to be handled reveal shortcomings in the existing HMI systems. In this investigation, an unconventional multimodal HMI based on Cognitive Info-communication methods was developed to provide remote support for industrial robot systems – dynamic, data-rich environments. Simple quantitative experiments, which were performed using the system, have shown that the HMI can be used to provide the necessary beneficial support. The experiments have also become a starting point for multidisciplinary analysis involving cognitive science, which has allowed the introduction of the concept of Cognitive HMI, and we have discussed a number of other issues related to modern HMIs. Among these, we have distinguished three Data Presentation Methods in HMI, analyzed data reduction methods, and finally proposed the structure and functionalities of Cognitive HMI. The basic idea lying behind cognitive HMI is maximization of human data comprehension efficiency by adaptive adjustments of the HMI with regard to estimated human and system states.

7. Acknowledgments

Thanks are due to EEA Grants and the Norwegian Research Council to partly supporting this research through the projects, "HU0045 HUNOROB" and "SkatteFUNN 40333 Cognitive man-machine communication for industrial robots", respectively. In addition, thanks to HERD ICT for the finance of the project BANOROB.

8. References

[1] "SMErobot," [Online]. Available: www.smerobot.org.

[2] S. Forge and C. Blackman, "A helping hand for Europe: the competitive outlook for the EU robotics industry," Publications Office of the United Europe, Luxembourg, 2010.

[3] S. Kinkel, "Potenziale und Nutzen von Robotersystemen für kleine und mitielständische," in *SMErobot Final Project Workshop*, Stuttgart Vaihingen, 2009.

[4] IFR International Federation of Robotics, "World Robotics 2011 Industrial Robots," IFR Statistical Department, Frankfurt am Main, 2011.

[5] J. N. Pires, "From the coworker to the cognitive factory scenario: new development for SME manufacturing," in *Congress on Welding Technology and Manufacturing*, Saltillo, 2008.

[6] IFR International Federation of Robotics, "European Robotics, a white paper on the status and opportunities of the European robotics industry," European Robotics Forum (IFR ERF) & European Robotics Research Network (EURON), 2001.

[7] P. Horst, G. Schreck and C. Willnow, "SME-service networks for cooperative operation of robot installations," in *Emerging Solutions for Future Manufacturing Systems*, vol. 159, L. M. Camarinha-Matos, Ed., Springer Science, 2005, pp. 339-346.

[8] V. B. Sunil and S. S. Pande, "WebROBOT: internet based robotic assembly planning system," *Computers in industry*, vol. 54, no. 2, pp. 191-207, 2004.

[9] J. Sales, R. Fernandez, J. M. Jimenez, R. Marin and P. J. Sanz, "Telecontrol of an industrial robot arm by means of a multimodal user interface: a case study," in *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, The Hauge, 2004.

[10] B. Rajapakse and N. Hildreth, "Remote access toolkit for industrial robots," *Journal of Engineering Manufacture*, vol. 218, no. 8, pp. 1017-1022, 2004.

[11] T. Thomessen and T. Kosicki, "Multimodal human-machine interface for remote operation of robot systems," in *10th IFAC Robot Symposium on Robot Control*, Dubrovnik, 2012.

[12] R. T. Bombang, "Development of architectures for internet telerobotics systems," *Journal of Bionic Engineering*, vol. 4, no. 4, pp. 291-297, 2007.

[13] A. Rodríguez, E. Nuño, L. Palomo and L. Basañez, "A multimodal teleoperation framework: implementation and experiments," in *40th IFR International Conference on Robotics and Automation*, Barcelona, 2009.

[14] E. Veras, K. Khokar, R. Alqasemi and R. Dubay, "Scaled telerobotic control of a manipulator in real time with laser assistance for ADL tasks," *Journal of the Franklin Institute*, vol. 349, no. 7, pp. 2268-2280, 2012.

[15] A. J. Álvares, G. C. de Carvalho, L. F. A. Paulinyi and S. C. A. Alfro, "Telerobotics: trough-the-internet teleoperation of the ABB IRB 2000 industrial robot," in *Proceedings of SPIE - the International Society for Optical Engineering*, Boston, 1999.

[16] P. Baranyi and Á. Csapó, "Cognitive Infocommunications: CogInfoCom," in *Computational Intelligence and Informatics (CINTI), 2010 11th International Symposium on*, Budapest, 2010.

[17] Á. Csapó and P. Baranyi, "A unified terminology for CogInfoCom applications," in *Cognitive Infocommunications (CogInfoCom), 2011 2nd International Conference on*, Budapest, 2011.

[18] M. Bates, "Models of natural language understanding," *Proceedings of the National Academy of Science of the United States of America,* vol. 92, no. 22, pp. 9977-9982, 1995.

[19] L. Billard and E. Diday, "From the statistics of data to the statistics of knowledge: symbolic data analysis," *Journal of the American Statistical Association,* vol. 98, no. 462, pp. 470-487, 2003.

[20] D. W. Schloerb, "A quantitative measure of telepresence," *Presence: Teleoperators and Virtual Environments,* vol. 4, no. 1, pp. 64-80, 1995.

[21] A. Ardila, "On the evolution of calculation abilities," *Frontiers in Evolutionary Neuroscience,* vol. 2, no. 7, pp. 1-7, 2010.

[22] P. Gordon, "Numerical cognition without words: evidence from Amazonia," *Science,* vol. 306, no. 5695, pp. 496-499, 2004.

[23] B. Wark, B. N. Lundstrom and A. Fairhall, "Sensory adaptation," *Current Opinion in Neurobiology,* vol. 17, no. 4, pp. 423-429, 2007.

[24] E. Rosch, "Principles of categorization," in *Cognition and Categorization*, E. Rosch and B. B. Lloyd, Eds., Hillsdale, Erlbaum Associates, 1978, pp. 27-48.

[25] J. Driver and R. S. J. Frackowiak, "Neurobiological measures of human selective attention," *Neuropsychologia,* vol. 39, no. 12, pp. 1257-1262, 2001.

[26] B. H. Ross, E. G. Taylor, E. L. Middleton and T. J. Nokes, "Learning and memory: a comprehensive reference," in *Concept and Category Learning in Humans*, J. H. Byrne, Ed., Oxford, Academic Press, 2008, pp. 535-556.

[27] A. Newell, J. C. Shaw and H. A. Simon, "Elements of a theory of human problem solving," *Psychological Review,* vol. 65, no. 3, pp. 151-166, 1958.

[28] T. B. Sheridan, "Telerobotics," *Automatica,* vol. 25, no. 4, pp. 487-507, 1989.

[29] G. C. Burdea and P. Coiffet, Virtual Reality Technology, 2nd ed., New Jersey: John Wiley & Sons, 2003.

[30] R. Vertegaal, J. S. Shell, D. Chen and A. Mamuji, "Designing for augmented attention: towards a framework for attentive user interfaces," *Computers in Human Behavior,* vol. 22, no. 4, pp. 771-789, 2006.

[31] E. N. Sokolov, "Higher nervous functions; the orienting reflex," *Annual Review of Physiology,* vol. 25, pp. 545-580, 1963.

[32] K. L. Slaney and T. P. Racine, "On the ambiguity of concept use in psychology: is the concept "concept" a useful concept?," *Journal of Theoretical and Philosophical Psychology,* vol. 31, no. 2, pp. 73-89, 2011.

[33] I. Biederman, "Recognition-by-components: a theory of human image understanding," *Psychological Review,* vol. 94, no. 2, pp. 115-147, 1987.

[34] P. Bach-y-Rita, "Sensory substitution and the human-machine interface," *Trends in Cognitive Sciences,* vol. 7, no. 12, pp. 541-546, 2003.

[35] A. L. Ostergaard and J. B. Davidoff, "Some effects of color on naming and recognition of objects," *Journal of Experimental Psychology: Learning, Memory, and Cognition,* vol. 11, no. 3, pp. 579-587, 1985.

[36] Á. Csapó and P. Baranyi, "A conceptual framework for the design of audio based cognitive infocommunication channels," *Studies in Computational Intelligence,* vol. 378, pp. 261-281, 2012.

[37] G. Rawlinson, "Reibadailty," *New Scientist,* vol. 162, no. 2188, p. 55, 1999.

[38] J. S. Adelman, S. J. Marquis and M. G. Sabatos-DeVito, "Letters in words are read simultaneously, not in left-to-right sequence," *Psychological Science,* vol. 21, no. 12, pp. 1799-1801, 2010.

[39] R. Agrawal, T. Imielinski and A. Swami, "Database mining: A performance perspective," *IEEE Transactions on Knowledge and Data Engineering,* vol. 5, no. 6, pp. 914-925, 1993.

[40] T. D'Orazio, M. Leo, C. Guaragnella and A. Distante, "A visual approach for driver inattention detection," *Pattern Recognition,* vol. 40, no. 8, pp. 2341-2355, 2007.

[41] "Driver Alcohol Detection System for Safety," [Online]. Available: http://www.dadss.org/.

[42] J.-H. Lee and H. Hashimoto, "Intelligent space - Concept and contents," *Advanced Robotics,* vol. 16, no. 3, pp. 265-280, 2002.