

A Unified Terminology for the Structure and Semantics of CogInfoCom Channels

Ádám Csapó, Péter Baranyi

¹ Budapest Univ. of Technology and Economics
Dept. of Telecommunications and Media Informatics
Magyar Tudósok Krt. 2
1117 Budapest, Hungary
E-mail: csapo@sztaki.hu

² Cognitive Informatics Research Laboratory
Computer and Automation Research Institute
Kende u. 13-17
1111 Budapest, Hungary

Abstract: Due to its multidisciplinary origins, the elementary concepts and terminology of Cognitive Infocommunications (CogInfoCom) would ideally be derived – among others – from the fields of informatics, infocommunications and cognitive sciences. The terminology used by these fields in relation to CogInfoCom are disparate not only because the same concepts are often referred to using different terms, but also because in many cases the same terms can refer to different concepts. For this reason, we propose a set of concepts to help unify the CogInfoCom-related aspects in these fields, from the unique viewpoint necessary for the human-oriented analysis and synthesis of CogInfoCom channels. Examples are given to illustrate how the terminology presents itself in engineering applications.

Keywords: Cognitive Infocommunications, CogInfoCom channels, icon layer, message layer, conceptual mapping

1 Introduction

Cognitive infocommunications (CogInfoCom) investigates the link between the research areas of infocommunications and cognitive sciences, as well as the various engineering applications which have emerged as the synergic combination of these sciences. According to its definition, the primary aim of CogInfoCom is to better

understand how natural cognitive processes can effectively co-evolve with processes in artificially cognitive systems [5].

CogInfoCom views any kinds of hardware or software component that encapsulates information and allows users to interact (communicate) with this information as infocommunication systems. Depending on the complexity required for an infocommunication system to obtain this information (e.g., through sensing or inference), it is said that the system can have various levels of cognitive capabilities (hence the term *artificially cognitive systems* [5]). In a large part of our research, our goal is to create CogInfoCom channels through which artificially cognitive systems can communicate information to the user in a way that appeals to the user's cognitive capabilities.

In this respect, CogInfoCom has common interests with other fields such as human-computer interaction (HCI), multimodal interaction, affective computing and sensory substitution. These fields all focus on novel forms of communication between humans and machines. However, a departure from these fields is often necessary for CogInfoCom, primarily because the types of information that must be communicated between artificial systems and the user may not always be directly perceptible by the available sensory and cognitive subsystems (hence the need for sensor-bridging). In other cases, the transformation of information from one sensory modality to another (sensor-bridging) and from one representation to another (representation-bridging) may provide more effective interaction between the user and the system. These two basic needs for the extension of scope compared to existing fields were eloquently highlighted by Bach-y-Rita and his colleagues [3]:

However, in the context of mediated reality systems, which may incorporate multiple modalities of both sensing and display, the use of one sense [...] to display information normally acquired via another human sense [...] or alternatively via a 'non-natural' sense such as sonar ranging, could be considered to be a form of sensory augmentation (i.e., addition of information to an existing sensory channel). [...] We therefore suggest that, at least in multimodality systems, new nomenclature may be needed to independently specify (a) the source of the information (type of environmental sensor, or virtual model); (b) the type of human information display (visual, auditory, tactual, etc.); and finally (c) the role of the information (substitutive or augmentative), all of which may play a role in reality mediation.

The definition of CogInfoCom addresses many of these propositions. More importantly, Bach-y-Rita and his colleagues clearly recognize that although sensory substitution is sufficient in describing many applications, it could be valuable to broaden the scope of sensory substitution so that it can be used to describe many forms of communication between humans and machines, even if the source or destination of the communication cannot be described using the traditional senses of the human nervous system.

From the point of view adopted in our research, the three major fields which are unified by CogInfoCom (informatics, infocommunications and cognitive sciences) have different roles. Thus, from our point of view, informatics deals with how information is represented in artificially cognitive systems, infocommunications deals with how this information can be transmitted physically, and cognitive sciences deal with how this information will be perceived and interpreted by humans, and whether or not – if presented in a different way – the information could be perceived and interpreted more easily.

In this paper, we propose a taxonomy of concepts which can be used to both analyze and synthesize CogInfoCom channels between users and infocommunication devices with different cognitive capabilities.

The paper is structured as follows. In section 2, we outline the proposed terminology for the description of various levels of communication between cognitive processes in humans and artificially cognitive systems. In section 3, the various levels are detailed further in terms of specific sensory modalities. The section also provides a model of how it might be possible to achieve effective mapping between different channels of multi-modal communication. Finally, in section 4, concrete applications of CogInfoCom are described in terms of the proposed terminology.

2 Unified Terminology

It is important to observe that the definition of CogInfoCom adopted an actor-based view of cognitive infocommunications [5]. Thus, CogInfoCom systems were categorized based on the relative level of cognitive capabilities (i.e., the categories of *inter-cognitive* and *intra-cognitive* infocommunication were defined based on this viewpoint) and the kinds of sensors and representations used on the two ends of communication (i.e., the categories of *sensor sharing*, *sensor bridging*, *representation sharing* and *representation bridging* were defined based on this viewpoint).

In this paper, a data-flow based view is adopted, so that the communication patterns between the actors on the two ends of communication can be better analyzed. The conceptual framework we propose has a layered structure and is composed, from bottom to top, of an icon layer, a compound icon layer, a message layer, a concept layer, and finally a concept mapping layer. The layers in the conceptual framework represent stations in information flow which fulfill a unique transformation in representation or interpretation so that various sensory channels can become semantically connected at various levels of abstraction. The layers are defined as follows:

Definition 1. The **icon layer** contains sensory percepts that give rise to immediate and unique semantic interpretations.

Definition 2. The percepts contained in the icon layer are referred to as **CogInfoCom icons**.

Remark. *Examples of CogInfoCom icons are Canfield-Smith's visual icons (tradi-*

tionally referred to as icons in graphical environments) [43], Gaver's auditory icons [17, 18] and Kaye's olfactory icons [30]. Gaver was the first to generalize Canfield-Smith's icon concept to modalities other than vision, by deriving the theory of auditory icons through a distinction between everyday listening and musical listening [18]. Most CogInfoCom icons in the icon layer were defined based on Gaver's work, however, we broaden Gaver's concept by using the term 'icon' to refer to any relatively short signal that is capable of evoking direct connotations which are not necessarily physical in nature.

Definition 3. The **compound icon layer** contains sensory data combined from several modalities which give rise to immediate and unique semantic interpretations.

Definition 4. The multi-modal percepts contained in the compound icon layer are referred to as **compound CogInfoCom icons**.

Remark. *An example of compound CogInfoCom icons is the haptic icon [34, 15]. Icons of the haptic sense are derived from more than one modality because more than a single, more basic sense can contribute to the haptic experience (namely, the cutaneous and kinesthetic/proprioceptive senses) [32, 47]. Depending on the research community, the exact meaning of these basic senses can vary (an effort was made in [36] to provide a comprehensive synthesis of definitions).*

Definition 5. The **message layer** contains abstract, sequential messages which are built up of elementary icons from the icon layer and/or compound icons of the compound icon layer.

Definition 6. The messages contained in the message layer are referred to as **CogInfoCom messages**.

Remark. *Examples of CogInfoCom messages are earcons [8] and smicons [30]. In general, the layer includes perceptual messages that are abstract in their interpretation (i.e., the meanings of the messages are not immediately obvious and have to be learned).*

Remark. *The notion that earcons, for example, are built up of auditory icons is not entirely new (a compositional relationship is briefly mentioned in [24]), however, the overwhelming majority of publications describe auditory icons and earcons as contradictory and mutually exclusive categories [8, 40, 33, 49]. From the point of view of CogInfoCom, it makes sense to regard the message layer (e.g., earcons) as a set of messages constructed out of elementary icons (e.g., auditory icons), because this view broadens the scope of CogInfoCom messages instead of restricting it. More specifically, from a syntactic point of view, earcons can be composed of any kinds of building blocks – or motives [8] – (e.g., sinusoidal notes, violin notes, flute notes, sounds of different kinds of doors being slammed, etc.) since every one of these building blocks can become auditory icons if heard often within the same context in a natural environment. From a semantic point of view, it is a viable solution to structure various kinds of CogInfoCom icons into well-defined messages, whenever the CogInfoCom system must communicate several kinds of iconic information types at the same time.*

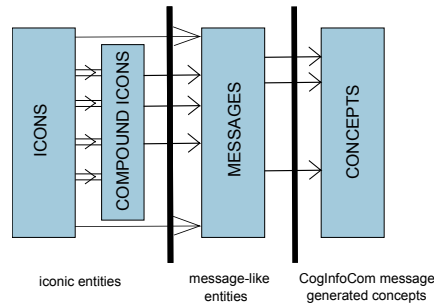


Figure 1

The first three layers of the proposed conceptual framework from bottom to top

Definition 7. The **concept layer** contains the set of all concepts which can be generated by the message layer in a given sensory modality.

Definition 8. The concepts within the concept layer are referred to as **message layer generated concepts**, or **CogInfoCom message generated concepts**. Thus, depending on the sensory modality, we may speak specifically of, e.g., **earcon generated concepts** and **smicon generated concepts**.

Remark. *Examples of earcon generated concepts are smoothness, raspiness, consonance and dissonance, because audio messages can be described using such concepts. Examples of smicon generated concepts include sweetness, stinkiness, and perfume smell.*

Definition 9. The **concept mapping layer** describes the associations which can be created – by the cognitive capabilities on the receiving end of communication – between message layer generated concepts of different sensory modalities. Through the associations contained in the concept mapping layer, concepts generated through different sensory modalities can be linked together at various levels of abstraction.

Remark. *The associations created in the concept mapping layer can be also viewed as occurring between different CogInfoCom channels (see Definition 10).*

Remark. *Whenever associations are formed between two different sensory modalities, the modality that is replaced by a new one is referred to as the **substituted modality**, while the new modality which interprets CogInfoCom messages is referred to as the **substituting modality** [1].*

The first four layers and their relationships in terms of composition and generation are shown in Figure 2. Based on the framework, it is possible to define CogInfoCom channels as follows:

Definition 10. A **CogInfoCom channel** is a group of CogInfoCom messages used to carry information on the state of a CogInfoCom message generated concept.

Remark. *The concept of CogInfoCom channels arises as an analogy to earcon families [8].*

Remark. *The framework we propose has both user-centered and system-centered aspects, in the sense described by Jokinen [29]. On the one hand, both CogInfoCom icons and messages are defined in terms of human sensory modalities. On the other hand, our definition of CogInfoCom channels is in agreement with the multimodal framework proposed by Vernier, Nigay and Coutaz [46] in the sense that two CogInfoCom channels can be different even if they are presented through the same sensory modality, as long as the language they use (e.g., the parameter types used to generate the set of messages) is different.*

3 Components of CogInfoCom layers

In the following subsections, we describe the components of each of the layers with respect to different sensory modalities. Further, a description is provided of the various forms of association which can be used by the concept mapping layer to create semantic links between different CogInfoCom channels.

3.1 The icon and compound icon layer

A large number of icons can be defined depending on the sensory modality that is being investigated. This is especially true because any number of compound modalities can be created if we consider different kinds of abstract and emotional information to be perceived / understood through modalities of their own (for example, if somebody is exceptionally sensitive to the various manifestations of emotional state in facial expressions or other forms of body language, it could be argued that the emotional information is perceived through an emotion-sensing modality; further, artificial modalities can be defined for abstract, aggregated data such as traffic patterns [48], ambient patterns such as lighting patterns [4], interaction scenarios between mobile ambients [44], etc.). In the following, we provide examples of CogInfoCom icons across a non-exhaustive list of sensory modalities. Wherever the icons have previously been defined, we provide reference to the earlier works.

- *Visual icons*: often abbreviated as "icons", these basic components of every graphical user interface were first defined by David Canfield-Smith, a PhD student working at Xerox in the early 1970's [43].
- *Auditory icons*: defined by Gaver in the context of every-day listening as caricatures of everyday sounds [17, 19].
- *Olfactory icons*: scent outputs used to convey information, where the scents are environmental and semantically related to the information which is being conveyed [30]

- *Haptic icons*: brief computer-generated signals, displayed to a user through force or tactile feedback to convey information such as event notification, identity, content or state [34]. As previously mentioned, we categorize haptic icons as compound CogInfoCom icons due to the compound nature of the haptic sense. It is interesting to note that MacLean and Enriquez admit that haptic icons are philosophically closer to earcons than auditory icons in that they are message-like [34], but in a different work, they write "*haptic icons, or 'hapticons', [are] brief programmed forces applied to a user through a haptic interface, with the role of communicating a simple idea in manner similar to visual or auditory icons*" [15]. Further, in [14], the same authors introduce the concept of haptic phonemes as "*the basic building blocks of haptic communication*". This clearly demonstrates the fact that compound icons are neither iconic, nor message-like, but lie between the two layers. Perhaps this is because the haptic sense itself is so multi-faceted. To resolve the contradiction, we propose to refer to icon-like haptic percepts as haptic icons (which we propose to mean the same thing as haptic phonemes), and to message-like abstract haptic percepts as hapticons.
- *Emotional icons*: Due to the compound nature of emotional sensing – in a way similar to haptic icons – we define emotional icons as concentrated representations of emotional information (the term "emoticon" is often used in the literature for what we may refer to as graphic emotional icons: however, it may be worthwhile to define similar icons through other modalities, such as speech-based emotional icons).
- *Kinesthetic icons*: We refer to icon-like feedback provided to the user through the kinesthetic sense as *kinesthetic icons*. An example of a kinesthetic icon would be various heart-rate patterns generated by different kinds of user activity (e.g., [22]).
- *Body icons*: We refer to elementary actions performed by the user of a system as *body icons*. The importance of body icons emerges from the fact that it is only through body icons that the user can stimulate the system (i.e., effect state changes in the system, or apparent behavior by the system). Body icons can be used to trigger other kinds of icons within the CogInfoCom icon layer in the natural and artificial world, as well as various kinds of messages within the CogInfoCom layer in artificial systems. The key difference between body icons and kinesthetic icons is that body icons are voluntary actions performed by the user, while kinesthetic icons are provided by the system (or the environment). It is interesting to note that the effects of kinesthetic icons may culminate in explicit body movements, although these movements are inadvertent on the user's part.

The unifying principle among CogInfoCom icons is that they are compact, iconic representations of information.

3.2 The message layer

The message layer contains messages created by temporally overlapping and/or juxtaposing a set of CogInfoCom icons. While CogInfoCom icons carry direct associations, CogInfoCom messages are more abstract in nature, although, as we will see later in the section on conceptual mapping, there are quite a few possibilities for making CogInfoCom messages easier to interpret. Again, due to the freedom we take in defining compound sensory modalities, we provide a non-exhaustive list of CogInfoCom messages:

- *Eyecons*: we define eyecons as visual messages which are used to convey information that has only an abstract relationship with the data it expresses
- *Earcons*: nonverbal audio messages used in the user-computer interface to provide information to the user about some computer object, operation, or interaction [8]
- *Smicons*: scents used to convey information that has only an abstract relationship with the data it expresses [30]
- *Hapticons*: we break up Enriquez and MacLean's terms of haptic icons and hapticons into two different levels. As described earlier, we use the term *haptic icon* when referring to haptic signals that are iconic (compound icon layer), and we use the term *hapticon* when referring to haptic messages that convey information which has only an abstract relationship with the data it represents. By adopting this view, it can be seen that Brewster's tactons can be considered as specific kinds of hapticons for the tactile sense [10].
- *Emoticons*: we define emoticons as abstract messages built up of emotional icons that convey information which has only an abstract relationship with the data it represents.
- *Bodicons*: we define bodicons as abstract messages conveyed to the system through both implicit and explicit body movements (comprised of kinesthetic icons and body icons) effected by the user.

3.3 Structure of the concept mapping layer

The mapping between various message-generated concepts can entail the emergence of a strong association between underlying CogInfoCom messages and icons. However, such mappings can be created in several different ways:

- *Direct mapping*: The conceptual mapping between two CogInfoCom channels is direct if direct representations of the data are transferred between the substituted and the substitute modalities, in the sense of directness first defined by

Norman [39]. We distinguish between *low-level* and *high-level* forms of direct mapping:

- *Low-level direct mapping*: the data is interpreted as a time series which directly controls the substituting CogInfoCom message. This mapping technique has been referred to as the lexical technique in earlier literature (e.g. [9]). Examples of low-level direct mapping include audification of data in the auditory domain [?], direct force/tactile feedback in the haptic domain, or the direct transmission of smells using electronic noses and scent generators. Thus, low-level direct mapping occurs at the level of the CogInfoCom icon layer.
 - *High-level direct mapping*: the substituting CogInfoCom message generates a concept which can also be generated by the substituted CogInfoCom message. The mapping scheme is direct because the substituting CogInfoCom message gives rise to direct interpretation based on the common generated concept it shares with the substituted CogInfoCom message. High-level direct mapping occurs at the level of the CogInfoCom message layer, since both local and global aspects of a CogInfoCom message can contribute to the generation of a concept. One example of high-level direct mapping is the use auditory representations of roughness that give rise to direct associations for conveying information on tactile roughness [12]. Another example would be to use different kinds of colors (i.e., visual icons) to convey information on temperature (specifically, colors can generate CogInfoCom concepts such as "coldness" and "warmth").
- *Analogical mapping*: the conceptual mapping between two CogInfoCom channels is analogical whenever the necessary associations are created through analogy. It is possible to distinguish the following sub-categories of analogical mapping:
 - *Structural mapping*: the structure (i.e., the dimensionality, speed and resolution) of the data is mapped to similar characteristics of the substituted message. Several researchers refer to this technique as *parameter mapping* (e.g. [26, 35, 42]). The term *structural mapping* is used in CogInfoCom because the structural parameters that are mapped to the CogInfoCom message are not necessarily parameters that are related to the generation of the transmitted concept. Structural mapping is also different from direct mapping techniques because the general qualities of the CogInfoCom message will usually have no resemblance to the transmitted concept, and the only resemblance will usually be related to the temporal patterns of occurrence, or speed and resolution-related aspects of the transmitted concept.A simple example of structural mapping would be to use a limited set of vibrations on a mobile device to convey a variety of information, such

that the concept to which the transmitted information refers would be obvious to the user only because of the timing, dimensionality and/or resolution of the vibrations.

Structural mapping can be considered to take place at the message level (it is the attributes of the message layer which serve as the basis for forming analogies between the substituted and substituting modalities).

- *Corroborative stimulation (co-stimulation)*: the necessary associations are formed through the simultaneous presentation of CogInfoCom messages in the substituted and the substituting modalities. Corroborative stimulation can rely both on natural associations (e.g., this is the case if the contact sounds of a remote probe are fed back to the teleoperator to provide information on tactile percepts), or on associations that are formed through training (the simplest example would be reinforcement learning, in which good practices become conceptually linked with the expectation of reward). Thus, co-stimulation based mapping takes place at the highest, conceptual level.

- *Scenario-based mapping*: the data flow within CogInfoCom messages is structured based on the structure of physical interaction. Scenario-based mapping can occur by having the receiver discover a virtual model through physical interactions (e.g., as in model-based sonification [23]), or by representing physical scenarios with virtual scenarios, so that the user receives CogInfoCom icons according to the same structural and temporal patterns as the concept is usually perceived in the physical world (e.g., an example of this approach is scenario-based orchestration [12]). Similar to co-stimulation, scenario-based mapping also takes place at the highest, conceptual level.

The structure of a modality in terms of CogInfoCom icons and CogInfoCom messages, as well as the communication points between the various mapping techniques and the different layers is shown in Figures 2 and 3. In the figure, CogInfoCom icons are generated based on a set of generation parameters, which are in turn dependent on a set of perceptual parameters stored in a perceptual gradation vector. The rationale behind this two-level indirection is that some forms of mapping – such as direct mapping and scenario-based mapping – occur during the conversion between the perceptual parameters and the icon-generating parameters (e.g., if the goal is to use direct mapping to convert distance information into the frequency of a beeping sound, then the distance information would be contained in the perceptual parameters, and this information would be converted into a frequency value which could then be used to generate the auditory icons).

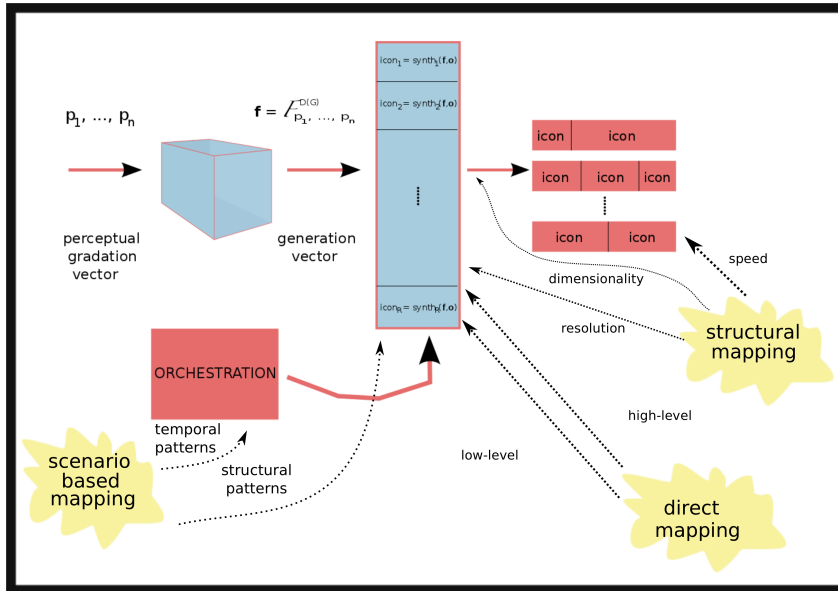


Figure 2

Structure of a single modality. The communication points between various mapping techniques and the modalities are depicted by dotted arrows. The CogInfoCom icon layer accepts a set of generation parameters and orchestration parameters based upon which the structural and temporal properties of CogInfoCom icons can be calculated. The succession of various CogInfoCom icons through time creates CogInfoCom messages. The communication points between various mapping techniques and the modalities are depicted by dotted arrows

4 Example applications in various domains

In this section, we give more detailed examples on how the proposed terminology applies to existing and future technologies, and how the various concept mapping techniques can help provide a meaningful link between various modalities.

In the first subsection, we briefly describe some very simple, everyday scenarios in which CogInfoCom plays a role. This is followed by a more detailed elaboration of three different domains in interactive computation which have a strong link with CogInfoCom.

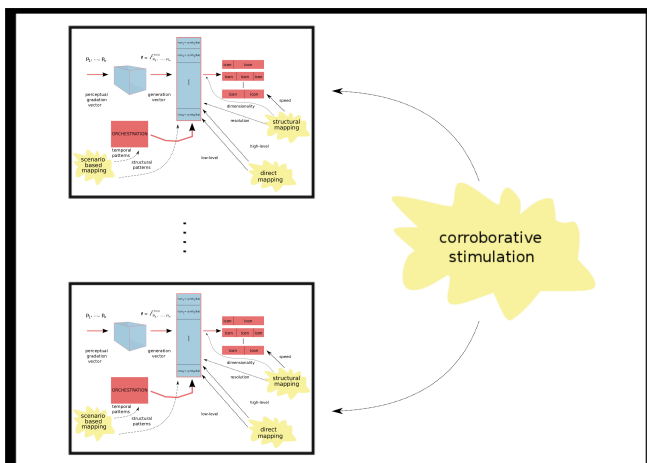


Figure 3

Structure of multiple modalities. The communication point between various mapping techniques and the modalities are depicted by dotted arrows

4.1 Rudimentary examples

4.1.1 Reverse driving

Most trucks and buses, and more recently, many cars are equipped with a *reverse radar system*. The goal of the system is to provide the driver with information on the distance between the vehicle and the car behind it while parking in reverse. The example of reverse radar systems concerns CogInfoCom because in most cases the distance information (which is therefore visual information) is transformed to the auditory modality. Thus, the distance information is represented by a single-pitched beeping sound, in which the frequency of the beeping correlates with the measured distance. While listening to the sounds to avert possible collision, the driver’s eyes are free to concentrate elsewhere in order to avoid collisions with objects or humans appearing suddenly.

The reverse radar system – while an extremely simple application of CogInfoCom – is a good example of low-level direct mapping, because the time series of distance values (normally acquired through eyecons) is used to modulate the temporal relationship between adjacent notes in an earcon.

4.1.2 CogInfoCom for cellphones

As a rudimentary example of CogInfoCom in mobile communications, let us consider the ability of users to recognize the identity of the caller.

In natural, face-to-face communication, there are both visual and auditory aspects of speaker identification. In addition, the visual and auditory cues enable the listener to gain an intuitive understanding of the speaker's psychological state. It is important to realize that if we are to design mobile communications devices capable of substituting for natural, face-to-face communication, all of these aspects of speaker recognition must be taken into consideration.

If we consider the history of speaker recognition in mobile communications, it becomes clear that the communication of caller information to the callee has evolved from an implicit representation (i.e., relying on the caller introducing himself or herself, and/or the capability of the callee to recognize the caller's voice) to visual icon and earcon-based representations (i.e., an image of the caller's face, and/or a special ringtone). The intermediate step of text-based caller representation (i.e., using eyecons) may also be mentioned.

This evolution of feedback types for caller information has interesting characteristics in terms of the directness of the display. The least direct representation was the implicit representation for caller identity. The text-based display which followed was more efficient, but because the name of a person is not always descriptive, this was not a direct way of representing caller information. The visual icon based representation is much more direct, because a person's face may be recognized faster, and it may also provide the callee with richer connotations. Finally, it can be argued that the use of earcons creates a new sensory channel through human audition for perceiving caller information, and a world of possibilities are available through this new modality. The mapping between the caller and a representative earcon can be based on:

- *corroborative stimulation*, – the mapping can be strengthened by the frequent experience of hearing the earcon and seeing the name and/or photograph of the caller displayed on the screen
- *high-level direct mapping*, – the earcon (e.g., the song or musical excerpt) may have connotations reminding the callee of events or situations experienced together with the caller

As psychophysical sensors become cheaper and more sophisticated in the future, cellphones may be augmented with the capability of sensing the psychological state of the caller and transmitting this information through different modulations of earcons or other CogInfoCom messages presented to the callee.

4.2 Detailed examples

4.2.1 Virtual collaboration

Virtual reality systems facilitating virtual collaboration have been a popular subject in research for a long time, but recent developments have contributed a great deal towards the proliferation of such systems in everyday use. Many examples can be cited of virtual environments facilitating virtual collaboration. Some of the existing systems emphasize communication with other people for business as well as recreational purposes (e.g. IBM Lotus Sametime [27], Second Life [41]), others facilitate the design of 3D objects using e.g. CAD-based tools (e.g., Airbus's VCEAD [13], VADE designed at Washington State Univ. [28]), while still others support a combination of human and robot collaboration (e.g. VirCA, designed at MTA SZTAKI [45]).

From a CogInfoCom perspective, such virtual collaboration systems must aim to resolve the contradiction between the two opposing tendencies of *unencumberment* (i.e., the tendency of freeing the user from having to wear burdensome equipment in order to receive feedback information [11]) and *parameter-rich, multi-modal feedback*. This problem of conflicting goals can be alleviated if alternative CogInfoCom channels are used to convey the necessary information to the user. Through the seminal work of Bach-y-Rita, it is well-known that sensory information from one modality can in many cases be conveyed through a different modality, and still be understood by the brain [3, 2]. Such applications are referred to as sensory substitution, sensory augmentation, and in some cases, sensorimotor extension [1]. Sensory substitution and augmentation can contribute significantly to CogInfoCom, because results in these fields can be used to analyze the dimensionality, speed and resolution at which the human brain can interpret information from various sensory modalities. However, in order to develop successful rendering across modalities in engineering applications, it is also important to take into consideration higher-level aspects such as the fact that some information types may be transferred more easily through iconic representations, while other information types may be more suitable for message-like transmission. Further, a balance is important between the different types of mapping used within the concept-mapping layer so that the level of information presentation (i.e., the locality or globality of the information, as described in [6, 7]) correspond with the level of the conceptual mapping.

At MTA SZTAKI, our goal was to develop applications which incorporate CogInfoCom channels in user interaction. The Virtual Collaboration Arena (VirCA) system, developed in our research group, is a component-based virtual collaboration environment (<http://virca.hu>). Although systems like VirCA are virtual environments in the traditional sense, from a CogInfoCom perspective, such systems are the most complex infocommunication systems designed for human communication in our day (systems such as these may become more prevalent as Future Internet technologies such as 3D Internet and Internet of Things become more widespread). During operation, the cognitive state of VirCA might include, among others, the state of different

objects in the virtual environment, and the status of various user interactions which are being performed. If the system is to communicate its cognitive state to the user in a way that is informative but not overwhelming, it is clear that a structured design procedure is necessary in order to design useful CogInfoCom channels.

Two examples of such channels include the use of earcons to convey tactile percepts (i.e., the surface properties of objects that are touched by the user in virtual space) [12], and the use of hapticons and a volumetric force concept to convey force feedback to the user during interactions [16].

The first example was developed in order to allow users to touch virtual objects in virtual space and feel the tactile qualities of their surface [12]. The application was developed with unencumbered and cost-effective interaction in mind (i.e., the goal was to allow the user to be able to perceive tactile qualities of objects without having to wear burdensome and expensive equipment). The earcons we have developed in order to achieve this goal contain information on the tactile dimensions of softness, roughness, stickiness and temperature, because psychophysical experiments have shown that these four dimensions contribute most significantly to human tactile perception [50, 25]. Because the concepts of softness and roughness are both *hapticon-generated concepts* and *earcon-generated concepts*, it was possible to create iconic representations (i.e., *auditory icons*) of different gradations across these dimensions. A form of *low-level, direct mapping* was used to convey information on the temperature of the surface, so that lower-pitched earcons with more rapid pulsations represented warmer surfaces. Due to the fact that the concept of stickiness is difficult to interpret in the auditory domain, a *scenario-based mapping technique* was used such that the stickiness of the surface was modeled as a time delay which presents itself when the user tries to move his or her hand across the surface in an imaginary, interactive scenario. More specifically, the scenario consisted of:

1. The user pressing down on the surface and hearing the auditory icon for softness
2. The user trying to move his/her hand across the surface and hearing the auditory icon for roughness. The length of the auditory icon for softness depended on the stickiness of the surface (the stickier the surface, the later the user could hear its roughness)
3. The user being able to feel the temperature of the surface throughout the whole process, and thus hearing the corresponding earcon

The earcon-based CogInfoCom channel for tactile percepts was developed to carry 10 gradations along each of the four tactile dimensions using 2-second long earcons (i.e., the whole interactive scenario was condensed into 2 seconds), and the channel was tested on a number of users to see whether or not they could accurately recall the sounds [12]. It is important to note, however, that the purpose of the application is not for users to be able to make nominal judgements. Instead, the channel was

developed to allow users to gain an intuitive understanding of what they are touching in the VirCA system without having to pay much explicit attention to the earcons. Further, even if users chose to pay explicit attention to the sounds, it would be to make relative, rather than nominal judgements.

The second example was developed in order to allow users to touch virtual objects in virtual space and be able to perceive force feedback through vibrotactile channels. Force feedback was computed based on the concept of *volumetric force*, which is equal to the volumetric overlap between the user's hand and the virtual object, multiplied by the volumetric stiffness of the virtual object (volumetric stiffness is a generalization of stiffness, and can be measured in $[N/m^3]$). A vibrotactile glove was developed, which contains shaftless vibration motors above each of the user's fingernails. The motors can be controlled independently using pulse width modulated signals. Six different feedback patterns – including various combinations of homogeneous and inhomogeneous pulse width, pulse frequency and pulse strength modulation – were proposed in order to allow the user to obtain perception of volumetric force feedback in virtual space [16]. In this application, haptic-based CogInfoCom channels are mapped to tacton-based CogInfoCom channels using low-level direct mapping.

4.2.2 Spatial memory in Intelligent Space

Intelligent Space (iSpace), proposed by Hashimoto in the early 2000s, is an intelligent environment capable of sensing activity through distributed sensors and acting on the distributed information it gathers [20, 21]. The success of iSpace relies on the fact that it is possible to develop more resilient and scalable systems if intelligence is implemented in a distributed fashion (e.g., using a large number of cheap components) rather than in a single, highly complex, autonomous system.

Since the concept of iSpace was proposed, several interesting applications have emerged. One such application, proposed by Niitsuma, is *spatial memory*. Spatial memory is a memory system in 3D space based on an analogy with random access memory in computers [37]. In typical interaction with spatial memory, users can store and invoke commands or other information in specific locations of the 3D environment using gestural commands (or bodicons in CogInfoCom terminology).

The concept of spatial memory introduces significant challenges for CogInfoCom. It is important to note that if the spatial memory system is viewed as an artificially cognitive system, two questions immediately arise: 1.) what kind of cognitive system does spatial memory have, and 2.) how should it communicate the information it encapsulates to the user?

The answer to the first question is that most probably the cognitive system of spatial memory systems is not fully developed yet, but in its current state the system is capable of sensing and understanding bodicons used to store and recall information, and representing bodicon-related and positional information which links the user's

position and actions to information at an address in memory [37].

Allowing users to communicate with the address (position) information is currently the most significant challenge, precisely because users are not well accustomed to dealing with spatial memories in 3D space (i.e., it is difficult for users to recall where in space they have stored a given command). There has been an effort to use 2D graphical displays as well as audio displays to help users localize themselves in terms of memory locations in 3D space [38]. A virtualized spatial memory was also developed in the VirCA system at MTA SZTAKI [31]. The system uses dedicated *virtual sensors* (i.e., sensors capable of obtaining information from virtual space) which are capable of sensing locations in the 3D space that have commands attributed to them through spatial memory (the locations also have eyecon-based representations). Whenever such a virtual sensor is attached to an actor in the VirCA system, the actor will be able to store and invoke commands through associated bodicons more easily.

The question as to how the information at various memory locations should be presented to the user depends on the application. CogInfoCom mapping layer techniques for concept mapping rely on knowledge of the application domain and the various kinds of connotations the application domain gives rise to.

5 Conclusion

Cognitive infocommunications (CogInfoCom) was created as a synergy of several fields – including informatics, infocommunications and cognitive sciences. However, the fields underlying CogInfoCom use concepts and terms that have some overlaps but often also contradict each other. For this reason, we proposed a set of concepts to help unify some aspects of these fields, from the unique viewpoint necessary for the human-oriented analysis and synthesis of CogInfoCom channels. The terminology we proposed enables engineers designing CogInfoCom channels to consider both the structural and semantic aspects of CogInfoCom channels. Through two rudimentary and two more complex engineering examples, the use and validity of the terminology was demonstrated.

Acknowledgments

The research was supported by the Hungarian National Development Agency, NAP project, NKTH-KCKHA005 (OMFB-01137/2008)

References

- [1] M. Auvray and E. Myin. Perception with compensatory devices: from sensory substitution to sensorimotor extension. *Cognitive Science*, 33:1036–1058, 2009.

- [2] P. Bach-y Rita. Tactile sensory substitution studies. *Annals of New York Academic Sciences*, 1013:83–91, 2004.
- [3] P. Bach-y Rita, M.E. Tyler, and K.A. Kaczmarek. Seeing with the brain. *International Journal of Human-Computer Interaction*, 15(2):285–295, 2003.
- [4] M. Badida, R. Kralikova, and E. Lumnitzer. Modeling and the use of simulation methods for the design of lighting systems. *Acta Polytechnica Hungarica*, 8(2):91–102, 2011.
- [5] P. Baranyi and A. Csapo. Cognitive infocommunications: Coginfocom. In *Computational Intelligence and Informatics (CINTI), 2010 11th International Symposium on*, pages 141–146, 2010.
- [6] S. Barrass. *Auditory Information Design*. PhD thesis, Australian National University, 1997.
- [7] J. Bertin. *Graphics and Graphic Information Processing*. Walter de Gruyter, 1981.
- [8] M. Blattner, D. Sumikawa, and R. Greenberg. Earcons and icons: Their structure and common design principles. *Human Computer Interaction*, 4(1):11–44, 1989.
- [9] M. M. Blattner, A.L. Papp, and E.P. Glinert. Sonic enhancement of two-dimensional graphics displays. In G. Kramer, editor, *Auditory Display: Sonification, Audification and Auditory Interfaces*, pages 447–470. Addison Wesley, Reading, Massachusetts, 1994.
- [10] S. Brewster and L.M. Brown. Tactons: structured tactile messages for non-visual information display. In *Proceedings of the fifth conference on Australasian user interface - Volume 28*, AUIC '04, pages 15–23, 2004.
- [11] A. Craig, W.R. Sherman, and J.D. Will. *Developing Virtual Reality Applications: Foundations of Effective Design*. Morgan Kaufmann, 2009.
- [12] A. Csapo and P. Baranyi. A conceptual framework for the design of audio based cognitive infocommunication channels. In *Recent Advances in Intelligent Engineering Systems*, volume 368 of *Studies in Computational Intelligence*, pages 261–281. Springer-Verlag Berlin Heidelberg, 2012.
- [13] M. Durstewitz, B. Kiefner, R. Kueke, H. Putkonen, P. Repo, and T. Tuikka. Virtual collaboration environment for aircraft design. *Information Visualisation, International Conference on*, page 502, 2002.
- [14] M. Enriquez, K. Maclean, and C. Chita. Haptic phonemes: Basic building blocks of haptic communication. In *in Proc. of 8th International Conference on Multimodal Interfaces (ICMI 2006)*. ACM Press, 2006.

- [15] M.J. Enriquez and K.E. MacLean. The haptic editor: A tool in support of haptic communication research. In *Proceedings of the 11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS'03)*, pages 356–. IEEE Computer Society, 2003.
- [16] P. Galambos and P. Baranyi. Vibrotactile force feedback for telemanipulation: Concept and applications. In *Proceedings International Conference on Cognitive Infocommunications 2011*, Budapest, Hungary, 2011.
- [17] W. Gaver. Auditory icons: Using sound in computer interfaces. *Human Computer Interaction*, 2(2):167–177, 1986.
- [18] W. Gaver. *Everyday Listening and Auditory Icons*. PhD thesis, University of California, San Diego, 1988.
- [19] W. Gaver. The sonicfinder: An interface that uses auditory icons. *Human Computer Interaction*, 4(1):67–94, 1989.
- [20] H. Hashimoto. Intelligent space – how to make spaces intelligent by using sound? In *IEEE International Conference on Systems, Man and Cybernetics*, pages 14–19, 2002.
- [21] H. Hashimoto. Intelligent space: Interaction and intelligence. *Artificial Life and Robotics*, 7(3):79–85, 2003.
- [22] K. Hercegfi. Heart rate variability monitoring during human-computer interaction. *Acta Polytechnica Hungarica*, 8(5):205–224, 2011.
- [23] T. Hermann. *Sonification for Exploratory Data Analysis*. PhD thesis, Bielefeld University, Bielefeld, 2002.
- [24] T. Hermann and H.J. Ritter. Listen to your data: Model-based sonification for data analysis. In G. E. Lasker, editor, *Advances in intelligent computing and multimedia systems*, page 189–194. Int. Inst. for Advanced Studies in System research and cybernetics, Int. Inst. for Advanced Studies in System research and cybernetics, 8 1999.
- [25] M. Hollins and S.R. Risner. Evidence for the duplex theory of tactile texture perception. *Perceptual Psychophysics*, 62:695–705, 2000.
- [26] A. Hunt, T. Hermann, and S. Pauletto. Interacting with sonification systems: Closing the loop. In Ebad Banissi and Katy Börner, editors, *IV '04: Proceedings of the Information Visualisation, Eighth International Conference on (IV'04)*, pages 879–884. IEEE Computer Society, 7 2004.
- [27] IBM lotus sametime advanced software: Breaking new ground for unified communications and collaboration. White paper, IBM, 2008. Available online (16 pages).

- [28] S. Jayaram, Y. Wang, U. Jayaram, K. Lyons, and P. Hart. A virtual assembly design environment. In *IEEE Virtual Reality '99, Houston*, pages 172–179, 1999.
- [29] K. Jokinen. User interaction in mobile navigation applications. In L. Meng, A. Zipf, and S. Winter, editors, *Map-based Mobile Services*, Lecture Notes in Geoinformation and Cartography, pages 168–197. Springer Berlin Heidelberg, 2008.
- [30] J. Kaye. Making scents: aromatic output for hci. *interactions*, 11:48–61, 2004.
- [31] P. Kovacs and M. Niitsuma. Spatial memory in virca (demonstration). In *International Conference on Cognitive Infocommunications 2011*, Budapest, Hungary, 2011.
- [32] S. Lederman. Haptic identification of common objects: Effects of constraining the manual exploration process. *Perception & Psychophysics*, 66(4):618–628, 2004.
- [33] P. Lemmens, M.P. Bussemakers, and A. De Haan. Effects of auditory icons and earcons on visual categorization: the bigger picture. In *In Proceedings of the International Conference on Auditory Display (ICAD '01)*, pages 117–125, 2001.
- [34] K. Maclean and M.J. Enriquez. Perceptual design of haptic icons. In *In Proceedings of Eurohaptics 2003*, pages 351–363, 2003.
- [35] T.M. Madhyastha and A. Reed. A framework for sonification design. In G. Kramer, editor, *Auditory Display*. Addison-Wesley, 1994.
- [36] M.R. McGee. *Investigating a Multimodal Solution for Improving Force Feedback Generated Textures*. PhD thesis, University of Glasgow, 2002.
- [37] M. Niitsuma and H. Hashimoto. Spatial memory as an aid system for human activity in intelligent space. *IEEE Transactions on Industrial Electronics*, 54:1122–1131, 2007.
- [38] M. Niitsuma, A. Shiraiishi, and H. Kobayashi. Extended user interface for improving usability of spatial memory. In *Proceedings International Conference on Cognitive Infocommunications 2011*, Budapest, Hungary, 2011.
- [39] D.A. Norman. *Cognitive Artifacts*, pages 17–38. Cambridge University Press, New York, NY, USA, 1991.
- [40] A. Pirhonen. Non-speech sounds as elements of a use scenario: a semiotic perspective. In *Proceedings of the 12th International Conference on Auditory Display (ICAD2006)*, 2006.
- [41] M. Rymaszewski, W.J. Au, M. Wallace, C. Winters, C. Ondrejka, and B. Batsone-Cunningham. *Second Life: The Official Guide*. Sybex, 2006.

- [42] C. Scaletti. Sound synthesis algorithms for auditory data representations. In G. Kramer, editor, *Auditory Display*. Addison-Wesley, 1994.
- [43] D.C. Smith. *Pygmalion: A Computer Program to Model and Stimulate Creative Thought*. PhD thesis, Stanford University, Dept. of Computer Science, 1975.
- [44] M. Tomasek. Encoding named channels communication by behavioral schemes. *Acta Polytechnica Hungarica*, 8(2):5–19, 2011.
- [45] A. Vamos, I. Fulop, B. Resko, and P. Baranyi. Collaboration in virtual reality of intelligent agents. *Acta Electrotechnica et Informatica*, 10:21–27, 2010.
- [46] F. Vernier and L. Nigay. A framework for the combination and characterization of output modalities. In *in Proceedings of DSV-IS2000, LNCS, Springer-Verlag*, pages 32–48. Springer-Verlag, 2000.
- [47] J. Voisin, Y. Lamarre, and C.E. Chapman. Haptic discrimination of object shape in humans: Contribution of cutaneous and proprioceptive inputs. *Experimental Brain Research*, 145(2):251–260, 2002.
- [48] Y-G. Wang, Xiang Huang, S.-S. W.-S., and Y-L. Pei. Multipattern road traffic crashes and injuries: a case study of xi'an city. *Acta Polytechnica Hungarica*, 8(4):171–181, 2011.
- [49] M.S. Wogalter. *Handbook of Warnings*. Lawrence Erlbaum Associates, 2006.
- [50] T. Yoshioka, S.J. Bensmaia, J.C. Craig, and S.S. Hsiao. Texture perception through direct and indirect touch: an analysis of perceptual space for tactile textures in two modes of exploration. *Somatosensory and Motor Research*, 24(1–2):53–70, 2007.