Structure of Cognitive Infocommunication Channels – Audio-Tactile Sensor Bridging

Thesis Booklet

Ádám B. Csapó
MSc in Computer Engineering

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Doctoral School of Informatics
Faculty of Electrical Engineering and Informatics
Budapest University of Technology and Economics

Supervisor:
Péter Baranyi, D.Sc.

Budapest, January 2014
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2. Historical Background</td>
<td>2</td>
</tr>
<tr>
<td>2.1 Convergence of the Cognitive Sciences and Infocommunications</td>
<td>2</td>
</tr>
<tr>
<td>2.2 Synergies of Disciplines Under CogInfoCom</td>
<td>3</td>
</tr>
<tr>
<td>2.3 Different Levels of Synergy behind CogInfoCom Channels</td>
<td>4</td>
</tr>
<tr>
<td>3. Motivations and Goals of the Dissertation</td>
<td>5</td>
</tr>
<tr>
<td>4. Research Methodology</td>
<td>7</td>
</tr>
<tr>
<td>5. Summary of Research Achievements and Theses</td>
<td>7</td>
</tr>
<tr>
<td>5.1 Theses 1 and 2</td>
<td>7</td>
</tr>
<tr>
<td>5.1.1 Summary of results</td>
<td>7</td>
</tr>
<tr>
<td>5.1.2 Thesis 1</td>
<td>12</td>
</tr>
<tr>
<td>5.1.3 Thesis 2</td>
<td>13</td>
</tr>
<tr>
<td>5.2 Thesis 3</td>
<td>14</td>
</tr>
<tr>
<td>5.2.1 Summary of results</td>
<td>14</td>
</tr>
<tr>
<td>5.2.2 Thesis 3</td>
<td>17</td>
</tr>
<tr>
<td>5.3 Thesis 4</td>
<td>17</td>
</tr>
<tr>
<td>5.3.1 Summary of results</td>
<td>17</td>
</tr>
<tr>
<td>5.3.2 Thesis 4</td>
<td>18</td>
</tr>
<tr>
<td>6. Applicability of Results</td>
<td>19</td>
</tr>
<tr>
<td>The Author’s Publications Relevant to this Dissertation</td>
<td>19</td>
</tr>
<tr>
<td>The Author’s Publications not Relevant to this Dissertation</td>
<td>22</td>
</tr>
<tr>
<td>Bibliography</td>
<td>24</td>
</tr>
</tbody>
</table>
1 Introduction

Cognitive infocommunications (CogInfoCom) is an emerging interdisciplinary research field that was created as a synergy between infocommunications and the cognitive sciences. CogInfoCom has both theoretical and industry-oriented motivations. The full definition of CogInfoCom, as well as the various modes (intra-cognitive, inter-cognitive) and types (sensor-bridging, sensor-sharing, representation-bridging, representation-sharing) of CogInfoCom systems can be found in [RJ-5, RC-14] and in the full text of this dissertation.

Based on trends in EU and other government supported research, the early forms of CogInfoCom emerged through the course of domestic and international projects. The conceptual framework of thinking in terms of CogInfoCom appeared as early as in 2005, and by 2008-2009, the term itself was mentioned in several papers and lectures. The definition of the scope and goals of CogInfoCom as an interdisciplinary research field was formulated by an international group of researchers in 2010, at the 1st International Workshop on Cognitive Infocommunications in Tokyo, Japan. Since then, CogInfoCom has established itself firmly among those modern research fields that promise to represent leading research directions in the near future. An international conference on CogInfoCom is held every year (the CogInfoCom 2012 and 2013 conferences were held under the umbrella of IEEE, with a growing number of participants). A scientific journal dedicated entirely to CogInfoCom is also expected to be initiated in the near future.

Although CogInfoCom itself is a newly established research discipline, its emergence is directly related to a past of several decades. During this time, both the cognitive sciences and infocommunications have seen considerable growth. As a result, it is now commonplace to talk about “cognitive machines”, “cognitive robotics” and “cognitive informatics” on the one hand, and about the “infocommunications revolution” and the “convergence” between informatics, media, infocommunications and regulation on the other. These trends are leading to the gradual appearance of artificial cognitive capabilities, i.e. capabilities that are directed towards a broadened scope of sensing and processing of unstructured data. Users are becoming accustomed to accessing these artificial cognitive capabilities through their infocommunications devices in a wide variety of contexts. However, more than just allowing users to access artificial cognitive capabilities, infocommunications devices of the future are also expected to allow users to extend their cognitive capabilities with artificial ones. This synergy of natural and artificial cognitive capabilities will be applicable in flexible and novel ways, both in physical contexts and in virtual worlds.

The general goal of this dissertation is to develop some of the theoretical aspects of CogInfoCom, specifically in relation to CogInfoCom channels. The term has been used extensively in the past few years in informal contexts to describe substitutive and augmentative communication through various sensory modalities. This dissertation primarily aspires to give a formal definition of CogInfoCom channels, and to answer questions regarding how such channels can best be designed and implemented.

**Definition of CogInfoCom**

**Definition.** Cognitive infocommunications (*CogInfoCom*) investigates the link between the research areas of infocommunications and the cognitive sciences, as well as the various engineering applications which have emerged as a synergic combination of these sciences.
The primary goal of CogInfoCom is to provide a systematic view of how cognitive processes can co-evolve with infocommunications devices so that the capabilities of the human brain may not only be extended through these devices, irrespective of geographical distance, but may also interact with the capabilities of any artificially cognitive system. This merging and extension of cognitive capabilities is targeted towards engineering applications in which artificial and/or natural cognitive systems are enabled to work together more effectively.

Consequences of the definition, and the definition of further concepts such as the modes and types of cognitive infocommunications can be found in [RJ-5, RC-14] and in the full text of this dissertation.

2 Historical Background

This section provides a brief overview of the research historical aspects which have led to the creation of CogInfoCom.

2.1 Convergence of the Cognitive Sciences and Infocommunications

Traditionally, the research fields of informatics, media, and communications were distinct research areas with different backgrounds. As a synthesis between pairs of these 3 disciplines, infocommunications, media informatics and media communications emerged in the latter half of the 20th century. The past evolution of these disciplines points towards their convergence in the near future, as modern network services aim to provide a more holistic user experience [34]. As a result of this convergence, the infocommunications sector now encompasses “all information processing and content management functions […] of information technology and electronic media” [34].

Parallel to these developments, with the enormous growth in scope and generality of the cognitive sciences in the past few decades, the new fields of cognitive media [31, 22], cognitive informatics [46, 45, 43] and cognitive communications [33, 15] are gradually emerging. These fields have also either fully made their way, or are steadily on their way into standard university curricula. In a way analogous to the evolution of media informatics, media communications and infocommunications, examples of research achievements which can be categorized as results in cognitive media informatics, cognitive media communications and cognitive infocommunications are gradually emerging, even if – as of yet – these fields are not always clearly defined (Figure 1).

\footnotetext{1}{Cognitive media is usually part of larger programs named interaction technologies or interactive media (see e.g., the Research Group Interaction Technologies at Fraunhofer Institute)}

\footnotetext{2}{A definition of cognitive informatics can be found on the official web page of a number of research organizations (e.g., the Pacific Northwest Laboratory funded by the U.S. Department of Energy). Today there are several research institutes dedicated to cognitive informatics, such as ICON in Sheffield, UK, which has an annual budget income of over 1 million GBP. An IEEE International Conference on Cognitive Informatics has been held every year since 2002.}

\footnotetext{3}{The term cognitive communications is used in two different ways in the literature. On the one hand, it can refer to the ways in which humans with different cognitive capabilities can communicate with each other (e.g., as in [33, 15]). On the other hand, the term is also used to refer to wireless devices and networks which can assign resources and functionalities in dynamic and intelligent ways (e.g., cognitive radio, which was defined by researchers at Virginia Polytechnic Institute and State University). It is possible that in the future, these two directions will reach common formulations for certain problems (see e.g., [90]).}
2.2 Synergies of Disciplines Under CogInfoCom

One of the research goals which is central to CogInfoCom research is the goal of creating CogInfoCom channels through which artificially cognitive systems and natural cognitive systems can communicate and collaborate (together: to co-evolve) with each other in a way that appeals to their respective cognitive capabilities. This section highlights three key disciplines leading to synergy between CogInfoCom and some other fields in computing which are strongly relevant to CogInfoCom channels (further disciplines are described in the full text of the dissertation).

**Future Internet and Virtual Reality** – Two major directions in Future Internet research are the Internet of Things and 3D Internet. The Internet of Things deals with the physical integration of everyday objects and their unique capabilities on the Internet [38, 39]. 3D Internet focuses on the growing expectation of users for “high-quality 3D imagery and immersive online experience” [19]. 3D Internet – with the appearance of (virtual and physical) everyday objects on the Internet – is leading to augmented 3D environments with which users expect to be able to interact in the same way as with the real physical world [32, 35]. A key challenge in augmented Internet-based (i.e., infocommunications-based) 3D environments is to ensure that the user does not feel like a ghost (besides being able to see and hear everything, the user will also wish to be able to touch and manipulate objects in much the same way as in real life). In an increasing number of cases, this is achieved by creating a virtual avatar for the user, which is ideally in at least partial sensory-motor synchrony with the user [47]. All of these goals require that the infocommunications device used to access virtual reality be able to have complex faculties in extending the user’s cognitive capabilities, which is one of the key topics of CogInfoCom.

**Sensory substitution** – Sensory substitution is a research direction since the late 1960s that provides room for synergy with sensor-bridging CogInfoCom. The key concepts of sensory substitution, and its utility was first described by Bach-y-Rita and his colleagues, who, in one of their more recent works, defined sensory substitution as “the use of one human sense to receive information normally received by another sense” [4].

The need to broaden the scope of sensory substitution, at least in engineering systems, was also recognized by Bach-y-Rita and his colleagues. This was eloquently highlighted as follows [4]:

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 above quote highlights the authors’ opinion that it would be worthwhile to extend the scope of sensory substitution research to encompass multimodal interactions in virtual environments, so as to address questions such as the source of information (i.e., information may equally be virtual and/or non-physical as well as real and/or physical), the human sensory modality used to transmit the information (this aspect is analogous to the notions of sensor-sharing and sensor-bridging), as well as the substitutive or augmentative nature of the application. Based on these aspects, CogInfoCom has common interests with sensory substitution.

### 2.3 Different Levels of Synergy behind CogInfoCom Channels

The main purpose of CogInfoCom channels is to describe sensory signals passing through the CogInfoCom interface from a syntactic and semantic point of view (Figure 2). In this regard, the large number of fields making synergic contributions to CogInfoCom have different motivations and use dif-

![Figure 2: Schematic view of a CogInfoCom system for inter-cognitive infocommunication. The action-reaction processes which influence human behavior are modeled as a set of interrelated components in artificially cognitive systems, referred to as CogInfoCom engines. CogInfoCom channels, in turn, are structured sensory messages which pass through the CogInfoCom interface between the communication systems.](image-url)
different approaches. The following list demonstrates, by way of example, how such differences appear in different fields:

- **Sensory substitution** – the field focuses specifically on the low-level mapping between sensory signals of different modalities, and its results have been used to provide e.g. visual information through the tactile [3, 4, 7, 41] and auditory modalities [2, 1, 20, 29].

- **Multi-modal interactions** – in contrast to sensory substitution, multi-modal interactions adopts a higher-level view and aims to map meaning to sensory signals at a higher level of abstraction. Multi-modal interactions must also take into consideration various effects which stem from the fact that multiple modalities are used – such as multi-sensory integration, cross-effects between sensory modalities, sensory dominance, etc. In the context of this dissertation, results obtained through such considerations have been applied e.g. to providing haptic feedback through vision (e.g., using pseudo-haptics [25, 24] and other graphical attributes [21]).

- **Future Internet** – if sensory substitution adopts a low-level view and multi-modal interactions adopts a higher-level view, Future Internet investigates interaction at an even higher level. Although much of the communication is still channel-based in Future Internet, the aspiration is neither to provide a direct link between the structure of sensory signals in different modalities, nor to convey specific meaning through the use of sensory illusions and multi-sensory effects. Instead, many applications in the Future Internet focus on helping the user understand what element in the interface can be used for what purpose. This is the case, for example, in tangible interfaces [40, 16], where the appearance of an object represents both feedback information, as well as information on how to effect changes to the system.

It can be seen from these examples that although the approach used in many of these fields are entirely different, all of them have aspects which focus on linking syntactic and semantic aspects of sensory signals.

### 3 Motivations and Goals of the Dissertation

The problems which motivated this research, and which are addressed in the dissertation can be summarized as follows:

- **The problem of iconic and message-like syntactic elements**: Much research has been carried out on how to design various syntactical elements for iconic and message-like sensory signals, and how they can be used to represent meaning. For example, in research on auditory display, there is considerable literature on how to design auditory icons (iconic audio signals) [11, 5, 12, 13] and earcons (message-like audio signals) [6, 26, 28]. The problem addressed by the dissertation in terms of iconic and message-like sensory signals is that the literature generally treats these concepts in an isolated way: the task of designing interfaces is usually formulated as choosing between iconic and message-like representations. Such debates miss the point: it was shown – at least in the context of auditory displays – that both auditory icons and earcons can be suitable, depending on the method used to learn them, the methods used to evaluate performance (i.e., the task at hand), and the liberty with which users can develop their own syntactic-semantic associations [8, 9].
In this respect, the goal of this dissertation is to provide a conceptual framework that unifies syntactic and semantic aspects of message-based interfaces.

- The problem of defining sensory modalities and treating them in a uniform way:
  The concept of sensory modality varies from field to field, and in cases even within a single field. For example, in multi-modal interactions, some authors adopt the view that a modality reflects the sensory system that is used to perceive information, while others focus on the “artificial” modality – i.e., the output device – that is used to present information (see e.g. [18]). Still others distinguish between different kinds of stimuli even in the same ‘sensory modality’ based on how the messages are encoded [42].

  Unification of ideas is further complicated by the fact that different design approaches are generally used when designing interfaces for different modalities.

  In this respect, the goal of this dissertation is to define the concept of modality in the context of CogInfoCom, and to provide a design framework for channel-based CogInfoCom systems which enables the unified (transparent) treatment of different sensory modalities.

- The lack of systematic and re-usable design strategies: Engineers are confronted with a lack of multi-modal interface design methodologies which are general enough to address a wide range of application scenarios, but rigorous enough to allow for parametric design. For example, two recent papers on auditory display design have observed and demonstrated that auditory displays are often designed in “ad-hoc” ways, and that designers of auditory interfaces often create “hand-crafted and prototyped” solutions that are based on unsupported design decisions [27, 10].

  In this respect, the goal of this dissertation is to provide a design methodology that allows for the formal, systematic and re-usable development of CogInfoCom interfaces and makes possible the comparison of competing solutions.

The goals of the dissertation can be summarized as follows:

- To provide a definition of the concept of CogInfoCom channels (i.e., channels of communication through sensory signals that provide users with feedback information in a structured and interpretable way). The definition should unify the overlapping – and in many cases contradictory – points of view of the various research fields that have synergic contributions to CogInfoCom. Thus, for instance, it is necessary to clarify distinctions between iconic and message-like sensory signals, to investigate how they can be unified (i.e., treated in a uniform way in applications), as well as how semantic meaning can be represented together with such syntactic elements.

- To provide a theoretical framework that will allow for the systematic and reusable development of CogInfoCom channels. This can be achieved if a formal representation of CogInfoCom channels is provided, so that researchers can share their own solutions, implement each other’s solutions, and provide formal support for their design decisions.

- To investigate whether or not the user-controlled tunability of CogInfoCom channels should be a feature of CogInfoCom systems. If differences in the users’ cognitive capabilities, variability of application scenarios, or any other adverse effects necessitate the user-controlled tunability of CogInfoCom channels, the goal is to provide a tuning model that fulfills all requirements dictated by such circumstances.
To provide a solution to the problem of remote tactile sensing based on the theoretical framework developed for CogInfoCom channels.

4 Research Methodology

Although CogInfoCom is a newly established field, as detailed earlier, there are number of different research areas that have synergic contributions to CogInfoCom. The research methodologies employed in this dissertation have foundations in some of those fields.

In particular, the dissertation makes use of the following formalisms and methodologies:

- A formal abstraction of CogInfoCom channels is given using concepts from higher-order linear algebra (e.g., the tensor concept, and higher-order singular value decomposition).
- The formal description of mapping techniques between CogInfoCom concepts and CogInfoCom channels is developed based on the Object-Attribute-Relation (OAR) model – which is a denotational language based on formal concept analysis.
- An interpretable tuning model – known as the Spiral Discovery Method (SDM) – is developed based on a generalization of the Tensor Product (TP) model of multivariate functions and concepts from higher-order linear algebra, such as higher-order orthogonal iteration.
- The sensor-bridging application introduced in the final part of the dissertation is validated through usability tests applied in ergonomics.

5 Summary of Research Achievements and Theses

A framework is proposed in the dissertation which defines the concept of CogInfoCom channels, and provides a formal approach for their design and implementation. In this section, a brief summary is given on the research results presented in the dissertation, and based on the results, 4 theses are formulated.

5.1 Theses 1 and 2

The first two theses are strongly interrelated, therefore the relevant results are presented together.

5.1.1 Summary of results

The first two theses are focused on the following results:

1. The proposed framework introduces the formal abstraction of CogInfoCom channels, which are ordered sets of CogInfoCom messages that carry information on different attribute values of the same CogInfoCom message generated concept. The concept has the following properties:
Figure 3: The first four layers of the proposed conceptual framework unifying iconic and message-like entities, from bottom (left) to top (right).

- It unifies CogInfoCom icons and CogInfoCom messages across different sensory modalities in one representational entity, which provides a high-level and generic view to resolve the dichotomy between iconic and message-like representations (Figure 3).
- It encompasses the complete set of possible CogInfoCom messages – relevant to a single high-level concept – as one formal entity (the channel). The formal representation links perceptual parameters (perceptual gradation vector \( \mathbf{p} \)) with a set of generation parameters (generation vector \( \mathbf{f} \)) used to synthesize CogInfoCom messages through a partially orderable multivariate parameter-generating function \( \mathbf{F} \), which can be represented by an \((N+1)\)-dimensional tensor \( \mathbf{F} \) such that:

\[
\mathbf{F}^{D(G)}_{p_1,\ldots,p_N,i} = (F(g_{p_1,p_2,\ldots,p_N}))(i), \forall i = 1..H
\]

where \((F(g_{p_1,p_2,\ldots,p_N}))(i)\) is the \(i\)th dimension of the \(H\)-dimensional output vector of \(F\) corresponding to the generation vector in point \(g_{p_1,p_2,\ldots,p_N}\) on discretization grid \(G\), and the generation vectors obtained as the output of \(F\) are represented along the \((N+1)\)-th dimension of tensor \(\mathbf{F}\). Each dimension in the tensor is indexed by the gradation specified along the corresponding dimension of grid \(G\), save for the last one, which is indexed by the position of a specific parameter in the generation vector (Figure 4).
- Its formal representation can be transformed into a canonical form based on the higher-order singular value decomposition (HOSVD) in higher-order linear algebra.

\[
\mathbf{F}^{D(G)} = \mathbf{S} \bigodot \sum_{n=1}^{N+1} \mathbf{X}_n
\]

where \(\bigodot\) refers to the tensor product operation defined in [23], and:

(a) \(\mathbf{X}_n = (x_1^{(n)}, \ldots, x_{I_n}^{(n)})\), \(n = 1..N\) is an orthonormed matrix of size \((P_n \times I_n)\)
(b) \(\mathbf{X}_{N+1} = (x_1^{(N+1)}, \ldots, x_{I_{N+1}}^{(N+1)})\) is an orthonormed matrix of size \((H \times I_{N+1})\)
(c) \(\mathbf{S}\) is a real tensor of size \(I_1 \times \ldots \times I_N \times I_{N+1}\), the subtensors \(\mathbf{S}_{i_n=\alpha}\) of which have the following properties:

- all-orthogonality: any pair of the subtensors of \(\mathbf{S}\) are orthogonal, i.e. for all possible values of \(n, \alpha\) and \(\beta\) subject to \(\alpha \neq \beta\):

\[
< \mathbf{S}_{i_n=\alpha}, \mathbf{S}_{i_n=\beta} >= 0
\]
Figure 4: Structure of a CogInfoCom message in a CogInfoCom channel, based on the perceptual gradation vector and the generation vector. 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 (synth refers to the synthesis algorithm used for CogInfoCom icon generation). The succession of various CogInfoCom icons through time creates CogInfoCom messages.

- ordering: All of the subtensors of $S$ along any given dimension $n$ are ordered according to their Frobenius norm, i.e. $\forall n = 1..N + 1$:

$$\|S_{i_n=1}\| \geq \|S_{i_n=2}\| \geq ... \geq \|S_{i_n=I_n}\| \geq 0$$ (4)

It is important in any scientific field for researchers to be able to compare results, and the canonical representation is a significant step in this direction, because:

- it is a candidate for a standard representation of CogInfoCom channels
- it is a representation that makes explicit the rank concept of higher-order algebra – which in turn is related to the complexity of the channel.
- it makes explicit how the CogInfoCom channel can be generated based on perceptual parameters, which also helps increase the reproducibility of results

The above points all represent important steps towards achieving design methodologies that are not ad-hoc, but systematic and reusable.

- It provides operations that allow for the transparent manipulation of individual CogInfoCom messages in the channel. The numerical foundations of these operations are provided by the HOSVD of the discretized parameter-generating function based tensor representation of CogInfoCom channels:

$$F^{D(G)}(g_{p_1,..,p_N}) = \left( S^{\underbrace{N+1}_{n=1}} X_n \right)(g_{p_1,..,p_N}) = S^{\underbrace{N+1}_{n=1}} \bigotimes_{n=1, n\neq k} x_{n,p_n} \times_k x_{k,p_k}$$ (5)

where the $p_k$-th row of matrix $X_k$ is denoted by $x_{k,p_k}$. It is obvious that if only row vector $x_{k,p_k}$ is changed, then only those output values of the parameter-generating function will
be changed which correspond to an input vector that is defined over the \( p_\kappa \)-th discretization point in the \( k \)-th dimension of hyper-rectangular grid \( G \). For this reason, the manipulation of vector \( x_{k,p_\kappa} \) is referred to as the local tuning of the CogInfoCom channel along the \( k \)-th input dimension. The values in vector \( x_{k,p_\kappa} \) in turn are referred to as tuning weights.

2. The proposed framework introduces a concept algebra based background of CogInfoCom channels together with various forms of conceptual mapping.

- The concept algebra based framework behind CogInfoCom channels is based on the Object-Attribute-Relation (OAR) model. The following basic concepts are used from the OAR model:
  - *formal contexts*: 3-tuples of a set of objects, a set of attributes, and a set of relations between them
  - *abstract concepts within a given formal context*: 5-tuples of objects, attributes and various kinds of relationships between them
  - *intensions*: intersects of all sets of attributes belonging to various objects within an abstract concept
  - *instant attributes*: unions of all attributes belonging to all objects within an abstract concept

The dissertation extends these concepts in the context of CogInfoCom with the following definitions:

- *CogInfoCom message-generated concepts of a sensory modality*: the set of concepts in which the objects are percepts obtained through the sensory modality and have a single attribute in their intension:

  \[
  \mathbb{C}_{\text{mg}}(x) = \left\{ c \mid f_{\text{ctx}}(c) = (f_{\text{perc}}(x), \mathbb{A}, \mathbb{R}), \exists! a : a \in c^\ast(c) \right\}
  \]  

  where \( f_{\text{ctx}}(c) \) is a function that returns the context of a concept, \( c \), and \( f_{\text{perc}}(x) \) is the totality of CogInfoCom icons and CogInfoCom messages (i.e. percepts) which can be perceived through sensory modality \( x \).

- *evaluation functions of CogInfoCom messages with respect to CogInfoCom message generated concepts*: functions having output values which represent the degree to which a given CogInfoCom message generated concept represents a given CogInfoCom message:

  \[
  f_{\text{eval}} : f_{\text{perc}}(x) \times \mathbb{C}_{\text{mg}}(x) \to [0, 1]
  \]

- *interactive concepts of a CogInfoCom message generated concept*: sets of concepts that have an attribute whose value changes during the course of interaction to a degree that is proportional to the value of the CogInfoCom message generated concept’s evaluation function:

  \[
  \mathbb{C}_{\text{iact}}(c_{\text{mg}}(x)) = \left\{ c = (O, A, R^c, R^i, R^o), \forall p \in f_{\text{perc}}(x) \mid \exists a \in A : \frac{a}{a} \propto f_{\text{eval}}(p, c_{\text{mg}}) \right\}
  \]

- The dissertation introduces the following forms of mapping between CogInfoCom message generated concepts and CogInfoCom channels:
– *direct mapping*: direct representations of CogInfoCom message generated concepts are transferred through a CogInfoCom channel of the substituting modality.

  * In the case of **low-level direct mapping**, attribute values in the intension of a CogInfoCom message generated concept are used to directly control the perceptual parameters of the substituting CogInfoCom message:

  \[
  \Delta a(cpt_1) = c \Rightarrow \Delta f_{eval}(msg; cpt_2) = \gamma c
  \]

  where \( a(cpt_1) \) is the attribute in the intension of the transmitted CogInfoCom concept \( (cpt_1) \), \( msg \) is the substituted CogInfoCom message, \( c \) is a constant, and \( \gamma \) is a scaling variable.

  * In the case of **high-level direct mapping**, the substituted CogInfoCom message generates a concept which can be also generated by the substituting CogInfoCom message. High-level direct mapping is possible if at least one of the instant attributes of the concept generated by the substituting CogInfoCom messages coincides with the intension of the substituted CogInfoCom concept:

  \[
  \exists a \in c^*_1 : a \in A^*_2
  \]

– **analogical mapping**: the necessary associations between CogInfoCom message generated concepts and CogInfoCom channels are created through analogy.

  * In the case of **structural mapping**, the structure (i.e., the dimensionality, speed and resolution) of the substituted CogInfoCom message generated concept is mapped to similar characteristics of the substituting message.

  * In the case of **corroborative stimulation**, the necessary associations are formed through the simultaneous presentation of CogInfoCom messages in the substituted and the substituting modalities.

  * In the case of **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 [14]), or by representing physical scenarios within virtual scenarios, so that the user receives CogInfoCom icons and messages according to the same structural and temporal patterns as the concept is usually perceived in the physical world (e.g., as in scenario-based orchestration [RB-1].

In the second case (i.e., physical-to-virtual scenario-based mapping), mapping can be formally modeled using the interactive concepts defined in the dissertation. More specifically, even if there is no opportunity to create a direct mapping between a concept and a CogInfoCom channel, it may still be possible to create a scenario-based mapping by mapping the attribute of an interactive concept onto the CogInfoCom message:

\[
\Delta a(c_{iact}(c_{msg})) = c \Rightarrow \Delta f_{eval}(msg_2; cpt) = \gamma c
\]

where \( a(c_{iact}(c_{msg})) \) is an attribute of an interactive concept \( (C_{iact}) \) of the transmitted concept \( (c_{msg}) \), \( msg_1 \) is the substituted CogInfoCom message, \( msg_2 \) is the
Figure 5: Structure of a CogInfoCom channel in terms of a single sensory modality, and in terms of the conceptual mapping techniques available to link the channel with CogInfoCom message generated concepts.

transmitted CogInfoCom message, and \( c \) is a constant, and \( \gamma \) is a scaling variable.

A conceptual schematic of various mapping techniques can be seen in Figure 5. The two aspects of the concept algebra based framework and the various forms of conceptual mapping provide a theoretical framework which may allow for the future development of automated mapping between CogInfoCom message generated concepts and CogInfoCom channels. This opens a wide range of perspectives for autonomous CogInfoCom systems.

5.1.2 Thesis 1

I proposed a novel conceptual framework and a corresponding terminology of CogInfoCom channels in order to provide a unified approach to the design, interpretation and analysis of CogInfoCom interfaces. I developed a structure for the proposed concepts, with the four levels of the CogInfoCom icon layer, the CogInfoCom message layer, the CogInfoCom message generated concept layer, and the CogInfoCom concept mapping layer. I showed that the proposed structure can capture the dichotomy between iconic and message-like stimuli in interface design. By incorporating layers for iconic stimuli, message-like stimuli, as well as the conceptual interpretations they generate, the concept of CogInfoCom channels provides a unified view of previously conflicting definitions in the various fields that have synergic contributions to CogInfoCom.

Subthesis 1.1 I defined the concepts of CogInfoCom icons and CogInfoCom messages in order to capture the dichotomy between icon-like and message-like stimuli. I demonstrated that multi-modal sensory signals in general can be categorized as belonging to the CogInfoCom icon layer or the CogInfoCom message layer. I proposed to adopt the view that CogInfoCom messages are composed
of CogInfoCom icons at a lower level, and showed that such a view is not restrictive in terms of the applicability of existing results in interface design if we consider CogInfoCom messages without CogInfoCom icons as a degenerate case. The proposed conceptual framework is unique in that it allows for a unified treatment of iconic and message-like signals across various sensory modalities.

Subthesis 1.2 I gave the definition of CogInfoCom message generated concepts (CogInfoCom concepts) in order to clarify the fact that the notion of semantic interpretation is not independent of the sensory modality used. I demonstrated that a distinction exists among the CogInfoCom concepts which can be generated using, e.g., auditory icons and earcons, haptic icons and hapticons, or olfactory icons and smicons. I gave a formal definition of CogInfoCom concepts based on the Object-Attribute-Relation (OAR) model.

Subthesis 1.3 I defined CogInfoCom channels as ordered sets of CogInfoCom messages which convey information on different attribute values of the same CogInfoCom message generated concept. By encapsulating both the sensory modality through which CogInfoCom messages are perceived, as well as the generation parameter types that are used to generate CogInfoCom messages, the concept of CogInfoCom channels helps to alleviate conflicts between user-centered, system-centered and encoding-centered aspects of multimodal interaction in a way that is transparent to the sensory modality being used.

Subthesis 1.4 I defined the CogInfoCom concept mapping layer, which is situated above the CogInfoCom message generated concept layer, and which describes the associations that can be created between CogInfoCom concepts that are generated by CogInfoCom messages of different sensory modalities. I introduced the problem of mapping between pairs of CogInfoCom channels as a key factor in creating cognitive associations between distinct sensory modalities.

Related publications: [RJ-4, RC-7, RC-11]

5.1.3 Thesis 2

I proposed a tensor algebra based canonical representation of CogInfoCom channels. The representation encapsulates the perceptual parameters that characterize CogInfoCom messages, the generation parameters that are used to create CogInfoCom messages, and the parameter-generating functions which link perceptual parameters and generation parameters in CogInfoCom channels. I proposed various direct and analogy-based techniques for mapping – and conceptually linking together – messages of a given CogInfoCom channel and CogInfoCom concepts normally received through different channels.

Subthesis 2.1 I proposed a tensor algebra based canonical representation of CogInfoCom channels that encapsulates both perceptual parameters and generation parameters that are used to synthesize CogInfoCom messages, as well as the discretized partially ordered multivariate parameter-generating function that creates a link between these two kinds of parameters. I proved that the representation allows for the comparison of different CogInfoCom channels in terms of a rank concept derived from higher-order linear algebra.
Figure 6: The schematic in this figure shows how SDM, the proposed interpretable tuning model works. Instead of using the original parameter space – denoted in this case by $p_1, p_2$ and $p_3$ (it should be noted that the number of parameters in general can be much more than 3) – it is always possible to describe a rough approximation of the original parameter space using the parameters of a discovery spiral. These parameters can be set transparently, without the user’s intervention (for example, the direction of the spiral can correspond to the direction of the principal component of a set of “control points” already specified by the user; parameters $r$, $d$ and $\alpha$ can be set so that the hyperspiral covers a representative portion of the parameter space, as required by the application, etc.). The user, in turn, is only required to manipulate 2 parameters: a “distance” parameter in terms of discrete steps along the hyperspiral (denoted by $s$ in the figure), and a “velocity” parameter which specifies the size of the steps (denoted by $v$ in the figure). An important idea behind SDM is that the user’s perceptual sensitivity will generally have some tolerance, thus it is sufficient to discover a structured subspace of the parameter space rather than its entirety.

**Subthesis 2.2** I defined the high-level and low-level direct mapping techniques, as well as the analogy-based mapping techniques of structural mapping, corroborative stimulation and scenario-based mapping. I gave formal definitions of these mapping techniques using the Object-Attribute-Relation (OAR) model. Through the formal definitions, I gave a referential framework to prove the viability of the direct and analogy-based mapping techniques.

*Related publications: [RJ-4, RB-1, RC-7, RC-11, RC-12]*

### 5.2 Thesis 3

#### 5.2.1 Summary of results

The dissertation investigates the necessity for the user-controlled tunability of CogInfoCom channels. In this regard, it reaches the conclusion that user-controlled tunability is important, due to the variability in cognitive capabilities of different users, due to the variability in application scenarios in which CogInfoCom channels are applied, and finally – but not less importantly – due to the high dimensionality of the generation parameter space of the majority of CogInfoCom channels. In order to allow users to manipulate an intuitive and possibly minimal set of tuning parameters, while still being able to control key variabilities in the generation parameter space, the dissertation introduces the Spiral Discovery Method (SDM). SDM operates through the following steps:
1. Assuming that the user would like to tune the system in the \( k \)-th dimension (as in Equation 5), the first step is to reduce its rank from \( \hat{I}_k \) to 1. This can be achieved using Higher-Order Orthogonal Iteration (HOOI) – a method which is proven to yield optimal rank-reduction [17] – to obtain:

\[
\arg\min_{\hat{S}, \{\hat{\mathbf{X}}_n\}_{n=1}^{N+1}} (\mathcal{F}^{D(G)} - \hat{\mathcal{F}}^{D(G)})
\]  

(12)

where

\[
\hat{\mathcal{F}}^{D(G)} = \hat{\mathcal{S}} \bigotimes_{n=1}^{N+1} \hat{\mathbf{X}}_n
\]

(13)

such that \( \hat{I}_k = 1, \hat{I}_n = I_n, \forall n \neq k \), and \( \hat{I}_k \) is the rank along the \( k \)-th dimension of the new system.

2. Having obtained this approximation, the goal is to increase the rank of the vector space which can be controlled by the single tuning parameter remaining in the \( k \)-th dimension. In order to achieve this, weighting matrices \( \tilde{\mathbf{X}}_n, n = 1..(N + 1) \), as well as the core tensor, \( \hat{\mathcal{S}} \) are expanded such that the obtained weighting matrices \( \tilde{\mathbf{X}}_n \) and core tensor \( \hat{\mathcal{S}} \) can be used to reconstruct the tensor representation of the original parameter-generating function:

\[
\mathcal{F}^{D(G)} = \hat{\mathcal{S}} \bigotimes_{n=1}^{N+1} \tilde{\mathbf{X}}_n
\]

(14)

where each \( \tilde{\mathbf{X}}_n, n = 1..(N + 1) \) has a rank of \( P_n \) (the number of discretization points along the given dimension, which is equivalent to saying that there are at least \( P_n \) columns in the \( n \)-th matrix), and \( \hat{\mathcal{S}} \) is augmented appropriately. After this step, even if the single tuning parameters (i.e., in the first column of \( \hat{\mathbf{X}}_k \)) are still only be used to control a single dimension within the system, the space that can be reconstructed by the system as a whole is \( \mathbb{R}^H \) once again (in other words, the original system can be fully reconstructed if the values in the augmented part of \( \hat{\mathcal{S}} \) are chosen appropriately).

3. By systematically modifying certain elements in the augmented part of the new core tensor, \( \hat{\mathcal{S}} \), as the first column of \( \tilde{\mathbf{X}}_k \) is modified by the user, it becomes possible for the user to traverse a subspace \( \mathcal{V} \subseteq \mathbb{R}^H \) such that \( \mathcal{V} \) is also \( H \)-dimensional.

A schematic explanation of the augmentation described in the previous points can be seen in Figure 7. SDM has the following properties:

- It is generic in the sense that it is not limited to a given sensory modality, and it can be used for any kind of CogInfoCom channel as long as the channel is parametric (there are other models in the literature for the tuning of e.g. audio signals, but they lack the generality of the proposed model)
Figure 7: Proposed augmentation method used to compensate for reduced interpretability after the tuning model is rank-reduced, assuming that $k = 1$ and $N = 2$. The light (yellow) shades contain fix values, and the dark (mauve and green) shades contain variables which can be chosen so as to compensate for the rank-reduction as well as to alter the slope of the hyperline that is traversed within $\mathbb{R}^H$ during tuning.

- It allows for the systematic traversal of parameter spaces while compromising precision. Due to the plasticity of the human brain, such an approach is viable and useful in cases where the goal is to look for solutions that are sufficiently good, and finding optimal solutions is not essential (Figure 6). When users are given a highly parametric, non-linear model to tune, it is impossible for them to understand and control the complexities of the situation without a structured approach like SDM.

The value of SDM is that it allows the user to trade off the complexity of the original system and the interpretability of a rank-reduced system. By iterating through the actions listed above in an interactive way, the user is allowed to either tune the set of sounds in the original parameter space, in a rank-reduced parameter space, or in a rank-reduced parameter space in which the principal components are systematically altered\textsuperscript{4}. Due to the fact that human perception is not crisp in the sense that small changes can in many cases be tolerated and adapted to, the SDM method can serve as a valuable cognitive artifact for tuning CogInfoCom channels.

In the final part of the dissertation, the usability of SDM was tested by asking test subjects to reorder and tune 20 gradations of an abstract dimension with 24 generation parameters. The meaning attributed to the dimension was left for the test subjects to decide, as a result of which a large variety of CogInfoCom messages were created. Results showed that a significant number of users were able to create and tune their own CogInfoCom channels and achieve test results which were well exceeded results which would have arisen by chance.

\textsuperscript{4}Note that the interactive scheme described here is somewhat relevant to the paradigm of Interactive Evolutionary Computation, as described by Takagi in [36], in the sense that the user’s subjective evaluations are used to guide an iterative search process.
5.2.2 Thesis 3

I showed that it is necessary to address the user-controlled tunability of CogInfoCom channels due to the variability of users’ cognitive capabilities, the high dimensionality of generation parameter spaces and the variability of application scenarios. I proved that the tensor algebra based canonical representation of CogInfoCom channels provides a generic tuning model if the weighting functions are interpreted as tuning parameters. I showed that while the generic tuning model is a minimal structure in the mathematical sense, its complexity is in the general case intractable when used as a tuning model. To address this problem, I developed the Spiral Discovery Method (SDM), which reduces the number of tuning parameters to two controls on the user interface (representing a distance along a hyper-spiral, and a sensitivity parameter), irrespective of the dimensionality of the generation parameter space. I verified that the numerical implementation of the Spiral Discovery Method yields effective trade-offs between the contradictory requirements of high complexity and high interpretability.

Subthesis 3.1 I introduced the problem of user-controlled tunability in the context of CogInfoCom channels. Based on the variability of users’ cognitive capabilities, the variability of the complexity of CogInfoCom channels, and the variability of the environments in which CogInfoCom channels are used, I argued that such tuning models will be necessary components in a large class of CogInfoCom systems. I proposed a generic tuning model of CogInfoCom channels based on their tensor algebra based orthonormed canonical form. I proved that the tuning model provides a minimal representation for the local tuning of individual CogInfoCom messages in CogInfoCom channels.

Subthesis 3.2 I showed that the dimensionality of the generation parameter space subject to tuning operations is generally large, and that its structure is in general highly non-linear. To address this problem, I developed the Spiral Discovery Method (SDM) based on the mathematical structure of the HOSVD-based generic tuning model. I proved that the Spiral Discovery Method creates a systematic relationship between the complete set of generation parameters and two characteristic control parameters. I provided a geometrical interpretation of the Spiral Discovery Method, and I proved that the two characteristic parameters control a discovery of the generation parameter space along a hyper-spiral structure. Through empirical usability tests, I verified that the tuning structure provided by the Spiral Discovery Method remains both complex and interpretable even in the case of high-dimensional and non-linear generation parameter spaces. I verified that users were able to control the discovery process using only the two characteristic parameters of the Spiral Discovery Method.

Related publications: [RJ-2, RC-7, RC-12, RC-13]

5.3 Thesis 4

5.3.1 Summary of results

The final part of the dissertation focuses on the validation of the proposed framework through experimental results and laboratory measurements. The part on practical applications is the result of the design and implementation of a large number of hardware and software components – based
on both open source and industrial standards – which can be flexibly configured for domain-specific CogInfoCom applications. All of these implementations are available at the 3DICC Laboratory, a joint laboratory of BME TMIT and MTA SZTAKI [37]. The following practical achievements are included in the dissertation:

- A sensor-bridging CogInfoCom application is described, in which audio-based CogInfoCom channels are used to provide feedback information on the 4 tactile dimensions of softness, roughness, stickiness and temperature. The motivation for the application is to allow users of immersive virtual realities to gain information on what the touch of a surface “feels” like, through low-cost audio equipment, without having to invest in expensive tactile feedback devices.

- Usability tests were performed, which demonstrate the test subjects’ ability to distinguish among 10 distinct gradations in each of the 4 tactile dimensions to various, but generally acceptable degrees. The variability of test subject results also supports the necessity to create user-controlled models (hence, test results also provide a posteriori motivation for the development of SDM).

5.3.2 Thesis 4

I proposed a solution to the problem of remote tactile sensing. I expressed the problem using the framework of inter-cognitive audio-tactile sensor-bridging CogInfoCom channels. I modeled tactile sensing in terms of the perceptual parameters of tactile softness, tactile roughness, tactile stickiness and temperature. I used high-level direct mapping and scenario-based analogical mapping to design earcons representing 10 gradations in each of the 4 perceptual dimensions. I verified the effectiveness of the sensor-bridging application through empirical usability tests.

Subthesis 4.1 I proposed a solution to the problem of remote tactile sensing. I modeled the problem along the 4 dimensions of tactile softness, tactile roughness, tactile stickiness and temperature. I proposed a mapping between 10 perceptual gradations along each of the dimensions and a set of earcons based on the mapping techniques of high-level direct mapping and scenario-based analogical mapping. I expressed the resulting earcons in analytical form. Based on the structure of the generation parameter space and the specific generation parameters, I expressed the proposed CogInfoCom channel in its tensor algebra based canonical form.

Subthesis 4.2 I carried out empirical usability tests with the help of 10 test subjects in order to validate the effectiveness of the audio-tactile sensor-bridging solution. The obtained error statistics demonstrated that while performance varied across test subjects, several test subjects were able to learn the mappings between audio-based CogInfoCom messages and the concepts of tactile softness, roughness, stickiness and temperature. Through these results, I have confirmed both the usability of the solution, as well as the necessity of user-controlled tunability for those users whose results were sub-optimal.

Related publications: [RB-1, RC-17]
6 Applicability of Results

The primary achievement of this dissertation is that it formalizes the concept of CogInfoCom channels and provides a set of tools for their design which is complete in the sense that it encompasses the entire design process. Thus, various methods are developed for creating CogInfoCom icons and messages, for linking them together with high-level CogInfoCom concepts at various cognitive levels, and for tuning them so as to allow for their conformity to any given user’s unique cognitive capabilities.

The results of the dissertation can be applied in a variety of contexts. Because CogInfoCom channels are capable of providing users with information on high-level concepts through sensory modalities other than the ones normally used for the same purpose (or through the same sensory modalities, but in different ways), they can be applied in environments where the user’s attention span or other cognitive capabilities are limited, or different from the normal case. Such scenarios can arise in e.g. virtual collaboration, Future Internet applications such as smart home applications, or during tasks which demand exclusive attention, such as driving a car.

Several results of the dissertation have been implemented in the Virtual Collaboration Arena (VirCA), which is a 3D virtual collaboration environment developed in the 3DICC laboratory – a joint laboratory of BME TMIT and MTA SZTAKI [37, 44]. VirCA is a distributed system which includes a modular architecture of loosely coupled components that can communicate with each other in transparent ways, and a 3D visualization engine. In particular, the following modules have been implemented in VirCA based on the dissertation:

- The mathematical set of tools used to describe and manipulate CogInfoCom channels.
- Sensor modules capable of transmitting data on the surface properties of 3D objects which influence tactile percepts. Thus, modules have been developed for the transmission of softness, roughness, stickiness and temperature.
- Sound-generating modules capable of mapping data on tactile percepts and the description of CogInfoCom channels onto a structured set of auditory icons and earcons.

Based on several years of research on virtual collaboration and 3D virtual reality, it has become clear to researchers that sensory modalities other than vision and audition must be involved in user-system interaction if the goal is to provide a truly immersive experience. The audio-tactile sensor-bridging application developed as part of the dissertation is one example in which the traditionally inaccessible tactile modality is artificially made accessible to the user.

The Author’s Publications Relevant to this Dissertation

Book chapters

Journal papers

impact factor: 4.529

independent citations: 2


impact factor: 0.385
independent citations: 5

impact factor: 0.385
independent citations: 49

Conference papers


independent citations: 1

independent citations: 3


independent citations: 3


independent citations: 30


independent citations: 3


independent citations: 1


The Author’s Publications not Relevant to this Dissertation

Journal papers


independent citations: 3


independent citations: 5


independent citations: 3

Conference papers


independent citations: 23

22

independent citations: 1


Bibliography


