Auditory Displays in Human–Machine Interfaces

From GUNNAR JOHANNSEN, Proc. IEEE Apr. 04 p. 742-750

Definitions

- Human–machine interaction
  - all aspects of interaction and communication between human users and their machine via a human–machine interface
- Human–machine system
  - human users, the human–machine interface (HMI), and the machine
- Machine
  - any kind of dynamic technical system (or real-time application)
- The automation components of the technical system are denoted as supervision and control systems
- Decision support systems
  - provide advice for the human users, e.g., in fault diagnosis tasks.
- Human supervisory control: guidance
  - of computerized and automated technological systems
  - of human cooperative systems, e.g., in music performance

Multimodal interaction

- The main objective of the multimedia approach with multimodal displays for process control is to make the best use also of other human sensory modalities, in addition to the visual one.
  - avoid the visual overload
  - bring the operator closer back to the process (e.g. noise)

- The general problem of the HMI design remains:
  - which information is needed by the human user,
  - when,
  - in which form,
  - why?

II. Functionalities of human-machine interfaces

III. Classification of auditory display

Classifications:
- Application orientation
- User orientation
- Sound orientation

Application-Oriented Classification of Auditory Displays

- Alarms: designed in order to attract the human’s attention
  - based on task-oriented transparency and understandability with respect to the context of a particular application domain.
- Sounds
  - Vehicle guidance
  - Medicine
  - Cinema
  - Industrial plans
  - Mobile phones
Application-Oriented Classification

- Functional objectives
  - State and internal information
  - Functioning of system components
  - Intent communication
  - Request of human inputs
- Alarms and warnings
  - Urgency
  - Distinctiveness
  - Arousal
- Appeal of products
  - Desired emotional impressions

User-Oriented Classification

- User classes
  - different jobs, task allocations and responsibilities
  - Operators, engineers, managers
  - Complementary sounds for team members
- Communications with individuals preferences and needs
  - Depending on capabilities, impairments, tasks, environments
  - Sighted and blind users
  - Transforming data and graphics into auditory display
- Individual user capabilities
  - Experts, novices, occasional users
  - Practice and quality of hearing sense
  - Non musicians vs. musicians
  - Human errors

Sound-Oriented Classification

- Speech
  - Artificial speech
    - Recording vs. synthesis
  - Non speech sounds
    - Natural
      - Characteristic for components in real environment
      - Reused as sound symbols (metaphors)
    - Artificial
      - Music
      - Sound effects
      - Sound symbols (Earcons)
        - Provide specific meaning and content
        - Created with computers and synthesizers

Cognitive systems life-cycle development of auditory displays

- Example: GUI design and evaluation (GUIDE) method

IV: User-centered development of AD

- Design aim of human–machine interaction
  - Human centeredness (user orientation)
    - Emphasizes the psycho-physiological and cognitive foundations of human behavior
    - Emphasizes the individual differences between users
    - User modeling
    - Views of designers and users may be different
    - Task orientation
      - Like a top-down principle for overall purposeful achievements.

Discussion

- User participation cannot yield design ideas.
- Artistic creativity with sound design is required
- Sound designers can only be successful
  - If they understand the systems engineering associations of the application domain equally well in all task situations and in all user views.
Sound and Meaning in auditory data display

From Herman-Ritter
Proc IEEE apr. 2004

Auditory data display
- Make data from a wide range of application domains accessible to auditory inspection, analysis and summarization
- devise mappings from data to sound patterns
- tool to uncover hidden structures
- aid and enhance the currently much wider established techniques of data visualization for the purpose of interactive, or exploratory data analysis
- additional route to meaning in data that is both synergistic and complementary to visualization.

Auditory data display
- Strengths
  - the capability of our auditory system to process several streams of information in parallel;
  - to offer a high temporal resolution;
  - its high sensitivity for structured motion, in particular, rhythm;
  - its ability to function well even in noisy contexts.

Sound and meaning in speech and music
- Ability to extract meaning from sound patterns, provided they are drawn from a certain family of "privileged" encoding schemes delineated by the phonetic and syntactic structure of human language.
- capacity genetically encoded in the brain areas that process language.
- Requires a sufficiently long prior listening experience of our native language.
- Encoded in the largely conventional association between phonetic patterns and their word meanings.
- Information encoded along further nonverbal dimensions.
- Classification of most voices as being male or female
- Recognition of individual persons in a highly selective way from their voice
- Infer important additional aspects of their emotional state and even their health or momentary condition, such as being tired or out of breath.

Sound and meaning in speech and music (cont.)
- Prosody is a major channel across which many of the above features become transmitted.
  - makes speech more impressive than writing
  - encode emotional information
- shares its major elements with music:
  - intensity, melody, articulation, and rhythm.
- Songs: language as a carrier medium for melodic sounds
- Prosody and music performance
  - see story telling
- Strong role of culture imprint for the constitution of meaning.
  - interpretation of musical meaning can be extremely subjective.

IIB. Meaning from the Perspective of Function
- Meaning usually is closely related with function
- Language and music
  - predominant functions may be seen as communication and enjoyment.
- Further acoustic experiences
  - The oldest function of sound is alerting.
  - capability of auditory localization.
  - Locating a prey or predator
- Ability to learn new sounds: iconic meaning
  - the slamming of a door,
  - the arrival of a particular person from the sound of her footsteps,
  - the starting of a car.
- Acoustic scenery: complex mixtures of natural or artificial acoustic events
  - telling us about the current weather, the situation on a busy city place, or what is happening in a forest.
- Function: aid coordination of actions.
  - E.g. the coordination of footsteps of marching soldiers.
IIC. Meaning from the Perspective of Listening

- Listening types:
  - Musical listening
  - pitch, melody, or harmonic organization or rhythmic patterns of a sound signal
  - Also rhythm of a bouncing ball
  - Everyday listening
  - identify the sound source
  - relative location of the sound source
  - Analytical everyday listening
  - learning about properties of the sound-producing process
  - When we shake an opaque box and try to guess its contents from the sound

IID. Meaning from the Perspective of Physics

- Many aspects of meanings in sound have their origin in conventions
  - e.g., language
- Meaning rooted in conveying information about important physical properties
  - interaction sounds.
    - allow us to discriminate a remarkable number of object properties
  - relative movements bringing the objects into contact
- Generate sound from first principles
  - link between the situational features and the emitted spatio-temporal sound pattern
  - can be computationally too heavy

III. Auditory display: audification

- Audification: use the data values directly as a series of sound pressure values.
  - e.g., seismic data
- Parameters are a time compression factor and a level scaling factor.
  - Filters are usually applied to preprocess the sound further.
- High-dimensional data display
  - by mixing different audifications together
  - by using a multichannel sound system.
- Makes a number of useful data properties directly accessible to the human ear:
  - the variance of the data becomes audible as sound level, data set size as duration,
  - and pitch and timbre can reflect many aspects of the detailed time-resolved variation of the series.
- Requires many data points to deliver reasonably long sounds
- Useful for data in which important regularities are already reflecting temporal variations

Parameter mapping

- Parameter mapping sonifications
  - Superposition of data-driven sound events, e.g., instrument sounds, according to given parameters like onset time, duration, pitch, and amplitude.
  - Each data point is mapped into the parameters of a separate sound event
  - Offers much more flexibility than audification
  - the underlying instrument sounds as well as the data-to-parameter mapping can be specified by the designer
  - Require explicit knowledge of the employed mapping
  - Non trivial the specification of a good mapping

Earcons

- Earcons: auditory patterns, usually composed of musical sounds, that represent a message in a short musical motive.
- Association from an earcon to its meaning has to be learned.
  - Each earcon represents an entire message of its own;
    - several earcons can be combined into a sequence to represent more complex messages
  - Suitable to convey symbolic messages, but limits their use for displaying continuous-valued or high-dimensional data items
  - Sonifications of such data by other means can benefit from earcons by embedding them as symbolic acoustic markers to annotate particular parts of the underlying continuous sonification.
Auditory icons

- Auditory icons follow the same purpose as earcons, to convey abstract symbolic messages by using nonspeech audio.
- They do not base their meaning on a mere convention (which can only be acquired by learning),
- but instead employ a crisp sound metaphor to encode their message.
- Ex.: a trash container sound to confirm the deletion of a file

The main problem with auditory icons
- for many messages (e.g., “silence”) it can be very difficult or even impossible to find an adequate sound pattern.
- Not suited for presenting high-dimensional data sets

Parameterized auditory icons

- Parameterized auditory icons
  - borrows some additional features from parameter mapping
  - in order to convey additional analog information by suitably controlling the parameters of the icon sound
- Example of trash container:
  - sound level \( \Rightarrow \) the size of the deleted file
  - sharpness \( \Rightarrow \) the elapsed time since the most recent modification date