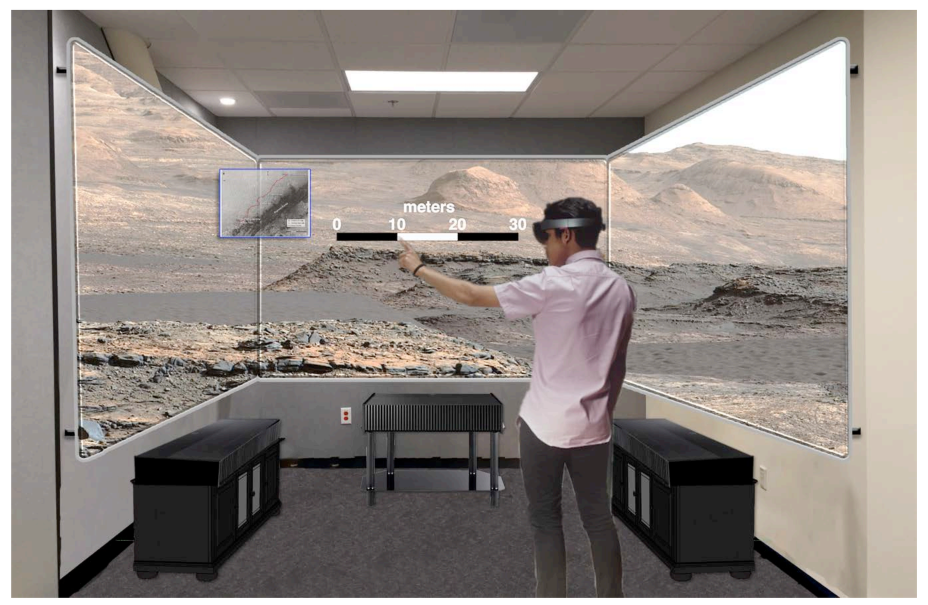
Advanced Multimodal Solutions for Information Presentation

Elizabeth M. Wenzel[1], Martine Godfroy-Cooper[2]
[1]NASA Ames Research Center, Moffett Field, CA 94035, [2]San Jose State University Foundation, Moffett Field, CA, 94035

Multimodal Display Rationale

High-workload, fast-paced, and degraded sensory environments (e.g., during EVA and telerobotic operations) are the likeliest to benefit from multimodal displays that can:

- Enhance situation awareness and task performance by maximizing the effectiveness of different sensory channels based on appropriate interaction between modalities.
- Play an important role informing interface guidelines for long duration exploration missions (LDEMs) requiring greater crew autonomy with increased dependence on spacecraft information systems for both routine and time/safety critical tasks.



A-VIEW Mixed Reality System: Mars Surface
NASA Ames Human Systems Integration Division (courtesy Kenji Kato)

Benefits of Multimodal Displays

- Increased bandwidth: increase in the amount of information that can be transmitted over a fixed time period.
- Redundancy: presentation of the same information in more than one sensory channel (e.g., critical alarms).
- Disambiguation: information from different sensory channels is combined to avoid ambiguity.
- Modality appropriateness: using sensory channels based on their suitability for presenting a particular kind of information (e.g., auditory channel for short commands, alarms).
- Complementarity: presenting related information in different sensory channels that should be merged to form a unified percept of an object or an event.
- Substitution: presenting information in an alternative modality when other sensory channels become temporarily or permanently unavailable.

Multisensory Integration

Evidence from both behavioral & neuroscience research demonstrates extensive cross-modal links between vision, audition, and touch:

- Modality expectations: a top-down influence on attention allocation. Leads to faster RTs and increased accuracy in the expected channel.
- Modality shifting effect: performance degrades when attention is shifted from the expected modality to an infrequently used channel. RTs slower for events in less frequent modalities.
- Cross-modal spatial and temporal links form the basis of the temporal and spatial "rules" that determine the likelihood and the strength of multisensory integration.
  - Crossmodal spatial links lead to enhanced facilitation of responses when simultaneous stimuli are at the same location or to response suppression for stimuli at different locations or that are separated in time.
  - Perfect spatial and temporal alignment is not required for multisensory integration to occur as long as the modalities are presented within close spatial and temporal proximity (e.g., as defined by psychophysical threshold studies).

Literature Review: Applied Studies of Multimodal Displays

Multimodal display research caveats:

- Often developed in a trial and error manner
- Don't consider basic mechanisms of human multisensory integration and cross-modal attention
- Equivalence between stimuli in each sensory channel (e.g., comparable detection thresholds) is rarely established prior to a study
- Performance measures not always directly compared among all possible combinations of unimodal, bimodal, and/or trimodal displays
- Clear inferences about relative multisensory benefits can be problematic

Examples of Applied Studies of Multimodal Displays

Automobile displays:

- Multimodal automobile interfaces such as collision event warnings (visual, auditory and/or tactile) display the source and spatial location of potential collisions
- Research results are inconsistent regarding whether multimodal displays produce better performance compared to unimodal displays. Performance varies considerably depending on factors like the spatial congruency (physical location) of the different display modalities.

Military & Aviation displays:

- A simulated ground combat vehicle study compared the effectiveness of unimodal, bimodal, and trimodal threat warning displays.
  - When all three cues were spatially congruent, threat acquisition was significantly better for a trimodal HUD display (visual HUD, 3D audio, and tactile belt) compared to bimodal displays (not all bimodal combinations tested).
- Similar benefits observed for multimodal directional alerting systems under conditions such as high acceleration in simulated aerial combat, in the presence of helicopter noise, and during air traffic control.

Effects of workload

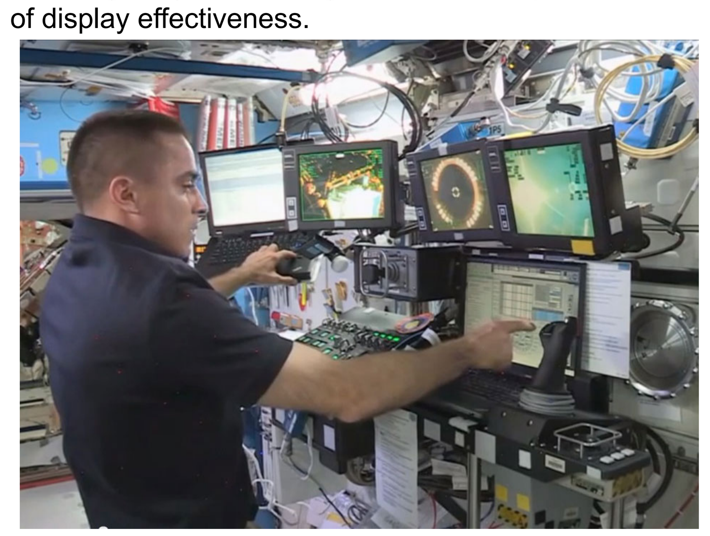
- Studies show multimodal displays produce differential effects depending on workload, e.g., multisensory cues may become significantly more effective than unimodal cues under high workload.

Adaptive Multimodal Information Displays

- Widespread agreement in the literature that fixed assignments of modalities to tasks or types of information are not desirable or even possible.
- Multimodal interfaces must accommodate possible changes in the needs, abilities and experience level of the user, the types of tasks being performed, the task environment, and the level of workload
- Adaptive displays presuppose adequate methods are developed for detecting and/or predicting operator state or task conditions.
- Considerable debate remains regarding whether multimodal interface flexibility should take the form of system-controlled adaptivity (automation) and/or user-controlled adaptability (user preference profiles).
- Simple approaches based on user preference are not likely to be sufficient in complex task environments such as LDEMs where crew members will require at least some degree of automated support.

Multimodal Displays in Space Environments

- Although the effect of microgravity is relatively well documented for individual sensory systems, less is known about interactions between the senses.
- The normal contribution (weight) of each sensory modality to multimodal perception experienced on Earth will not be relevant in space, since the reliability of the different senses will change. For example, the usual dominance of visual cues in multisensory perception may decrease, and the role of auditory and tactile cues may increase.
- Some conclusions may be drawn from analog studies or conditions mimicking altered gravity but investigation in a true space environment will remain the best test of display effectiveness.



H-II Transfer Vehicle (HTV)-4 Capture and Berthing: Robotics Ground Controllers powered up the Space Station Remote Manipulator System (SSRMS), August 2013

Current Standards & Guidelines

Surprisingly little consistent guidance directly addressing multimodal displays is currently available in the form of standards and guidelines. These vary considerably in terms of their specific focus and level of abstraction. Guidelines:

- May be essentially unimodal, addressing the properties and preferable uses of the individual sensory channels.
- May focus on a very specific type of display for a very specific task such as multimodal warning signals for driver-vehicle interfaces.
- Are high-level design principles that can apply independent of modality, such as complementarity, consistency, and redundancy of information presentation either within or across sensory channels.
- May be more general guidelines concerned with the effective combination and integration of sensory channels, but they are primarily based on research using bimodal information and few directly address trimodal (or beyond) information integration.
- NASA 3001 standards for crew interfaces are either very high level or very specific. HIDH handbook recommendations/ guidelines also tend to be high level. Neither specifically address multimodal displays.

Examples of Current Standards & Guidelines

- MIL-STD-1472G (2012): Update of the widely used DoD Design Criteria Standard: Human Engineering that provides a few standards for combining visual and auditory modalities that emphasize the primacy of visual information.
- ISO 14915-3 (2002): International standard of **multimedia design principles** that focuses on the impact of media selection per se, rather than modality selection, based on supporting thematic congruence, manageable information loading, complementarity, consistency & redundancy.
- Sarter (2006, 2013) reviews and critiques current guidelines for multimodal displays and outlines questions that should addressed in formulating guidelines.
- Reeves, et al. (2004) is an oft-cited review article that discusses six categories of guidelines: Requirements Specification, Designing Multimodal Input and Output, Adaptivity, Consistency, Feedback, Error Prevention/Handling.
- Giang, et al. (2010) is a comprehensive document published by Defence Research and Development Canada that addresses a variety of topics related to the design of multimodal interfaces.

Recommendations for NASA

- Incorporate multimodal (MM) guidelines into HIDH as mid-level guidelines that elaborate higher level recommendations (this should be done for a number of topic areas in HIDH).
- Clarify what are true multimodal guidelines and their underlying rationale.
- Provide pointers to other HIDH sections for guidelines that are not specific to multimodal interface design or that focus on the choice of individual sensory channels for given tasks and contexts.
- Assemble a team to formulate, critique & reiterate guidelines:
  - Experts with specific experience in the areas of multimodal displays and cross-modal perception
  - NASA, DoD, industry personnel with interface development experience in both applied research & operational/mission contexts
  - Experts in MM software/hardware architecture, e. g., with experience in virtual/augmented displays
- Significantly invest in research to validate proposed guidelines & multimodal prototypes

Reference

Wenzel, E.M. & Godfroy-Cooper, M. (Nov., 2017). Advanced Multimodal Solutions for Information Presentation. Final Report to the Human Research Program.