Multi-Modality Trends and Strategies in Automotive MMI

Chris Wild
W3C
21 - June 2005
Toulouse
Agenda

• The Problem
• The Devices
• Ergonomics
• MMI Architecture
• Further Challenges - ADAS
Information and Control for Drivers

- **Basic Driving controls and Displays**
  - Driving Cluster (Dashboard)

- **Security related information**
  - Alarms for car failures
  - Seat belts/Reverse gear
  - Maintenance

- **Entertainment control and distribution**
  - Video/TV
  - Audio

- **Navigation and Advanced Driver Assist (ADAS)**
  - Route planning and guidance
  - Speed and obstruction management
  - Enhanced driving security

- **Professional and lifestyle**
  - Phone
  - Email
  - Telematics/Fleet services
Trend 1: Massive Increase in Information Flows

Example:

BMW Series 5 Infotainment system
- 300+ Screens
- > 1000 user functions
Device Possibilities
Graphic Devices

Infotainment screens

Head up display

Instruments
Audio Devices

Audio out (Loudspeakers, mixer, amplifier)

Speech recognition
- Discrete
- Continuous
- Phoneme-based
- Natural language

Media Player
Sound Generator
Haptic Devices

Programmable Haptic feedback devices (rotaries and joysticks)
- e.g. BMW Series 5 iDrive – Context dependent configuration and feedback
Gesture

**Character recognition**
- Finger writing on pressure pad

**Virtual Interface**
- Control gestures
- Driver monitoring

![Camera](image1)

![Touchpad + Rotary](image2)
Trend 2: Device Diversity
Which Devices and When: Ergonomy

- Acuity
- Expectations
- Attention management
- Facility
- Semantic Load
- Design
- Legality
MMI Architecture For Multi-Modality
Device to Event

Keypad

Screen

Phone
Traditional State Machine Model

- **Driver**
- **Events**
- **Commands**

**Graphics**

**State machine**
Add Another Device – with Context
Mixing Devices – Step 1 (State of the Art)

Intermediate abstractions needed to allow source of events to be hidden from the state machines and applications.
For "context-free" devices this allows for mixed mode input.
From Events to Dialogue

- Context implies dialogue – prompt, input, confirmation, errors etc. Speech recognition is the classic example but within a multi-modal system dialogue become the basis for all user management

BUT
- Not a solved problem
What is Envisaged

Dialogues

Device Abstraction

Dialog Engine

Context Engine

Device 1

Device 2

Device ...

Select

Run

Task/Application

Select
Trend 3: Rethink of Ergonomy Concepts

- Increased treatment of dialogues for devices (not just Speech)
- Multi-modal possibilities for single tasks (entering a phone number for example)
- Increased research on Frameworks and concepts (big step for the automotive industry)

BUT

Legal issues may have the greater impact.
Further On: Augmented Reality for Advanced Driver Assist

**Combination of**

- Real-time images
- Database Images
- Graphics
- Sound
- Animation
- Haptic Feed-back

=> Attention management for Security
Route Presentation
Navigation databases can be used to identify problem areas such as dangerous junctions.

Combinations of enhanced reality, animation and 'appropriate' sounds can be used to ensure that the driver's attention is pulled onto the approaching risk. (similar to the image management in the film "MATRIX")