Intelligent Driving Data Analysis

Kari Torkkola

Motorola Labs

Human Interface Lab

Tempe, AZ, USA

Contents of the Presentation

- What is Motorola Driver Advocate™?
- Driving Simulator
  - Data collection
- Characteristics of driving data
- Data mining / machine learning problems in driving domain

MOTOROLA LABS

Driver Advocate™

- Problem:
  - Productivity, entertainment in the automobile environment
  - Driver distraction due to these and other factors
- Solution:
  - Use intelligent systems to control distraction, reduce cognitive load, and aid the driver in his tasks
  - Create an assistant to aid the driver -- not a substitute for the driver -- using artificial intelligence technologies (machine learning discussed here)

MOTOROLA LABS

Driver Advocate™

The **right** information
at the **right** time
delivered the **right** way

“Right” is personal.

MOTOROLA LABS

Driver Advocate™ Defined

Vision

Manage driving and non-driving information to enhance road safety

Concept

Intelligent system controller integrates, prioritizes, and manages information from sensors and devices, and delivers through a multimodal user interface

- Roadway: weather, location, adjacent vehicles...
- Vehicle: tires, speed, braking, steering, yaw rate...
- Cockpit: personal UI, occupant sensors, infotainment, navigation...
- Driver: driving performance, distraction, drowsiness...
- Communications: cell phone, web browsing…

Goals

- **Improves driving safety by enhancing driver situational awareness**
- **Reduces distraction by directing driver's attention to critical tasks**
- **Alerts the driver to potential road hazards**

How about Pilot's Assistance Systems

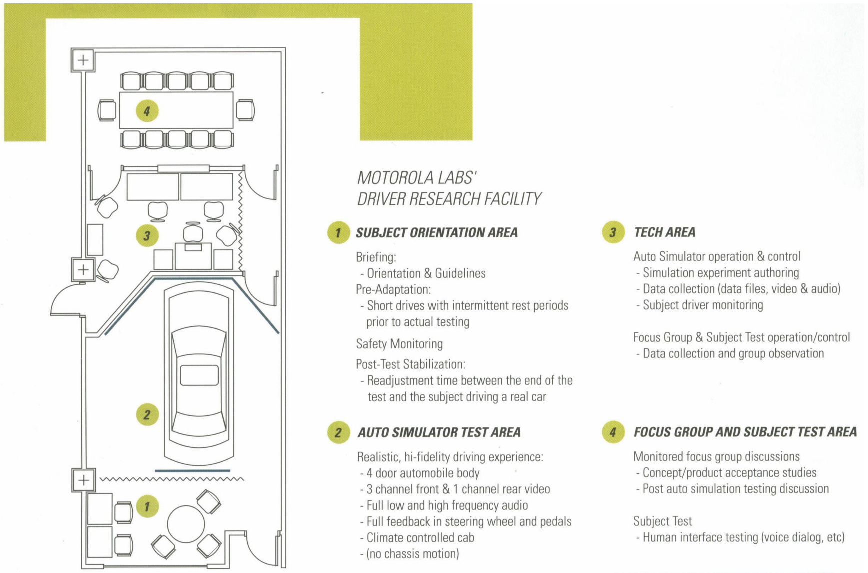
- **Pilot vs. Driver (training received is different)**
- **Response time (sky vs. road)**
- **Traffic conditions**
- **Systems with different sophistication**
- **Ranges of interference**
- **Degrees of mission Criticality**

How to get there?

- **Target: Intelligent Driving Assistance Systems**
- **Research Task**
  - **To investigate information presentation to vehicle operators to improve safety in the face of distraction and workload**
  - **In other words: better behavior with a *non-autonomous vehicle***
- **Our Starting Place**

  An automobile simulator which is currently being used for Human Factors research

Motorola Labs Driver Research Facility



Auto Simulator Test Area

Tech Area

Driver Monitoring

Demo Video Clips

- **Channel 12 News**
- **Eye Tracker**

Data Collection- ~40GB/HF experiment

KQ Sim
76 Variables
60 samples/sec
1MB/min

Driver Advocate™
244 Variables
60 samples/sec
0.1MB/min (compressed)

faceLAB
88 Variables
30
samples/sec
1.2 MB/min

Time	Frame	Velocity	Lane Pos	Steer	Brake	Subject X	Subject Y	LatAccel	Long Accel	VehAhead	Headway Time	Headway Dist	TTC
586.70	35200	20.82	0.741	-4.00	0.0350	97.46	844.79	-0.254	-0.693	LeadCar1	2.137	44.502	7.538
586.71	35201	20.81	0.741	-4.00	0.0540	97.46	844.45	-0.254	-0.792	LeadCar1	2.134	44.404	7.475
586.73	35202	20.79	0.741	-3.80	0.0650	97.46	844.12	-0.251	-0.997	LeadCar1	2.131	44.322	7.419
586.75	35203	20.78	0.741	-3.80	0.0750	97.46	843.75	-0.251	-1.161	LeadCar1	2.128	44.205	7.362
586.76	35204	20.76	0.741	-3.70	0.0770	97.46	843.44	-0.249	-1.204	LeadCar1	2.126	44.140	7.313
586.78	35205	20.74	0.740	-3.50	0.0680	97.46	843.08	-0.243	-1.184	LeadCar1	2.123	44.022	7.260
586.80	35206	20.72	0.740	-3.30	0.0750	97.46	842.73	-0.241	-1.215	LeadCar1	2.120	43.922	7.209
586.81	35207	20.70	0.740	-3.10	0.0770	97.46	842.41	-0.235	-1.205	LeadCar1	2.118	43.837	7.160
586.83	35208	20.68	0.739	-2.90	0.0770	97.46	842.04	-0.226	-1.224	LeadCar1	2.114	43.718	7.108
586.85	35209	20.66	0.739	-2.60	0.0780	97.46	841.70	-0.220	-1.249	LeadCar1	2.112	43.616	7.059
586.86	35210	20.64	0.738	-2.40	0.0830	97.46	841.37	-0.211	-1.272	LeadCar1	2.110	43.531	7.012
586.88	35211	20.61	0.738	-2.10	0.0850	97.46	841.01	-0.199	-1.311	LeadCar1	2.106	43.410	6.964

Audio/Video
Stereo/Quad Screen
MPEG-2 30MB/min

Personal Data,
Workload, Value Judgments
129 Variables x 20

Collected Data

- Data is
  - Multimedia, mixed type
  - Sequential, streaming
  - Abundant
  - High temporal resolution
- Three main purposes
  - Analysis of human factors experiments (standard statistics of a couple of variables)
  - Machine Learning
  - Data Mining

Machine Learning

- Learn driver models from the data, for example,
  - Steering response
  - Driver's cognitive workload
  - Attentional model (from eye-gaze)
- Learn classifiers for driving states from the data
  - High-workload traffic vs. leisurely cruising (do we let the cell-phone call through or not?)
  - Learn building-blocks for modeling driving state sequences ("drivemes")
- Learn to give advice

Demos

- Annotation tool demo
- Learning steering response demo

Why Learning to Give Advice?

- Avoid programming the response of the DA to every imaginable situation, but rather …
- Learn the response of the DA from either simulated or collected data, or on-line, and …
- Learn in a way that generalizes well to unseen situations.
- Learn user behavior/desired response through interaction.

More intelligence for intelligent systems!

Steps to an Intelligent Driver Advocate

- Technology Survey:
  - Key players, trends
- Machine Learning Approach:
  - An architecture for modeling driver distraction
  - Components of the architecture
  - Key research needed
- Where we are now:
  - Simulator facility
  - Driver monitoring
  - Data collection
  - Machine learning tool development
  - Experimentation

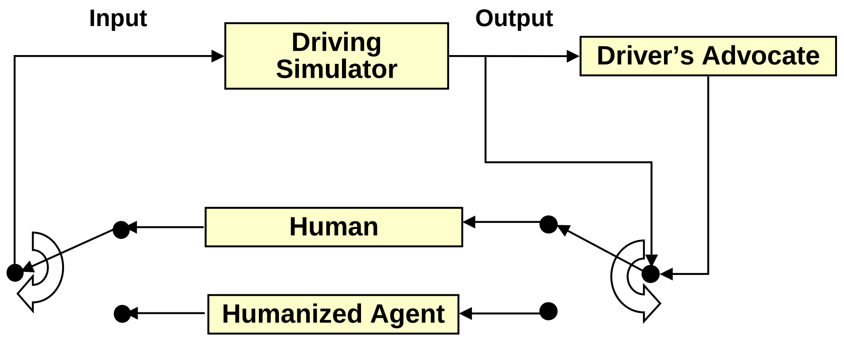
Technology Scan: Machine Learning for Driving

- Top players (young field, no established players)
  1. Carnegie Mellon
  2. Daimler Chrysler
  3. Cambridge Basic Research (Nissan)
  4. MIT
  5. University of Michigan
- Observations
  - Almost half of the published papers were concerned with pattern recognition and feature extraction, that is, generating something more intelligent from the raw sensory input
  - User modeling is also a dominant area

Machine Learning Architecture (long term!)

- Learn how to advise a driver to avoid distraction
  - Can be viewed as a human “control” problem
  - Limit response to alerts; do not control the car
- Learn how a user reacts
  - Model human response with sufficient fidelity
  - Create a humanized agent
- Use a simulator platform to simulate interaction
  - Avoids issues of safety
  - Can be made faster than real time

Machine Learning Architecture

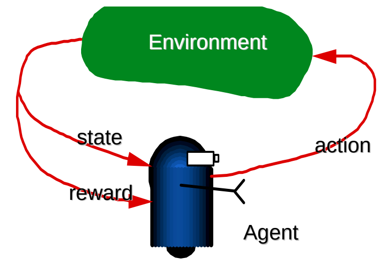


Humanized driving agent: An aide to help learn the Driver's Advocate, bootstrapped from user models .

Agents may interact faster than real time and explore simulated driving incidents.

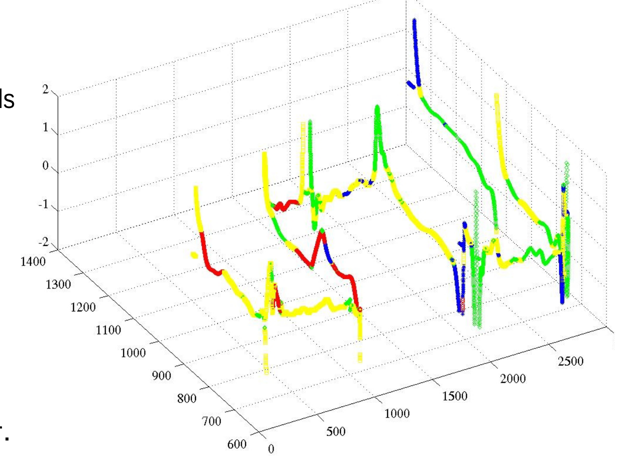
Reinforcement Learning (to Advise)

- **Humanized Agent**
  - **Data collection, incl. eye/head tracking**
  - **Supervised learning to simulate human response**
  - **Create fidelity only in areas needed**
- **Advising**
  - **Handcrafted rule system done**
  - **Bootstrap from those**
  - **Use reinforcement learning to learn how to advise a humanized agent**
  - **Difficulties: large state spaces, need for generalization**



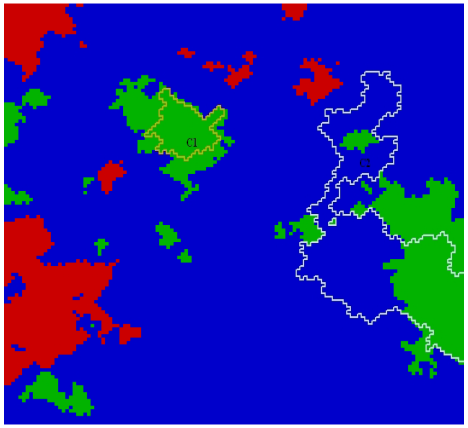
Driving State Modeling and Segmentation

- Can we divide world into **states** for which actions are learned?
- Unsupervised state segmentation using Hidden Markov Models (discovering "drivemes")
- Example: Four-state HMM segmenting a driving path. X,Y coordinates are the driving coordinates, Z is accelerator, segmentation is indicated by the color.



Data Mining Techniques for Driving Data

- Visualization and rule induction tools in the simulator environment
- Simple example: In what situations does driver use the left turn signal?
- Visualization of Signal=L outlined over the steering angle (green=left, red=right)



Signal=L, 13 rules covers 3072 items 73.2% (2248:0)

```
0: [accuracy:100.0%,coverage:15.1%,confidence: 99.8%]
                LaneIndex = LT1
1: [accuracy:100.0%,coverage:9.3%,confidence: 99.7%]
                Steer <= -243.50 (-125)
2: [accuracy:100.0%,coverage:3.7%,confidence: 99.3%]
                Slip = T
3: [accuracy:100.0%,coverage:8.2%,confidence: 97.4%]
                Velocity > 11.07 (-40)  AND
                Velocity <= 12.62 (-18)  AND
                LongAccel > -1.23 (-92)  AND
                LongAccel <= 0.73 (56)
 4: [accuracy:100.0%,coverage:8.1%,confidence: 97.4%]
                Steer > -211.00 (-125)  AND
                Steer <= -64.30 (-94)  AND
                LongAccel > -0.68 (-51)  AND
                LongAccel <= 0.83 (64)
5: [accuracy:100.0%,coverage:6.3%,confidence: 96.7%]
                Steer > -25.10 (-34)  AND
                Steer <= -1.80 (2)  AND
                SubjectPitch > -0.60 (-125)  AND
                SubjectPitch <= -0.10 (-73)
```

Conclusions

- Exciting area to explore:
  - Advising a driver to mitigate distraction
  - Analyze massive streaming multimedia databases
- Opportunities
  - Learn driver models from data
  - Learn driving models from data
  - Learn to advise from the data
  - In human factors experiments, instead of relationships between 2-3 variables, explore relationships between ~400 variables