Initiatives on Signal Processing in Smart Ambiance

Prof. Dr. Hüseyin Abut

FERP Professor of Engineering, ECE Department
And
Faculty of Engineering and Natural Science

Email: abut@anadolu.sdsu.edu abut@sabanciuniv.edu
URL: http://anadolu.sdsu.edu/abut

San Diego State University
San Diego, CA 92182, U.S.A.
Phone: (619) 594-3702
FAX: (619) 594-2654

Sabanci University
Tuzla, Istanbul, Turkey
(0216) 483-9505
(0216) 483-9550

Research Expertise & Interest Areas

- Next generation communication and information processing systems in wired/wireless platforms, systems, consumer products.
- Intelligence in vehicular systems.
- Intelligence in work place/house/school/health services.
- Speaker identification in access control and in electronic commerce.
- Data fusion for biometric identification.
- Medium-to-low rate speech coding for wireless communication.
- Signal processing in telemedicine.
- Intelligence in hands-free crane systems.

Select Recent Initiatives & Activities

- **DRIVE-SAFE:** Signal Processing and Advanced Information Technologies for improving Driver/Driving Prudence and Accident Reduction. Consortium formed by OTAM, Sabancı University, Koç University, Technical University of Istanbul, Ford Otosan A.Ş., Tofaş A.Ş., Renault A.Ş., Turkey. (Under Review)
  (February *2004 -)*

- **Turkey 20^3:** Establishment of a national initiative called **Turkey 20^3** with the purpose of 20 percent savings on energy consumption per capita no later than the year 2020 by development and deployment of devices, systems, and personal behavior retrofitted with tools of "living in smart ambiance."
  (July 2005 -)

- Emergency/Crisis Management Education Using DIstributed Sensor Networks ." Murat Tekalp, Reha Civanlar, H. Abut (Possible Leonardo Project Proposal) *(July 2005 -)*

- **Biennial** on DSP for In-Vehicle and Mobile Systems *(September 2-3, 2005, Sesimbra, Portugal)*

- **International Alliance for Advanced Studies on In-Car Human Behavioral Signals**
  (April 2003 -)

- **SWIM:** Special Workshop in Maui *(January 2004)*

- ***DSP for In-Vehicle and Mobile Systems***, H. Abut, J.H.L. Hansen, and K. Takeda (Editors), Springer, New York, NY, 2004, IBSN: 0-387-22978-7; Online Order Information: http://www.springeronline.com
  (December 2004)

- ***DSP for In-Vehicle and Mobile Systems 2***, H. Abut, J.H.L. Hansen, and K. Takeda (Editors), Springer, New York, NY *(Summer 2006)*

- **Distinguished Lecturer**, Class of 2002-2003, IEEE Signal Processing Society
  (January 2002-June 2003)

Selected Recent Publications

- ***DSP for In-Vehicle and Mobile Systems***, H. Abut, J.H.L. Hansen, and K. Takeda (Editors), Springer, New York, NY, 2005, IBSN: 0-387-22978-7; http://www.springeronline.com
- ***DSP for In-Vehicle and Mobile Systems 2***, H. Abut, J.H.L. Hansen, and K. Takeda (Editors), Springer, New York, NY *(Scheduled for Summer 2006)*
- ***Communication Systems: Signals, Systems, and Emerging Applications***, H. Abut, KB Books, San Diego, CA. *(September 2003, available on-line, free, September 2004)*
- E. Erzin, Y. Yemez, A.M. Tekalp, A. Erçil , H. Erdogan, and H. Abut, "Multimodal Person Identification for Human Vehicle Interaction," submitted for publication in the IEEE Multi-Media Magazine *(April 2005)*
- H. Erdoğan, A. Erçil, H. K. Ekenel, S. Y. Bilgin, İ. Eden, M. Kirişçi and H. Abut, "Multi- modal Person Recognition for Vehicular Applications," Chapter in *Multiple Classifier Systems*, Oza, Polikar, Kittler and Roli (Editors), LNCS Volume: 3541, Springer-Verlag, Berlin, Heidelberg, 2005.
- H. Erdoğan, A. Erçil, and H. Abut, "Driver-Identification Using Fusion," to be presented at the 2005 Biennial on DSP for In-Vehicle and Mobile Systems, Sesismbra, Portugal. *(September 2-3, 2005)*
- W. Abdul, C. K. Tan, H. Abut and K. Takeda, "Driver Recognition using FNN and Statistical Methods," to be presented at the 2005 Biennial on DSP for In-Vehicle and Mobile Systems, Sesismbra, Portugal. *(September 2-3, 2005)*
- "Biometric Identification Using Driving Behavior," K. Igarashi, C. Miyajima, K. Itou, K. Takeda, H. Abut and F. Itakura, Proceedings IEEE ICME June 2004, Taipei, Taiwan.
- W. Abdul and H. Abut, "Self-Organizing Fuzzy Neural Network for Speech Enhancement," Proceedings of the SWIM: Lectures by Masters in Speech Processing, January 2004, Maui, HI.
- H. Abut, "Tutorial on Digitized and Digital Signatures for Biometric Identification," Keynote Address at the 2003 SIU Meeting in Istanbul, Turkey.
- H. Abut, "Tutorial on Digitized and Digital Signatures for Biometric Identification," Special Tutorial Session Presentation at ICASSP2003, Hong Kong, P.R. China.
- Y. Öztürk and H. Abut, "System of Associative Relations", Signal Processing: Image Communication, Vol. 17, Issue: 3, Pp. 261-276, Elsevier Science.

Selected Recent Publications (Continued)

- "Speech Enhancement for Vehicles based on Multi-Dimensional Cerebellar Machine Computer Architecture (MCMAC)," H. Abut, Keynote Address, Proceedings of WOSPA2002, Brisbane, Australia.
  (December 2002)
- K. Igerashi, K. Takeda, F. Itakura, and H. Abut, "Is our Driving Behavior Unique?" Proceedings of the 2003 Workshop on DSP in Vehicular and Mobile Systems, Nagoya, Japan.
  (April 3-4, 2003)
- W. Abdul and H. Abut, "MCMAC as an Amplitude Spectral Estimator for Speech Enhancement," Proceedings of the 2003 Workshop on DSP in Vehicular and Mobile Systems, Nagoya, Japan.
  (April 3-4, 2003)
- W. Abdul, Q. Chai, and H. Abut, "Self-Organizing Fuzzy Neural Network for Echo Cancellation," Proceedings of the 2003 Workshop on DSP in Vehicular and Mobile Systems, April 3-4, 2003, Nagoya, Japan.
  (April 3-4, 2003)
- W. Abdul, E.C. Tan and H. Abut, "Simple and Fast Algorithm for Noise Cancellation using Stereo Microphone," Proc. of IEEE region 10 Conference, TENCON200.
  (December 2001)
- W. Abdul, E.C. Tan, and H. Abut, "Higher order CMAC for Signal Enhancement for Vehicular Environment," Proc. of ISSPA2001, Kuala Lumpur, Malaysia.
  (October 2001)
- A. Wahab, E. C. Tan, and H. Abut, "Robust Speech Enhancement Using Amplitude Spectral Estimator," Proceedings of the IEEE ICASSP2000 Silver Anniversary, Istanbul, Turkey.
  (June 2000)
- C.F. Buman, H. Abut, M. S. Al-Ghazi, and R.H. Yakoob, "Applications of Signal Processing to Conformal Radiation Therapy Dose Optimization," Proceedings of the IEEE ICASSP2000 Silver Anniversary, Istanbul, Turkey.
  (June 2000)

Güvenli- Sürüş

Sinyal İşleme ve İleri Teknolojiler Kullanarak Araç, Sürüş ve Sürücü Güvenliğinin Arttırılması ve Kazaların Azaltılması

Drive-Safe

Signal Processing & Advanced Information Technologies for improving Driving Prudence & Accident Reduction

Partners:

1. **Sabanci University, Istanbul, Turkey:** Erçil, (P.I.), Abut, Erdoğan, Yanıkoğlu, Keskinoz, Ates, Saygın, Gürbüz, Bozkurt
2. **Koc University, Istanbul, Turkey:** Erzin (Co-PI), Tekalp, Yemez
3. **OTAM Automative R&D Center, Istanbul, Turkey:** Göktan (Co-PI), Akalin, Atabay
4. **Faculty of Mechanical Engineering, ITU, Istanbul, Turkey:** L. Guvenc (Co-PI) and A.B. Guvenc)
5. **Ford Otosan A.S., Istanbul, Turkey**
6. **TOFAS (Fiat), Bursa, Turkey**
7. **Renault, Bursa, Turkey**

Current Funding Status:

1. EU Sixth Frame Project at Sabanci University (January 2005- December 2007).
2. EU Sixth Frame Project at ITU (January 2005- December 2007).
3. Turkish Development Agency funding of Drive-Safe (August 2005-Dec. 2006)

Projenin Gerekçesi

1. **Ülkemizde 2002 yılında karayollarında meydana gelen 400.000'den fazla trafik kazasında:**
   - **Yaralı: 116.045 kişi**
   - **Ölü: 4.169 kişi**

2. **Kazaların büyük bir coğunluğu sürücü yorgunluğu ve diğer davranış sorunlarından kaynaklanmaktadır.**

Özellikle, sarhoş ve yorgun sürücülerin sebep olduğu kazalardaki can ve mal kaybı daha ciddi boyutlardadır.

ÖRNEKLER

- **Amerika: "U.S. National Highway Traffic Safety Administration" tarafından verilen rakamlara göre şöför yorgunluğu her yıl Amerika'da 240,000 kazaya neden olmaktadır. Uyku ilintili kazalar can kaybının ötesinde Amerikan hükümetine ve iş dünyasına her yıl $46 milyar dolar maliyet getirmektedir.**

- **Avustralya: Kazaların nedeni: %30 yorgunluk/uyku**

- **Finlandiya: Trafik kazaları sonucu ölenlerin %26'sının alkolle bağlantılı olduğu gösterilmiştir.**

ÖRNEKLER (devam):

- **Kanada (1998): Kazalarda ölen sürücülerde yapılan alkol muayenesi sonucu %33'ünde alkol oranının yasal sınırın üzerinde olduğu tespit edilmiştir. (Kandaki Alkol Yoğunluğu-BAC>80 mg)**
- **'UNC Highway Safety Research Center' Bulgusu (1999): 1403 şöförle yapılan araştırmalara göre günde 6-7 saat uyuyan kişilerin 8 saat veya daha fazla uyuyan kişilere göre kaza yapma olasılıklarının iki katı olduğu gözlemlenmiştir. 24 saat uyanık kalmak kanuni olarak sarhoş olmakla eşdeğerdir.**
- **Alkol limitinin 2 katı üzerinde alkol almış bir sürücünün ölümlü bir kazada yer alma olasılığı 50 kat daha fazladır.**

DPT 8. Kalkınma Planındaki Bu Konudaki Devlet Politikasi:

Trafik kazalarına bağlı ölüm, yaralanma ve özürlü kalma durumlarının önlenmesi ve azaltılması için etkili tedbirler alınacak, acil sağlık ve kurtarma hizmetlerinin ülke genelinde yaygınlaştırılması ve geliştirilmesi sağlanacaktır.

Ülkemizde önemli can ve mal kaybına sebep olan bu kazaların azaltılması çabaları arasına teknolojik çözümler de dahil edilmelidir.

TÜRKİYE'DE ARGE SİNERJİSİ VE AĞI OTOSEKTÖR- 2003

TUBİTAK-MAM TARAFINDAN KOORDİNE EDİLEN, OSD, ÜNİVERSİTELER, TAYSAD VE OTOMOBİL KURULUŞLARININ KATILDIĞI " OTO-SEKTÖR" ÇALIŞTAYLARI VE ULUSAL VE/VEYA AB 6. ÇP KAPSAMINDA PROJELER YAPILMASI PLANLANMIŞTIR.

PROJELER:

- YAKIN GELECEKTE MOTOR PROJELERİ
- UZAK GELECEK MOTOR TEKNOLOJİLERİ
- AL VE MG ALAŞIMLARININ OTOMOTİV UYGULAMALARI
- ARAÇTA AKTİF VE PASİF GÜVENLİK
- AKILLI KOKPİT, ERGONOMİ

BELİRLENEN GÜÇLÜ YANLAR:

- ÜRETİMLE İLGİLİ BİLGİ VE DENEYİM
- GELİŞMİŞ ÜLKELERE İHRACAT
- YETİŞMİŞ VE GİRİŞİMCİ İNSAN GÜCÜ
- REKABETÇİ İŞ GÜCÜ MALİYETİ
- ULUSLARARASI STANDARDA UYGUN VE KALİTELİ ÜRETİM VE
- AR-GE YE YAPILAN YATIRIMLARIN BÜYÜMESİ

TÜRKİYE'DE ARGE SİNERJİSİ VE AĞI OTOSEKTÖR- 2003

GELİŞMEYE AÇIK YANLAR:

- ÖZGÜN TEKNOLOJİ VE ÜRÜN GELİŞTİRMEDE YETKİNLİK
- YAN SANAYİNİN TASARIMA KATILIMI VE TASARIM YAPMA YETENEĞİ
- MÜHENDİSLİK VE TASARIM KONULARINDA UZMANLAŞMIŞ ŞİRKETLER
- HAMMADDE VE YARDIMCI MALZEME KONUSUNDA YETERLİ ALT YAPI

BAŞLICA AR-GE ALANLARI:

- MEKATRONİK
- YENİ MALZEMELER
- ARAÇ GÖVDESİ TASARIM YETENEĞİ
- İÇ FİNİSYONA YÖNELİK PLASTİK PARÇANIN
- DÜŞÜK YAKIT TÜKETİMİ VE C02 EMİSYONU
- ELEKTRONİK KONTROL
- YENİ YAKITLAR VE TAHRİK SİSTEMLERİ
- ARAÇ DİNAMİĞİ
- SÜRÜCÜ KONFORU
- ORTAK ARGE PLATFORMU GELİŞTİRME
- MÜŞTERİ ODAKLI TASARIM VE ARGE
- ARAÇ AĞIRLIĞINI AZALTMA

Projenin Amacı:

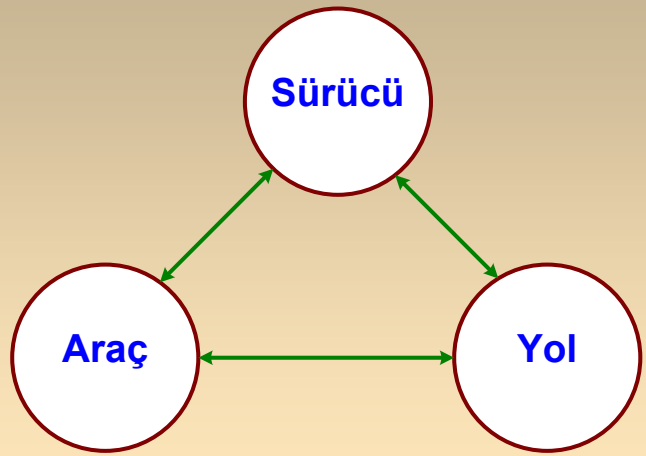
Sinyal işleme ve iletişim teknolojileri kullanılarak sürüş ve sürücü güvenliğinin arttırılması ve kazaların azaltılması hedeflenmiştir.

Kamera, mikrofon, ve diğer sensör teknolojileri gibi donanımlarla zenginleştirilecek “bilisim ve iletişim teknolojisi yüklü” araçlarda

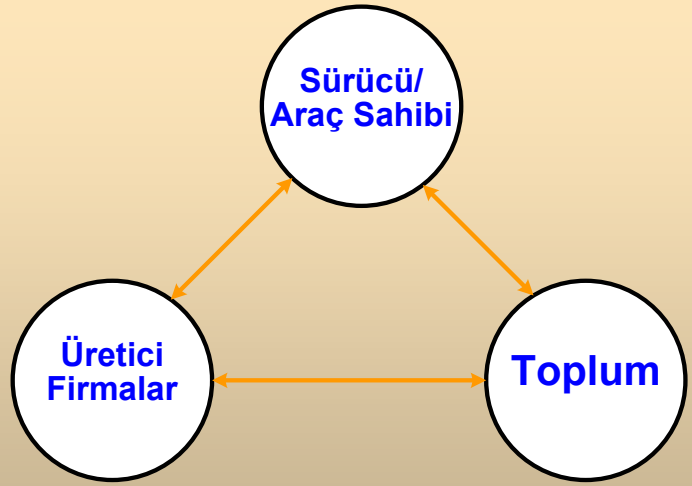
- sürücünün davranışı,
- sürücü-araç etkileşimi,
- aracın durumu ve
- ulaşım koşulları

konusunda veri toplanması; bu verilerin incelenerek sürüş ve sürücü kosullarının otomatik olarak tespit edilmesi ve gerektiğinde uyarı veya yardımda bulunulması önerilmektedir.

- **Yararlar (Fiziksel Katman):**



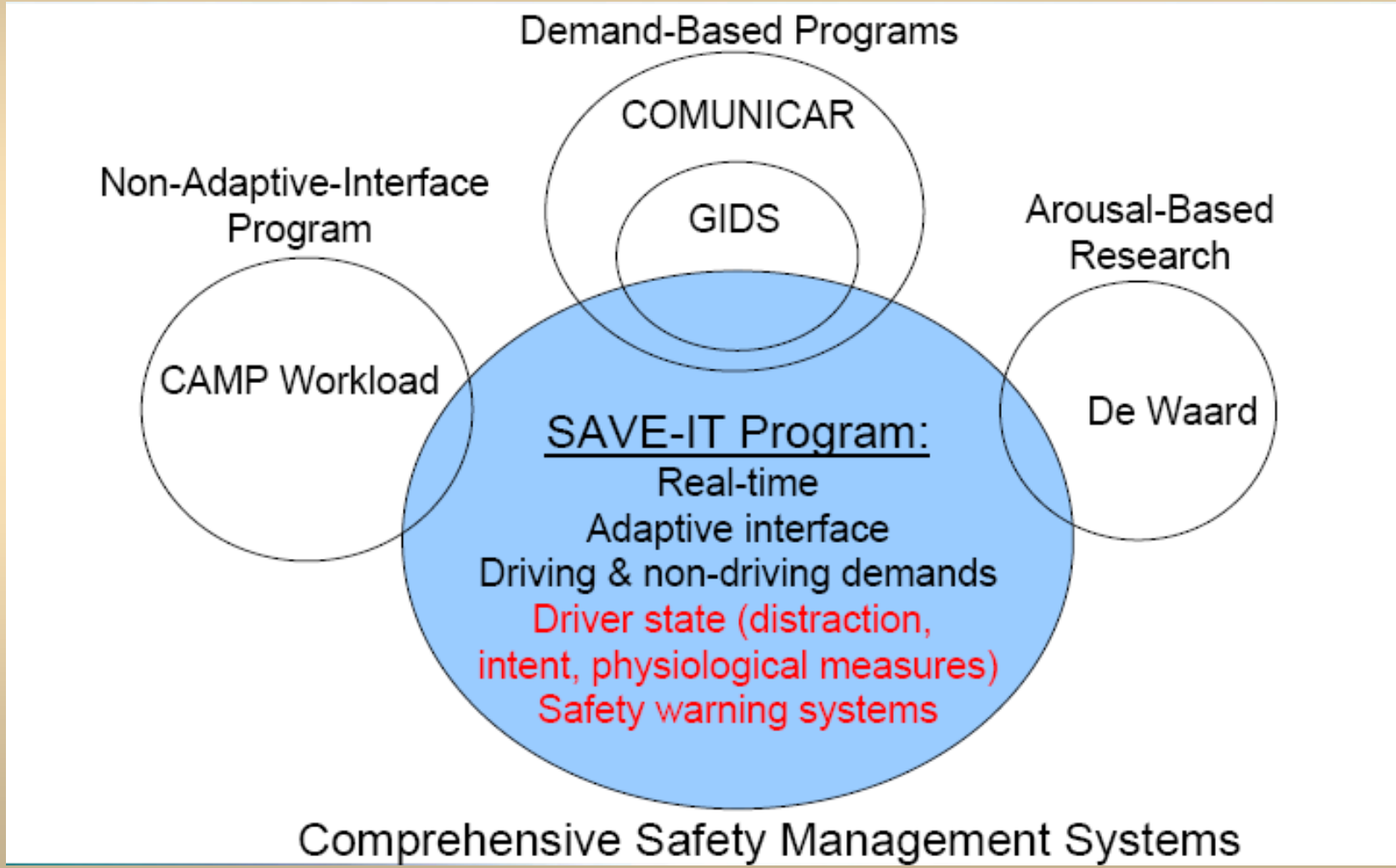
- **Yararlar (Ekonomik Katman)**



Similar Study Model

Safety Vehicle using adaptive Interface Technology (SAVE-IT)

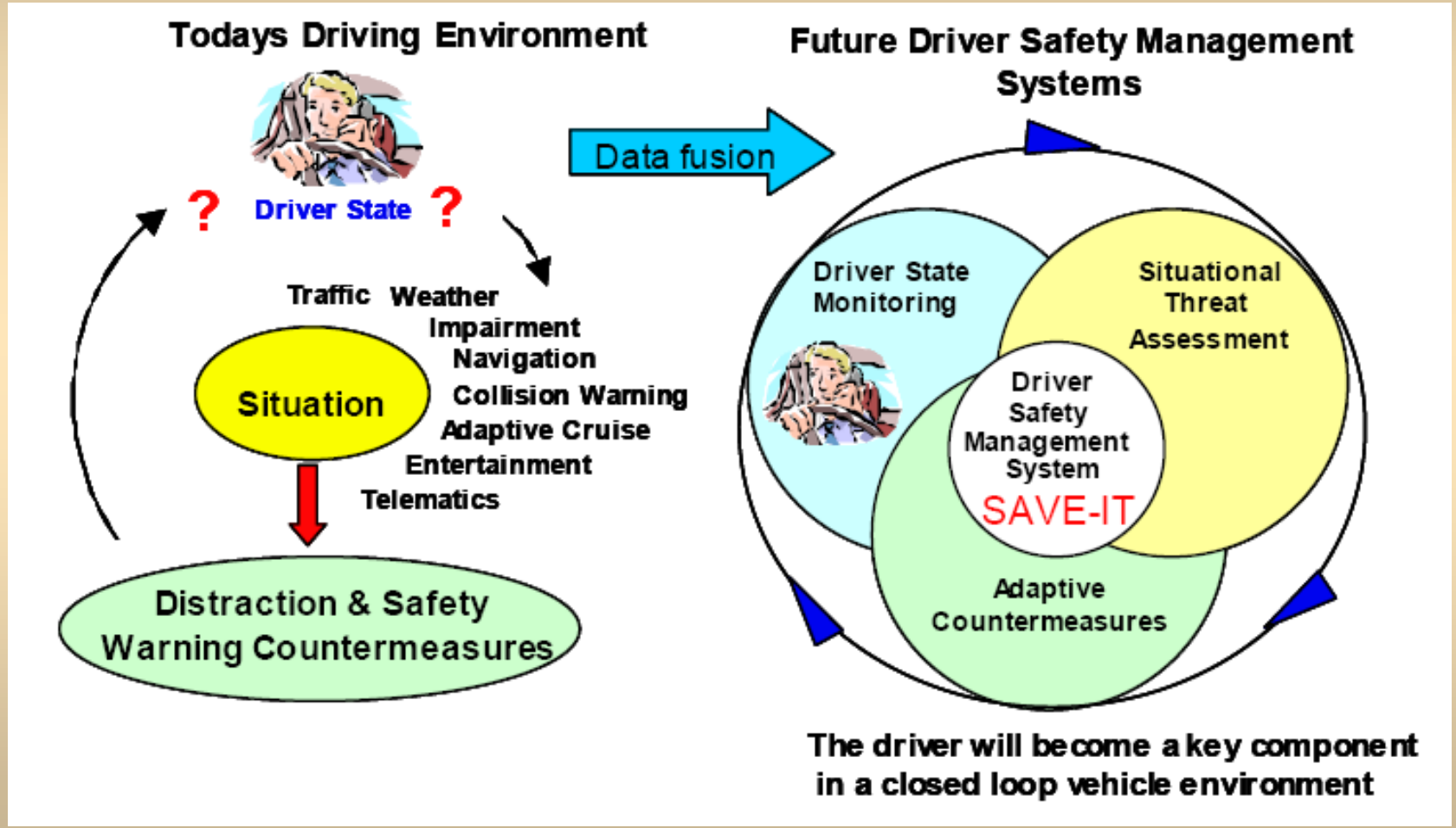
from Gerald J. Witt, Harry Zhang, Matthew R. H. Smith (Delphi)



Similar Study Model

Safety Vehicle using adaptive Interface Technology (SAVE-IT)

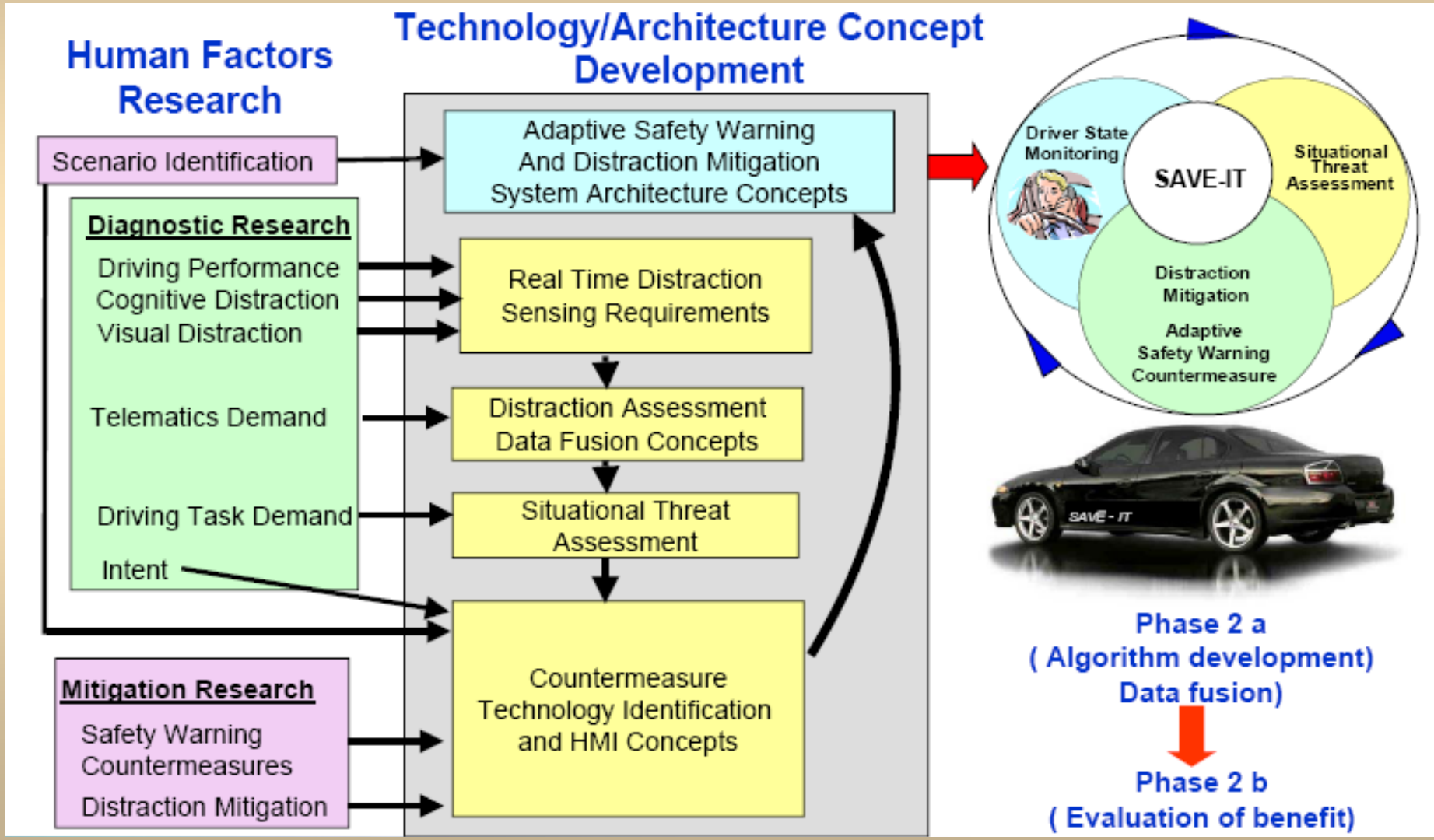
Gerald J. Witt, Harry Zhang, Matthew R. H. Smith (Delphi)



Similar Study Model

Safety Vehicle using adaptive Interface Technology (SAVE-IT)

Gerald J. Witt, Harry Zhang, Matthew R. H. Smith (Delphi)



CIAIR

Center for Integrated Acoustic Information Research
(20th Century Center of Excellence at Nagoya University)

IME

Integrated Media Engineering
(21st Century Center of Excellence at Nagoya University)

An International Alliance for Advanced Studies on In-Car Human Behavioral Signals

Research Challenge: How we can built signal processing technologies for various human behavioral signals that are implicitly associated with human intention and the environmental realities?

Sample Applications:

- Predictive driving interface
- Personalization of car
- Robust Communication for Safer Driving

Fundamental Technologies

- Speech/image signal processing
- Interactive systems
- Improved navigational systems

An International Alliance for Advanced Studies on In-Car Human Behavioral Signals: Current Alliance Membership

- **Coordinator:** (H. Abut, San Diego State/Nagoya/Sabanci Universities)

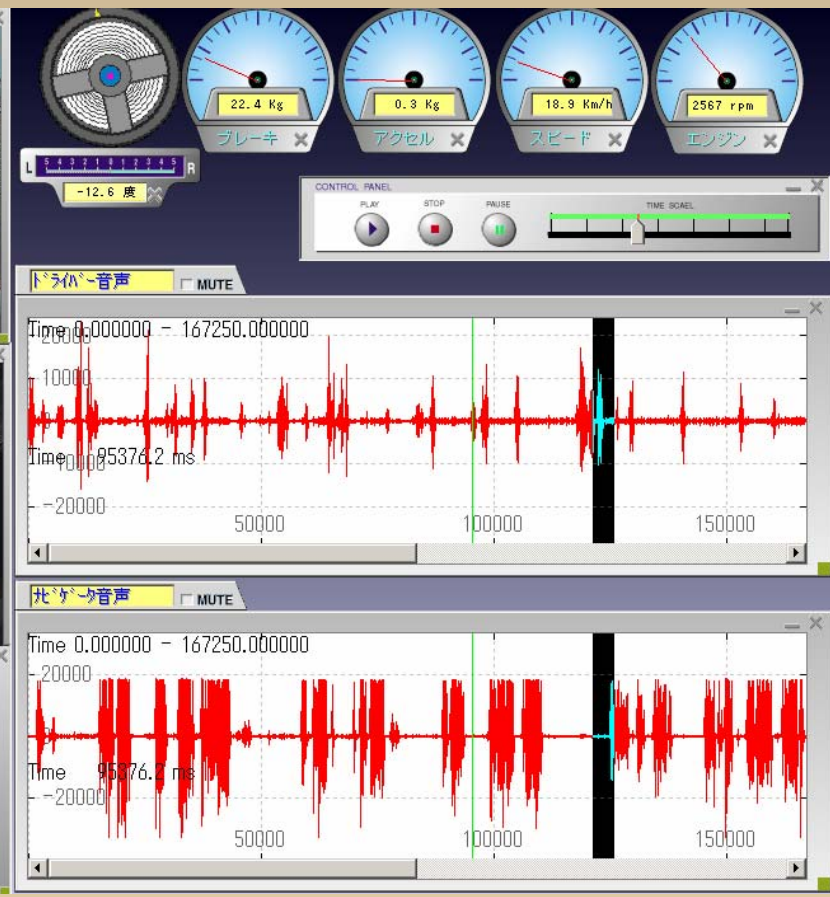
- Center for Integrated Acoustic Information Research (CIAIR) Nagoya University (Itakura, Takeda)
- Furui Laboratory, Tokyo Institute of Technology, Japan (Furui)
- Center for Speech and Hearing Science, University of Colorado (Hansen)
- HR Laboratories, Malibu, CA (Kadambe)
- ITC-IRST, Trento, Italy (Omologo)
- IEIIT-CNR, Torino, Italy (De Martin)
- Koc University, Turkey (Erzin, Tekalp, Yemez)
- Sabanci University, Turkey (Ercil, Erdogan, Keskinoz, Abut)
- Nanyang Technological University, Singapore (Abdul Wahab, Ser Wee, Tan)
- INESC ID, Lisboa, Portugal (Trancoso)
- National Taipei University of Technology, Taiwan (Hwang, S.-H.)
- National Cheng Kung University, Taiwan (Wang, J.-F.)
- Institute for Human-Machine Communication,Munich, Germany (Rigoll)
- Laboratoire Informatique d'Avignon, France (Nocera, associate member)
- MIT Lincoln Laboratory (Campbell, Reynolds, associate members)

Research Roadmap

- Establishment of an Alliance for on In-Car Multi-Language Multi-Sensor Corpus Formation Effort
  - Studies on In-car Robust ASR
  - **Studies on In-Car Enhanced Speech Systems**
  - **Studies on In-Car Enhanced Video Systems**
    - Integration & Interfacing with Navigational Systems for Improved Safety
      - Driver Identification & Applications
        - Modeling Driving Behavior
          - Field Tests of Emerging Systems
            - Recommendations to Scientific Community and Industry

An International Alliance for Advanced Studies on In-Car Human Behavioral Signals: DATABASE

- **Large Database with more than 800 drivers**
  - **(99% Japanese Speakers)**
- **Real driving conditions with Subjects driving on public streets while holding dialogues.**
- **Multi-mode dialogues**
  - **Dialogues are recorded using:**
    - **Human navigator,**
    - **WOZ system, and**
    - **ASR system.**
- **Multi-media recordings**
  - **Recorded data include multi-channel audio, multi-channel video, vehicle-related information (speed, pedals, steering handle etc..), location.**



Alliance Projects: 1 CIAIR-IME Speech Effort

COMPARED STUDIES ON SINGLE-CHANEL DENOISING SCHEMES FOR IN-CAR SPEECH ENHANCEMENT

Weifeng Li[♯], *Katunobu Itou*[†], *Kazuya Takeda*[†] *and Fumitada Itakura*[‡]

Graduate School of Engineering[♯], Graduate School of Information Science[†], Nagoya University
Faculty of Science and Technology[‡], Meijo University
Nagoya, 464–8603 Japan

ABSTRACT

This paper describes a new single-channel in-car speech enhancement method that estimates the log spectra of speech at a close-talking microphone based on the nonlinear regression of the log spectra of noisy signal captured by a distant microphone and the estimated noise. We compare the speech enhancement performance of proposed method to those of *spectral subtraction* (SS) and *short-time spectral attenuation* (STSA) based methods. The proposed method provides significant overall quality improvements in our subjective evaluation on the regression-enhanced speech. Based on our isolated word recognition experiments conducted under 15 real car environments, the proposed adaptive nonlinear regression approach shows an advantage in average relative word error rate (WER) reductions of 54.2% and 16.5%, respectively, compared to original noisy speech and ETSI advanced front-end.

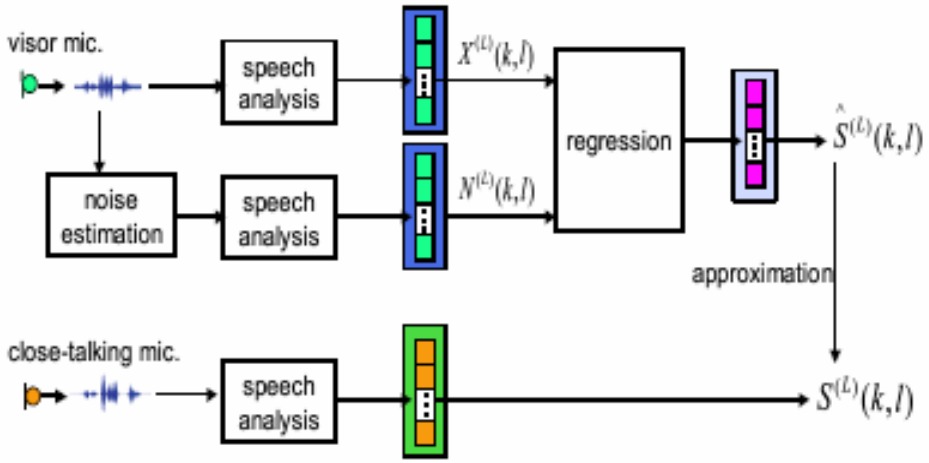


Fig. 1. Concept of regression-based speech enhancement.

Alliance Projects: 2 CIAIR-IME Speech Effort

TOWARDS ROBUST SPOKEN DIALOGUE SYSTEM USING LARGE-SCALE IN-CAR SPEECH CORPUS

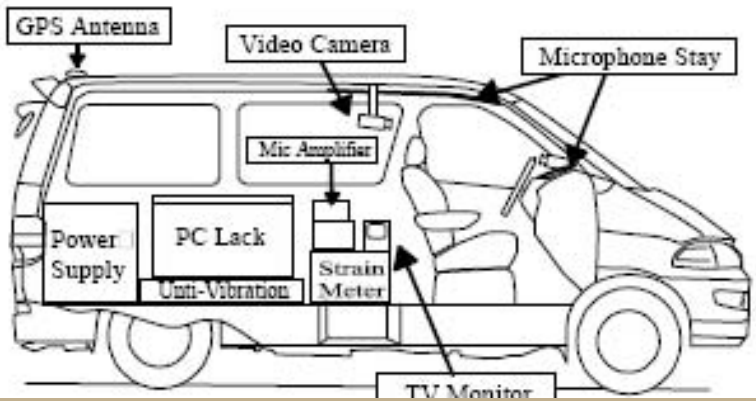
Yukiko YAMAGUCHI†, *Keita HAYASHI*†, *Takahiro ONO*†, *Shingo KATO*†, *Yuki IRIE*†,
Tomohiro OHNO†, *Hiroya MURAO*‡, *Shigeki MATSUBARA*†,
Nobuo KAWAGUCHI†, *Kazuya TAKEDA*†,

†Nagoya University Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan
‡SANYO Electric Co., Ltd. 1-18-13 Hashiridani, Hirakata-shi, Osaka, 573-8534, Japan
yamaguchi@itc.nagoya-u.ac.jp

ABSTRACT

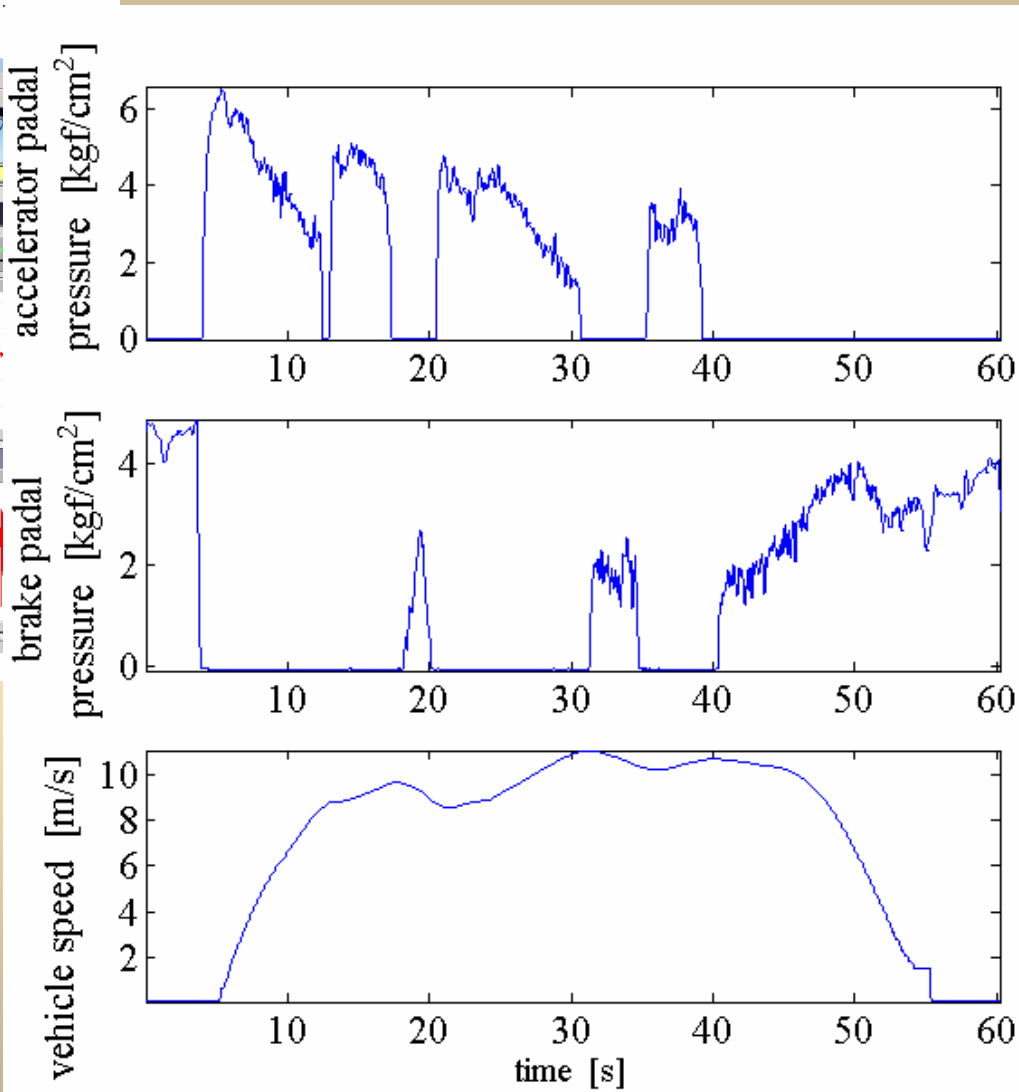
We have been studying various topics by using a large-scale corpus, which was built at CIAIR, to construct a robust and practical spoken dialogue system. The CIAIR project has developed a data collection vehicle and collected about 179 hours of multi-modal data in total.

We have transcribed the speech data by about 800 subjects, and annotated speech intentions, dependency structures, dialogue structures to the text data. We are continuing various research using the annotated data, such as speech intention understanding and speaker's knowledge acquisition. In this paper, we introduce our research activities, and present the various fruits of the in-car speech corpus.



Alliance Projects: 3 Driving Behavior

CIAIR in-car DB Browser

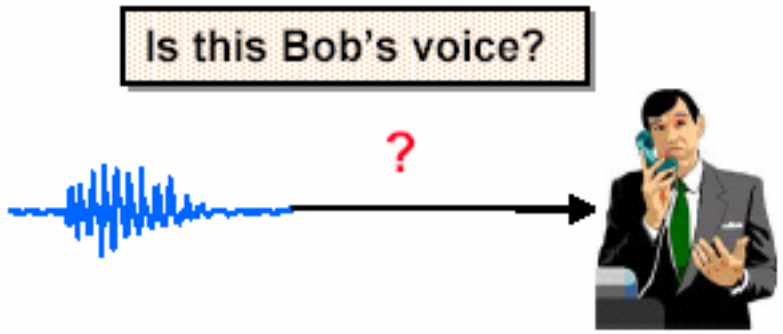


Alliance Projects: 4 Driving Behavior

Driver Verification (Sabanci, Koc, Nagoya, SDSU)

Verification: Determine if the person is who they claim to be (Authentication)

- One-to-one mapping
- Unknown data can come from any impostor; open-set
- Adding "none of the above" makes it closed-set



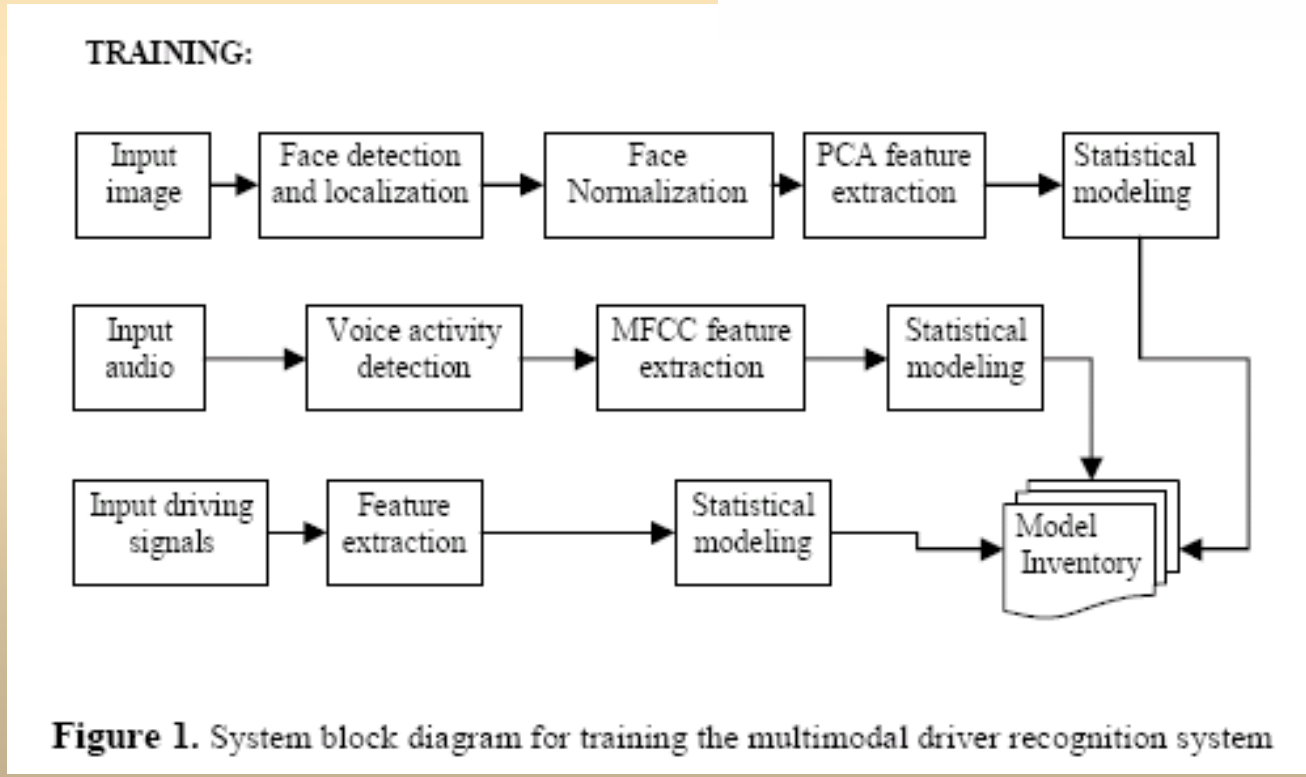


Figure 1. System block diagram for training the multimodal driver recognition system

Alliance Projects: 4 Driving Behavior

Driver Verification (Cont.) (Sabanci, Koc, Nagoya, SDSU)

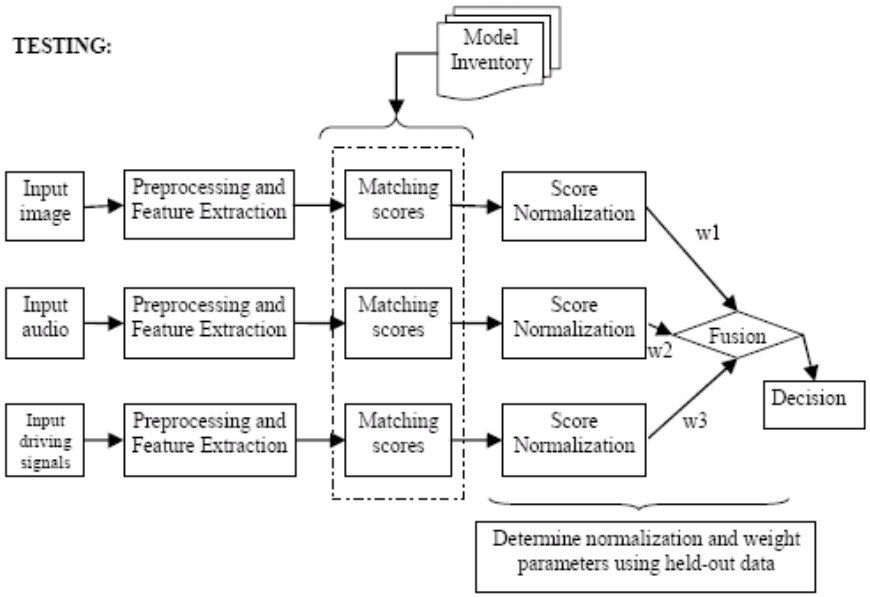


Figure 2. System block diagram for testing the multimodal person recognition system

Table 1. Closed-set speaker identification, speaker verification and open-set speaker identification results are shown. Tests are performed on a 20 speaker subset of the CIAIR database using various modalities and their combinations. Multimodal decision fusion was performed using weighted summation of normalized modality scores.

Modality	Weights	Closed-set ID (Accuracy %)	Verification (EER %)	Open-set ID (EER %)
A	Audio only	98.00	2.15	2.37
F	Face only	89.00	6.08	11.00
D	Driving only	88.25	4.00	12.00
A+D	(.62,.38)	99.25	0.83	1.10
F+D	(.43,.57)	98.00	1.62	2.25
A+F	(.63,.37)	99.75	0.50	0.50
A+F+D	(.47,.33,.20)	100.0	0.25	0.25

Alliance Projects: 5 In-Car Interaction System

ITC irst

Centro per la Ricerca Scientifica e Tecnologica

On the development of an in-car speech interaction system at Irst

omologo@itc.it

Maurizio Omologo
ITC-irst, Trento, Italy

http:///shine.itc.it

Alliance Projects: 5 In-Car Interaction System (Cont.)

ITC irst

In-car speech interaction

VICO

DIPLODOC

General framework of VICO (Virtual Intelligent COdriver) and DIPLODOC projects:

Natural and spontaneous hands-free speech interaction (with synthetic prompts) to provide tourist information, hotel and restaurant reservation, navigation assistance.

Scenario:

A tourist who is driving in Trentino

Three languages:

Italian, English, German

Functionalities:

- Hotel Information Retrieval (address, prices, services, phone number, etc)
- Hotel and Restaurant Reservation
- Points Of Interest (POI) Retrieval (churches, petrol stations, etc.)
- Simple Route Query
- Traffic information access

Main issues:

- Robustness at different levels
- Spontaneous speech
- Barge-in
- Medium-large vocabularies
- Parallel recognition units
- Dialogue strategies
- Retrieval from real databases
- *WOZ experiments*
- *Real prototype experiments*

Alliance Projects: 6 University of Colorado

CU-MOVE: IN-VEHICLE SPEECH SYSTEMS FOR NAVIGATION

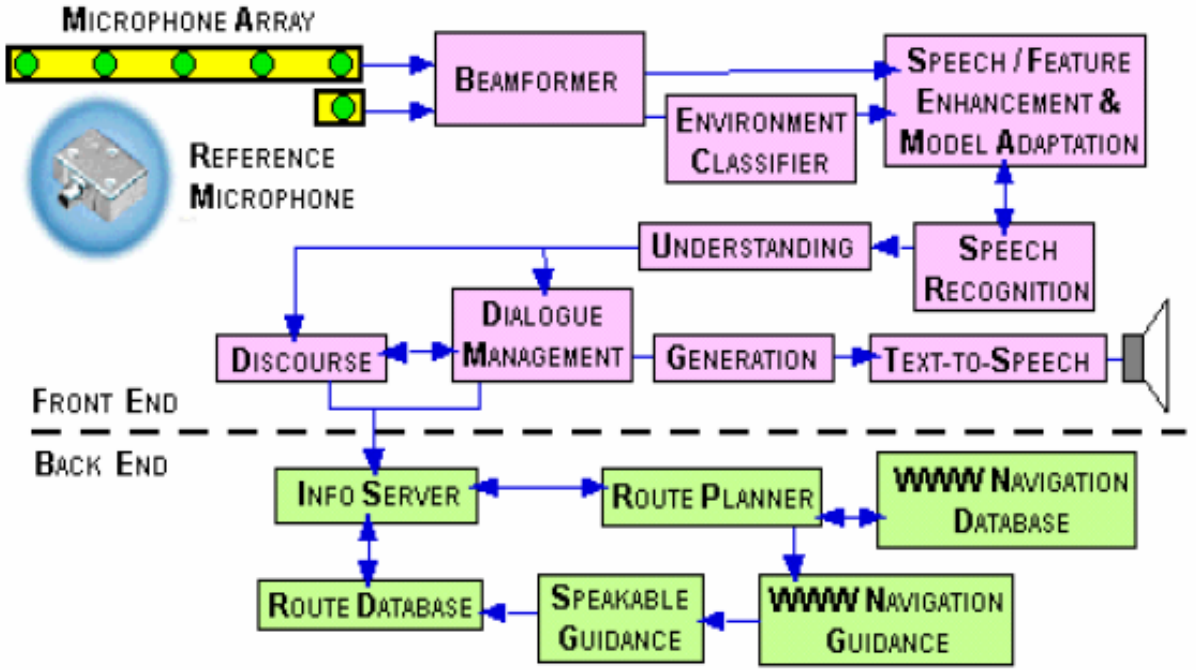


Figure 2-1. Flow Diagram of CU-Move Interactive Dialogue System for In-Vehicle Route Navigation

Alliance Projects: 7 Politecnico di Torino

ADAPTIVE H.264 VIDEO TRANSMISSION OVER 802.11 INTER-VEHICULAR AD HOC NETWORKS

*Paolo Bucciol, Enrico Masala, Juan Carlos De Martin**

Dipartimento di Automatica e Informatica / *IEIIT–CNR
Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
Phone: +39-011-564-7036 / -5421, fax: +39-011-564-7099
E-mail: [paolo.bucciol | masala | demartin]@polito.it

ABSTRACT

This paper focuses on video communications in inter-vehicular environments using the 802.11 ad hoc network protocol. In the first part of the work we present the results of transmission experiments between two cars equipped with 802.11 devices in two typical driving scenarios, urban and highway. Various video bitrates and packetization policies have been tested. The results show that the two scenarios differ in terms of link availability and SNR. Moreover, the video quality measured at the receiver by means of the PSNR value shows that the best packetization policy depends on the scenario. Building on these results, we design an algorithm which adapts the video packet size to the current driving conditions to improve the efficiency of the video transmission. Consistent perceptual quality gains in terms of PSNR value (up to about 3 dB) are achieved with respect to a fixed-policy transmission technique.

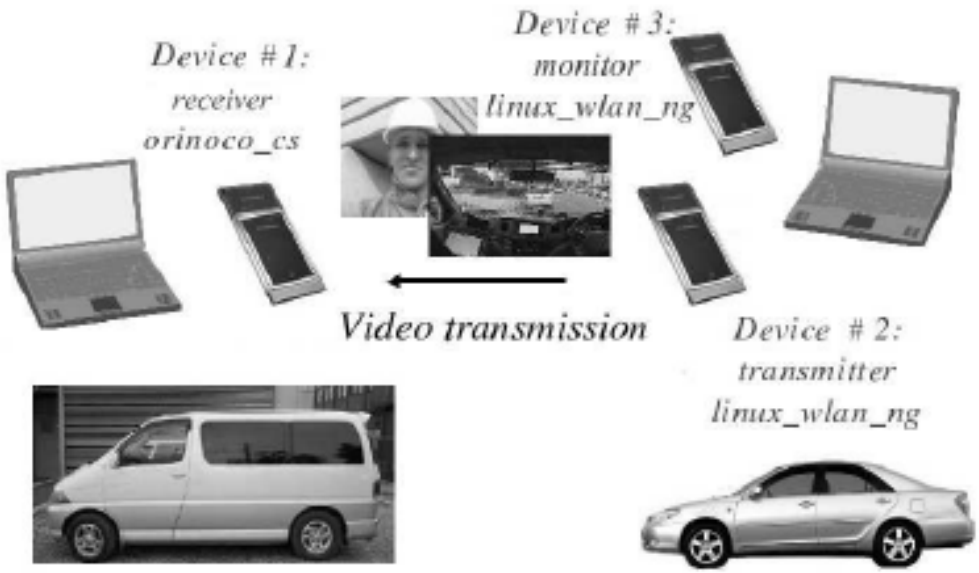


Fig. 2. The experimental testbed. The video flow is transmitted from the car to the van.

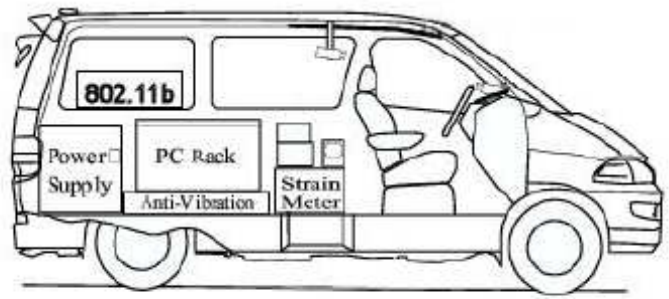


Fig. 1. Data collection vehicle used during the experiment.

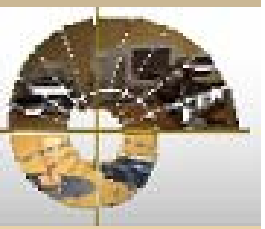
AllianceProjects: 8

Multimodal Audio-Visual Speaker Identification

Engin Erzin, Yücel Yemez, A. Murat Tekalp
Multimedia, Vision and Graphics Laboratory
College of Engineering, Koç University, Istanbul
http://mvgl.ku.edu.tr

Summary

- Each modality or group of correlated modalities are associated with individual classifiers
  - Audio and Audio-lip (HMM based), Face Image (Eigenface based)
- Multimodal identification system requires combining classifiers
  - How to combine the individual classifiers?
    - Sum weighted log-likelihoods over modalities.
    - If reliability of each modality is same, then uniform weighting is optimal.
    - However under varying acoustic noise, lighting, face-image backgrounds, reliabilities do also vary.
    - A novel adaptive weighting strategy is proposed for the sum rule and it's observed to be robust under all environmental conditions.



THE USE OF LIP INFORMATION FOR ROBUST SPEAKER IDENTIFICATION AND SPEECH RECOGNITION

H. E. Çetingül, E. Erzin, Y. Yemez and A. M. Tekalp

Multimedia, Vision and Graphics Laboratory
College of Engineering, Koç University, Sarıyer, Istanbul, 34450, Turkey
ecetingul,eerzin,yyemez,mtekalp@ku.edu.tr

ABSTRACT

This study investigates the benefits of multimodal fusion of audio, lip motion and lip texture modalities for speaker and speech recognition. The audio modality is represented by the well-known mel-frequency cepstral coefficients (MFCC) along with the first and second derivatives, whereas lip texture modality is represented by the 2D-DCT coefficients of the luminance component within a bounding box about the lip region. A new lip motion modality representation based on *discriminative analysis* of the dense motion vectors within the same bounding box is employed for speaker/speech recognition. The fusion of audio, lip texture and lip motion modalities is performed by the so-called *Reliability Weighted Summation* (RWS) decision rule. Experimental results show that inclusion of lip motion and lip texture modalities provides further performance gains in both speaker identification and speech recognition scenarios.

İlgilerinize Teşekkürler

4 Temmuz 2005 – İzmir Yüksek Teknoloji Enstitüsü

H. Abut
abut@anadolu.sdsu.edu
http://anadolu.sdsu.edu/abut/DL/Abut_August_2005.pdf

OTAM – Otomotiv Teknoloji ARGE Merkezi