IEEE

IEEE Signal Processing Society

2025 ICASSP HYDERABAD, INDIA

Celebrating Signal Processing

GOLDEN JUBILEE 50

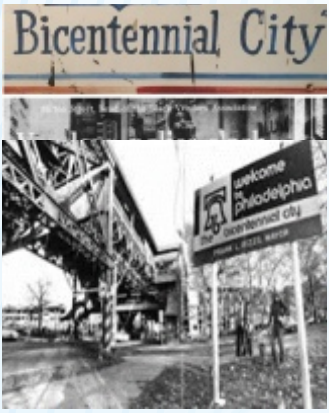
ICASSP@50

Welcome to the ICASSP@50 Pavilion. Here, we have collected a little bit of history, a few stories, and an attempt through some statistics and graphics to give you a sense of how ICASSP has evolved from its early days to the modern-day extravaganzas.

Take a look at the word clouds and the most cited/most popular/award-winning papers to see how our field has evolved in the last 50 years. Look at charts depicting the number of papers and attendees to see how ICASSPs have grown. Yet, when we look at the anecdotes and tidbits of information we have about each event, it looks like very little has changed!

This is not meant to be a comprehensive history of ICASSPs. We have just begun. If you have more information that you would like to share, please drop us a note. In the meantime, enjoy the panels that are waiting for you.

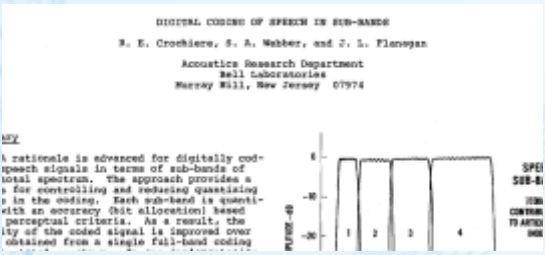
ICASSP 1976 - PHILADELPHIA, USA
APRIL 12 - 14



"In the future, the ASSP Society plans to hold similar conferences on an annual basis. It is hoped that this conference will serve as a platform available to engineers and scientists for the dissemination of the latest developments in this rapidly changing technology."

Charles Teacher, General Chair, ICASSP 1976

Most Cited Paper



DIGITAL CODING OF SPEECH IN SUB-BANDS

R. E. Crochiere, S. A. Webber, and J. L. Flanagan

Acoustics Research Department
Bell Laboratories
Murray Hill, New Jersey 07974

Most Downloaded Paper

ORTHOGONAL TRANSFORMS FOR DIGITAL SIGNAL PROCESSING

K. R. Rao
Dept. of Electrical Engineering
University of Texas at Arlington
Arlington, Texas

N. Ahmed
Dept. of Electrical Engineering
Kansas State University
Manhattan, Kansas

Abstract

600 attendees from 16 countries

In comparison, the bicentennial concert by Peter Frampton and Yes attracted 130000 people to the JFK Stadium

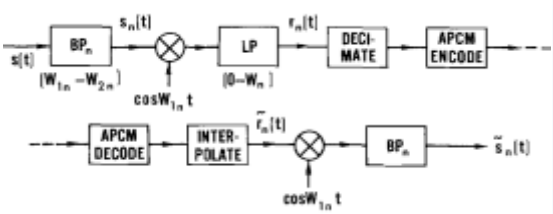


Figure from Crochiere, Webber, Flanagan

Also in 1976:
On July 20th, NASA's Viking 1 became the first spacecraft to land on Mars and send back pictures.
Apple Inc. was founded on April 1st.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

IEEE Signal Processing Society

ICASSP 1977 - HARTFORD, USA
MAY 9 - 11

Most Cited Paper

APPLICATION OF QUADRATURE MIRROR FILTERS TO SPLIT BAND VOICE CODING SCHEMES

D. Esteban and C. Galand

IBM Laboratory
06610 La Gaude, France

2) QMF band splitting

Principle

Let us consider for explanation in which we describe the decomposed signal in two contiguous ...

. H_0 is a sampled half band low ... impulse response $h_0(n)$.

Most Downloaded Paper

FREQUENCY DOMAIN BEAMFORMING TECHNIQUES

Gordon L. DeMuth

IBM Federal Systems Division
9500 Godwin Drive, Manassas, VA 22110

ACT

well known that sensor signals must be sampled at a rate much higher than the Nyquist rate to minimize degradations of side lobe levels and ... responses and gain in sampled time domain ... Various frequency domain techniques ... employed to form beams without requiring excessively high data rate that can cause ... management problems in programmable signal ... Techniques applicable to line, ring, ... and spherical arrays are developed.

$$B_0(k) = \sum_{p=0}^{N-1} W_p X_p (k - S_{m,p})$$

where W_p is the aperture shading coefficient

X_p is the sensor signal.

$S_{m,p}$ represents the beamforming delay for the mth beam and the pth sensor expressed as the nearest integer number of sample intervals.

the conference account; the monies collected for over-length page charges, 2.9K, just about paid for the coffee, danish, and cookies served during the session breaks.

Source: ASSP AdCom Minutes; 1977

An excellent idea for raising coffee funds!

"In 1977, I served as the general chairman of the first – some will call it the second – ICASSP in Hartford, CT."
Harvey Silverman, General Chair, ICASSP 1977

Also in 1977:
Rings of Uranus were discovered on March 10th.
The first MRI body scan performed on a human occurred on July 3rd.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1978 - TULSA, USA
APRIL 10 - 12

IEEE 1978 ICASSP

Most Cited and Downloaded Paper

AN ADAPTIVE LATTICE STRUCTURE FOR NOISE-CANCELLING APPLICATIONS

Lloyd J. Griffiths

Electrical Engineering Dept.
University of Colorado
Boulder, Colorado 80309

This paper describes an adaptive filter structure which may be used in multi-channel noise-cancelling applications. The proposed structure corrupted by the presence of muscle noise and 60 Hz power line pickup. In the majority of cases, the weak, desired fetal signal is not evident upon vis-

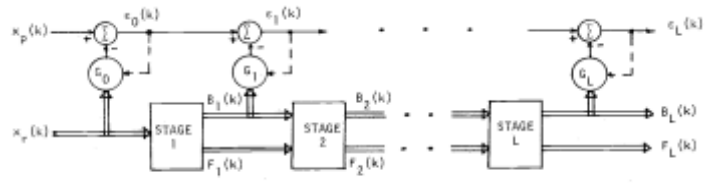


Fig. 3. Lattice form implementation of noise-cancelling filter.

Figure from Griffiths, ICASSP 1978.

Rao Yarlagadda, General Chair, ICASSP 1978

The choice of Tulsa as a venue for ICASSP raised some discussions.

Headquarters. The discussion centered on the advisability of having the meeting in the mid-south. H. Helms reported that the Arden House survey on meeting site preference gave little guidance. H. Silverman commented that 65 percent of the attendees are speech people and 60 percent of these reside on the east coast. Availability of travel funds was questioned, but it was felt that travel to Oklahoma would be easier to obtain than to the West Coast. R. Ashley mentioned the increasing pressure from exhibitors for opportunities in Middle America, such as Chicago or Nashville.

Source: ASSP AdCom Minutes; 1977

Also in 1978:
The first Global Positioning System (GPS) satellite Navstar 1 was launched on February 22nd.
Charon, the first moon of Pluto to be identified, was discovered on June 22nd.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1979 - WASHINGTON D.C., USA
APRIL 2 - 4

ICASSP 79

Most Cited and Downloaded Paper

ENHANCEMENT OF SPEECH CORRUPTED BY ACOUSTIC NOISE*

M. Berouti, R. Schwartz, and J. Makhoul

Bolt Beranek and Newman Inc.
Cambridge, Mass.

ABSTRACT

This paper describes a method for enhancing speech corrupted by broadband noise. The method is based on the spectral noise subtraction method. The original method entails subtracting an estimate of the noise power spectrum from the speech power spectrum, setting negative differences to zero, recombining the new power spectrum with the original phase, and then reconstructing the time

from that reported by others is two major ways: first, we subtract a factor (α) times the noise spectrum, where α is a number greater than unity and varies from frame to frame. Second, we prevent the spectral components of the processed signal from going below a certain lower bound which we call the spectral floor. We express the spectral floor as a fraction β, of the original noise power spectrum $P_d(w)$.

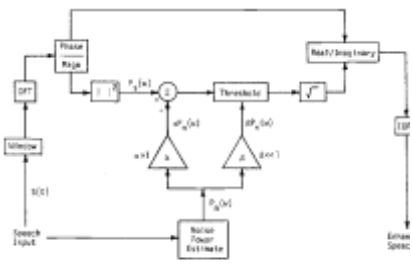


Fig. 1 Modified spectral noise subtraction method with spectral floor.

The Conference will be devoted to the areas of:

- *General Signal Processing* — (Spectral analysis, filter design, software and systems, hardware and architecture, structures and quantization, multidimensional processing, high speed algorithms, etc.).
- *Speech Processing* — (Analysis, wideband and narrowband speech communication and evaluation, aids for the handicapped, automatic recognition/understanding of continuous speech, automatic speech segmentation and phoneme recognition, speech production and synthesis, etc.).
- *Underwater Acoustics, Seismic Signals and Radar* — (Hardware, theory and experimentation, detection and localization, adaptive filtering and beamforming, simulation, etc.).
- *Psychoacoustics, Electroacoustics* — (Loud speakers, environmental noise, physiological noise, recording and reproduction.)

Source: ICASSP 1979 Call for Papers

Also in 1979:
Sony began selling the first model of the Walkman, TPS-L2 on July 1st.
VisiCalc, the first spreadsheet program for personal computers was released on October 17th.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1980 - DENVER, USA
APRIL 9 - 11

ICASSP 80

Most Cited Paper

A Filter Family Designed for Use in Quadrature Mirror Filter Banks

J. D. Johnston

Bell Laboratories
Murray Hill, New Jersey 07974

STRACT

family of filters that have been Mirror Filter[1] (QMF) Banks. significant improvement over uiripple and window designs

QMF Definition

Figure 1 shows the QMI Given $h_1(n)$, a linear phase number of taps and a 3dB poin

Most Downloaded Paper

Frequency Domain Data Transmission using Reduced Computational Complexity algorithms

Abraham Peled and Antonio Ruiz

Computer Sciences Department, IBM T.J. Watson Research Center, P.O. Box 218, Yorktown Heights, N.Y. 10598

Abstract

we describe a frequency domain data transmission used for digital data transmission over analog tele-

Let us also define the complex sequence

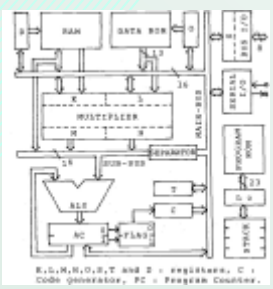
$$\begin{cases} 0 & n = \\ A_n + jB_n & 0 < \end{cases}$$

Even in a span of just 5 years, we can see some changes in the word cloud. The word "adaptive" did not appear in the 1976 word cloud.

Compare the list of exhibitors (source: Proceedings of ICASSP 1980) at ICASSP 1980 to that in 2025. This is yet another indication of how our field has evolved over the years.

Analogic Corporation
CONTAL
CSPI
EPC Laboratories
Floating Point Systems
INTEL
Kay Elemetrics
Martin Marietta Aerospace
Nicolet Scientific
Precision Filters, Inc.
Signal Technology, Inc.
Spectral Dynamics Corporation
Springer-Verlag N.Y. Inc.
TRW LSI Products

NEC (formerly Nippon Electric Company) developed the first commercially available, programmable digital signal processor chip (μPD7720). NEC presented their work at ICASSP 1980.



Block diagram from T. Nishitani, et al., "LSI Signal Processor Development for Communication Equipment," ICASSP 1980.

Also in 1980:
K. Fukushima published "Neocognitron: A Self-organizing Neural Network Model ...", generally considered to be the first paper to discuss convolutional neural networks.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1981 - ATLANTA, USA
MARCH 30 - APRIL 1

Most Cited Paper

Array Design for MEM and MLM Array Processing*

R. W. Lang, G. L. Duckworth, and J. H. McClellan

Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, Ma. 02139

ABSTRACT

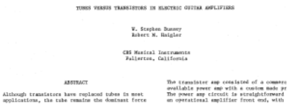
Much work has been done on designing arrays for use with conventional beamforming techniques. However, high resolution array processing techniques such as the maximum entropy (MEM) and maximum likelihood (MLM) methods impose different constraints on the array. MEM and MLM operate on measured samples of the correlation function and so are expected to perform better when the correlation function is measured at more and more evenly distributed lags. A uni-

directions. The sensor outputs are processed, typically by linear filtering and summing, to extract the desired signal from the noise. Some conventional beamforming techniques as well as Capon's[1] maximum likelihood method (MLM) are formulated in this manner.

II. History and Applications

Arsac[2] proposed the use of thinned or nonredundant arrays in a radio-astronomical application. Barber[3], Haubrich[4], and Moffet[5] elaborated on Arsac's work, incorporating certain results in the theory of numbers[6] and proposing various design techniques in

Most Downloaded Paper



TUBES VERSUS TRANSISTORS IN ELECTRIC GUITAR AMPLIFIERS

W. Stephen Bussey
Robert M. Haigler

CBS Musical Instruments
Fullerton, California

ABSTRACT

Although transistors have replaced tubes in most applications, the tube remains the dominant force

The transistor amp consisted of a commerc available power amp with a custom made pr The power amp circuit is straightforward on operational amplifier front end, with

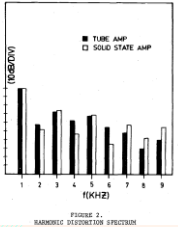


Figure from Bussey and Hagler, ICASSP 1981.

General Chair: Ron Schafer

Technical Program Chair: Russell Mersereau

Number of attendees : 950
Papers accepted : 295
Conference surplus : $17,853.93

Also in 1981:
Solar challenger, an electric aircraft made a 162-mile flight on June 7th using only solar power from wing-mounted photovoltaic cells.
The CNTR+ALT+DEL command was invented by David Bradley.

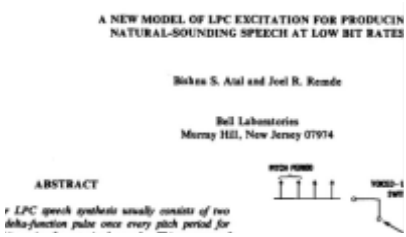
2025 ICASSP HYDERABAD, INDIA
Celebrating Signal Processing

IEEE Signal Processing Society

ICASSP 1982 - PARIS, FRANCE
MAY 3 - 5

ICASSP 82

Most Cited Paper



A NEW MODEL OF LPC EXCITATION FOR PRODUCIN
NATURAL-SOUNDING SPEECH AT LOW BIT RATES

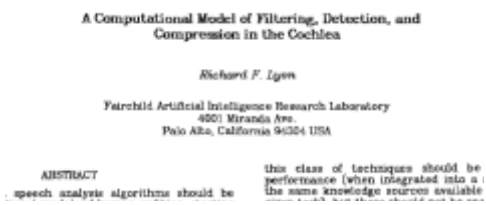
Bishnu S. Atal and Joel R. Remde

Bell Laboratories
Murray Hill, New Jersey 07974

ABSTRACT

LPC speech synthesis usually consists of two delta-function pulse once every pitch period for

Most Downloaded Paper



A Computational Model of Filtering, Detection, and Compression in the Cochlea

Richard F. Lyon

Fairchild Artificial Intelligence Research Laboratory
4001 Miranda Ave.
Palo Alto, California 94304 USA

ABSTRACT

speech analysis algorithms should be

this class of techniques should be performance (when integrated into a the same knowledge sources available

The ASSP Society made, for the first time at an ICASSP, small travel grants totaling $12K to 100 delegates.

In 1982, ICASSP was held outside of the USA for the first time.

More than 1600 delegates attended the conference, a record at that time and 60% more than in 1981. The number of accepted papers also saw a similar increase.

Papers were presented in 6 parallel lecture sessions and 2 poster sessions. This was the first time papers were presented in poster format at an ICASSP.

Also in 1982:
Barney Clark became the first recipient of a "permanent" artificial heart on Dec. 2nd.
Sony released the World's first commercial CD player, CDP-101 on Oct. 1st.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1983 - BOSTON, USA
APRIL 14 - 16

Most Cited and Downloaded Paper

IMPROVING THE RESOLUTION PERFORMANCE OF
EIGENSTRUCTURE-BASED DIRECTION-FINDING ALGORITHMS*

Arthur J. Barabell

Dept. of Electrical Engineering and Computer Science
Massachusetts Institute of Technology, Cambridge, MA 02139
and
M.I.T. Lincoln Laboratory, Lexington, MA 02173

ABSTRACT

Recently there has been much interest in algorithm which form a direction-finding spectrum based on the eigenstructure of the sensor covari-

where "H" denotes the conjugate transpose operation, and P is the signal-in-space covariance matrix. For the remainder of this paper, we will

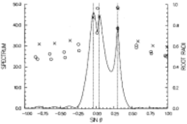


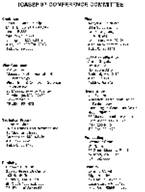
Fig. 2. Proposed method spectrum and associated poles and zeroes (NOTE: all poles are "triple" poles).

Figure from Barabell, ICASSP 1983.

Can you imagine mailing conference programs via postal service?

Non USA mailings of the advance program were late. The concensus of the ADCOM was for the ICASSP-83 organizers to offer special consideration (in terms of reduced advanced registration fees) to those who had received late programs. The printer thought they had sent them out, but they had not. Plans for

Source: ASSP AdCom Minutes; April 1983



ICASSP 83 CONFERENCE COMMITTEE

Registration Fee: USD 80.

Source: Proceedings of ICASSP 1983

Also in 1983:
Because ARPANET officially adopted TCP/IP as its standard protocol starting January 1, many consider 1/1/1983 as the birthday of the internet.

2025 ICASSP HYDERABAD, INDIA
Celebrating Signal Processing

IEEE Signal Processing Society

ICASSP 1984 - SAN DIEGO, USA
MARCH 19 - 21

ICASSP 84

Most Cited and Downloaded Paper

SCALE-SPACE FILTERING:
A New Approach To Multi-Scale Description

Andrew P. Witkin
Fairchild Laboratory for Artificial Intelligence Research

ABSTRACT

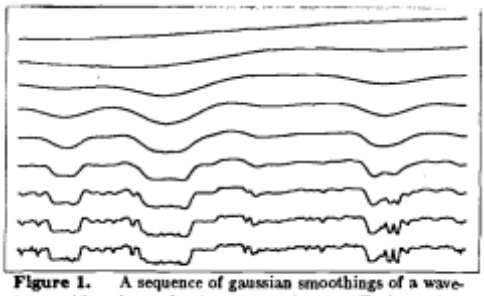


Figure 1. A sequence of gaussian smoothings of a waveform, with σ decreasing from top to bottom. Each graph is a constant-σ profile from the scale-space image.

Figure from Witkin, ICASSP 1984.

"The health, strength, and general well-being of the society and the profession are well served by this geographic broadening of the society's activities."
Stanley A. White, General Chair, ICASSP 1984

Source for both quotes: Proceedings of ICASSP 1984

There were plenary events at ICASSP 2004. According to general chair White, "Because of the vast technical breadth and depth of this conference there is no plenary session this year so that more time can be given to the technical specialty areas."

Also in 1984:
On July 25th, cosmonaut Svetlana Savitskaya became the first woman to perform a space walk.
The first ever TED Talk was given on February 23rd.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1985 - TAMPA, USA
MARCH 26 - 29

Most Cited and Downloaded Paper


CODE-EXCITED LINEAR PREDICTION (CELP): HIGH-QUALITY SPEECH AT VERY LOW BIT RATES

Manfred R. Schroeder
Bishnu S. Atal
AT&T Bell Laboratories

ABSTRACT

SPEECH SYNTHESIS MODEL

General Chair: Rex Dixon
Technical Program Chair: Vijay Jain

The duration of the main conference was extended to four days to accommodate the growth of the conference.

"In addition, session chairmen have been urged to help each presenter in the creation and use of audiovisual aids for the convenience and interests of attendees."

"We have also provided an opportunity for those attendees who want the convenience of advance registration, but who cannot pay in advance, to have the convenience at a slightly higher registration fee payable when they arrive."
- General Chair Rex Dixon

Source: Proc. ICASSP 1985

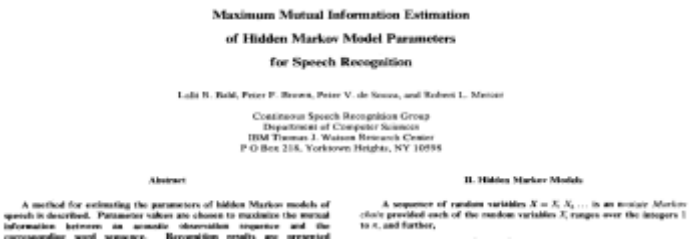
Also in 1985:
The first ever Microsoft Windows operating system was released on Nov. 20th.
The domain name symbolics.com was registered on March 15th, This was the first .com domain name.

2025 ICASSP HYDERABAD, INDIA
Celebrating Signal Processing

IEEE Signal Processing Society

ICASSP 1986 - TOKYO, JAPAN
APRIL 7 - 11

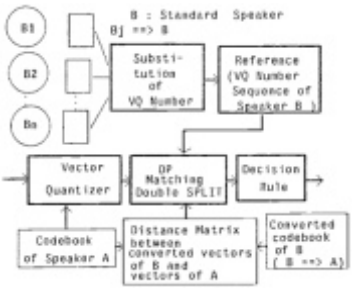
Most Cited and Downloaded Paper



Maximum Mutual Information Estimation of Hidden Markov Model Parameters for Speech Recognition

Lalit R. Bahl, Peter F. Brown, Peter V. de Souza, and Robert L. Mercer

Continuous Speech Recognition Group
Department of Computer Science
IBM Thomas J. Watson Research Center
P O Box 218, Yorktown Heights, NY 10598



Shikano, Fu, Reddy, "Word Recognition System Through Speaker Adaptation," ICASSP 1986

ICASSP was held in Asia for the first time in 1986.

"Number of abstracts submitted > 1300

Authors from more than 40 countries submitted abstracts

Number of accepted paper: 785

Number of attendees: 1350

General Chair
Hiroya Fujisaki

"I would wish that everyone of our young men could visit Japan once at least in his lifetime to see how Japanese talk and work more .... Buddha has more influence on Japan than India
- Hiroya Fujisaki

Source: Interview in Star of Mysore newspaper on April 12, 2018

Also in 1986:
On February 9th,– Halley's Comet reached its perihelion, the closest point to the Sun. This was the comet's second visit to the Solar System in the 20th century. The first was in 1910.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1987 - DALLAS, USA
APRIL 6 - 9

ICASSP 87

Most Cited and Downloaded Paper

A SPEECH ENHANCEMENT METHOD BASED ON KALMAN FILTERING

K.K. Paliwal and Anjan Basu

Computer Systems and Communication Group
Tata Institute of Fundamental Research
Homi Bhabha Road, Bombay 400005, India

ABSTRACT

In this paper, the problem of speech enhancement when only corrupted speech signal is available for processing is considered. For this, the Kalman filtering method is studied and compared with the Wiener filtering method. Its performance is found to be significantly better than the Wiener filtering method. A delayed-Kalman filtering method is also proposed which improves the speech enhancement performance of Kalman filter further.

Wiener filter is designed for each short-time speech segment (duration: 20-30 msec) using a least-squares procedure.

Though the nonstationary Wiener filter is optimum for a given segment in a least-squares-error sense, it does not exploit the knowledge about speech production process. In the present paper, we propose Kalman filtering method which allows for the nonstationarity of speech

Welcome to ICASSP 87! Texas is always thought of as "big," and Dallas is known as "Big D." So we extend a big welcome to y'all. Texas just celebrated its 150th birthday (sesquicentennial, for those who can pronounce the word). It still preserves the spirit of a frontier state, but is now a leader in such areas as energy and high technology. Dallas in particular has one of the higher concentrations of high-tech industries in the country. Companies active in telecommunications, electronics, computers, and related areas as well as the local excellent Universities make "Big D" an ideal location to hold ICASSP 87.

- Message from the General Chair Panos Papamichalis

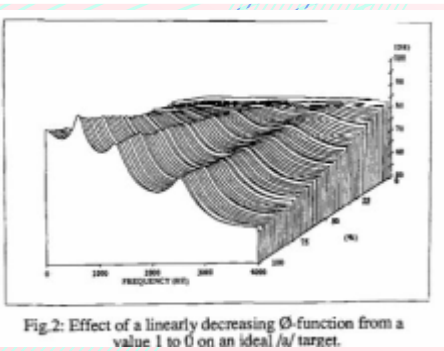


Fig.2: Effect of a linearly decreasing Ø-function from a value 1 to 0 on an ideal /a/ target.

Figure from G. Ahlbom et al., "Modeling Spectral Speech Transitions Using Temporal, Decomposition Techniques" ICASSP 1987

"Number of abstracts submitted > 1200

Number of accepted paper: 615

Number of attendees: 1400

The technical program was organized into 54 sessions arranged into 5 parallel lecture sessions and two poster sessions.

Also in 1987:
Thomas Knoll and John Knoll developed the first version of Photoshop.
Perl programming language was developed.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1988 - NEW YORK, USA
APRIL 11 - 14

ICASSP 88

Most Cited Paper

Voice conversion through vector quantization

Masanobu Abe, Satoshi Nakamura, Kiyohiro Shikano, and Hisao Kuwabara

ATR Interpreting Telephony Research Laboratories,

ABSTRACT

In this paper, we propose a new voice conversion technique through vector quantization and spectrum mapping. The basic idea of this technique is to make

codebooks carry all in individuality in terms o A conversion of acoustic another is, therefore, red the codebooks of the t method of making mapp

PAPER CATEGORIES

Most Downloaded Paper

E3.6

MUSIC, MAXIMUM LIKELIHOOD AND CRAMÉR-RAO BOUND

Petre Stoica
Department of Automatic Control
Polytechnic Institute of Bucharest
Splaiul Independentei 313
R-77 206 Bucharest
Romania

Arye Nehorai
Department of Electrical Engineering
Yale University
New Haven, CT 06520
USA

Abstract

The problem of finding the directions of multiple plane waves with linear arrays of sensors and the related one of estimating the parameters of multiple superimposed exponential signals in noise, have attracted considerable interest recently. Several methods, such as the MUSIC and

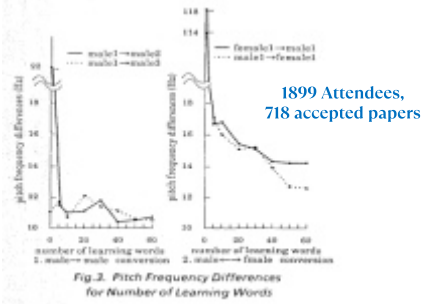


Fig.3. Pitch Frequency Differences for Number of Learning Words

1899 Attendees,
718 accepted papers

Figure from , Abe et al., ICASSP 1988

Jont Allen
General Chair, ICASSP 1988

Also in 1988:
Disney's "Who Framed Roger Rabbit" was released.
The first well known computer worm,
Morris Worm, was released on Nov. 2nd.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1989 - GLASGOW, UK
MAY 23 - 26

ICASSP·89

Most Cited and Downloaded Paper

D4.2

PRACTICAL FAST 1-D DCT ALGORITHMS WITH 11 MULTIPLICATIONS

Christoph Loeffler*, Adriaan Ligtenberg**, and George S. Moschytz*

* Institute for Signal and Information Processing, ETH Zürich, CH-8092 Zürich, Switzerland.
** AT&T Bell Laboratories, Crawfords Corner Rd., Holmdel N.J. 07733

ABSTRACT

A new class of practical fast algorithms is introduced for the Discrete Cosine Transform (DCT), an important transform that is of particular interest in image compression.

method	Chen	Wang	Lee	Vetterli	Suehiro	Ran	our Alg.
ref.	[4]	[5]	[6]	[7]	[8]	[9]	Alg.
mult.	16 (13)	13	12	12	12	11	11
add.	26 (19)	29	29	29	29	29	29

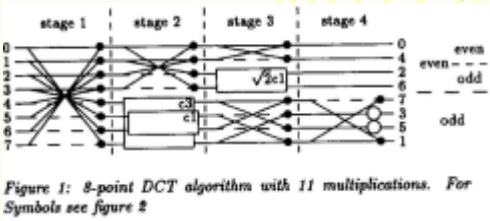


Figure 1: 8-point DCT algorithm with 11 multiplications. For Symbols see figure 2

Figure from Loeffler et al., ICASSP 1989.

Tariq Durrani, General Chair, ICASSP 1989, honored with the title, Order of the British Empire "for services to electronics research and higher education."

1617 Attendees
711 accepted papers

"... the main thing was the people involved in society were very prescient, they realized that this is an emerging field. More importantly, there is a need to provide a platform for people who are working in this area, a coherent platform. That was really what the Signal Processing Society became, and the annual IEEE ICASSP."

- Tariq Durrani, in his Oral History Project interview. https://ethw.org/Oral-History:Tariq_Durrani

Also in 1989:
Japan aired the very first daily broadcast of a high-definition television program
In 1989, Nintendo released the Game Boy, and gaming was changed forever!

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1990 - ALBUQUERQUE, USA
APRIL 3 - 6

ICASSP 90

On October 6, 1989, the IEEE Acoustics, Speech and Signal Processing Society changed its name to IEEE Signal Processing Society. However, the Society decided to keep the ICASSP brand for its flagship conference and never changed the conference name.

Most Cited and Downloaded Paper

On a simple algorithm to calculate the 'energy' of a signal

James F. Kaiser

Bell Communications Research, Inc., Room 2A-281
445 South Street, Morristown, NJ 07960-1910

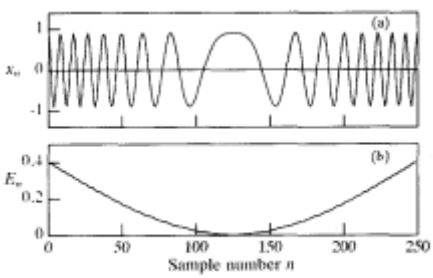


Fig. 2 The application of the algorithm to a chirp signal.

The Technical Program Chair for ICASSP 1990 was Nasir Ahmed, who invented the discrete cosine transform, His story was featured in an episode of This is Us, an NBC television drama series, in 2021.

Delores Etter was the first woman to be the General Chair of an ICASSP. She was also the first woman to become president of the IEEE Acoustics, Speech and Signal Processing (1988 - 1989)

Also, in 1990:
The first ever website was launched on August 6th,
The human genome project was launched on October 1st.

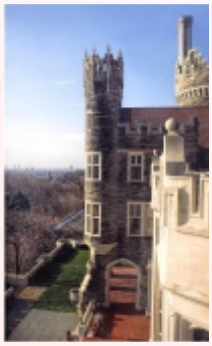
2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

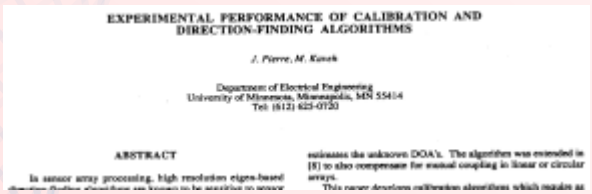
IEEE Signal Processing Society

ICASSP 1991 - TORONTO, CANADA
MAY 14 - 17

ICASSP 91

Most Cited and Downloaded Paper



EXPERIMENTAL PERFORMANCE OF CALIBRATION AND DIRECTION-FINDING ALGORITHMS

J. Pierre, M. Kaveh

Department of Electrical Engineering
University of Minnesota, Minneapolis, MN 55414
Tel: (612) 625-0720

ABSTRACT

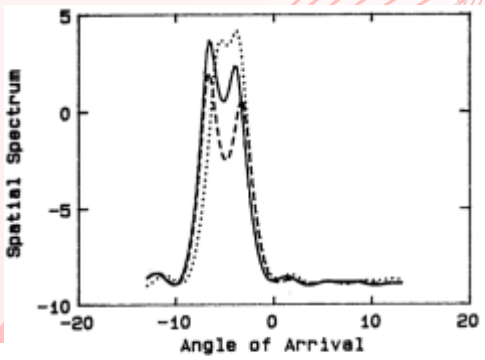


Figure 1. MUSIC's spatial spectrum for sources at -4° and -6°: calibration matrix $\hat{G}_1$ - dotted; calibration matrix $\hat{G}_2$ - solid; and calibration matrix $\hat{G}_3$ - dashed.

Figure from J. Pierre , M. Kaveh ICASSP 1991

Number of attendees – 1949
No of papers accepted – 933

Yiu-Tong Chan
General Chair, ICASSP 1991

AN. Venetsanopoulos
TPC Chair, ICASSP 1991

Also In 1991:
The first-ever website was created!
It was just a page explaining what the internet was.
Soviet Union collapsed, ending the Cold War era.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1992 - SAN FRANCISCO, USA
MARCH 23 - 30

ICASSP-92

Most Cited and Downloaded Paper

SWITCHBOARD: Telephone Speech Corpus for Research and Development

John J. Godfrey Edward C. Holliman Jane McDaniel

Texas Instruments, Inc., Dallas, TX 75265

Abstract

SWITCHBOARD is a large multispeaker corpus of conversational speech and text which should be of interest to researchers in speaker authentication and large vocabulary speech recognition. About 2500 conversations by 500 speakers from around the U.S. were collected automatically over T1 lines at Texas Instruments. Designed for train-

2 SWITCHBOARD CORPUS

SWITCHBOARD, which is about 65% complete at the time of this writing, will include 2500 conversations of three to ten minutes' duration, carried on by about 500 paid volunteers of both sexes from every major dialect of American English. In round numbers, this amounts to over 250 hours of speech and nearly 3 million words of text.

	Sessions					
	Training	Test Sets				
Spkr	1 2 ... 15	16	17	...	25	Size
1	x x ... x	x	x	...	x	
2	x x ... x	x	x	...	x	625
⋮						Development
25	x x ... x	x	x	...	x	tokens
26	x x ... x	x	x	...	x	
27	x x ... x	x	x	...	x	625
⋮						Evaluation
50	x x ... x	x	x	...	x	tokens
51	x x ... x					
52	x x ...					
⋮		Imposter Set(s)				
i	x x					
i + 1	x					
⋮		3750 tokens				
499	x					
500	x					

Figure 1: A possible configuration of SWITCHBOARD for speaker authentication. Each token (x) is one talker's side of one conversation.

Figure from Godfrey et al, ICASSP 1992

General Chair: Marcia A. Bush

Technical Program Committee Chairs: Michael R. Portnoff, Gary Kopec

Number of attendees: 2000
Number of accepted papers: 806

Also in 1992:
Disneyland Paris opened, and Mickey Mouse learned how to say "Bonjour".
IBM Simon, a touchscreen mobile phone and PDA, considered the first smartphone, was introduced.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

IEEE Signal Processing Society

ICASSP 1993 - MINNEAPOLIS, USA
APRIL 27 - 30

ICASSP-93

Most Cited Paper

RECOVERY OF LOST OR ERRONEOUSLY RECEIVED MOTION VECTORS

Wai-Man Lam†, Amy R. Reibman‡, and Bede Liu†
†Department of Electrical Engineering
Princeton University, NJ 08544

‡AT&T Bell Laboratories
Holmdel, NJ 07733

ABSTRACT

a motion compensated video coding, if motion
ors are lost or received with errors, not only
the current frame be corrupted, but also the

fact that motion vectors are usually cod
entially. The use of an intraframe mec
past motion vector has been proposed f
covery of lost or erroneously received mo
tors [1]. In this paper, we propose a

Most Downloaded Paper

SOME USEFUL PROPERTIES OF TEAGER'S ENERGY OPERATORS

James F. Kaiser

Department of Electrical and Computer Engineering
Rutgers University, Piscataway, New Jersey 08855

ABSTRACT

Teager's energy operators are defined in both the continuous and discrete domains and are very useful 'tools' for analyzing single component signals from an energy point-of-view. A number of important properties of these operators are shown that make it possible to determine the energy functions of quite complicated functions provided these functions can be expressed as products of simpler functions, this operation of function multiplication being typical of a modulation process. Some of the signal-action properties of the energy operator are also given that illustrate the special role of the trigonometric, Gaussian, and single soliton functions.

This energy function is a very local property of the signal depending only on the signal and its first two time derivatives.

It is also useful to define a second energy-like function representing an interaction between two time functions such as $y(t)$ and $h(t)$; let

$$\Psi_c[y,h] \triangleq \dot{y}\dot{h} - y\ddot{h} \quad \text{and its inverse} \quad \Psi_c[h,y] \triangleq \dot{h}y - h\dot{y} \,, \tag{D.2}$$

where the order of the functions is important.

From (D.2) there follows the identity

$$\Psi_c[y,y] \equiv \Psi[y] \,. \tag{D.3}$$

Mos Kaveh, the General Chair of ICASSP 1993, has attended 40 ICASSPs, starting with 1981.

Also in 1993:
Mosaic, the first general-use internet browser, was released.
The movie "Jurassic Park" was released.
It made pioneering use of computer-generated imagery to depict prehistoric animals.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1994 - ADELAIDE, AUSTRALIA
APRIL 19 - 22

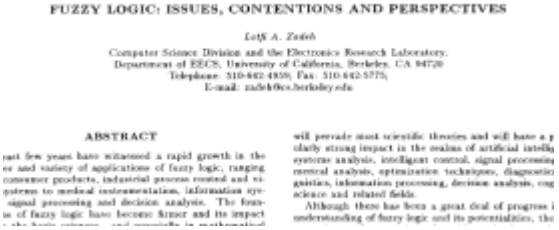
First ICASSP held in Australia

Most Cited Paper

BLIND FRACTIONALLY-SPACED EQUALIZATION, PERFECT-RECONSTRUCTION FILTER BANKS AND MULTICHANNEL LINEAR PREDICTION

ABSTRACT

Most Downloaded Paper



FUZZY LOGIC: ISSUES, CONTENTIONS AND PERSPECTIVES

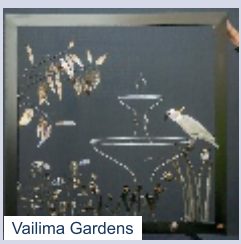
Lotfi A. Zadeh

Computer Science Division and the Electronics Research Laboratory, Department of EECS, University of California, Berkeley, CA 94720

ABSTRACT

Portrait of the Artist as a young Nerd

Robert Bogner, General Chair of ICASSP 1994, circa 1980. Bob took up smithy after retirement. He does artistic smithy-work, and at 90, is now preparing for his next art exhibition. (Pictures reproduced with permission.)



Vailima Gardens

Aboriginal dance during banquet.

Also in 1994:
Amazon was launched on July 5th.
QR code was invented

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1995 - DETROIT, USA
MAY 9 - 12

ICASSP-95
DETROIT MICHIGAN

Most Cited and Downloaded Paper

IMPROVED BACKING-OFF FOR M-GRAM LANGUAGE MODELING

Reinhard Kneser

Philips GmbH
Research Laboratories
D-52066 Aachen, Germany
kneser@pfa.philips.de

Hermann Ney

Lehrstuhl für Informatik VI
RWTH Aachen, University of Technology
D-52056 Aachen, Germany
ney@informatik.rwth-aachen.de

ABSTRACT

stochastic language modeling, backing-off is a widely d method to cope with the sparse data problem. In

models, for example, a $(M-1)$-gram distribution be used for backing-off.

Usually the normal probability distribution of

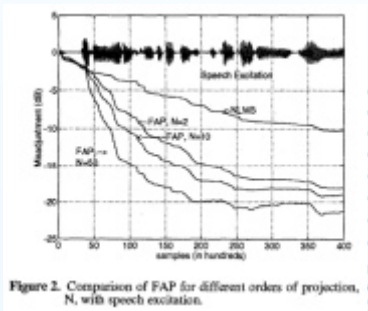


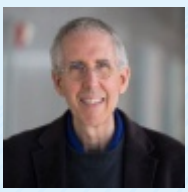
Figure 2. Comparison of FAP for different orders of projection, N, with speech excitation.

Figure from Steven L. Gay and Sanjeev Tavathia, "The Fast Affine Projection Algorithm" ICASSP, 1995.

General Chair: Alfred Hero

Technical Program Chair: William J. Williams and Andrew Yagle

Number of attendees: 1941
Number of accepted papers: 917



"At the end of the conference, the SP Soc. president Tariq Durrani gave the organizers a gift of a bottle of Scottish Whiskey – this was consumed with pleasure at the final post-ICASSP-95 committee meeting in July 1995!"

- Al Hero, General Chair

General Chair Al Hero:

Also in 1995:
The first-ever Toy Story movie was released, making us believe that toys really come to life!
DVD was invented.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1996 - ATLANTA, USA
MAY 7 - 10

Most Cited Paper

UNIT SELECTION IN A CONCATENATIVE SPEECH SYNTHESIS SYSTEM USING A LARGE SPEECH DATABASE

Andrew J. Hunt and Alan W. Black
ATR Interpreting Telecommunications Research Labs
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02, Japan
ashunt,awb@hip.atr.co.jp

ABSTRACT

One approach to the generation of natural-sounding synthesized speech waveforms is to select and concatenate units from a large speech database. Units (in the current work, phonemes) are selected to produce a natural realisation of a target phoneme sequence predicted from text which is annotated with prosodic and phonetic context information. We propose that the units in a synthesis database can be considered as a state transition network in which the state occupancy cost is the distance between a database unit and a target, and the transition cost is an estimate of the quality of concatenation of two consecutive units. This framework has many similarities to HMM-based speech recognition. A to be flexible. CHATR has been used to synthesize speech from a wide range of databases including male and female speakers, Japanese and English, and isolated words and continuous read speech [1]. Three databases have varied in size from 10 minutes to 150 minutes.

The most important issue to be solved to make this concatenative synthesis approach effective is the selection of appropriate units from the database. This paper advances previous research on CHATR by characterising the selection procedure as Viterbi decoding of a state transition network consisting of all units in a synthesis database, as described in Section 3. This new view of the synthesis database permits automated training using natural speech. In Section 5,

Most Downloaded Paper

SPEECH ENHANCEMENT BASED ON A PRIORI SIGNAL TO NOISE ESTIMATION

Pascal SCALART[1], Jaque VIEIRA FILHO[2,3]

[1]France Telecom - CNET/LAA/TSS/CMC, 2 Avenue Pierre Marzin 22307 Lannion Cedex, France
[2]Universidade Estadual Paulista DEE/FEIS/UNESP, Av. Brasil Centro 56, Ilha solteira- SP, Brasil
[3]Universidade Estadual de Campinas (DECOM/FEE/UNICAMP), SP, Brasil
E-mail : scalart @lannion.cnet.fr

ABSTRACT

This paper addresses the problem of single microphone frequency domain speech enhancement in noisy environments. The main characteristics of available frequency domain noise reduction algorithms are firstly presented. We have confirmed that the A Priori SNR estimation leads to the best subjective results. According to these conclusions, a new approach is then developed which achieves a trade-off between effective noise reduction and low computational load for real-time

II - SPEECH ENHANCEMENT SYSTEMS

To date, conventional single microphone frequ domain speech enhancement techniques have b proposed on more or less an ad hoc basis. A com feature of these techniques is that the noise reduc process can be related to the estimation of a *Short-T Suppression Factor*. Since the spectral components assumed to be statistically independent, this factor adjusted individually as a function of the relative lev *Posteriori Signal to Noise Ratio* on each frequency

Monson Hayes, General Chair of ICASSP 1996, was general chair for two more ICASSPs, in 2018 in Montreal and in 2024 in Seoul. No one else has led the organization of three ICASSPs.

Also in 1996:
Dolly the sheep, the first cloned mammal, was born.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1997 - MUNICH, GERMANY
APRIL 21 - 24

ICASSP 97

Most Cited and Downloaded Paper

TWO-DIMENSIONAL PILOT-SYMBOL-AIDED CHANNEL ESTIMATION BY WIENER FILTERING

Peter Hoeher, Stefan Kaiser, and Patrick Robertson
Institute for Communications Technology
German Aerospace Research Establishment (DLR)
P.O. Box 1116, D-82230 Oberpfaffenhofen, Germany
E-mail: Firstname.Lastname@dlr.de

ABSTRACT

clude the upcoming European Terrestrial Digital Vid
Broadcasting and Digital Audio Broadcasting standard

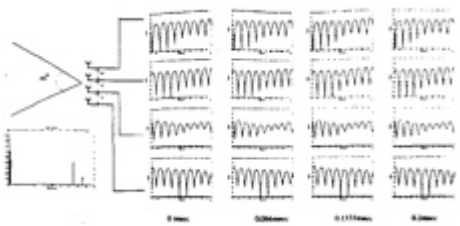


Figure 3. Channel frequency response at four different antennas for GSM in a typical hilly terrain at 1800 MHz. Mobile speed is 100 KPH. The response is plotted at four time instances spaced 1200 μ secs.

Figure from Paulraj, "Space-Time Processing for Wireless Communications," ICASSP, 1997.

General Chair: Manfred Lang

Technical Program Committee Chair: Josef A. Nossek

No. of attendee: 1800.
No of papers accepted: 1050

Also in 1997:
IBM's Deep Blue supercomputer defeated Garry Kasparov, the reigning world chess champion. Titanic became the highest-grossing movie of all time at the time.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1998 - SEATTLE, USA
MAY 12 - 14

Most Cited and Downloaded Paper

FFTW: AN ADAPTIVE SOFTWARE ARCHITECTURE FOR THE FFT

Matteo Frigo

MIT Laboratory for Computer Science
545 Technology Square NE43-203
Cambridge, MA 02139
athena@theory.lcs.mit.edu

Steven G. Johnson

Massachusetts Institute of Technology
77 Massachusetts Avenue, 12-104
Cambridge, MA 02139
stevenj@alum.mit.edu

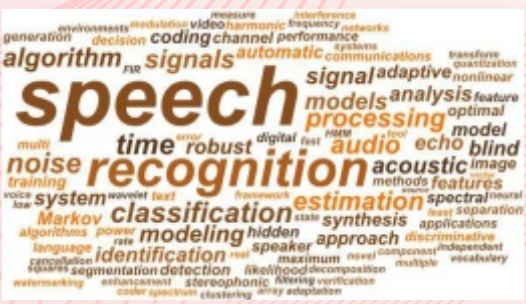
ABSTRACT

FFT literature has been mostly concerned with minimizing the number of floating-point operations performed by an algorithm. Unfortunately, on present-day microprocessors this measure is far


```
fftw_plan plan;
COMPLEX A[n], B[n];

/* plan the computation */
plan = fftw_create_plan(n);
```

For the first time, the technical program was based on reviews of full papers submitted by authors. Until 1997, only abstracts or short summaries of papers were reviewed.



The Banquet Speaker at ICASSP 1998 was Nathan Myhrvold, Chief Technology Officer, Microsoft. He referred to humans of the future as 'meat-based computers' during the talk.

Also in 1998:
Google, Inc. was founded in Menlo Park, CA on September 4th.
The first module of the International Space Station, named Zaryam was launched on Nov, 20th

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 1999 - PHOENIX, USA
MARCH 15 - 19

Most Cited and Downloaded Paper

METHOD OF OPTIMAL DIRECTIONS FOR FRAME DESIGN

Kjersti Engan, Sven Ole Aase, and John Håkon Husøy

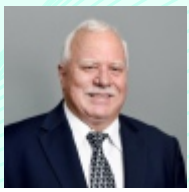
Høgskolen i Stavanger
Department of Electrical and Computer Engineering
P. O. Box 2557 Ullandhaug, N-4004 Stavanger, Norway
Phone: +47 51 83 20 08, Fax: +47 51 83 17 50
E-mail: Kjersti.Engan@tn.his.no

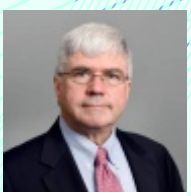
ABSTRACT

A frame design technique for use with vector selection algorithms, for example Matching Pursuits (MP), is presented. The design al- ... significantly improved version of the frame design algorithm, and we call it the Method of Optimal Directions (MOD).

The Industry Technology Track was organized at ICASSP for the first time in 1999. The Society established a Standing Committee for Industry Technology at the same time.

There were two General Chairs, Andreas Spanias (left) and Doug Cochran (right) , in 1999. This was a first in ICASSP history.





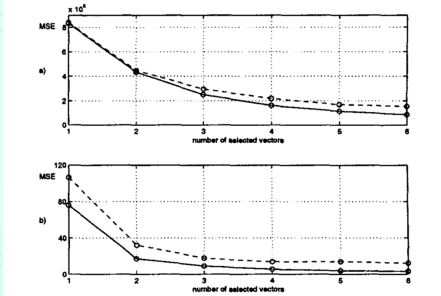


Figure 2: MSE is plotted as a function of different numbers of vectors in an approximation. Test signal is used. Dotted: frames optimized by the old frame design algorithm, solid: frames optimized using MOD a) speech signal, b) ECG signal, MIT100.

Figure from Engan et al, ICASSP 1999

Also in 1999:

Euro was established as the electronic currency for financial transactions in 11 countries of the European Community on January 1st.

On May 29th, Discovery became the first space shuttle to dock with the International Space Station,

Celebrating Signal Processing

ICASSP 2000 - ISTANBUL, TURKEY
JUNE 5 - 9

Most Cited and Downloaded Paper

SPEECH PARAMETER GENERATION ALGORITHMS FOR HMM-BASED SPEECH SYNTHESIS

Keiichi Tokuda[1], Takayoshi Yoshimura[1], Takashi Masuko[2], Takao Kobayashi[2], Tadashi Kitamura[1],

[1]Department of Computer Science, Nagoya Institute of Technology, Nagoya, 466-8555 Japan
[2]Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, Yokohama, 226-8502 Japan
Email: {tokuda,yossie,kitamura}@ics.nitech.ac.jp, {masuko,takao}@ip.titech.ac.jp

This was the 25th ICASSP. The silver jubilee was celebrated throughout the conference.

Professor Huseyin Abut wanted to organize ICASSP at the turn of the millennium in Istanbul, Turkey so much that he proposed the city as the ICASSP 2000 venue in 1992. Istanbul was selected to host ICASSP 2000 at the Conference Board meeting at ICASSP 1992. This was the earliest selection of a venue in ICASSP history, a record not challenged since then either.

A "Humor in DSP" competition was organized at ICASSP 2000

Also in 2000:
The year 2000 was a leap year not because the year was divisible by 4, but because it was divisible by 400. January 1 came and went without causing much of the chaos predicted because of the Y2K problem.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2001 – SALT LAKE CITY, USA
MAY 7 - 11

DSP Everywhere

2001 ICASSP

Salt Lake City

Most Cited Paper

PERCEPTUAL EVALUATION OF SPEECH QUALITY (PESQ) – A NEW METHOD FOR SPEECH QUALITY ASSESSMENT OF TELEPHONE NETWORKS AND CODECS

Antony W. Rix[1], John G. Beerends[2], Michael P. Hollier[1] and Andries P. Hekstra[2]
[1]PsyTechnics, BT-4/96 Adastral Park, Ipswich IP5 3RE, United Kingdom
[2]Royal PTT Nederland NV, NL-2260 Leidschendam, The Netherlands
E-mail: antonyr@psytechnics.com

ABSTRACT

Various objective speech quality assessment models, such as ...

Most Downloaded Paper

THE SQUARE-ROOT UNSCENTED KALMAN FILTER FOR STATE AND PARAMETER-ESTIMATION

Rudolph van der Merwe and Eric A. Wan
Oregon Graduate Institute of Science and Technology
20000 NW Walker Road, Beaverton, Oregon 97006, USA
{rvdmerwe,ericwan}@ece.ogi.edu

ABSTRACT

Over the last 20-30 years, the extended Kalman filter (EKF) has become the algorithm of choice in numerous nonlinear estimation and machine learning applications. These include estimating the state of a nonlinear dynamic system as well estimating parameters for nonlinear system identification (e.g., learning the weights of a neural network). The EKF applies the standard linear Kalman filter methodology to a linearization of the true nonlinear system. This approach is sub-optimal, and can easily lead to divergence. Julier et al. [5] proposed the unscented Kalman filter (UKF) as a derivative-free alternative to the extended Kalman filter in the framework of state estimation. This was extended to parameter-

ICASSP 2001 offered a student forum where students presented their work on their own.

Full registration included admission to the banquet.

V John Mathews, General Chair of ICASSP 2001, is one of the two General Chairs of ICASSP 2025.

Also in 2001:
Wikipedia, the free, online, open–source encyclopedia, was launched on Jan. 15th.
Claude Shannon passed away on Feb. 24th.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2002 - ORLANDO, USA
MAY 11 - 17

Most Cited and Downloaded Paper

WHY IS IMAGE QUALITY ASSESSMENT SO DIFFICULT?

Zhou Wang and Alan C. Bovik

Lab for Image and Video Engi., Dept. of ECE
Univ. of Texas at Austin, Austin, TX 78703-1084
zhouwang@ieee.org, bovik@ece.utexas.edu

Ligang Lu

IBM T. J. Watson Research Center
Yorktown Heights, NY 10598
lul@us.ibm.com

ABSTRACT

Image quality assessment plays an important role in various image processing applications. A great deal of effort has been made in re-

However, the MOS method is too inconvenient, slow and expensive for practical usage. The goal of objective image and video quality assessment research is to supply quality metrics that can

General Chairman Fred Taylor (Top) and Technical Program Chair Jose Principe.

Figure from Bovik et al, ICASSP 2002

Also In 2002:
NASA's Mars Odyssey detected signs of water on Mars.
Steve Fossett completed the first solo, around-the-world balloon flight.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2003 - HONG KONG
APRIL 6 - 10

ICASSP 2003 Hong Kong

Most Cited Paper

KERNEL INDEPENDENT COMPONENT ANALYSIS

Francis R. Bach

Computer Science Division
University of California
Berkeley, CA 94720, USA
fbach@cs.berkeley.edu

Michael I. Jordan

Computer Science Division
and Department of Statistics
University of California
Berkeley, CA 94720, USA
jordan@cs.berkeley.edu

ABSTRACT

We present a class of algorithms for independent component analysis (ICA) which use contrast functions based on canonical correlations in a reproducing kernel Hilbert space. On the one hand, we show that our contrast functions are related to mutual information and have desirable mathematical properties as measures of

"kernelization" of an extant ICA algorithm. Rather, it is a new approach to ICA based on novel kernel-based measures of dependence. We introduce two such measures. In Section 3, we define a kernel-based contrast function in terms of the first eigenvalue of a certain generalized eigenvector problem, and show how this function relates to probabilistic independence. In Section 4.3, we introduce an alternative kernel-based contrast function based on

Most Downloaded Paper

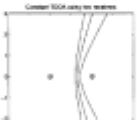
POSITIONING USING TIME-DIFFERENCE OF ARRIVAL MEASUREMENTS

Fredrik Gustafsson and Fredrik Gunnarsson

Department of Electrical Engineering
Linköping University, SE-581 83 Linköping, Sweden
Email: fredrik@isy.liu.se,fredrik@isy.liu.se

ABSTRACT

The problem of position estimation from Time Difference Of Arrival (TDOA) measurements occurs in a range of applications from wireless communication networks to electronic warfare positioning. Correlation analysis of the transmitted signal to two receivers gives rise to one hyperbolic function. With more than two receivers, we can compute more hyperbolic functions, which ideally intersect in one unique point. With TDOA measurement uncertainty, we



ICASSP 2003 was cancelled because of the outbreak of the Severe Acute Respiratory Syndrome (SARS). More than 1260 accepted papers were published as the Proceedings of ICASSP 2003 and can be downloaded from IEEE Xplore.

Also in 2003:
The Human Genome Project to map the entire human genome was completed.
Apple launched iTune Stores in April.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2004 - MONTREAL, CANADA
MAY 17 - 21

ICASSP 2004 Montreal

Most Cited and Downloaded Paper

HIGH-QUALITY LINEAR INTERPOLATION FOR DEMOSAICING OF BAYER-PATTERNED COLOR IMAGES

Henrique S. Malvar, Li-wei He, and Ross Cutler

Microsoft Research
One Microsoft Way, Redmond, WA 98052, USA

ABSTRACT

This paper introduces a new interpolation technique for demosaicing of color images produced by single-CCD digital cameras. We show that the proposed simple linear filter can lead to an ... count, either with better linear filters [4], or with nonlinear filters that adapt interpolation smoothness to a measure of image activity or edginess [1]–[4].

In this paper we present a simple linear demosaicing filter, with better performance and lower complexity than that in [4].

Demosaicing results for various interpolation algorithms. Figure from Malvar et al., ICASSP 2004.

Doug O'Shaughnessy, the General Chair of ICASSP 2004, has attended 40 ICASSPs, starting with the very first one in 1976.

Cirque du Soleil performed at the conference banquet.

One of the lunch speakers at ICASSP 2004 was Julie Payette, a famous Canadian astronaut. She later became the Governor General of Canada.

Also in 2004:
Facebook was launched on Feb. 4th.
Google launched Gmail on April 1st.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2005 - PHILADELPHIA, USA
MARCH 18 - 23

ICASSP 2005 Philadelphia

Most Cited Paper

IDENTIFYING USERS OF PORTABLE DEVICES FROM GAIT PATTERN WITH ACCELEROMETERS

Jani Mäntyjärvi, Mikko Lindholm, Elena Vildjiounaite, Satu-Marja Mäkelä, Heikki Ailisto

VTT Electronics, P.O.Box 1100, FI-90571 Oulu, Finland
phone: +358885512111, fax: +358885512320,
email: Firstname.Surname@vtt.fi

ABSTRACT

Identifying users of portable devices from gait signal acquired with three-dimensional accelerometers was studied. Three approaches, correlation, frequency domain

or burdened by the technology she is using. Gait, i.e., walking style, is fairly characteristic for individuals [3], whereas deliberate imitation of other person's walking style is difficult.

Gait recognition has been studied as a behavioral

Most Downloaded Paper

A MULTI USER BEAMFORMING SCHEME FOR DOWNLINK MIMO CHANNELS BASED ON MAXIMIZING SIGNAL-TO-LEAKAGE RATIOS

Alireza Tarighat, Mitrette Sadek, and Ali H. Sayed

Electrical Engineering Department
University of California
Los Angeles, CA 90095
Email: {tarighat,mitrette,sayed}@ee.ucla.edu

ABSTRACT

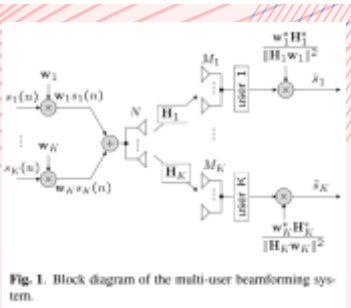


Fig. 1. Block diagram of the multi-user beamforming system.

Figure from Tarighat et al, ICASSP 2005

The conference banquet was held at the Philadelphia Museum of Art and featured a private viewing of a Salvador Dalí exhibit.

ICASSP 2005 introduced a student paper competition, travel grants for students, and a panel on "Women in Signal Processing.

General Chair Athina Petropulu

Also In 2005:
YouTube was officially launched on Dec. 15th.
Google Maps made its debut on Feb. 8th.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2006 - TOULOUSE, FRANCE
MAY 14 - 19

Most Cited Paper

SVM BASED SPEAKER VERIFICATION USING A GMM SUPERVECTOR KERNEL AND NAP VARIABILITY COMPENSATION

W. M. Campbell, D. E. Sturim, D. A. Reynolds, A. Solomonoff

MIT Lincoln Laboratory
Lexington, MA 02420
E-mail: {wcampbell,sturim,dar,als}@ll.mit.edu

ABSTRACT

Gaussian mixture models with universal backgrounds (UBMs) have become the standard method for speaker recognition. Typically, a speaker model is constructed by MAP

in this potentially high-dimensional space. The main design component in an SVM is the kernel, which is an inner product in the SVM feature space. Since inner products induce distance metrics and vice versa, the basic goal in SVM kernel design is to find an appropriate metric in the SVM feature

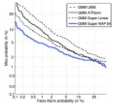


Figure from Campbell et al, ICASSP 2006

Figure 3: Matching result for image pair Graffiti (from INRIA) with viewpoint changes.

Figure 4: Matching result for image pair Car (from our lab) with translation, rotation and scale changes.

Figures from Zhao et al, ICASSP 2006

Most Downloaded Paper

IMAGE MATCHING BY NORMALIZED CROSS-CORRELATION

Feng Zhao, Qingming Huang, Wen Gao

Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
Graduate School of the Chinese Academy of Sciences, Beijing, China
{fzhao, qmhuang, wgao}@jdl.ac.cn

ABSTRACT

Correlation is widely used as an effective similarity measure in matching tasks. However, traditional correlation based matching methods are limited to the short baseline case. In this paper we propose a new correlation based method for matching two images with large camera motion. Our

robust against rotation and scale changes. There are also generalized versions of cross-correlation, which calculate the cross-correlation for each assumed geometric transformation of the correlation windows [8,9]. Although they are able to handle more complicated imaging conditions, the computational load grows very fast in the mean time.

General Chair

Technical Program Chairs

Tutorials

Special Sessions

Plenaries

Number of paper submissions > 3,000
Number of papers accepted: 1572
Number of attendees: 2169

Also In 2006:
Twitter was born, and people started sharing random thoughts in 140 characters.
Pluto's status was relegated to that of a dwarf planet.

2025 ICASSP HYDERABAD, INDIA
Celebrating Signal Processing

IEEE Signal Processing Society

ICASSP 2007 - HONOLULU, USA
APRIL 15 - 19

Hawaii 2007 ICASSP

Most Cited and Downloaded Paper

APPROXIMATING THE KULLBACK LEIBLER DIVERGENCE BETWEEN GAUSSIAN MIXTURE MODELS

John R. Hershey and Peder A. Olsen

IBM T. J. Watson Research Center

ABSTRACT

The Kullback Leibler (KL) Divergence is a widely used tool in statistics and pattern recognition. The KL divergence between two Gaussian Mixture Models (GMMs) is frequently needed as the basis of speech and image recognition. Unfortunately the KL divergence be-

where π_a is the prior probability of each state, and $\mathcal{N}(x;\mu_a,\Sigma_a)$ is a gaussian in x with mean μ_a and variance Σ_a.

We will frequently use the shorthand notation $f_a(x) = \mathcal{N}(x;\mu_a,\Sigma_a)$ and $g_b(x) = \mathcal{N}(x;\mu_b,\Sigma_b)$. Our estimate of $D(f\|g)$ will make use of the KL-divergence between individual components

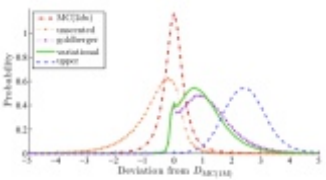


Fig. 2. Distribution of leading approximations to KL divergence relative to the reference estimate $D_{MC(1|M)}$.

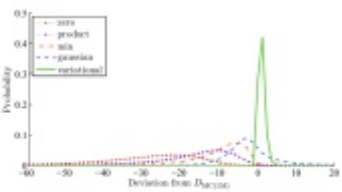


Fig. 3. Distribution of the simple/closed-form approximations to KL divergence relative to the reference estimate $D_{MC(1|M)}$. A trivial lower bound of zero is also included for reference.

Figures from Hershey, ICASSP 2007

One of the gifts every delegate received was a gift card to a nearby shopping center so that they could buy souvenirs.

"Before 2007 ICASSP, we had mostly ONE single coffee and/or food table with a long line due to traffic. We were the first to implement distributed tables to ease the traffic jam."

-General Chair Ray Liu

Also in 2007:
Apple released the first-ever iPhone on June 29th.
Netflix introduced streaming services, allowing subscribers to access content online.

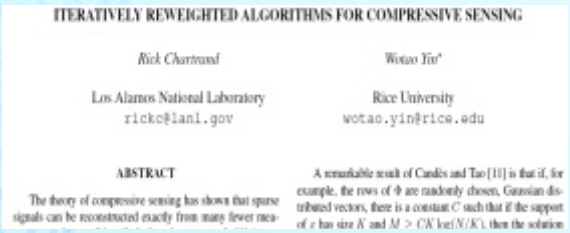
2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2008 - LAS VEGAS, USA
MARCH 30 - APRIL 4

Most Cited and Downloaded Paper



ITERATIVELY REWEIGHTED ALGORITHMS FOR COMPRESSIVE SENSING

Rick Chartrand — Los Alamos National Laboratory — rickc@lanl.gov

Wotao Yin* — Rice University — wotao.yin@rice.edu

ABSTRACT

The theory of compressive sensing has shown that sparse signals can be reconstructed exactly from many fewer mea-

A remarkable result of Candès and Tao [11] is that if, for example, the rows of Φ are randomly chosen, Gaussian distributed vectors, there is a constant C such that if the support of x has size K and $M > CK \log(N/K)$, then the solution

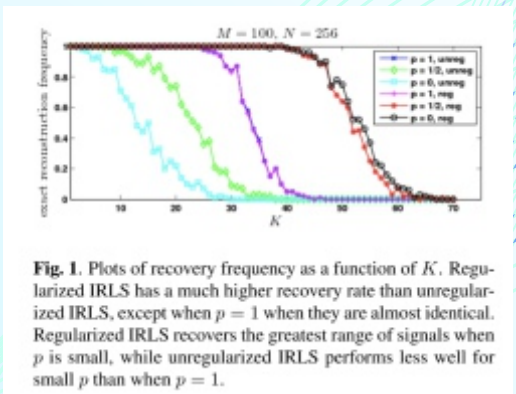
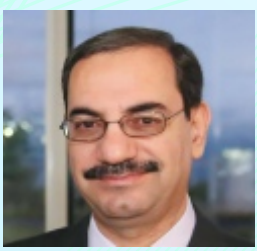


Fig. 1. Plots of recovery frequency as a function of K. Regularized IRLS has a much higher recovery rate than unregularized IRLS, except when $p = 1$ when they are almost identical. Regularized IRLS recovers the greatest range of signals when p is small, while unregularized IRLS performs less well for small p than when $p = 1$.

Figure from Chartrand et al., ICASSP 2008



ICASSP 2008 Program gave special attention to trends in online education.

Ali H .Sayed, General Chair

IEEE Student registration fee was kept at US$ 150.

Instead of the regular banquet program, ICASSP 2008 organized a poolside reception with entertainment with no additional cost. However, the number of attendees at the reception was limited to 1000.

Also In 2008:
The first Android phone was released.
Spotify was launched, changing the way people listen to music.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2009 - TAIPEI, TAIWAN
APRIL 19 - 24

Most Cited Paper

INTERFERENCE ALIGNMENT VIA ALTERNATING MINIMIZATION

Steven W. Peters and Robert W. Heath, Jr.

The University of Texas at Austin
Department of Electrical & Computer Engineering
1 University Station C0806
Austin, TX 78712-0240
email {peters,rheath}@ece.utexas.edu

ABSTRACT

Using interference alignment, it has been shown that the number of degrees of freedom in the interference channel scales linearly with the number of users. Unfortunately, closed-form solutions for interference alignment over constant-coefficient channels with more than 3 users are

and interference subspaces are constrained to be orthonormal and, with the optimization used, will be shown to lie on the Grassmann manifold. The gradient of the objective function on this manifold has a closed-form solution so an alternating minimization approach can be applied. We establish convergence of the algorithm, although convergence to a

Most Downloaded Paper

VIBE: A POWERFUL RANDOM TECHNIQUE TO ESTIMATE THE BACKGROUND IN VIDEO SEQUENCES

Olivier Barnich and Marc Van Droogenbroeck

University of Liège
Montefiore Institute, INTELSIG Group
Liège, Belgium

ABSTRACT

Background subtraction is a crucial step in many automatic video content analysis applications. While numerous acceptable techniques have been proposed so far for background extraction, there is still a need to produce more efficient algorithms in terms of adaptability to multiple environments, noise resilience, and computation efficiency. In this paper, we

environment (changing time of day, clouds, tree leafs...). However since its sensitivity cannot be accurately tuned, its ability to successfully handle high- and low-frequency changes in the background is debatable, as detailed in [4]. Furthermore the estimation of the parameters of the model (especially the variance) can become problematic for noisy images.

Lin-shan Lee, General Chair

"It is apparent that the frontiers of science as well as various new technologies have created a completely new horizon of signal processing. The ICASSP 2009 theme is hence "Signals over the Horizon." It is our hope that signals about such frontiers and new technologies and beyond will be transmitted from the conference and received all over the world."

- Lin-shan Lee in "Message from the General Chair".

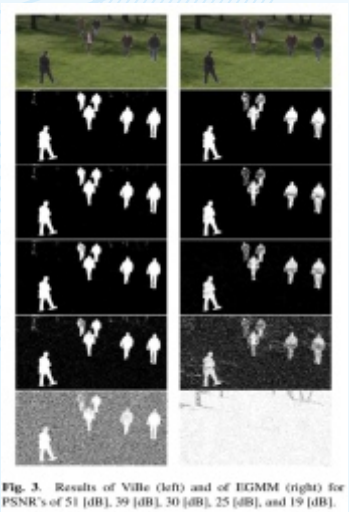


Fig. 3. Results of Ville (left) and of EGMM (right) for PSNR's of 51 [dB], 39 [dB], 30 [dB], 25 [dB], and 19 [dB].

Figure from Barnich, ICASSP 2009

Also In 2009:

Barack Obama became the first African American to be the President of the USA.

Researchers at NIST created the first small-scale device that could be described as a quantum computer.

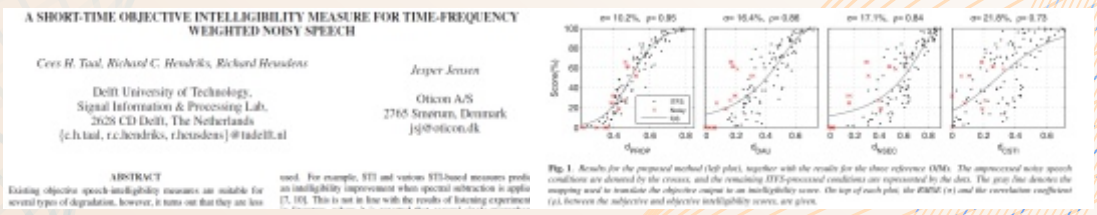
2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2010 - DALLAS, USA
MARCH 14 - 19

Most Cited and Downloaded Paper



A SHORT-TIME OBJECTIVE INTELLIGIBILITY MEASURE FOR TIME-FREQUENCY WEIGHTED NOISY SPEECH

Cees H. Taal, Richard C. Hendriks, Richard Heusdens

Delft University of Technology, Signal Information & Processing Lab, 2628 CD Delft, The Netherlands {c.h.taal, r.c.hendriks, r.heusdens}@tudelft.nl

Jesper Jensen

Oticon A/S, 2765 Smørum, Denmark jsj@oticon.dk

ABSTRACT

Existing objective speech intelligibility measures are suitable for several types of degradation, however, it turns out that they are less used. For example, STI and various STI-based measures predi an intelligibility improvement when spectral subtraction is applie [7, 10]. This is not in line with the results of listening experiment

Fig. 1. Results for the proposed method (left plot), together with the results for the three reference (MM). The unprocessed noisy speech conditions are denoted by the crosses, and the remaining DTFS-processed conditions are represented by the dots. The gray line denotes the mapping used to translate the objective output to an intelligibility score. On top of each plot, the RMSE (σ) and the correlation coefficient (ρ), between the subjective and objective intelligibility scores, are given.

Figure from Taal et al, ICASSP 2010

"The tour that ICASSP takes from continent to continent underscores the international nature and global impact of the signal processing field."

- General Chair Scott Douglas

The banquet entertainment at ICASSP 2010 included a rodeo performance.

Ron Schafer gave a plenary talk entitled "A Celebration of the Science and Technology of DSP," presented by Ronald W. Schafer of Hewlett-Packard Laboratories. This presentation honored the significant advancements in Digital Signal Processing over the previous six decades, and its widespread impact across various applications.

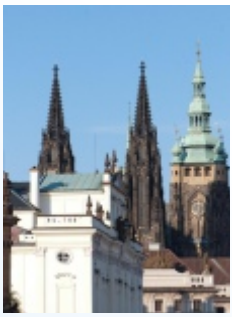
Also in 2010:
Google's self-driving car completed its first successful test drive.
Apple released the first-ever iPad.

ICASSP 2011 – PRAGUE, CZECH REPUBLIC MAY 22 - 27

ICASSP

May 22–27, 2011
Prague
Czech Republic

2011 International Conference
on Acoustics, Speech and Signal Processing



Most Cited and Downloaded Paper

A COMPLETE ENSEMBLE EMPIRICAL MODE DECOMPOSITION WITH ADAPTIVE NOISE

*María E. Torres *, Marcelo A. Colominas *, Gastón Schlotthauer *, Patrick Flandrin ‡*

* Laboratorio de Señales y Dinámicas no Lineales, Universidad Nacional de Entre Ríos, Argentina
‡ Laboratoire de Physique (UMR CNRS 5672), École Normale Supérieure de Lyon, France

ABSTRACT

In this paper an algorithm based on the ensemble empirical mode decomposition (EEMD) is presented. The key idea on the EEMD relies on averaging the modes obtained by EMD situations, in this paper we propose a variation of the EEMD algorithm that provides an exact reconstruction of the original signal and a better spectral separation of the modes, with a lower computational cost.

Prior to ICASSP in Prague, the conference was held outside North America no more than once every three years. This was changed with ICASSP 2011 with ICASSP 2009 organized in Taipei. In fact, ICASSP 2012 was also held outside North America, in Kyoto, Japan.

ICASSP 2011 introduced overview talks by experts called "trends" as part of the technical program. Such overview talks have been included in many ICASSPs since then, often under different names.

For the first time, and perhaps the only time ever, all lecture presentations at ICASSP 2011 were recorded and made available on the internet.

General Chairs

Petr Tichavský
Institute of Information Theory and Automation, Prague, Czech Republic

Jan Černocký
Brno University of Technology, Brno, Czech Republic

Aleš Procházka
Institute of Chemical Technology, Prague, Czech Republic

Technical Program Chairs

Jonathon A. Chambers
Loughborough University, UK

Alle-Jan van der Veen
Delft Univ. of Technology, Delft, Netherland

Also in 2011:
In a televised Jeopardy! contest in February, IBM's Watson computer defeated the TV quiz show's two foremost all-time champions, Brad Rutter and Ken Jennings.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

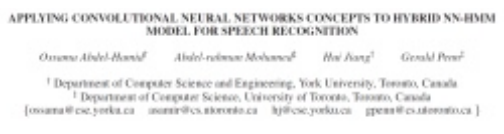
IEEE Signal Processing Society

ICASSP 2012 - KYOTO, JAPAN
MARCH 25 - 30

ICASSP 2012

IEEE International Conference on Acoustics, Speech, and Signal Processing

Most Cited and Downloaded Paper



APPLYING CONVOLUTIONAL NEURAL NETWORKS CONCEPTS TO HYBRID NN-HMM MODEL FOR SPEECH RECOGNITION

Ossama Abdel-Hamid[1] Abdel-rahman Mohamed[2] Hui Jiang[3] Gerald Penn[2]

[1] Department of Computer Science and Engineering, York University, Toronto, Canada
[2] Department of Computer Science, University of Toronto, Toronto, Canada
[{oasamir@cse.yorku.ca asamir@cs.toronto.ca hj@cse.yorku.ca gpenn@cs.toronto.ca }]

ABSTRACT

Convolutional Neural Networks (CNN) have shared masses in achieving translation invariance for many image processing tasks.

other units sensitive to a subset of the input features that are significant to recognition.

In this paper we propose to use some concepts developed in Con-

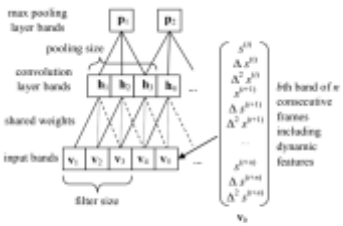


Fig. 1. Diagram to shown a pair of CNN convolution layer and max-pooling layers, where weights represented by the same line style are shared among all convolution layer bands.

Figure from Abdel-Hamid, et al., ICASSP 2012

General Chairs Hideaki Sakai (Left) and Takao Nishitani (Right).

The conference venue was the Kyoto International Conference Center. The buildings were designed as a combination of trapezoidal and inverted triangular forms reminiscent of the old Japanese architecture "gassho-zukuri," resulting in few vertical walls or columns.

Also in 2012:
Curiosity Rover landed on Mars, expanding space exploration.
The world was supposed to end on Dec 21, 2012, but here we are today at the 50th ICASSP.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

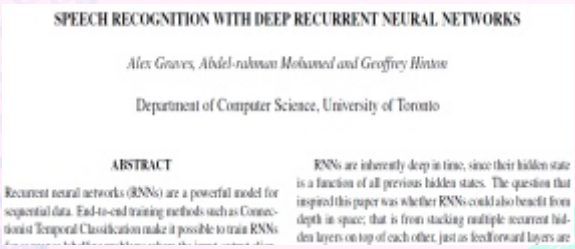
50

IEEE Signal Processing Society

ICASSP 2013 - VANCOUVER, CANADA
MAY 26 - 31

ICASSP 2013

Most Cited and Downloaded Paper



SPEECH RECOGNITION WITH DEEP RECURRENT NEURAL NETWORKS

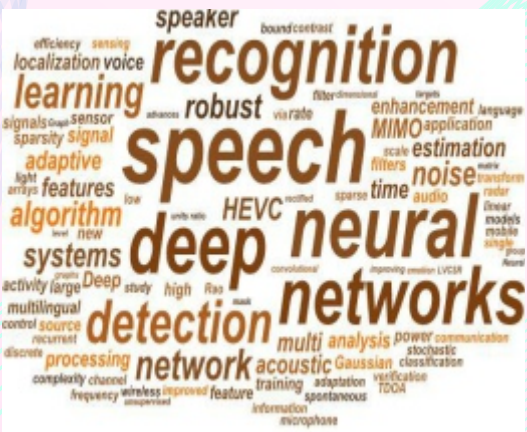
Alex Graves, Abdel-rahman Mohamed and Geoffrey Hinton

Department of Computer Science, University of Toronto

ABSTRACT

Recurrent neural networks (RNNs) are a powerful model for sequential data. End-to-end training methods such as Connectionist Temporal Classification make it possible to train RNNs

RNNs are inherently deep in time, since their hidden state is a function of all previous hidden states. The question that inspired this paper was whether RNNs could also benefit from depth in space; that is from stacking multiple recurrent hidden layers on top of each other, just as feedforward layers are



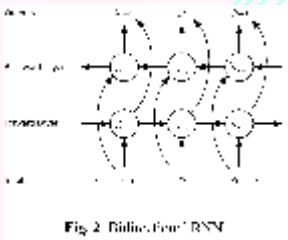
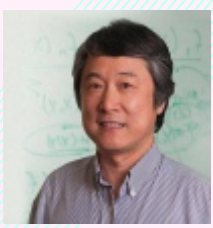
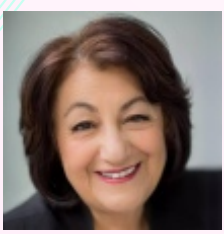


Fig. 2. Bidirectional RNN

Figure from Graves, et al., ICASSP 2013

Geoffrey Hinton, one of the authors of the most cited as well as the most downloaded papers presented at ICASSP 2013, was also a plenary speaker in this event. He went on to win the Turing Award in 1918 and the Nobel Prize for Physics in 2024.





General Chairs Li Deng (left) and Rabab Ward (right)

Also in 2013:
Selfies became a huge trend, and was named Oxford Dictionary's Word of the Year.
Bitcoin crossed the $1,000 mark for the first time ever!

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2014 - FLORENCE, ITALY
MAY 4 - 9

ICASSP 2014
IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING

Most Cited Paper

DEEP NEURAL NETWORKS FOR SMALL FOOTPRINT TEXT-DEPENDENT SPEAKER VERIFICATION

Ehsan Variani[1], Xin Lei[2], Erik McDermott[2], Ignacio Lopez Moreno[3], Javier Gonzalez-Dominguez[2,3]*

[1]Johns Hopkins Univ., Baltimore, MD USA
[2]Google Inc., USA
[3]ATVS-Biometric Recognition Group, Universidad Autonoma de Madrid, Spain

variani@jhu.edu {xinlei,erikmcd,elnota,ypd}@google.com

ABSTRACT

In this paper we investigate the use of deep neural networks (DNNs) for a small footprint text-dependent speaker verification task. At development stage, a DNN is trained to classify speakers at the frame-level. During speaker enrollment, the trained DNN is used to extract

A wide variety of SV systems have been studied using different statistical tools for each of the three phases in verification. The state-of-the-art SV systems are based on i-vectors [5] and Probabilistic Linear Discriminant Analysis (PLDA). In these systems, IFA is used as a feature extractor to extract a low-dimensional i-vector as the

Most Downloaded Paper

DEEP LEARNING OF FEATURE REPRESENTATION WITH MULTIPLE INSTANCE LEARNING FOR MEDICAL IMAGE ANALYSIS

*Yan Xu[1,2], Tao Mo[2,3], Qiwei Feng[2,4], Peilin Zhong[2,4], Maode Lai[5], Eric I-Chao Chang[2]**

[1]State Key Laboratory of Software Development Environment, Key Laboratory of Biomechanics and Mechanobiology of Ministry of Education, Beihang University
[2]Microsoft Research, Beijing, China
[3]Department of Computer Science and Technology, Tsinghua University, Beijing, China
[4]Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing, China
[5]Department of Pathology, School of Medicine, Zhejiang University, China

{eric.chang, v-tamo, v-qifen,v-penzho}@microsoft.com, xuyan04@gmail.com, lmd@zju.edu.cn

ABSTRACT

This paper studies the effectiveness of accomplishing high-level tasks with a minimum of manual annotation and good feature representation for medical images. In medical image analysis, objects like cells are characterized by significant visual feature

ages. Fully supervised feature learning [8] requires a large amount of accurately annotated data. Obtaining such annotated data is time-consuming, labor-intensive, and unhygienic. Unsupervised feature learning [12, 13, 16, 15] is based on unlabeled data. It can learn intrinsic and subtle features from the statistics of the real data. In this

General Chairs Fulvio Gini (Left) and Marco Luise (Right)

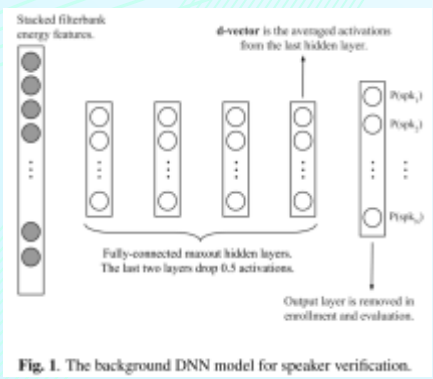


Fig. 1. The background DNN model for speaker verification.

Figure from Ehsan Variani et al, ICASSP 2014

Also in 2014:
Malaysia Airlines Flight MH370 disappeared, a major aviation mystery.
Facebook bought WhatsApp for $19 billion!

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2015 - BRISBANE, AUSTRALIA
APRIL 19 - 24

Most Cited Paper

LIBRISPEECH: AN ASR CORPUS BASED ON PUBLIC DOMAIN AUDIO BOOKS

Vassil Panayotov, Guoguo Chen*, Daniel Povey*, Sanjeev Khudanpur*

*nter for Language and Speech Processing & Human Language Technology Center of Excellen The Johns Hopkins University, Baltimore, MD 21218, USA

{vassil.panayotov,dpovey}@gmail.com, {guoguo,khudanpur}@jhu.edu

ABSTRACT

per introduces a new corpus of read English speech, suitable ning and evaluating speech recognition systems. The Lib h corpus is derived from audiobooks that are part of the Lib roject, and contains 1000 hours of speech sampled at 16 kHz

the language models, which we make available with this corp rally in Section 5 we present experimental results on models t on this data set, using both the LibriSpeech dev and test set Wall Street Journal (WSJ) [5] test sets.

2. AUDIO ALIGNMENT

Most Downloaded Paper

CONVOLUTIONAL, LONG SHORT-TERM MEMORY, FULLY CONNECTED DEEP NEURAL NETWORKS

Tara N. Sainath, Oriol Vinyals, Andrew Senior, Haşim Sak

Google, Inc., New York, NY, USA
{tsainath, vinyals, andrewsenior, hasim}@google.com

ABSTRACT

rolutional Neural Networks (CNNs) and Long Short-Term (LSTM) have shown improvements over Deep Neural Net- (DNNs) across a wide variety of speech recognition tasks. STMs and DNNs are complementary in their modeling CNNs are good at reducing frequency variations

nonlinear hidden layers. If factors of variations in the hidd could be reduced, then the hidden state of the model could ize the history of previous inputs more efficiently. In turn, make the output easier to predict. Reducing variations in th states can be modeled by having DNN layers after the LST This is similar in spirit to the hidden to output model pr

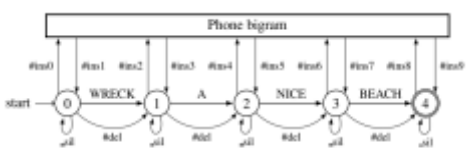


Fig. 1. Example grammar (G) acceptor for the second stage of the alignment algorithm

Figure from Sainath et al, ICASSP 2015

In his plenary talk, Neil Gordon discussed the application of Bayesian time series methods in the search for Malaysia Airlines Flight MH370.

Dr. Neil Gordon

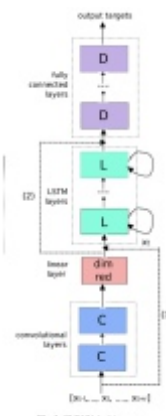


Fig. 1. CLDNN Architecture

Figure from Panayotov et al, ICASSP 2015

Also In 2015:
Apple released the first-generation Apple Watch in April.
The world's first self-driving taxis were introduced in Singapore.

Celebrating Signal Processing

ICASSP 2016 - SHANGHAI, CHINA
MARCH 20 - 25

Shanghai, China
ICASSP · 2016

Most Cited Paper

Listen, Attend and Spell

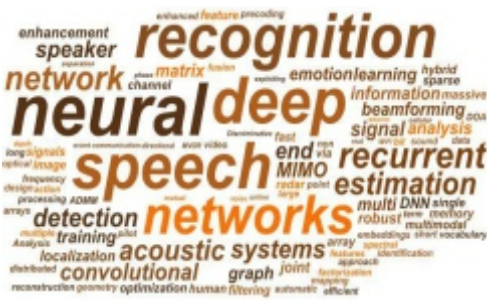
William Chan
Carnegie Mellon University
williamchan@cmu.edu

Navdeep Jaitly, Quoc V. Le, Oriol Vinyals
Google Brain
[ndjaitly,qvl,vinyals]@google.com

Abstract

We present Listen, Attend and Spell (LAS), a neural network that learns to transcribe speech utterances to characters. Unlike traditional DNN-HMM models, this model learns all the components of a speech recognizer jointly. Our system has

This was the first time an ICASSP was held in China.



General Chairs Zhi Ding (Left), Tom Luo (Bottom right) and Wenjun Zhang (Top).

Most Downloaded Paper

LISTEN, ATTEND AND SPELL: A NEURAL NETWORK FOR LARGE VOCABULARY CONVERSATIONAL SPEECH RECOGNITION

William Chan
Carnegie Mellon University

Navdeep Jaitly, Quoc Le, Oriol Vinyals
Google Brain

ABSTRACT

We Listen, Attend and Spell (LAS), a neural speech recognizer that transcribes speech utterances directly to characters without pronunciation models, HMMs or other components of traditional speech recognizers. In LAS, the neural network architecture subsumes the acoustic, pronunciation and language models making it

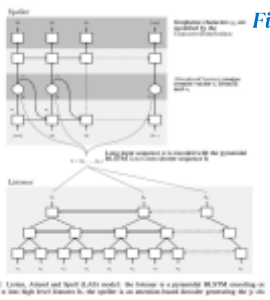


Figure from Chan et al, ICASSP 2016

Also in 2016:
Pokémon Go made people walk miles to catch a virtual Charizard.
China's Giant Pandas are no longer considered endangered

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2017 - NEW ORLEANS, USA
MARCH 5 - 9

Most Cited Paper

AUDIO SET: AN ONTOLOGY AND HUMAN-LABELED DATASET FOR AUDIO EVENTS

Jort F. Gemmeke, Daniel P. W. Ellis, Dylan Freedman, Aren Jansen, Wade Lawrence, R. Channing Moore, Manoj Plakal, Marvin Ritter

Google, Inc., Mountain View, CA, and New York, NY, USA

ABSTRACT

Audio event recognition, the human-like ability to identify and relate sounds from audio, is a nascent problem in machine perception. Comparable problems such as object detection in images have reaped enormous benefits from comprehensive datasets – principally ImageNet. This paper describes the creation of *Audio Set*, a large-...

among 50 environmental sounds. LeMaitre and Heller [8] proposed a taxonomy of sound events distinguishing objects and actions, and used identification time and priming effects to show that listeners find a "middle range" of abstraction most natural.

Engineering-oriented taxonomies and datasets began with Gaver [9] who used perceptual factors to guide the design of synthetic sound effects conveying different actions and materials (tapping,...

Most Downloaded Paper

CNN ARCHITECTURES FOR LARGE-SCALE AUDIO CLASSIFICATION

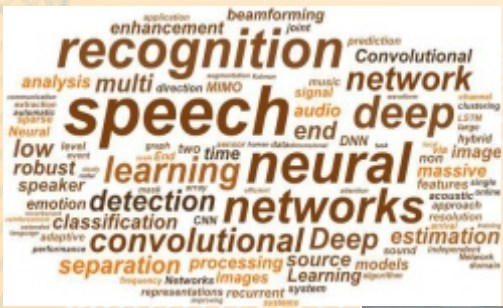
Shawn Hershey, Sourish Chaudhuri, Daniel P. W. Ellis, Jort F. Gemmeke, Aren Jansen, R. Channing Moore, Manoj Plakal, Devin Platt, Rif A. Saurous, Bryan Seybold, Malcolm Slaney, Ron J. Weiss, Kevin Wilson

Google, Inc., New York, NY, and Mountain View, CA, USA

ABSTRACT

Convolutional Neural Networks (CNNs) have proven very effective in image classification and show promise for audio. We use various CNN architectures to classify the soundtracks of a dataset of 70M training videos (5.24 million hours) with 30,871 video-level labels. We examine fully connected Deep Neural Networks (DNNs), AlexNet [1], VGG [2], Inception [3], and ResNet [4]. We investigate ...

Eghbal-Zadeh et al. [19] recently won the DCASE 2016 Acoustic Scene Classification (ASC) task, which, like soundtrack classification, involves assigning a single label to an audio clip containing many events. Their system used spectrogram features feeding a VGG classifier, similar to one of the classifiers in our work. This paper, however, compares the performance of several different architectures. To our knowledge, we are the first to publish results of Inception and ResNet networks applied to audio.



General Chair Magdy Bayoumi also organized ICIP in Cairo in 2011.

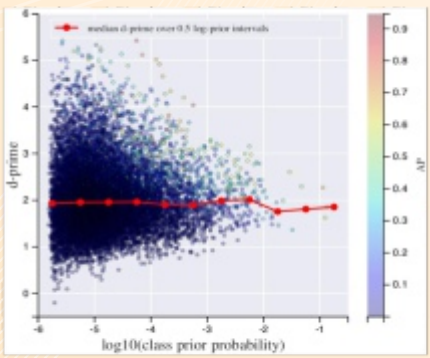


Fig. 1: Scatter plot of ResNet-50's per-class d-prime versus log prior probability. Each point is a separate class from a random 20% subset of the 30K set. Color reflects the class AP.

Figure from Gemmeke, ICASSP 2017

Also in 2017:

On March 30th, SpaceX launched a previously flown Falcon 9 rocket, the first time a payload carrying orbital class rocket had re-flown.

NASA Discovered Seven Earth-Like Planets!

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2018 - CALGARY, CANADA
APRIL 15 - 19

Most Cited and Downloaded Paper

X-VECTORS: ROBUST DNN EMBEDDINGS FOR SPEAKER RECOGNITION

David Snyder, Daniel Garcia-Romero, Gregory Sell, Daniel Povey, Sanjeev Khudanpur

Center for Language and Speech Processing & Human Language Technology Center of Excellence
The Johns Hopkins University, Baltimore, MD 21218, USA

ABSTRACT

In this paper, we use data augmentation to improve performance of deep neural network (DNN) embeddings for speaker recognition. The DNN, which is trained to discriminate between speakers, maps

Alternatively, neural networks can be directly optimized to discriminate between speakers. This has potential to produce powerful, compact systems [13], that only require speaker labels to train. In early systems, neural networks are trained to separate speakers, and frame-level representations are extracted from the network and

Best Industry Paper Award: Prabhavalkar, Sainath, Nguyen, Chen, Chiu, and Kannan, "Minimum Word Error Rate Training for Attention-Based Sequence-to-Sequence Models."

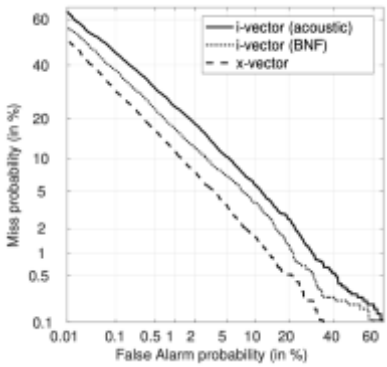


Figure from Huang, ICASSP 2020

ICASSP 2018 was originally awarded to Seoul, Korea. However, because of concerns about geopolitical tension in the region, the conference was moved to Calgary, Canada. The organizing committee remained the same.

General Chairs Monty Hayes (left) and Hanseok Ko (right).

Also in 2018:
A Tesla Roadster was launched into space by SpaceX.
A 125-million-year-old dinosaur fossil was found with perfectly preserved feathers!

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2019 - BRIGHTON, UK
MAY 12 - 17

Most Cited Paper

EXPOSING DEEP FAKES USING INCONSISTENT HEAD POSES

Xin Yang, Yuezun Li* and Siwei Lyu*

University at Albany, State University of New York, USA

ABSTRACT

this paper, we propose a new method to expose AI-erated fake face images or videos (commonly known he *Deep Fakes*). Our method is based on the observations Deep Fakes are created by splicing synthesized face re-i into the original image, and in doing so, introducing rs that can be revealed when 3D head poses are estimated as the neural network synthesis algorithm does not guara the original face and the synthesized face to have consis facial landmarks, as shown in Fig. 1.

The errors in landmark locations may not be visible rectly to human eyes, but can be revealed from head po (i.e., head orientation and position) estimated from 2D la marks in the real and faked parts of the face. Specifically, compare head poses estimated using all facial landmarks

Most Downloaded Paper

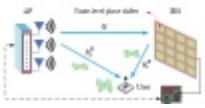
BEAMFORMING OPTIMIZATION FOR INTELLIGENT REFLECTING SURFACE WITH DISCRETE PHASE SHIFTS

Qingqing Wu and Rui Zhang

Department of Electrical and Computer Engineering, National University of Singapore, Singapore
Email: {elewuqq, elezhang}@nus.edu.sg

ABSTRACT

Intelligent reflecting surface (IRS) is a cost-effective solution for achieving high spectrum and energy efficiency in future wireless communications systems by leveraging massive low-cost passive elements that are able to reflect the signals with adjustable phase shifts. Prior works on IRS mostly consider continuous phase shifts at each reflecting element, which however, is practically difficult to realize due to the hardware limitation. In contrast, we study in this paper an



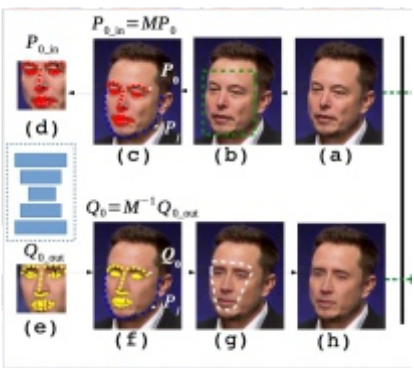


Fig. 1. *Overview of Deep Fake work-flow (Left) and our method (Right). It face in the image.* **(c)** *Detected 2D facial landmarks.* **(d)** *Cropped face in (a*

Figure from Yang, et al., ICASSP 2019

The theme of ICASSP 2019 was, "Signal Processing: Empowering Science and Technology for Humankind."

Also in 2019:
Scientists released the first-ever image of a black hole.
Water was detected on an exoplanet (K2-18b),
raising hopes of finding alien life.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2020 - (Virtual Event)
MAY 4 - 8

Most Cited and Downloaded Paper

UNET 3+: A FULL-SCALE CONNECTED UNET FOR MEDICAL IMAGE SEGMENTATION

Huimin Huang[1], *Lanfen Lin[1], Ruofeng Tong[1], *Hongjie Hu[2], Qiaowei Zhang[2], Yutaro Iwamoto[3], Xianhua Han[3], *Yen-Wei Chen[1,4], Xiaoi Wu[1]

[1]College of Computer Science and Technology, Zhejiang University, China
[2]Department of Radiology, Sir Run Run Shaw Hospital, China
[3]College of Information Science and Engineering, Ritsumeikan University, Japan
[4]Research Center for Healthcare Data Science, Zhejiang Lab, Hangzhou, China
*Corresponding Author: Lanfen Lin (lanfen@zju.edu.cn); Hongjie Hu (hongjiehu@zju.edu.cn); Yen-Wei Chen (chen@is.ritsumei.ac.jp)

ABSTRACT

Recently, a growing interest has been seen in deep learning-based semantic segmentation. UNet, which is one of deep

connections, aiming at reducing the semantic gap between the encoder and decoder. Despite achieving good performance, this type of approach is still incapable of exploring sufficient information from full scales

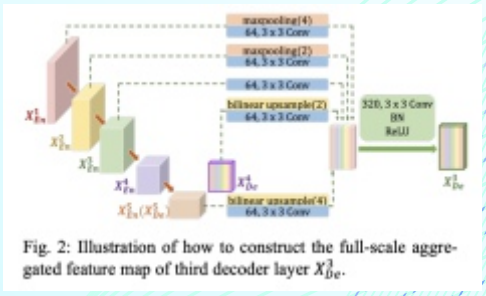


Fig. 2: Illustration of how to construct the full-scale aggregated feature map of third decoder layer X_{De}^3.

Figure from Huang, ICASSP 2020

ICASSP 2020 was originally allocated to Barcelona, Spain. The Society had to change the conference format to an online conference because of COVID-19. Registration was free – Over 16000 people participated.

General Chairs Ana Peres Neira (left) and Xavier Mestre (right)

ICASSP 2020 may have been the first to give a best paper award. Student paper awards and Industry paper awards was in place already,

Also in 2020:
In 2020, COVID-19 lockdowns began, and suddenly, baking banana bread became a hobby. The fastest-ever vaccine development happened, with multiple COVID-19 vaccines rolled out.

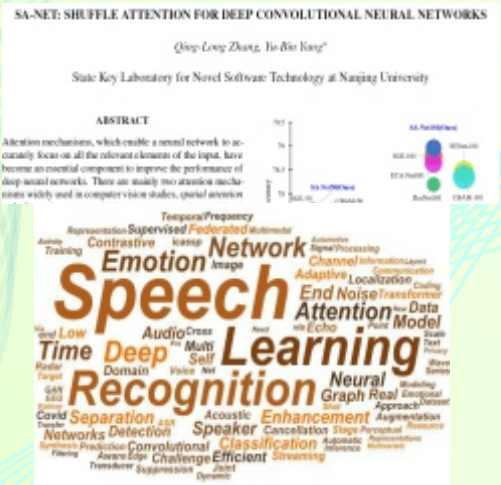
2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2021 (Virtual Event)
June 6 - 11

Most Cited Paper



Most Downloaded Paper

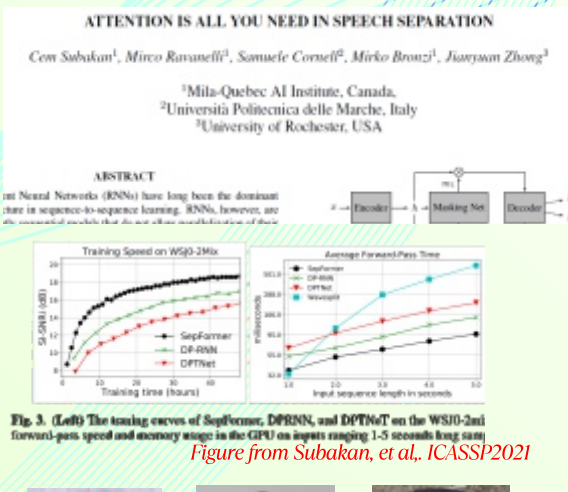


Fig. 3. (Left) The training curves of SepFormer, DPRNN, and DPTNeT on the WSJ0-2mi
forward-pass speed and memory usage in the GPU on inputs ranging 1-5 seconds long sam

Figure from Subakan, et al,. ICASSP2021

ICASSP 2021 was a fully virtual conference because of the COVID-19 pandemic. 1732 papers were published and 2184 delegates registered

Microsoft researchers organized the Acoustic Echo Cancellation and Deep Noise Suppression Challenges to stimulate research in the area of speech enhancement and its applications.



General Chairs Dimitry Androutsos, Kostas Plataniotis and Steven Zhang.

Also in 2021:
NASA flew a small helicopter on Mars for the first time!
El Salvador Made Bitcoin Legal Tender!

Celebrating Signal Processing

ICASSP 2022 – SINGAPORE
MAY 7 - 13

icassp 2022 Singapore

Most Cited Paper

MIXED TRANSFORMER U-NET FOR MEDICAL IMAGE SEGMENTATION

Mengxi Wang[1] Shixia Xie[1] *Lanjin Lin[2] Yasuro Iwamoto[3] Xian-Hua Han[3] *Yen-Wei Chen[3,4,5] Ruofeng Tong[1,4]

[1]College of Computer Science and Technology, Zhejiang University, China
[2]College of Information Science and Engineering, Ritsumeikan University, Japan
[3]Artificial Intelligence Research Center, Yamaguchi University, Japan
[4]Research Center for Healthcare Data Science, Zhejiang Lab, China

ABSTRACT

Though U-Net has achieved tremendous success in medical image segmentation tasks, it lacks the ability to explicitly model long-range dependencies. Therefore, Vision Transformers have emerged as alternative segmentation structures

Recently, many works try to solve this problem by using Transformer encoder [5, 6, 7]. Transformer is an attention-based model originally designed for sequence-to-sequence prediction [8]. Self-Attention (SA) is the key component of Transformer. It can model correlations among all the input

Most Downloaded Paper

MULTI-SCALE TEMPORAL FREQUENCY CONVOLUTIONAL NETWORK WITH AXIAL ATTENTION FOR SPEECH ENHANCEMENT

Guochang Zhang, Libiao Yu, Chunlang Wang, Jianqiang Wei

Department of Speech Technology, Baidu Inc, Beijing, 100085, China

ABSTRACT

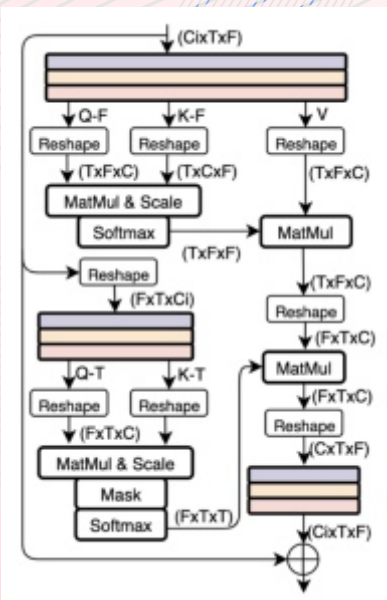
Speech quality is often degraded by acoustic echoes, background noise, and reverberation. In this paper, we propose a system consisting of deep learning and signal processing to simultaneously suppress echoes, noise, and reverberation. For the deep learning, we design a novel speech dense-prediction

[6] with PNLMS adaptive filter [7]. For the DNN part, we propose a novel backbone for speech dense-prediction called multi-scale temporal frequency convolutional network with axial self-attention (MTFAA-Net). In this work, our contributions include:

General Chair Haizhou Li (left) and Co-General Chair Sadaoki Furui (right)

Best Industry Paper Award: "VARARRAY: Array-Geometry-Agnostic Continuous Speech Separation" by Takuya Yoshioka, et al.

ICASSP 2022 organization was arguably among the most complex undertaking in the 50-year history of ICASSPs. The world was just coming out of the COVID-19 pandemic, and the organizing committee implemented an in-place event and a hybrid event in Singapore along with a satellite event in China.



Also in 2022:
Axiom-1 became the first privately funded crewed mission to dock with the International Space Station
Rut Linnéa Ingegärd Larsson from Sweden became the oldest person to skydive at 103 years old.

2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2023 - RHODES ISLAND, GREECE
JUNE 4 - 10



Most Cited Paper

Efficient Multi-Scale Attention Module with Cross-Spatial Learning

Daliang Ouyang, Su He, Guozhong Zhang, Mingzhu Luo, Huaiyong Guo, Jian Zhan, Zhijie Huang

AEROSPACE SCIENCE & INDUSTRY SHENZHEN (GROUP) CO., LTD., Shenzhen, China

Abstract

Most Downloaded Paper

HIGH-SPEED DRONE DETECTION BASED ON YOLO-V8

Ao-Hyeo Kim, Nuwoo Kim, Chee Sun Won

Department of Electronics Electrical Engineering, Dongguk University, Seoul, Korea

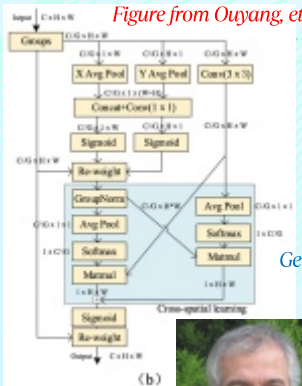
ABSTRACT

The theme of ICASSP 2023 was "Signal Processing in the AI era," promoting the synergy between signal processing and machine learning.

Best Paper Award: N. Ito and M. Sugiyama, "Audio signal enhancement with learning from positive and unlabelled data."

Treble's cloud-based sound system secured the Best Startup Pitch award.

Figure from Ouyang, et al., ICASSP2023.



General Chair Petros Maragos

General Co-chairs Kostas Berberidis (left) and Petros Boufounos (right)

Also in 2023:

India became the first country to land near the Moon's south pole with the Chandrayaan-3 mission. NASA's OSIRIS-Rex mission brought a sample of asteroid Bennu to earth on Sep. 24th.

2025 icassp HYDERABAD, INDIA

Celebrating Signal Processing

50

IEEE Signal Processing Society

ICASSP 2024 - SEOUL, SOUTH KOREA
APRIL 15 - 19

ICASSP 2024 KOREA



Most Cited Paper

BUILDING LANE-LEVEL MAPS FROM AERIAL IMAGES

Xuexi Xie[1*] Xiaochao Pan[2*] Tong Wu[1] Xiaofeng Zhang[3]

[1] University of Washington [2] Taiyuan University of Technology
[3] Shanghai Jiao Tong University

ABSTRACT

Detecting lane lines from sensors is becoming an increasingly significant part of autonomous driving systems. However, less development has been made on high-definition lane-level mapping based on aerial images, which could automatically build and update offline maps for auto-driving systems. To this end, our work focuses on extracting fine-level detailed lane lines together with their topological structures. This task is challenging since it requires large amounts of data covering different lane types, terrain and regions. In this paper, we introduce a new benchmark for lane detection using aerial images for planning [13, 12], or building a coarse-level map of the road network or curb map [14, 16]. However, building detailed lane maps offline from top-view aerial image data is feasible but not under-studied, and the accuracy and effectiveness of such high-definition mapping are still significantly underperformed. There is a lack of a benchmark and baseline for offline lane recognition from aerial images task.

Previous lane detection datasets such as SDLane [20], ONCE-3DLanes [19], and OpenLane [17] are all designed for lane detection from a vehicle perspective. In this paper, we introduce a new benchmark for lane detection using

Downloaded Paper

SMALL-FOOTPRINT AUTOMATIC SPEECH RECOGNITION SYSTEM USING TWO-STAGE TRANSFER LEARNING BASED SYMMETRIZED TERNARY WEIGHT NETWORK

Xiaobao Zhang, Wei Kou, Chenyle Xia, Hao Cai, Bo Liu*

School of Integrated Circuits, Southeast University, Nanjing, China

ABSTRACT

Traditional automatic speech recognition (ASR) models face challenges when deployed on edge devices due to their high computational requirements and storage demands. To address this issue, we present a novel ASR system specifically designed for edge applications, encompassing both keyword spotting (KWS) and speaker verification (SV) functionalities with on-chip learning for speaker registration. Our proposed system employs a compact model trained using a two-stage transfer learning method for on-the-fly small-sample speaker registration. In the proposed model, sparsity-controllable weights are symmetrically ternary quantized to further exploit data sparsity. Additionally, we introduce a Huffman-coding based weight lossy compression method to achieve efficient storage compaction.

Compared with BWN, in ternary weight network (TWN) [12, 13, 14], weights are quantized to 2 bits {0, +1, -1}, which brings about an increase in complexity and sparsity. TWN can also eliminate multiplication operations, while the storage of weights is doubled. There are prior works to exploit the symmetry of weight distribution in TWN for weight compression and reuse of intermediate computation results [15, 16, 17]. In this paper, an ASR system which uses a compact model for both KWS and SV is proposed. The main contributions are:

- A two-stage transfer learning method is proposed for on-the-fly speaker registration on edge devices. By employing this method, users can conveniently register speakers' voices on

Approximately 4400 people participated in ICASSP 2024. This is by far the largest number of attendees at an ICASSP for a fully onsite conference.

The original Psy (right) performed Gangnam Style during the welcome reception at ICASSP2024. Performing with him is Hanseok Ko, Conference General Chair.

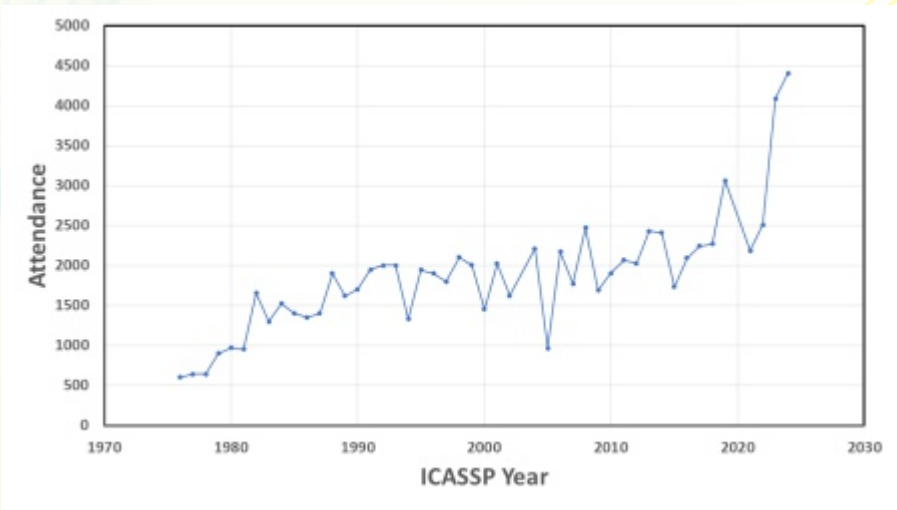
Also in 2024:
Three moons in the solar system were discovered, one around Uranus and two around Neptune.
The European Union passed the Artificial Intelligence Act, a comprehensive regulatory framework for AI.

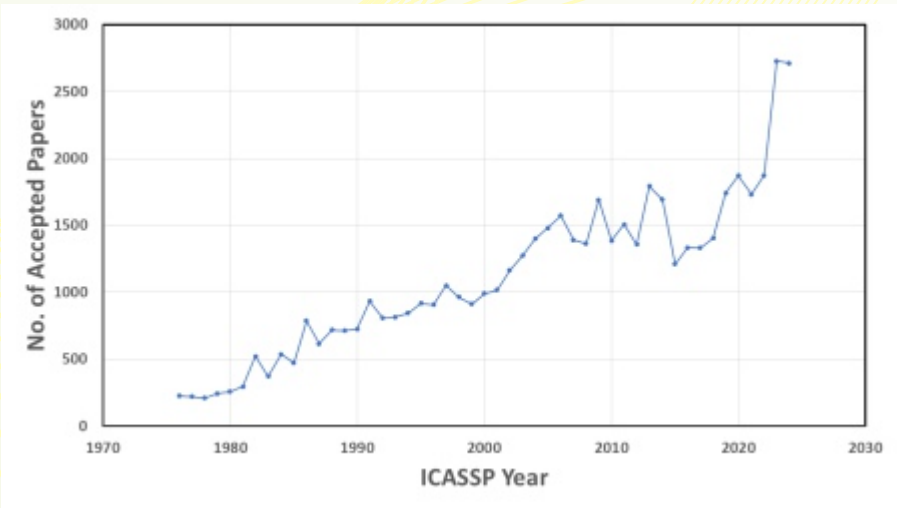
2025 icassp HYDERABAD, INDIA
Celebrating Signal Processing

50

IEEE Signal Processing Society

A QUICK SUMMARY





WHAT NEXT?

What is the next big thing?

Can we sustain the growth?

Where do we go next?
Africa?
South America?

Many thanks to:

Sahana Gowda
Kiran Mathews
Om Kathalkar
Baly Naganjani
Saneesh Cleatus