ELYSIUMPRO
INSPIRING THE LEADING EDGE TECHNOLOGIES

ELYSIUMPRO

A UNIT OF ELYSIUM GROUPS

FINAL YEAR PROJECTS

SIGNAL PROCESSING 2017-2018

TITLES WITH ABSTRACTS

ELYSIUMPRO

A UNIT OF ELYSIUM GROUPS

ELYSIUMPRO

INSPIRING THE LEADING EDGE TECHNOLOGIES

www.elysiumpro.in

f / elysiumtechnologiespvtltd / elysiumind in / elysium-technologies-pvt

#227, Church Road, Anna Nagar, Madurai, Tamil Nadu, India

CALL US @ 994-479-3398, (0452) 4390702

ETPL SP - 001	Improved facial expression recognition using graph signal processing

A novel method of facial expression recognition using graph signal processing (GSP) is proposed. Here the dimensionality of the feature vector of the facial expression based on the histogram of oriented gradients and discrete wavelet transform is reduced using the GSP and then applied to a classifier. The proposed scheme is compared with the existing methods on the JAFFE dataset in terms of the accuracy and the length of the feature vector. It is seen from the simulation results that the proposed scheme provides better accuracy than the existing methods.

ETPL SP - 002	Sizing Energy Storage to Mitigate Wind Power Forecast Error Impacts by Signal Processing Techniques

This paper proposes to use discrete Fourier transform (DFT) and discrete wavelet transform (DWT) methods to schedule grid-scale energy storage systems to mitigate wind power forecast error impacts while considering energy storage properties. This is accomplished by decomposing the wind forecast error signal to different time-varying periodic components to schedule sodium sulfur (NaS) batteries, compressed air energy storage (CAES), and conventional generators. The advantage of signal processing techniques is that the resultant decomposed components are appropriate for cycling of each energy storage technology. It is also beneficial for conventional generators, which are more efficient to operate close to rated capacity. The tradeoff between installing more energy storage units and decreasing the wind spillage, back-up energy, and the standard deviation of residual forecast error signal is analyzed. The NaS battery life cycle analysis and CAES contribution on increasing NaS battery lifetime are studied. The impact of considering the frequency bias constant to allow small frequency deviations is also investigated. To showcase the applicability of the proposed approach, a simulation case study based on a real-world 5-min interval wind data from Bonneville Power Administration (BPA) in 2013 is presented.

ELYSIUMPRO

A UNIT OF ELYSIUM GROUPS

ETPL SP - 003	Detection of Paroxysms in Long-Term, Single-Channel EEG-Monitoring of Patients with Typical Absence Seizures

Absence seizures are associated with generalized 2.5-5 Hz spike-wave discharges in the electroencephalogram (EEG). Rarely are patients, parents, or physicians aware of the duration or incidence of seizures. Six patients were monitored with a portable EEG-device over four times 24 h to evaluate how easily outpatients are monitored and how well an automatic seizure detection algorithm can identify the absences. Based on patient-specific modeling, we achieved a sensitivity of 98.4% with only 0.23 false detections per hour. This yields a clinically satisfying performance with a positive predictive value of 87.1%. Portable EEG-recorders identifying paroxystic events in epilepsy outpatients are a promising tool for patients and physicians dealing with absence epilepsy. Albeit the small size of the EEG-device, some children still complained about the obtrusive nature of the device. We aim at developing less obtrusive though still very efficient devices, e.g., hidden in the ear canal or below the skin.

ETPL SP - 004	Wireless Power Transfer Strategies for Implantable Bioelectronics: Methodological Review

Neural implants have emerged over the last decade as highly effective solutions for the treatment of dysfunctions and disorders of the nervous system. These implants establish a direct, often bidirectional, interface to the nervous system, both sensing neural signals and providing therapeutic treatments. As a result of the technological progress and successful clinical demonstrations, completely implantable solutions have become a reality and are now commercially available for the treatment of various functional disorders. Central to this development is the wireless power transfer (WPT) that has enabled implantable medical devices (IMDs) to function for extended durations in mobile subjects. In this review, we present the theory, link design, and challenges, along with their probable solutions for the traditional near-field resonant inductively coupled WPT, capacitively coupled short ranged WPT, and more recently developed ultrasonic, mid-field and far-field coupled WPT technologies for implantable applications. A comparison of various power transfer methods, based on their power budgets and WPT range follows. Power requirements of specific implants like cochlear, retinal, cortical and peripheral are also considered and currently available IMD solutions are discussed. Patient's safety concerns with respect to electrical, biological, physical, electromagnetic interference and cyber security from an implanted neurotech device are also explored in this review. Finally, we discuss and anticipate future developments that will enhance the capabilities of current-day wirelessly powered implants and make them more efficient and integrable with other electronic components in IMDs.

ETPL SP - 005	Epileptic Focus Localization Using Discrete Wavelet Transform Based on Interictal Intracranial EEG

Over the past decade, with the development of machine learning, discrete wavelet transform (DWT) has been widely used in computer-aided epileptic electroencephalography (EEG) signal analysis as a powerful time-frequency tool. But some important problems have not yet been benefitted from DWT, including epileptic focus localization, a key task in epilepsy diagnosis and treatment. Additionally, the parameters and settings for DWT are chosen empirically or arbitrarily in previous work. In this work, we propose a framework to use DWT and support vector machine (SVM) for epileptic focus localization problem based on EEG. To provide a guideline in selecting the best settings for DWT, we decompose the EEG segments in seven commonly used wavelet families to their maximum theoretical levels. The wavelet and its level of decomposition providing the highest accuracy in each wavelet family are then used in a grid search for obtaining the optimal frequency bands and wavelet coefficient features. Our approach achieves promising performance on two widely-recognized intrancranial EEG datasets that are also seizure-free, with an accuracy of 83.07% on the Bern-Barcelona dataset and an accuracy of 88.00% on the UBonn dataset. Compared with existing DWT-based approaches in epileptic EEG analysis, the proposed approach leads to more accurate and robust results. A guideline for DWT parameter setting is provided at the end of the paper.

ETPL SP - 006	Adaptive processing of image using DWT and FFT OFDM in AWGN and Rayleigh channel

Wireless communication has a very important role in the day to day life for everyone. The wireless technology is growing rapidly and there is a growing demand of high performance, capacity and larger bit rate wireless communication systems which cope wireless communication services such as high speed data, video and voice signals. Multicarrier modulation scheme like OFDM provides an efficient solution to overcome this problem. In this paper a gray scale image processing is done using a LMS algorithm with wavelet based OFDM system using the QPSK modulation scheme in AWGN and Rayleigh channel in SISO environment and results are compared with the conventional adaptive FFT based OFDM system. We reconstruct our transmitted signal at receiver; in both systems by minimizing the error by adaptive filter but the computational complexity of FFT based system is more as compared to DWT based system. Results are compared in term of SNR vs BER which shows that adaptive DWT based OFDM system perform better as compared to the conventional adaptive FFT OFDM system.

ETPL SP - 007	Epileptic seizure classification using statistical features of EEG signal

Epilepsy detection is enough time consuming and requires thorough observation to determine epilepsy type and locate the responsible area of the cerebral cortex. This paper proposes an effortless epilepsy classification method for straightforward epilepsy detection and investigates the classification accuracy of multiclass EEG signal during epilepsy. To accomplish our research work we exploit DWT MATLAB toolbox to obtain responsible features to accumulate feature vectors. Afterwards feature vectors are given in the input layer of the NN classifiers to differentiate normal, interictal and ictal EEG periods. Accuracy rate is calculated based on the confusion matrix. Proposed method can be utilized to monitor and detect epilepsy type incorporating with alarm system.

ETPL SP - 008	Spoofing Speech Detection Using Modified Relative Phase Information

The detection of human and spoofing (synthetic or converted) speech has started to receive an increasing amount of attention. In this paper, modified relative phase (MRP) information extracted from a Fourier spectrum is proposed for spoofing speech detection. Because original phase information is almost entirely lost in spoofing speech using current synthesis or conversion techniques, some phase information extraction methods, such as the modified group delay feature and cosine phase feature, have been shown to be effective for detecting human speech and spoofing speech. However, existing phase information-based features cannot obtain very high spoofing speech detection performance because they cannot extract precise phase information from speech. Relative phase (RP) information, which extracts phase information precisely, has been shown to be effective for speaker recognition. In this paper, RP information is applied to spoofing speech detection, and it is expected to achieve better spoofing detection performance. Furthermore, two modified processing techniques of the original RP, that is, pseudo pitch synchronization and linear discriminant analysis based full-band RP extraction, are proposed in this paper. In this study, MRP information is also combined with the Mel-frequency cepstral coefficient (MFCC) and modified group delay. The proposed method was evaluated using the ASVspoof 2015: Automatic Speaker Verification Spoofing and Countermeasures Challenge dataset. The results show that the proposed MRP information significantly outperforms the MFCC, modified group delay, and other phase information based features. For the development dataset, the equal error rate (EER) was reduced from 1.883% of the MFCC, 0.567% of the modified group delay to 0.013% of the MRP. By combining the RP with the MFCC and modified group delay, the EER was reduced to 0.003%. For the evaluation dataset, the MRP obtained much better performance than the magnitude-based feature and other phase-based features, except for S10 spoofing speech.

ETPL SP - 009	Extraction of Fundamental Frequency from Degraded Speech Using Temporal Envelopes at High SNR Frequencies

In this paper we propose a method for extracting the fundamental frequency (f_o) from degraded speech signals using single frequency filtering (SFF) approach. The SFF of frequency-shifted speech signal gives high signal-to-noise ratio (SNR) segments at some frequencies and hence the SFF approach can be exploited for f_o extraction using autocorrelation function of those segments. Since the f_o is computed from the envelope of a single frequency component of the signal, the vocal tract resonances do not affect the f_o extraction. The use of the high SNR frequency component in a given segment helps in overcoming the effects of degradations in the speech signal, without explicitly estimating the characteristics of noise. The proposed method of f_o extraction is shown to give better performance for several types of real and simulated degradations, in comparison with some of the methods reported recently in the literature.

ETPL SP - 010	Combining Binaural and Cortical Features for Robust Speech Recognition

The segregation of concurrent speakers and other sound sources is an important ability of the human auditory system, but is missing in most current systems for automatic speech recognition (ASR), resulting in a large gap between human and machine performance. This study combines processing related to peripheral and cortical stages of the auditory pathway: A physiologically motivated binaural model estimates the positions of moving speakers to enhance the desired speech signal. Second, signals are converted to spectro-temporal Gabor features that resemble cortical speech representations and which have been shown to improve ASR in noisy conditions. Spectro-temporal Gabor features improve recognition results in all acoustic conditions under consideration compared with Mel-frequency cepstral coefficients. Binaural processing results in lower word error rates (WERs) in acoustic scenes with a concurrent speaker, whereas monaural processing should be preferred in the presence of a stationary masking noise. In-depth analysis of binaural processing identifies crucial processing steps such as localization of sound sources and estimation of the beamformer's noise coherence matrix, and shows how much each processing step affects the recognition performance in acoustic conditions with different complexity.

ETPL SP - 011	Continuous Wavelet Transform-Based Frequency Dispersion Compensation Method for Electromagnetic Time-Reversal Imaging

The invariance of wave equations in lossless media allows the time reversal (TR) technique to spatiotemporally refocus back-propagated signals in a given ultrawideband imaging scenario. However, the existence of dispersion and loss in the propagation medium breaks this invariance and the resultant TR focusing exhibits frequency and propagation duration dependent degradation. We propose an algorithm based on the continuous wavelet transform that tackles this degradation to improve focusing resolution under such conditions. The developed algorithm has been successfully applied to the scenario for localization of lung cancer.

ETPL SP – 012	An Algorithm for Real-Time Pulse Waveform Segmentation and Artifact Detection in Photo plethysmograms

Photoplethysmography has been used in a wide range of medical devices for measuring oxygen saturation, cardiac output, assessing autonomic function, and detecting peripheral vascular disease. Artifacts can render the photoplethysmogram (PPG) useless. Thus, algorithms capable of identifying artifacts are critically important. However, the published PPG algorithms are limited in algorithm and study design. Therefore, the authors developed a novel embedded algorithm for real-time pulse waveform (PWF) segmentation and artifact detection based on a contour analysis in the time domain. This paper provides an overview about PWF and artifact classifications, presents the developed PWF analysis, and demonstrates the implementation on a 32-bit ARM core microcontroller. The PWF analysis was validated with data records from 63 subjects acquired in a sleep laboratory, ergometry laboratory, and intensive care unit in equal parts. The output of the algorithm was compared with harmonized experts' annotations of the PPG with a total duration of 31.5 h. The algorithm achieved a beat-to-beat comparison sensitivity of 99.6%, specificity of 90.5%, precision of 98.5%, and accuracy of 98.3%. The interrater agreement expressed as Cohen's kappa coefficient was 0.927 and as F-measure was 0.990. In conclusion, the PWF analysis seems to be a suitable method for PPG signal quality determination, real-time annotation, data compression, and calculation of additional pulse wave metrics such as amplitude, duration, and rise time.

ETPL SP - 013	Improved Computational Efficiency of Locally Low Rank MRI Reconstruction Using Iterative Random Patch Adjustments

This paper presents and analyzes an alternative formulation of the locally low-rank (LLR) regularization framework for magnetic resonance image (MRI) reconstruction. Generally, LLR-based MRI reconstruction techniques operate by dividing the underlying image into a collection of matrices formed from image patches. Each of these matrices is assumed to have low rank due to the inherent correlations among the data, whether along the coil, temporal, or multi-contrast dimensions. The LLR regularization has been successful for various MRI applications, such as parallel imaging and accelerated quantitative parameter mapping. However, a major limitation of most conventional implementations of the LLR regularization is the use of multiple sets of overlapping patches. Although the use of overlapping patches leads to effective shift-invariance, it also results in high-computational load, which limits the practical utility of the LLR regularization for MRI. To circumvent this problem, alternative LLR-based algorithms instead shift a single set of non-overlapping patches at each iteration, thereby achieving shift-invariance and avoiding block artifacts. A novel contribution of this paper is to provide a mathematical framework and justification of LLR regularization with iterative random patch adjustments (LLR-IRPA). This method is compared with a state-of-the-art LLR regularization algorithm based on overlapping patches, and it is shown experimentally that results are similar but with the advantage of much reduced computational load. We also present theoretical results demonstrating the effective shift invariance of the LLR-IRPA approach, and we show reconstruction examples and comparisons in both retrospectively and prospectively undersampled MRI acquisitions.

ETPL SP - 014	Dual-Tree Complex Wavelet Transform-Based Control Algorithm for Power Quality Improvement in a Distribution System

This paper presents a dual tree-complex wavelet transform-based control algorithm for a distribution static compensator (DSTATCOM) to improve the power quality (PQ) in a distribution system. PQ disturbances like harmonics and starting as well as ending of unbalancing in all phase load currents are also assessed simultaneously. The distorted load current of each phase is decomposed into various frequency levels with this technique to extract respective line frequency component for the estimation of the reference active power component. The deviations of respective sensed load currents from these estimated reference components are used to generate the reference currents for the control of voltage source converter used as DSTATCOM. Simulated performance of DSTATCOM is presented at varying load conditions. The proposed control algorithm is also validated experimentally on a laboratory prototype of DSTATCOM. The total harmonic distortion (THD) of supply current is obtained below 5 percent with unity power factor under different load conditions which is satisfactory as per IEEE-519 standard.

ETPL SP - 015	Signal Processing for Temporal Spectrum Sharing in a Multi-radar Environment

Regulators, aware of the significant underutilization of spectrum reserved for radar operation, are starting to open these bands for sharing with commercial services. In this paper, we provide the signal processing techniques necessary to apply temporal sharing to reduce radar exclusion zones and increase spectral efficiency. Our approach directly extends to the fairly common scenario of multiple radars operating at relatively close distance in the same frequency and allows a secondary user to transmit without exceeding a stipulated level of interference at any radar. We require only that radars behave periodically; our secondary users apply adaptive sensing to track radar behavior in real-time without a priori information. To accomplish this, we introduce a pulse deinterleaving mechanism to separate multiple radar emissions in real-time, with no batch or offline processing. We show that our approach to temporal sharing is applicable to static or low mobility sharing scenarios, where the interference channel displays quasi-periodic features.

ETPL SP - 016	Tensor Decomposition for Signal Processing and Machine Learning

Tensors or multiway arrays are functions of three or more indices (i, j, k, . . .)-similar to matrices (two-way arrays), which are functions of two indices (r, c) for (row, column). Tensors have a rich history, stretching over almost a century, and touching upon numerous disciplines; but they have only recently become ubiquitous in signal and data analytics at the confluence of signal processing, statistics, data mining, and machine learning. This overview article aims to provide a good starting point for researchers and practitioners interested in learning about and working with tensors. As such, it focuses on fundamentals and motivation (using various application examples), aiming to strike an appropriate balance of breadth and depth that will enable someone having taken first graduate courses in matrix algebra and probability to get started doing research and/or developing tensor algorithms and software. Some background in applied optimization is useful but not strictly required. The material covered includes tensor rank and rank decomposition; basic tensor factorization models and their relationships and properties (including fairly good coverage of identifiability); broad coverage of algorithms ranging from alternating optimization to stochastic gradient; statistical performance analysis; and applications ranging from source separation to collaborative filtering, mixture and topic modeling, classification, and multilinear subspace learning.

ETPL SP - 017	Improving the efficiency of noise resistance processing of speech signal

The results of noise-resistant speech signal processing, with the use of linear adaptive filters. To enhance the filtration efficiency, the filtration is proposed to use a cascade signal. This approach improves the noise resistance of the signal. A comparative analysis of the performance of the developed algorithm with known methods of filtration. The effectiveness of the proposed filter is demonstrated modeling. Practical application of this method can significantly enhance the ability adequately decode the information contained in the experimental data.

ETPL SP - 018	Multichannel Signal Processing With Deep Neural Networks for Automatic Speech Recognition

Multichannel automatic speech recognition (ASR) systems commonly separate speech enhancement, including localization, beamforming, and postfiltering, from acoustic modeling. In this paper, we perform multichannel enhancement jointly with acoustic modeling in a deep neural network framework. Inspired by beamforming, which leverages differences in the fine time structure of the signal at different microphones to filter energy arriving from different directions, we explore modeling the raw time-domain waveform directly. We introduce a neural network architecture, which performs multichannel filtering in the first layer of the network, and show that this network learns to be robust to varying target speaker direction of arrival, performing as well as a model that is given oracle knowledge of the true target speaker direction. Next, we show how performance can be improved by factoring the first layer to separate the multichannel spatial filtering operation from a single channel filterbank which computes a frequency decomposition. We also introduce an adaptive variant, which updates the spatial filter coefficients at each time frame based on the previous inputs. Finally, we demonstrate that these approaches can be implemented more efficiently in the frequency domain. Overall, we find that such multichannel neural networks give a relative word error rate improvement of more than 5% compared to a traditional beamforming-based multichannel ASR system and more than 10% compared to a single channel waveform model.

ETPL SP - 019	An Energy Efficient ECG Signal Processor Detecting Cardiovascular Diseases on Smartphone

A novel disease diagnostic algorithm for ECG signal processing based on forward search is implemented in Application Specific Integrated Circuit (ASIC) for cardiovascular disease diagnosis on smartphone. An ASIC is fabricated using 130-nm CMOS low leakage process technology. The area of our PQRST ASIC is 1.21 mm^2. The energy dissipation of PQRST ASIC is 96 pJ with a supply voltage of 0.9 V. The outputs from the ASIC are fed to an Android application that generates diagnostic report and can be sent to a cardiologist via email. The ASIC and Android application are verified for the detection of bundle branch block, hypertrophy, arrhythmia and myocardial infarction using Physionet PTB diagnostic ECG database. The failed detection rate is 0.69%, 0.69%, 0.34% and 1.72% for bundle branch block, hypertrophy, arrhythmia and myocardial infarction respectively. The AV block is detected in all the three patients in the Physionet St. Petersburg arrhythmia database. Our proposed ASIC together with our Android application is the most suitable for an energy efficient wearable cardiovascular disease detection system.

ETPL SP - 020	An Algorithm for Real-Time Pulse Waveform Segmentation and Artifact Detection in Photo plethysmograms

Photoplethysmography has been used in a wide range of medical devices for measuring oxygen saturation, cardiac output, assessing autonomic function, and detecting peripheral vascular disease. Artifacts can render the photoplethysmogram (PPG) useless. Thus, algorithms capable of identifying artifacts are critically important. However, the published PPG algorithms are limited in algorithm and study design. Therefore, the authors developed a novel embedded algorithm for real-time pulse waveform (PWF) segmentation and artifact detection based on a contour analysis in the time domain. This paper provides an overview about PWF and artifact classifications, presents the developed PWF analysis, and demonstrates the implementation on a 32-bit ARM core microcontroller. The PWF analysis was validated with data records from 63 subjects acquired in a sleep laboratory, ergometry laboratory, and intensive care unit in equal parts. The output of the algorithm was compared with harmonized experts' annotations of the PPG with a total duration of 31.5 h. The algorithm achieved a beat-to-beat comparison sensitivity of 99.6%, specificity of 90.5%, precision of 98.5%, and accuracy of 98.3%. The interrater agreement expressed as Cohen's kappa coefficient was 0.927 and as F-measure was 0.990. In conclusion, the PWF analysis seems to be a suitable method for PPG signal quality determination, real-time annotation, data compression, and calculation of additional pulse wave metrics such as amplitude, duration, and rise time.

ETPL SP - 021	Improved Audio Scene Classification Based on Label-Tree Embeddings and Convolutional Neural Networks

In this paper, we present an efficient approach for audio scene classification. We aim at learning representations for scene examples by exploring the structure of their class labels. A category taxonomy is automatically learned by collectively optimizing a tree-structured clustering of the given labels into multiple metaclasses. A scene recording is then transformed into a label-tree embedding image. Elements of the image represent the likelihoods that the scene instance belongs to the metaclasses. We investigate classification with label-tree embedding features learned from different low-level features as well as their fusion. We show that the combination of multiple features is essential to obtain good performance. While averaging label-tree embedding images over time yields good performance, we argue that average pooling possesses an intrinsic shortcoming. We alternatively propose an improved classification scheme to bypass this limitation. We aim at automatically learning common templates that are useful for the classification task from these images using simple but tailored convolutional neural networks. The trained networks are then employed as a feature extractor that matches the learned templates across a label-tree embedding image and produce the maximum matching scores as features for classification. Since audio scenes exhibit rich content, template learning and matching on low-level features would be inefficient. With label-tree embedding features, we have quantized and reduced the low-level features into the likelihoods of the metaclasses, on which the template learning and matching are efficient. We study both training convolutional neural networks on stacked label-tree embedding images and multistream networks. Experimental results on the DCASE2016 and LITIS Rouen datasets demonstrate the efficiency of the proposed methods.

ETPL SP - 022	Comparative Validation of Polyp Detection Methods in Video Colonoscopy: Results From the MICCAI 2015 Endoscopic Vision Challenge

Colonoscopy is the gold standard for colon cancer screening though some polyps are still missed, thus preventing early disease detection and treatment. Several computational systems have been proposed to assist polyp detection during colonoscopy but so far without consistent evaluation. The lack of publicly available annotated databases has made it difficult to compare methods and to assess if they achieve performance levels acceptable for clinical use. The Automatic Polyp Detection sub-challenge, conducted as part of the Endoscopic Vision Challenge (http://endovis.grand-challenge.org) at the international conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) in 2015, was an effort to address this need. In this paper, we report the results of this comparative evaluation of polyp detection methods, as well as describe additional experiments to further explore differences between methods. We define performance metrics and provide evaluation databases that allow comparison of multiple methodologies. Results show that convolutional neural networks are the state of the art. Nevertheless, it is also demonstrated that combining different methodologies can lead to an improved overall performance.

ETPL SP - 023	A Two-Dimensional Framework of Multiple Kernel Subspace Learning for Recognizing Emotion in Speech.

As a highly active topic in computational paralinguistics, speech emotion recognition (SER) aims to explore ideal representations for emotional factors in speech. In order to improve the performance of SER, multiple kernel learning (MKL) dimensionality reduction has been utilized to obtain effective information for recognizing emotions. However, the solution of MKL usually provides only one nonnegative mapping direction for multiple kernels; this may lead to loss of valuable information. To address this issue, we propose a two-dimensional framework for multiple kernel subspace learning. This framework provides more linear combinations on the basis of MKL without nonnegative constraints, which preserves more information in the learning procedures. It also leverages both of MKL and two-dimensional subspace learning, combining them into a unified structure. To apply the framework to SER, we also propose an algorithm, namely generalised multiple kernel discriminant analysis (GMKDA), by employing discriminant embedding graphs in this framework. GMKDA takes advantage of the additional mapping directions for multiple kernels in the proposed framework. In order to evaluate the performance of the proposed algorithm a wide range of experiments is carried out on several key emotional corpora. These experimental results demonstrate that the proposed methods can achieve better performance compared with some conventional and subspace learning methods in dealing with SER.

ETPL SP - 024	Joint Segmentation of Retinal Layers and Focal Lesions in 3-D OCT Data of Topologically Disrupted Retinas

Accurate quantification of retinal structures in 3-D optical coherence tomography data of eyes with pathologies provides clinically relevant information. We present an approach to jointly segment retinal layers and lesions in eyes with topology-disrupting retinal diseases by a loosely coupled level set framework. In the new approach, lesions are modeled as an additional space-variant layer delineated by auxiliary interfaces. Furthermore, the segmentation of interfaces is steered by local differences in the signal between adjacent retinal layers, thereby allowing the approach to handle local intensity variations. The accuracy of the proposed method of both layer and lesion segmentation has been evaluated on eyes affected by central serous retinopathy and age-related macular degeneration. In addition, layer segmentation of the proposed approach was evaluated on eyes without topology-disrupting retinal diseases. Good agreement between the segmentation performed manually by a medical doctor and results obtained from the automatic segmentation was found for all data types. The mean unsigned error for all interfaces varied between 2.3 and 11.9 μm (0.6-3.1 pixels). Furthermore, lesion segmentation showed a Dice coefficient of 0.68 for drusen and 0.89 for fluid pockets. Overall, the method provides a flexible and accurate solution to jointly segment lesions and retinal layers.

ETPL SP - 025	Optoacoustic Dermoscopy of the Human Skin: Tuning Excitation Energy for Optimal Detection Bandwidth With Fast and Deep Imaging in vivo

Accurate quantification of retinal structures in 3-D optical coherence tomography data of eyes with pathologies provides clinically relevant information. We present an approach to jointly segment retinal layers and lesions in eyes with topology-disrupting retinal diseases by a loosely coupled level set framework. In the new approach, lesions are modeled as an additional space-variant layer delineated by auxiliary interfaces. Furthermore, the segmentation of interfaces is steered by local differences in the signal between adjacent retinal layers, thereby allowing the approach to handle local intensity variations. The accuracy of the proposed method of both layer and lesion segmentation has been evaluated on eyes affected by central serous retinopathy and age-related macular degeneration. In addition, layer segmentation of the proposed approach was evaluated on eyes without topology-disrupting retinal diseases. Good agreement between the segmentation performed manually by a medical doctor and results obtained from the automatic segmentation was found for all data types. The mean unsigned error for all interfaces varied between 2.3 and 11.9 μm (0.6-3.1 pixels). Furthermore, lesion segmentation showed a Dice coefficient of 0.68 for drusen and 0.89 for fluid pockets. Overall, the method provides a flexible and accurate solution to jointly segment lesions and retinal layers.

ETPL SP - 026	Multi-Modality Imaging Enables Detailed Hemodynamic Simulations in Dissecting Aneurysms in Mice

A multi-modality imaging-based modeling approach was used to study complex unsteady hemodynamics and lesion growth in a dissecting abdominal aortic aneurysm model. We combined in vivo ultrasound (geometry and flow) and in vitro optical coherence tomography(OCT) (geometry) to obtain the high resolution needed to construct detailed hemodynamic simulations over large portions of the murine vasculature, which include fine geometric complexities. We illustrate this approach for a spectrum of dissecting abdominal aortic aneurysms induced in male apolipoprotein E-null mice by high-dose angiotensin II infusion. In vivo morphological and hemodynamic data provide information on volumetric lesion growth and changes in blood flow dynamics, respectively, occurring from the day of initial aortic expansion. We validated the associated computational models by comparing results on time-varying outlet flows and vortical structures within the lesions. Three out of four lesions exhibited abrupt formation of thrombus, though different in size. We determined that a lesion without thrombus formed with a thickened vessel wall, which was resolvable by OCT and histology. We attribute differences in final sizes and compositions of these lesions to the different computed flow and vortical structures we obtained in our mouse-specific fluid dynamic models. Differences in morphology and hemodynamics play crucial roles in determining the evolution of dissecting abdominal aortic aneurysms. Coupled high resolution in vivo and in vitro imaging approaches provide much-improved geometric models for hemodynamic simulations. Our imaging-based computational findings suggest a link between perturbations in hemodynamic metrics and aneurysmal disease heterogeneity.

ETPL SP - 027	Segmentation of Locally Varying Numbers of Outer Retinal Layers by a Model Selection Approach

Extraction of image-based biomarkers, such as the presence, visibility, or thickness of a certain layer, from 3-D optical coherence tomography data provides relevant clinical information. We present a method to simultaneously determine the number of visible layers in the outer retina and segment them. The method is based on a model selection approach with special attention given to the balance between the quality of a fit and model complexity. This will ensure that a more complex model is selected only if this is sufficiently supported by the data. The performance of the method was evaluated on healthy and retinitis pigmentosa (RP) affected eyes. In addition, the reproducibility of automatic method and manual annotations was evaluated on healthy eyes. Good agreement between the segmentation performed manually by a medical doctor and results obtained from the automatic segmentation was found. The mean unsigned deviation for all outer retinal layers in healthy and RP affected eyes varied between 2.6 and 4.9 μm. The reproducibility of the automatic method was similar to the reproducibility of the manual segmentation. Overall, the method provides a flexible and accurate solution for determining the visibility and location of outer retinal layers and could be used as an aid for the disease diagnosis and monitoring.

ETPL SP - 028	Monte Carlo Simulations of Diffusion Weighted MRI in Myocardium: Validation and Sensitivity Analysis

A model of cardiac microstructure and diffusion MRI is presented, and compared with experimental data from ex vivo rat hearts. The model includes a simplified representation of individual cells, with physiologically correct cell size and orientation, as well as intra- to extracellular volume ratio. Diffusion MRI is simulated using a Monte Carlo model and realistic MRI sequences. The results show good correspondence between the simulated and experimental MRI signals. Similar patterns are observed in the eigenvalues of the diffusion tensor, the mean diffusivity (MD), and the fractional anisotropy (FA). A sensitivity analysis shows that the diffusivity is the dominant influence on all three eigenvalues of the diffusion tensor, the MD, and the FA. The area and aspect ratio of the cell cross-section affect the secondary and tertiary eigenvalues, and hence the FA. Within biological norms, the cell length, volume fraction of cells, and rate of change of helix angle play a relatively small role in influencing tissue diffusion. Results suggest that the model could be used to improve understanding of the relationship between cardiac microstructure and diffusion MRI measurements, as well as in testing and refinement of cardiac diffusion MRI protocols.

ETPL SP - 029	An Efficient Reconstruction Algorithm Based on the Alternating Direction Method of Multipliers for Joint Estimation of $R*2$ and Off-Resonance in fMRI

R^*_2 mapping is a useful tool in blood-oxygen-level dependent fMRI due to its quantitative-nature. However, like T^*_2-weighted imaging, standard R^*_2 mapping based on multi-echo EPI suffers from geometric distortion, due to strong off-resonance near the air-tissue interface. Joint mapping of R^*_2 and off-resonance can correct the geometric distortion and is less susceptible to motion artifacts. Single-shot joint mapping of R^*_2 and off-resonance is possible with a rosette trajectory due to its frequent sampling of the k-space center. However, the corresponding reconstruction is nonlinear, ill-conditioned, large-scale, and computationally inefficient with current algorithms. In this paper, we propose a novel algorithm for joint mapping of R^*_2 and off-resonance, using rosette k-space trajectories. The new algorithm, based on the alternating direction method of multipliers, improves the reconstruction efficiency by simplifying the original complicated cost function into a composition of simpler optimization steps. Compared with a recently developed trust region algorithm, the new algorithm achieves the same accuracy and an acceleration of threefold to sixfold in reconstruction time. Based on the new algorithm, we present simulation and in vivo data from single-shot, double-shot, and quadruple-shot rosettes and demonstrate the improved image quality and reduction of distortions in the reconstructed R^*_2 map.

ETPL SP - 030	Weight Multispectral Reconstruction Strategy for Enhanced Reconstruction Accuracy and Stability with Cerenkov Luminescence Tomography

Cerenkov luminescence tomography (CLT) provides a novel technique for 3-D noninvasive detection of radiopharmaceuticals in living subjects. However, because of the severe scattering of Cerenkov light, the reconstruction accuracy and stability of CLT is still unsatisfied. In this paper, a modified weight multispectral CLT (wmCLT) reconstruction strategy was developed which split the Cerenkov radiation spectrum into several sub-spectral bands and weighted the sub-spectral results to obtain the final result. To better evaluate the property of the wmCLT reconstruction strategy in terms of accuracy, stability and practicability, several numerical simulation experiments and in vivo experiments were conducted and the results obtained were compared with the traditional multispectral CLT (mCLT) and hybrid-spectral CLT (hCLT) reconstruction strategies. The numerical simulation results indicated that wmCLT strategy significantly improved the accuracy of Cerenkov source localization and intensity quantitation and exhibited good stability in suppressing noise in numerical simulation experiments. And the comparison of the results achieved from different in vivo experiments further indicated significant improvement of the wmCLT strategy in terms of the shape recovery of the bladder and the spatial resolution of imaging xenograft tumors. Overall the strategy reported here will facilitate the development of nuclear and optical molecular tomography in theoretical study.

ETPL SP - 031	A Two-Dimensional Framework of Multiple Kernel Subspace Learning for Recognizing Emotion in Speech.

As a highly active topic in computational paralinguistics, speech emotion recognition (SER) aims to explore ideal representations for emotional factors in speech. In order to improve the performance of SER, multiple kernel learning (MKL) dimensionality reduction has been utilized to obtain effective information for recognizing emotions. However, the solution of MKL usually provides only one nonnegative mapping direction for multiple kernels; this may lead to loss of valuable information. To address this issue, we propose a two-dimensional framework for multiple kernel subspace learning. This framework provides more linear combinations on the basis of MKL without nonnegative constraints, which preserves more information in the learning procedures. It also leverages both of MKL and two-dimensional subspace learning, combining them into a unified structure. To apply the framework to SER, we also propose an algorithm, namely generalised multiple kernel discriminant analysis (GMKDA), by employing discriminant embedding graphs in this framework. GMKDA takes advantage of the additional mapping directions for multiple kernels in the proposed framework. In order to evaluate the performance of the proposed algorithm a wide range of experiments is carried out on several key emotional corpora. These experimental results demonstrate that the proposed methods can achieve better performance compared with some conventional and subspace learning methods in dealing with SER.

ETPL SP - 032	Automated Anatomy-Based Tracking of Systemic Arteries in Arbitrary Field-of-View CTA Scans

We propose an automated pipeline for vessel centerline extraction in 3-D computed tomography angiography (CTA) scans with arbitrary fields of view. The principal steps of the pipeline are body part detection, candidate seed selection, segment tracking, which includes centerline extraction, and vessel tree growing. The final tree-growing step can be instantiated in either a semi- or fully automated fashion. The fully automated initialization is carried out using a vessel position regression algorithm. Both semi-and fully automated methods were evaluated on 30 CTA scans comprising neck, abdominal, and leg arteries in multiple fields of view. High detection rates and centerline accuracy values for 38 distinct vessels demonstrate the effectiveness of our approach.

ETPL SP - 033	Ultrafast Synthetic Transmit Aperture Imaging Using Hadamard-Encoded Virtual Sources with Overlapping Sub-Apertures

The development of ultrafast ultrasound imaging offers great opportunities to improve imaging technologies, such as shear wave elastography and ultrafast Doppler imaging. In ultrafast imaging, there are tradeoffs among image signal-to-noise ratio (SNR), resolution, and post-compounded frame rate. Various approaches have been proposed to solve this tradeoff, such as multiplane wave imaging or the attempts of implementing synthetic transmit aperture imaging. In this paper, we propose an ultrafast synthetic transmit aperture (USTA) imaging technique using Hadamard-encoded virtual sources with overlapping sub-apertures to enhance both image SNR and resolution without sacrificing frame rate. This method includes three steps: 1) create virtual sources using sub-apertures; 2) encode virtual sources using Hadamard matrix; and 3) add short time intervals (a few microseconds) between transmissions of different virtual sources to allow overlapping sub-apertures. The USTA was tested experimentally with a point target, a B-mode phantom, and in vivo human kidney micro-vessel imaging. Compared with standard coherent diverging wave compounding with the same frame rate, improvements on image SNR, lateral resolution (+33%, with B-mode phantom imaging), and contrast ratio (+3.8 dB, with in vivo human kidney micro-vessel imaging) have been achieved. The f-number of virtual sources, the number of virtual sources used, and the number of elements used in each sub-aperture can be flexibly adjusted to enhance resolution and SNR.

ETPL SP - 034	Passive BCI in Operational Environments: Insights, Recent Advances, and Future Trends

Goal: This minireview aims to highlight recent important aspects to consider and evaluate when passive brain-computer interface (pBCI) systems would be developed and used in operational environments, and remarks future directions of their applications. Methods: Electroencephalography (EEG) based pBCI has become an important tool for real-time analysis of brain activity since it could potentially provide covertly-without distracting the user from the main task-and objectively-not affected by the subjective judgment of an observer or the user itself-information about the operator cognitive state. Results: Different examples of pBCI applications in operational environments and new adaptive interface solutions have been presented and described. In addition, a general overview regarding the correct use of machine learning techniques (e.g., which algorithm to use, common pitfalls to avoid, etc.) in the pBCI field has been provided. Conclusion: Despite recent innovations on algorithms and neurotechnology, pBCI systems are not completely ready to enter the market yet, mainly due to limitations of the EEG electrodes technology, and algorithms reliability and capability in real settings. Significance: High complexity and safety critical systems (e.g., airplanes, ATM interfaces) should adapt their behaviors and functionality accordingly to the user' actual mental state. Thus, technologies (i.e., pBCIs) able to measure in real time the user's mental states would result very useful in such “high risk” environments to enhance human machine interaction, and so increase the overall safety.

ETPL SP - 035	Real-Time Model-Based Fault Detection of Continuous Glucose Sensor Measurements

Objective: Faults in subcutaneous glucose concentration readings with a continuous glucose monitoring (CGM) may affect the computation of insulin infusion rates that can lead to hypoglycemia or hyperglycemia in artificial pancreas control systems for patients with type 1 diabetes (T1D). Methods: Multivariable statistical monitoring methods are proposed for detection of faults in glucose concentration values reported by a subcutaneous glucose sensor. A nonlinear first principle glucose/insulin/meal dynamic model is developed. An unscented Kalman filter is used for state and parameter estimation of the nonlinear model. Principal component analysis models are developed and used for detection of dynamic changes. K-nearest neighbor classification algorithm is used for diagnosis of faults. Data from 51 subjects are used to assess the performance of the algorithm. Results: The results indicate that the proposed algorithm works successfully with 84.2% sensitivity. Overall, 155 (out of 184) of the CGM failures are detected with a 2.8-min average detection time. Conclusion: A novel algorithm that integrates data-driven and model-based methods is developed. The proposed method is able to detect CGM failures with a high rate of success. Significance: The proposed fault detection algorithm can decrease the effects of faults on insulin infusion rates and reduce the potential for hypo- or hyperglycemia for patients with T1D.

ETPL SP - 036	Fast Numerical Simulation of Focused Ultrasound Treatments During Respiratory Motion With Discontinuous Motion Boundaries

Objective: Focused ultrasound (FUS) is rapidly gaining clinical acceptance for several target tissues in the human body. Yet, treating liver targets is not clinically applied due to a high complexity of the procedure (noninvasiveness, target motion, complex anatomy, blood cooling effects, shielding by ribs, and limited image-based monitoring). To reduce the complexity, numerical FUS simulations can be utilized for both treatment planning and execution. These use-cases demand highly accurate and computationally efficient simulations. Methods: We propose a numerical method for the simulation of abdominal FUS treatments during respiratory motion of the organs and target. Especially, a novel approach is proposed to simulate the heating during motion by solving Pennes' bioheat equation in a computational reference space, i.e., the equation is mathematically transformed to the reference. The approach allows for motion discontinuities, e.g., the sliding of the liver along the abdominal wall. Results: Implementing the solver completely on the graphics processing unit and combining it with an atlas-based ultrasound simulation approach yields a simulation performance faster than real time (less than 50-s computing time for 100 s of treatment time) on a modern off-the-shelf laptop. The simulation method is incorporated into a treatment planning demonstration application that allows to simulate real patient cases including respiratory motion. Conclusion: The high performance of the presented simulation method opens the door to clinical applications. Significance:The methods bear the potential to enable the application of FUS for moving organs.

ETPL SP - 037	Computerized Lung Sound Screening for Pediatric Auscultation in Noisy Field Environments

Goal: Chest auscultations offer a non-invasive and low-cost tool for monitoring lung disease. However, they present many shortcomings including inter-listener variability, subjectivity, and vulnerability to noise and distortions. The current work proposes a computer-aided approach to process lung signals acquired in the field under adverse noisy conditions, by improving the signal quality and offering automated iden- tification of abnormal auscultations indicative of respiratory pathologies. Methods: The developed noise-suppression scheme eliminates ambient sounds, heart sounds, sensor artifacts and crying contamination. The improved high-quality signal is then mapped onto a rich spectro-temporal feature space before being classified using a trained support-vector machine clas- sifier. Individual signal frame decisions are then combined using an evaluation scheme, providing an overall patient-level decision for unseen patient records. Results: All methods are evaluated on a large data set with > 1,000 children enrolled, 1-59 months old. The noise suppression scheme is shown to significantly improve signal quality; and the classification system achieves an accuracy of 86.7% in distinguishing normal from pathological sounds, far surpassing other state-of-the art methods. Conclusion: Computerized lung sound processing can benefit from the enforcement of advanced noise-suppression. A fairly short processing window size (< 1 s) combined with detailed spectro-temporal features is recommended, in order to capture transient adventitious events without highlighting sharp noise occurrences. Significance: Unlike existing methodologies in the literature, the proposed work is not limited in scope or confined to laboratory-settings:

ETPL SP - 038	eCurves: A Temporal Shape Encoding

Objective: This paper presents a framework for temporal shape analysis to capture the shape and changes of anatomical structures from 3D+t(ime) medical scans. Method: We first encode the shape of a structure at each time point with the spectral signature, i.e., the eigenvalues and eigenfunctions of the Laplace operator. We then expand it to capture morphing shapes by tracking the eigenmodes across time according to the similarity of their eigenfunctions. The similarity metric is motivated by the fact that small shape deformations lead to minor changes in the eigenfunctions. Following each eigenmode from the beginning to end results in a set of eigenmode curves (eCurves) representing the shape and its changes over time. Results: We apply our encoding to a cardiac data set consisting of series of segmentations outlining the right and left ventricles over time. We measure the accuracy of our encoding by training classifiers on discriminating healthy adults from patients that received reconstructive surgery for Tetralogy of Fallot (TOF). The classifiers based on our encoding significantly surpasses deformation based encodings of the right ventricle, the structure most impacted by TOF. Conclusion: The strength of our framework lies in its simplicity: it only assumes pose invariance within a time series but does not assume point-to-point correspondence across time series or a (statistical or physical) model. In addition, it is easy to implement and only depends on a single parameter, i.e., the number of curves.

ETPL SP - 039	An Integrated Circuit for Simultaneous Extracellular Electrophysiology Recording and Optogenetic Neural Manipulation

Objective: The ability to record and to control action potential firing in neuronal circuits is critical to understand how the brain functions. The objective of this study is to develop a monolithic integrated circuit (IC) to record action potentials and simultaneously control action potential firing using optogenetics. Methods: A low-noise and high input impedance (or low input capacitance) neural recording amplifier is combined with a high current laser/lightemitting diode (LED) driver in a single IC. Results: The low input capacitance of the amplifier (9.7 pF) was achieved by adding a dedicated unity gain stage optimized for high impedance metal electrodes. The input referred noise of the amplifier is 4.57 μV_{rms}, which is lower than the estimated thermal noise of the metal electrode. Thus, the action potentials originating from a single neuron can be recorded with a signal-to-noise ratio of at least 6.6. The LED/laser current driver delivers a maximum current of 330 mA, which is adequate for optogenetic control. The functionality of the IC was tested with an anesthetized Mongolian gerbil and auditory stimulated action potentials were recorded from the inferior colliculus. Spontaneous firings of fifth (trigeminal) nerve fibers were also inhibited using the optogenetic protein Halorhodopsin. Moreover, a noise model of the system was derived to guide the design. Significance:

ETPL SP - 040	Features for Masking-Based Monaural Speech Separation in Reverberant Conditions

Monaural speech separation is a fundamental problem in speech and signal processing. This problem can be approached from a supervised learning perspective by predicting an ideal time-frequency mask from features of noisy speech. In reverberant conditions at low signal-to-noise ratios (SNRs), accurate mask prediction is challenging and can benefit from effective features. In this paper, we investigate an extensive set of acoustic-phonetic features extracted in adverse conditions. Deep neural networks are used as the learning machine, and separation performance is evaluated using standard objective speech intelligibility metrics. Separation performance is systematically evaluated in both nonspeech and speech interference, in a variety of SNRs, reverberation times, and direct-to-reverberant energy ratios. Considerable performance improvement is observed by using contextual information, likely due to temporal effects of room reverberation. In addition, we construct feature combination sets using a sequential floating forward selection algorithm, and combined features outperform individual ones. We also find that optimal feature sets in anechoic conditions are different from those in reverberant conditions.

Thank you!