

# Papers Published in Technical Journals and Conference Proceedings

## Escaped-Huffman and Adaptive Recursive Rice Coding for Lossless Compression of the Mapped Domain Linear Prediction Residual

N. Harada, Y. Kamamoto, and T. Moriya

Proc. of ICASSP 2010, IEEE, pp. 4646–4649, Dallas, Texas, USA.

ITU-T Recommendation G.711.0 has just been established. It defines a lossless and stateless compression for G.711 packet payloads (for both A-law and  $\mu$ -law). This paper introduces some coding technologies proposed and applied to the G.711.0 codec, such as Plus-Minus zero mapping for the mapped domain linear predictive coding and escaped-Huffman coding combined with adaptive recursive Rice coding for lossless compression of the prediction residual. Performance test results for those coding tools are shown in comparison with the results for the conventional technology. The performance is measured on the basis of the figure of merit (FoM), which is a function of the trade-off between compression performance and computational complexity. The proposed tools improve the compression performance by 0.16% in total while keeping the computational complexity of the encoder/decoder pair low (about 1.0 WMOPS on average and 1.667 WMOPS in the worst case).

## Emerging ITU-T Standard G.711.0—Lossless Compression of G.711 Pulse Code Modulation

N. Harada, Y. Kamamoto, T. Moriya, Y. Hiwasaki, M. A. Ramalho, L. Netsch, J. Stachurski, L. Miao, H. Taddei, and F. Qi

Proc. of ICASSP 2010, IEEE, pp. 4658–4661, Dallas, Texas, USA.

ITU-T Recommendation G.711 is the benchmark standard for narrowband telephony. It has been successful for many decades because of its proven voice quality, ubiquity, and utility. A new ITU-T recommendation, denoted G.711.0, has recently been established defining lossless compression for G.711 packet payloads typically found in IP networks. This paper presents a brief overview of technologies employed within the G.711.0 standard and summarizes the compression and complexity results. It is shown that G.711.0 provides greater than 50% average compression in typical service provider environments while keeping low computational complexity for the encoder/decoder pair (1.0 WMOPS average, <1.7 WMOPS worst case) and low memory footprint (about 5 k octets of RAM, 5.7 k octets of ROM, and 3.6 k of program memory measured in the number of basic operators).

## Voice Activity Detection Using Frame-wise Model Re-estimation Method Based on Gaussian Pruning with Weight Normalization

M. Fujimoto, S. Watanabe, and T. Nakatani

Proc. of INTERSPEECH 2010, Makuhari, Chiba, Japan.

This paper proposes a frame-wise model re-estimation method based on Gaussian pruning with weight normalization for noise robust voice activity detection (VAD). Our previous work, switching-Kalman-filter-based VAD, sequentially estimates a non-stationary-noise Gaussian mixture model (GMM) and constructs GMMs of observed noisy speech signals by composing pre-trained silence and clean GMMs and sequentially estimated noise GMMs. However, the

composed models are not optimal because they do not fully reflect the characteristics of the observed signal. Thus, to ensure the optimality of the composed models, we investigate a method for re-estimating the composed model. Since our VAD method works under frame-wise sequential processing, there are insufficient re-training data for re-estimation of all the model parameters. To solve this problem, we propose a model re-estimation method that involves the extraction of reliable information using Gaussian pruning with weight normalization. Namely, the proposed method re-estimates the model by pruning non-dominant Gaussian distributions in expressing the local characteristics of each frame and by normalizing the Gaussian weights of the remaining distributions.

## Increase of Fundamental Oscillation Frequency in Resonant Tunneling Diode with Thin Barrier and Graded Emitter Structure

S. Suzuki, A. Teranishi, M. Asada, H. Sugiyama, and H. Yokoyama

Proc. of Infrared Millimeter and Terahertz Waves (IRMMW-THz), Rome, Italy, 2010.

We obtained an increase in the oscillation frequency of resonant tunneling diodes (RTDs) using a graded emitter and thin barriers for reduced transit and tunneling times. The fundamental oscillation frequency of 1.04 THz was achieved with this structure. The dependence of output power on oscillation frequency is also shown. The output power was around 10  $\mu$ W in the 0.9–1 THz region.

## Wide-range and Fast-tracking Frequency Offset Estimator for Optical Coherent Receivers

T. Nakagawa, K. Ishihara, T. Kobayashi, R. Kudo, M. Matsui, Y. Takatori, and M. Mizoguchi

Proc. of ECOC 2010, Torino, Italy.

A blind spectrum-based frequency offset estimator with fast tracking time is presented. A receiver using the proposed estimator to eliminate the frequency ambiguity of the  $m^{\text{th}}$  power algorithm can achieve precise estimation over a wide frequency range.

## Efficient Secure Auction Protocols Based on the Boneh-Goh-Nissim Encryption

T. Mitsunaga, Y. Manabe, and T. Okamoto

Proc. of SCIS 2010, IEICE, Takamatsu, Japan.

This paper presents efficient secure auction protocols for first price auction and second price auction. Previous auction protocols are based on a generally secure multi-party protocol called the mix-and-match protocol. However, the time complexity of the mix-and-match protocol is large, although it can securely calculate any logical circuits. The proposed protocols reduce the number of times the mix-and-match protocol is used by replacing them with Boneh-Goh-Nissim encryption, which enables calculation of 2-DNF (disjunctive normal form) of encrypted data.

### Spatial Pooling of One-dimensional Second-order Motion Signals

K. Maruya and S. Nishida

Journal of Vision, Vol. 10, No. 13:24, pp. 1–18, 2010.

We can detect visual movements not only from luminance motion signals (first-order motion) but also from non-luminance motion signals (second-order motion). It has been established for first-order motion that the visual system pools local one-dimensional motion signals across space and orientation to solve the aperture problem and to estimate two-dimensional object motion. In this study, we investigated (i) whether local one-dimensional second-order motion signals are also pooled across space and orientation into a global 2D motion and, if so, (ii) whether the second-order motion signals are pooled independently of, or in cooperation with, first-order motion signals. We measured the direction-discrimination performance and the rating of a global circular translation of four oscillating bars, each defined either by luminance or by a non-luminance attribute, such as flicker or binocular depth. The results showed evidence of motion pooling both when the stimulus consisted only of second-order bars and when it consisted of first-order and second-order bars. We observed global motion pooling across first- and second-order motions even when the first-order motion was not accompanied by trackable position changes. These results suggest the presence of a universal pooling system for first- and second-order one-dimensional motion signals.

### Creation and Analysis of a Japanese Speaking Style Parallel Database for Expressive Speech Synthesis

H. Nakajima, N. Miyazaki, A. Yoshida, T. Nakamura, and H. Mizuno

Proc. of Oriental COCOSDA 2010, COCOSDA Asia Section, Kathmandu, Nepal.

This paper describes a newly developed database for expressive speech synthesis. This speech database is characterized by two features: i) the sentences are taken from real domains such as sales talk, storytelling, and telephone conversations, where speech is uttered in expressive (or conversational) style, so sentences are domain-dependent and ii) each sentence is uttered in both reading style and expressive style, so this database stores parallel speaking style speech. This database is designed to capture the acoustic and prosodic differences between parallel styles and to elucidate the domain-dependent linguistic characteristics that cause those differences. This paper describes both the concept of this speech database and the issues raised by its implementation. We detail the basic characteristics and preliminary results of two style comparisons to elucidate the linguistic characteristics that contribute to the establishment of expressive speech synthesis.

### Effect of Speech Sound Naturalness on the Neural Basis of Format Frequency Discrimination

S. Hiroya and F. H. Guenther

Proc. of Neuroscience 2010, the Society for Neuroscience, San Diego, 2010.

Few previous imaging studies of speech perception have investigated the neural mechanisms underlying discriminability of format frequencies for less natural vowel sounds. First, we developed a novel method for controlling vowel naturalness on the basis of band-limited finite impulse response filters related to the first four formats. The results showed that naturalness was significantly reduced for sounds with decreasing filter bandwidths. Next, we performed a functional

magnetic resonance imaging study that investigated the neural basis of formant frequency discrimination in less natural vowel sounds. The result showed that the left-lateralized premotor cortex was more activated for less natural sounds, and the area overlapped that of vowel production. This suggests that the involvement of the premotor cortex varies depending on speech sound naturalness.

### Fast Template Matching Based on Normalized Cross Correlation Using Adaptive Block Partitioning and Initial Threshold Estimation

M. Mori and K. Kashino

Proc. of 2010 IEEE International Symposium on Multimedia, pp. 196–203, Taichung, Taiwan.

This paper proposes a fast template matching method based on normalized cross correlation (NCC). NCC is more robust against image variations such as illumination changes than the widely used sum of absolute difference (SAD). A problem with NCC has been its high computation cost. To deal with this problem, we use adaptive block partitioning and initial threshold estimation to extend the multilevel successive elimination algorithm. Adaptive block partitioning provides efficient sub-block partitioning and tighter boundaries. Initial threshold estimation yields a larger boundary threshold. They greatly suppress the number of search points at an earlier level from the beginning of search. The proposed method is exhaustive and robust with respect to template position and size. Experiments show that our method is up to 400 times faster than the brute force method and is significantly faster than conventional methods.

### Coded Packet Immediate Access for Contention-based Wireless Relay Networks

D. Umehara, S. Denno, M. Morikura, and T. Sugiyama

Proc. of the 4th International Conference on Signal Processing and Communication Systems (ICSPCS), Vol. 1, No. 1, pp. 1–9, Gold Coast, Australia, 2010.

This paper proposes a medium access control (MAC) protocol with network coding on relay nodes for contention-based multihop wireless relay networks. The proposed protocol is called coded packet priority access (CPPA) protocol in which coded packets have higher transmission opportunity than non-coded native packets at relay nodes. In this paper, the performance of coded packet immediate access (CPIA) protocols, which are a subclass of CPPA protocols, is evaluated for single-relay bidirectional symmetric traffic and upper and lower bounds of analytical throughput are derived for any given node traffic. It is shown that the lower bound approximates to the throughput obtained from computer simulations with high accuracy. The conventional slotted ALOHA protocol with network coding (S-ALOHA/NC) is required to adapt the transmission probability of a relay node to a rational function of node traffic so as to maximize the throughput whereas the CPIA protocol achieves the maximize throughput only if the relay node transmits no native packets. Furthermore it is clarified that the CPIA protocol is superior to the S-ALOHA/NC protocol in delay for given retransmission probabilities of user nodes.

### Enhancement of IEEE 802.11 and Network Coding for Single-relay Multi-user Wireless Networks

D. Umehara, C.-H. Huang, S. Denno, M. Morikura, and T. Sugiyama

Proc. of the 4th International Conference on Signal Processing and Communication Systems (ICSPCS), Vol. 1, No. 1, pp. 1–9, Gold Coast, Australia, 2010.

Network coding is a promising technique for improving system performance in wireless multihop networks. In this paper, the throughput and fairness in single-relay multiuser wireless networks are evaluated. The carrier sense multiple access with collision avoidance (CSMA/CA) protocol and network coding are used in the medium access control (MAC) sublayer in such networks. The fairness of wireless medium access among stations (STAs), the access point (AP), and the relay station (RS) results in asymmetric bidirectional flows via the RS; as a result, the wireless throughput decreases substantially. To overcome this problem, an autonomous optimization of the minimum contention window size is developed for CSMA/CA and network coding to assign appropriate transmission opportunities to both the AP and RS. By optimizing the minimum contention window size according to the number of STAs, the wireless throughput in single-relay multi-user networks can be improved and fairness between bidirectional flows via the RS can be achieved. Numerical analysis and computer simulations enable us to evaluate the performances of CSMA/CA and network coding in single-relay multi-user wireless networks.

---

#### Efficient Data Gathering for Hierarchical Sensor Networks

Y. Kishino, Y. Sakurai, K. Kamei, T. Maekawa, Y. Yanagisawa, and T. Okadome

Information Processing Society of Japan, Vol. 3, No. 4, pp. 82–93, 2010.

In this paper we propose an efficient data gathering method using a hierarchical tree topology in a high-density sensor network. The proposed method gathers sensor data using Singular Value Decomposition (SVD) for each cluster by taking advantage of periodicity and correlation among sensor data. It can reduce the amount of data in wireless communication and errors and achieve efficient data gathering. Our experimental result shows that the hierarchical network topology and data gathering by SVD can reduce the amount of data and errors when the level of network topology is high.

---

#### Single-electron Devices based on Si Nanoscale FETs

K. Nishiguchi and A. Fujiwara

Proc. of Workshop on Innovative Devices and Systems, WINDS, Hawaii, USA, 2010.

A MOSFET-based circuit utilizing single electrons is demonstrated at room temperature. A nano-wire MOSFET can transfer single electrons one after another to a storage node thanks to the extremely small current leakage of the MOSFET. The electrons transferred to the node are detected by another wire MOSFET, which is located near the node. The combination of these nano-wire MOSFETs allows real-time monitoring of the single-electron transfer, which helps microscopic understanding of individual electron movement in the MOSFET. While the MOSFET can control the average movement of single electrons, each movement of an individual electron is completely random. Such controllability and randomness of electron movement is used for high-quality random-number generation (RNG) suitable for data processing that stochastically extracts the most preferable pattern among various ones. The MOSFET-based RNG allows fast operation as well as high controllability, which leads to flexible extraction of the preferable pattern. This stochastic data processing promises high efficiency, fast operation, and low power

consumption like the human brain.

---

#### Risk Management and Intelligence Management during Emergency

M. Higashida, Y. Maeda, and H. Hayashi

Journal of Disaster Research, Vol. 5, No. 6, pp. 636–637, 2010.

In the 15 years since Kobe's Hanshin-Awaji Earthquake, awareness is growing that simply gathering information may not be enough for preparing systems, executing emergency responses, and making decisions rapidly and precisely. The question has become how—and whether—emergency response information can be used effectively and efficiently for rapid disaster response, recovery, and rebuilding. We analyzed emergency response decision making from the perspective of information processing, looking for the features organizations need to process information efficiently. We also propose how to continuously improve emergency response performance.

---

#### Query Graph Pattern Optimization for Efficient Discovery of Implicit Information within Linked Data

R. Sakai, K. Iiduka, H. Sato, T. Murayama, T. Kobayashi, H. Hattori, and T. Ishida

Information Processing Society of Japan, Vol. 51, No. 12, pp. 2298–2309, 2010.

Spreading the use of the Semantic Web and Linked Data has made it possible to explore new connections between data which were not linked before. The use of various query graph patterns allows us to explore such connections and leads us to find potentially useful information; however, this information may not be reliable owing to its linkages; for example, if the linkage between the data is weak. Moreover, in a huge set of Linked Data, there are too many patterns to follow and many of them may contain similar semantic connections. Our goal is to select query graph patterns that allow us to find reliable information efficiently from a huge number of patterns. We propose a method to reduce the number of query graph patterns while maintaining the pattern variations to meet different users' needs. To achieve high reliability for the information reached, we select the patterns made of two paths. To ensure wide variations, we classify existing patterns into some representative categories and keep all the categories while we reduce the number. By verifying with real datasets, we confirmed that we could reduce the number by up to 9.1% of its original and also underpinned the reliability of the resulting information.

---

#### Bounce Hardness Index of Gravitational Waves

F. Ishiyama and R. Takahashi

Classical and Quantum Gravity, Institute of Physics, Vol. 27, No. 24, p. 245021, 2010.

We present a method of mode analysis to search for signals with frequency evolution and limited duration in a given data stream. Our method is a natural expansion of Fourier analysis, and we can obtain information about frequency evolution with high frequency precision and high time resolution. Applications of this method to the analysis of inspiral and burst signals show that the signals are characterized by an index that we name 'bounce hardness'. The index corresponds to the growth rate of the signals.

**Fundamental Oscillation of Resonant Tunneling Diodes above 1 THz at Room Temperature**

S. Suzuki, M. Asada, A. Teranishi, H. Sugiyama, and H. Yokoyama

Appl. Phys. Lett., Vol. 97, No. 24, p. 242102, 2010.

Fundamental oscillations up to 1.04 THz were achieved in resonant tunneling diodes at room temperature. A graded emitter and thin barriers were introduced in GaInAs/AlAs double-barrier resonant tunneling diodes to reduce the transit time in the collector depletion region and the resonant tunneling time, respectively. Output powers were 7  $\mu$ W at 1.04 THz and around 10  $\mu$ W in the 0.9–1 THz region. A change in oscillation frequency of about 4% with bias voltage was also obtained.

**Performance Analysis of Slotted ALOHA and Network Coding for Single-relay Multi-user Wireless Networks**

D. Umehara, S. Denno, M. Morikura, and T. Sugiyama

Ad Hoc Networks, Vol. 9, No. 2, pp. 164–179, 2011.

Deployment of wireless relay nodes can enhance system capacity, extend wireless service coverage, and reduce energy consumption in wireless networks. Network coding enables us to mix two or more

packets into a single coded packet at relay nodes and improve performances in wireless relay networks. In this paper, we succeed in developing analytical models of the throughput and delay on slotted ALOHA (S-ALOHA) and S-ALOHA with network coding (S-ALOHA/NC) for single-relay multi-user wireless networks with bidirectional data flows. The analytical models involve the effects of queue saturation and unsaturation at the relay node. The throughput and delay for each user node can be extracted from the total throughput and delay by using the analytical models. One can formulate various optimization problems in traffic control in order to maximize the throughput, minimize the delay, or achieve fairness of the throughput or the delay. In particular, we clarify that the total throughput is enhanced in the S-ALOHA/NC protocol on condition that the transmission probability at the relay node is set to the value on the boundary between queue saturation and unsaturation. Our analysis provides achievable regions in throughput on two directional data flows at the relay node for both the S-ALOHA and S-ALOHA/NC protocols. As a result, we show that the achievable region in throughput can be enhanced by using network coding and traffic control.