

Top Downloads in IEEE Xplore

Each “Reader’s Choice” column focuses on a different publication of the IEEE Signal Processing Society (SPS). This month we are highlighting articles featured at the IEEE International Conference on Image Processing (ICIP).

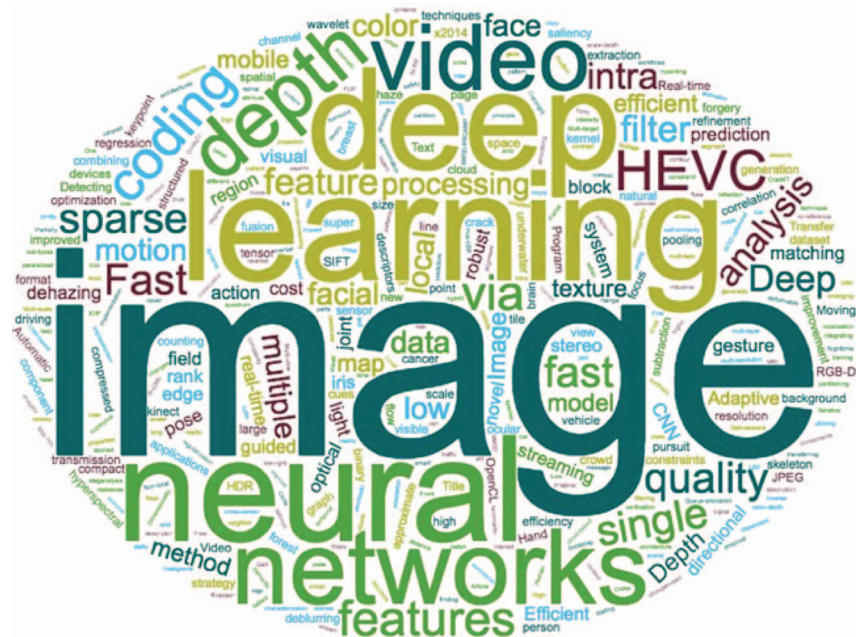
The ICIP, sponsored by the IEEE SPS, is the premier forum for the presentation of technological advances and research results in the fields of theoretical, experimental, and applied image and video processing.

This issue's "Reader's Choice" column lists the top 15 articles from ICIP 2014, ICIP 2015, and ICIP 2016, indicated by the year listed after the abstract, that were the most downloaded from January 2015 to June 2017. Please send your suggestions and comments on this column to Associate Editor Changshui Zhang (zcs@mail.tsinghua.edu.cn).

Hand Gesture Recognition with Leap Motion and Kinect Devices

Marin, G.; Dominio, F.; Zanuttigh, P. This paper proposes a novel hand gesture recognition scheme explicitly targeted to leap motion data. An adhoc feature set based on the positions and orientation of the fingertips is fed into a multiclass support vector machine classifier to recognize the performed gestures. A set of features is also ex-

Digital Object Identifier 10.1109/MSP.2017.2743111
 Date of publication: 13 November 2017



tracted from the depth computed from the Kinect and combined with the leap motion ones to improve the recognition performance.

2014

Local Binary Pattern Network: A Deep Learning Approach for Face Recognition

Xi, M.; Chen, L.; Polajnar, D.;
Tong, W.

In this paper, a novel unsupervised deep-learning-based methodology, named *local binary pattern network (LBPNet)*, is proposed to efficiently extract and compare high-level overcomplete features in a multilayer

hierarchy. The LBPNet retains the same topology of the convolutional neural network, whereas the trainable kernels are replaced by the off-the-shelf computer vision descriptor (i.e., LBP).

2016

Road Crack Detection Using Deep Convolutional Neural Network

Zhang, L.; Yang, F.; Zhang, Y.D.;
Zhu, Y.J.

A deep-learning-based method for crack detection is proposed in this paper. A supervised deep convolutional neural network is trained to classify each image patch in the collected images.

2016



GRAPHICSTOCK

A Deep Neural Network for Image Quality Assessment

Bosse, S.; Maniry, D.; Wiegand, T.; Samek, W.

This paper presents a no-reference image quality assessment method based on a deep convolutional neural network (CNN). The CNN takes un-preprocessed image patches as an input and estimates the quality without employing any domain knowledge. By that, features and natural scene statistics are learned purely data driven and combined with pooling and regression in one framework.

2016

Fast Multidimensional Image Processing with OpenCL

Oliveira Dantas, D.; Danilo Passos Leal, H.; Oliveira Barros Sousa, D.

VisionGL is an open-source library that provides a set of image processing functions and can help the programmer by automatically generating code. The objective of this work is to augment VisionGL by adding multidimensional image processing support with OpenCL for high performance through use of graphic processing units.

2016

Dimensionality Reduction of Brain Imaging Data Using Graph Signal Processing

Rui, L.; Nejati, H.; Cheung, N.-M.

This paper presents a new dimensionality reduction method based on the recent graph signal processing theory for the task of classifying the brain imaging signals recording the cortical activities in response to visual stimuli. Authors propose using the resting-state measurements (i.e., before onset of the stimulus) of the subjects to build a connectivity graph. The graph Laplacian and graph-based filtering are then applied to learn the low-dimensional linear subspace for the task-state measurements (i.e., after onset of the stimulus).

2016

Moving Object Segmentation Using Depth and Optical Flow in Car Driving Sequences

Kao, J.-Y.; Tian, D.; Mansour, H.; Vetro, A.; Ortega, A.

In this paper, based on an analysis of motion vanishing points of the scene and estimated depth, a geometric model that relates extracted two-dimensional (2-D) motion to a three-dimensional (3-D) motion field relative to the camera is

derived. A constrained optimization problem that considers group sparsity is formulated to recover the 3-D motion field from the 2-D motion. The recovered 3-D motion field is then clustered to provide the segmentation of moving objects.

2016

ORB-SLAM Map Initialization Improvement Using Depth

Fujimoto, S.; Hu, Z.; Chapuis, R.; Aufrère, R.

Map initialization and scale ambiguity are well-known challenging problems for visual simultaneous localization and mapping. In this paper, a triangulation is used on red, green, and blue feature points for getting three-dimensional points from out of the limited area in depth. The authors combined both advantages of triangulation and depth to improve the performance of robustness to initialization and tracking.

2016

Deep Learning Network for Blind Image Quality Assessment

Gu, K.; Zhai, G.; Yang, X.; Zhang, W.

The authors in this paper introduce a new deep-learning-based image quality index (DIQI) for blind quality assessment. Extensive studies are conducted on the new TID2013 database and confirm the effectiveness of their DIQI relative to classical full-reference and state-of-the-art reduced- and no-reference IQA approaches.

2014

Depth Augmented Stereo Panorama for Cinematic Virtual Reality with Focus Cues

Thatte, J.; Boin, J.-B.; Lakshman, H.; Wetzstein, G.; Girod, B.

Cinematic virtual reality aims to provide immersive visual experiences of real-world scenes on head-mounted displays. The authors propose a new content representation, depth augmented stereo panorama, which permits generating light fields across the observer's pupils, achieving an order of magnitude reduction in data requirements compared to the existing techniques.

2016

ICIP 2016 Competition on Mobile Ocular Biometric Recognition

Rattani, A.; Derakhshani, R.; Saripalle, S.K.; Gottemukkula, V.

The aim of this competition is to evaluate and compare the performance of mobile ocular biometric recognition schemes in visible light on a large scale database (VISOB Data Set ICIP 2016 Challenge Version) using standard evaluation methods. Four different teams from universities across the world participated in this competition, submitting five algorithms altogether. The best results were obtained by a team from Norwegian Biometrics Laboratory (NTNU, Norway).

2016

Semantic Context and Depth-Aware Object Proposal Generation

Zhang, H.; He, X.; Porikli, F.; Kneip, L.

This paper presents a context-aware object proposal generation method for stereo images. The authors propose to incorporate additional geometric and

high-level semantic context information into the proposal generation.

2016

Super-Resolution of Compressed Videos Using Convolutional Neural Networks

Kappeler, A.; Yoo, S.; Dai, Q.; Katsaggelos, A.K.

In this paper, for the problem of compressed video superresolution, the authors propose a CNN that is trained on both the spatial and the temporal dimensions of compressed videos to enhance their spatial resolution. Consecutive frames are motion compensated and used as input to a CNN that provides superresolved video frames as output.

2016

Classification of Hyperspectral Image Based on Deep Belief Networks

Li, T.; Zhang, J.; Zhang, Y.

In this paper, deep-learning frameworks, the restricted Boltzmann machine

model, and its deep structure deep belief networks are introduced in hyperspectral image processing as the feature extraction and classification approach.

2014

Image Character Recognition Using Deep Convolutional Neural Network Learned from Different Languages

Bai, J.; Chen, Z.; Feng, B.; Xu, B.

This paper proposes a shared-hidden-layer deep convolutional neural network (SHL-CNN) for image character recognition. In SHL-CNN, the hidden layers are made common across characters from different languages, performing a universal feature extraction process that aims at learning common character traits existing in different languages, such as strokes, while the final softmax layer is made language dependent, trained based on characters from the destination language only.

2014

SP

PANEL AND FORUM (continued from page 13)

Steve Young (sjy@eng.cam.ac.uk) is a professor of information engineering at the University of Cambridge, United Kingdom, and a senior member of technical staff at Apple. His main research interests lie in the area of statistical spoken language systems, including speech recognition, speech synthesis, and dialog management. He is the recipient of a number of awards including an IEEE Signal Processing Society Technical Achievement Award and the IEEE James L. Flanagan Speech and Audio Processing Award. He is a Fellow of the IEEE and the U.K. Royal Academy of Engineering. In addition to his academic career, he has also founded a number of successful start-ups in the speech technology area.

References

- [1] L. Deng and D. Yu, "Deep learning: Methods and applications," in *Foundations and Trends in Signal Processing Series*. Boston, MA: NOW Publishers, 2014.
- [2] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.
- [3] L. Deng and X. Li, "Machine learning paradigms for speech recognition: An overview," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, pp. 1060–1089, May 2013.
- [4] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, pp. 83–98, May 2013.
- [5] A. Sandryhaila and J. M. F. Moura, "Big data processing with signal processing on graphs," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 80–90, Sept. 2014.
- [6] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: Sampling theory," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6510–6523, Dec. 2015.
- [7] F.-L. Luo and C. Zhang, Eds. *Signal Processing for 5G: Algorithms and Implementations*. Chichester, U.K.: Wiley-IEEE Press, 2016.
- [8] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," *J. Mach. Learn. Res.*, vol. 12, pp. 2493–2537, Aug. 2011.
- [9] L. Deng, "Industrial technology advances: Deep learning: From speech recognition to language and multimodal processing," *APSIPA Trans. Signal Inform. Process.*, vol. 5, pp. 1–15, Jan. 2016.
- [10] N. Mrksic, I. Vulic, D. O. Seaghdha, I. Leviant, R. Reichart, M. Gasic, A. Korhonen, and S. Young, "Semantic specialization of distributional word vector spaces using monolingual and cross-lingual constraints," *Trans. Assoc. Computat. Linguistics*, vol. 5, pp. 309–324, Sept. 2017.
- [11] J. A. Fessler. (2017). Medical image reconstruction: A brief overview of past milestones and future directions. [Online]. Available: <http://arxiv.org/abs/1707.05927>
- [12] Y. Liu, J. Chen, and L. Deng, "Unsupervised sequence classification using sequential output statistics," in *Proc. NIPS*, Dec. 2017.
- [13] G. Hinton, et al., "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.

SP