

Editorial

Image Perception

Gloria Menegaz¹ and Guang-Zhong Yang²

¹ *Department of Information Engineering, University of Siena, 53100 Siena, Italy*

² *Department of Computing, Imperial College, London SW7 2AZ, UK*

Received 2 January 2007; Accepted 2 January 2007

Copyright © 2007 G. Menegaz and G.-Z. Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Perception is a complex process that involves brain activities at different levels. The availability of models for the representation and interpretation of the sensory information opens up new research avenues that cut across neuroscience, imaging, information engineering, and modern robotics. The goal of the multidisciplinary field of perceptual signal processing is to identify the features of the stimuli that determine their “perception,” namely “a single unified awareness derived from sensory processes while a stimulus is present,” and to establish associated computational models that can be generalized and exploited for designing a human-centered approach to imaging. In the case of vision, the stimuli go through a complex analysis chain of visual pathways, starting with the encoding by the photoreceptors in the retina (low-level processing) and ending with cognitive mechanisms (high-level processes) that depend on the task being performed. Accordingly, low-level models are concerned with image representation and aim at emulating the way that the visual stimulus is encoded by the early stages of the visual system, as well as at capturing the varying sensitivity to the features of the input stimuli, whereas high-level models are related to image interpretation and allow to predict the performance of a human observer in a predefined task. A global model that accounts for both such bottom-up and top-down approaches would enable an automatic interpretation of the visual stimuli based on both low-level features and semantic contents.

In image processing, methods that take advantage of such models include feature extraction, content-based image description and retrieval, model-based coding, and the emergent domain of medical image perception.

This special issue gives a flavor of the scope and potential of perception-based image and video processing by providing an overview of the way that visual mechanisms at different levels can be modeled and exploited. In particular, the

eleven selected papers span the following fields:

- (1) perceptually plausible mathematical bases for the representation of visual information;
- (2) nonlinear processes and their exploitation in the imaging field (compression, enhancement, and restoration);
- (3) beyond early vision: investigating the pertinence and potential of cognitive models, and semantics.

The majority of the papers in this special issue follow the bottom-up approach. The first group of six papers deal with image representation and propose models for both linear and nonlinear mechanisms to solve classical image processing problems based on early vision. The next three papers take a slightly different perspective, aiming at extracting saliency based on low-level features, whereas the last two papers of this special issue pursue the complementary path and focus on semantics first.

In the paper entitled “Sparse approximation of images inspired from the functional architecture of the primary visual areas,” Sylvain Fischer et al. present a sparse approximation scheme that models the receptive fields of both simple and complex cells while accounting for inhibition and facilitation interactions between neighboring neurons. This allows the handling of classical issues like denoising, compression, and edge detection in an unified framework. It also provides a novel tool for probing cortical functionality.

Along the same line, in the paper “A biologically motivated multiresolution approach to contour detection” by Giuseppe Papari et al., the authors present a contour detection algorithm that combines a Bayesian denoising step with surround inhibition at each level of multiscale image decomposition to solve the problem of oversegmentation which affects classical edge detectors in the presence of textures.

An example of modeling nonlinear processes in the visual system, such as light adaptation and frequency masking, is presented in the paper “Simulating visual pattern detection and brightness perception based on implicit masking,” by Jian Yang, where the author proposes a computational model of the behavior of the contrast sensitivity function (CSF) at varying mean luminance based on a quantitative model of implicit masking. Visual processing is simulated by a front-end lowpass filter, a retinal local compressive nonlinearity, a cortical representation of the stimulus in the Fourier domain, and a frequency-dependent compressive nonlinearity model. The model allows qualitative reproduction of the effects of simultaneous contrast, assimilation, and crispening, demonstrating its potential as a general model for visual processing.

The issue of light adaptation is also addressed in the paper “Pushing it to the limit: adaptation with dynamically switching gain control,” by M. S. Keil and J. Vitrà which presents a model simulating the functional aspects of light adaptation in retinal photodetectors. Given a two-dimensional normalized stimulus, the membrane potential is assumed to be controlled by a differential equation linking its temporal variations with the driving potential, the excitatory input (i.e., the conductance) and the leakage, or passive, conductance. A “dynamically switching gain control” mechanism is controlled by the membrane potential being above or below a given threshold. This leads to an adaptation mechanism mapping luminance values spread over several orders of magnitude onto a fixed target range, typically of one or two orders of magnitude without affecting contrast strength and introducing tedious compression artifacts. Results show that the model is comparative to other state-of-the-art methods in rendering of high-dynamic range images, whilst being faster and more computationally effective.

A different approach to image representation and modeling is presented in the paper “Logarithmic adaptive neighborhood image processing (LA-NIP): introduction, connections to human brightness perception and application issues,” where J.-C. Pinoli and J. Debayle follow the general adaptive neighborhood image processing (GANIP) framework. An interesting aspect of this framework is that it is consistent with several human visual characteristics like intensity range inversion, saturation, Weber’s and Fechner’s laws, psychophysical contrast, and spatial adaptivity, and it leads to competitive results in many image processing tasks like segmentation and denoising.

In the paper “A feedback-based algorithm for motion analysis with application to object tracking,” S. Shah and P. S. Sastry propose a method for selecting regions featuring coherent motion in image sequences. The problem is solved by integrating a feedback mechanism for evidence segregation based on a cooperative dynamical system whose states at each time point represent the current motion. This functional model of object segregation through motion features is plausible for representing neural processing and can lead to robust object tracking even in the presence of dynamic occlusions.

A review of computational vision is presented in the paper “A survey of architecture and function of the primary visual cortex (V1).” In this paper, Jeffrey Ng et al. provide a review of the structure and functionality of neurons in V1, as well as some of the most responded models of early vision and their applications in image processing. They also propose a model for preattentive saliency computation that accounts for intra-cortical interactions related to the “bottom-up” approach of image segmentation in vision.

The same “bottom-up” approach to the extraction of saliency is followed by the paper entitled “An attention-driven model for grouping similar images with image retrieval applications” by Oge Marques et al. In this contribution, two different saliency-based visual attention models (the Stentiford and the Itti models) are combined to derive a biologically plausible algorithm for extracting regions of interest from images. Clustering based on the features extracted from the identified regions are used for grouping. Images containing perceptually similar objects are assigned to the same cluster in a way that is closely related to the users’ expectations.

The exploitation of low-level features for image classification is the subject of the paper “Indoor versus outdoor scene classification using probabilistic neural networks” by Lalit Gupta et al. The authors propose a fully automatic content-based image retrieval (CBIR) system using low-to-mid-level features to distinguish indoor from outdoor scenes. An unsupervised segmentation step based on fuzzy C-means clustering is employed to partition the input image into a suitable number of segments. To this end, the mean and variances of the lowpass versions of the rectified output of a discrete wavelet transform are used. Subsequently, feature vectors are built for each segment by extracting the shape, color, and texture descriptors, and are used as input to a probabilistic neural network. Results show that the most effective feature in this respect is texture, and that the proposed system provides a good classification accuracy.

The last two papers of the special issue follow the complementary “topdown” approach, which starts with the identification of the semantic visual primaries. Accordingly, they are concerned with higher levels of processing of the visual information that deals with perceptual organization. They both focus on the issue of *categorization*.

In the first paper, “A discrete model for color naming,” G. Menegaz et al. propose a discrete computational model for color categorization and naming. The 424-color specimens of the OSA-UCS set are used as the anchor points in the CIELAB color space that is partitioned by a 3D Delaunay triangulation. Each of the 11 *basic* color categories identified by Berlin and Kay is modeled as a fuzzy set. The class membership functions of each OSA-UCS sample are estimated by using the categorization data from the first naming experiment. Linear interpolation is used to predict the membership values of other points in the color space. Automatic naming is obtained by assigning a given test color a *label* that corresponds to the maximum among the

associated membership values. The model is validated both directly via the second naming experiment, and indirectly, through the analysis of its suitability for image segmentation.

Finally, the paper “On the perceptual organization of image databases using cognitive discriminative biplots” by Christos Theoharatos et al. proposes a human-centered approach to image database organization. Instead of deriving image or region descriptors from low-level features, they used a categorization experiment aiming at identifying prototypical images from a set of predefined categories. This transforms the problem to that of learning the structure of class prototypes, which is solved by representing the results in the form of *biplots*, where perceptual similarity is expressed by the distance between points. This simplifies the categorization problem and enables the organization of the entire image database by using the *appending technique*.

ACKNOWLEDGMENTS

On behalf of all the guest editors, we would like to express our sincere gratitude to all those who have contributed to this special issue, all the authors who share with us the vision of human-centric approach to signal processing, and the reviewers who have provided many critical comments and constructive suggestions to the manuscripts submitted. We are sorry that many of the papers submitted are not able to be included in this special issue due to the time constraints and the amount of modification required. We do hope that this special issue will boost the interest of both the image processing and the vision sciences communities whose cross-fertilization is the key to the success of this exciting field of research.

*Gloria Menegaz
Guang-Zhong Yang*

Gloria Menegaz received the M.S. degree in electrical engineering from the Polytechnic University of Milan, Milan, Italy, in 1993, the Postgrade M.S. degree in information technology from the Research and Education Center in Information Technology (CEFRIEL), Milan, in 1995, and the Ph.D. degree in applied sciences from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in July 2000. She is currently an Adjunct Professor at the Department of Information Technology of the University of Siena, Italy. She is a Member of the IEEE, a Member of the scientific committees of several international conferences, and Associate Editor of the EURASIP Journal on Advances in Signal Processing. She has published more than 40 papers and she is the author of two book chapters. Her research interests are primarily in the area of perception-based image processing, modeling of vision and multi-sensory processing, and model-based coding.



Guang-Zhong Yang received the Ph.D. in computer science from Imperial College London and served as a Senior and then as a Principal Scientist of the Cardiovascular Magnetic Resonance Unit of the Royal Brompton Hospital prior to assuming his current full-time academic post at Imperial. He is Director of Medical Imaging and Robotics, Institute of Biomedical Engineering, founding Director of the Royal Society/Wolfson Medical Image Computing Laboratory, and cofounder of the Wolfson Surgical Technology Laboratory at Imperial College. He was also Chairman of the Imperial College Imaging Sciences Centre. His research has been focused on biomedical imaging, robotics, and sensing. He received a number of international awards in medical imaging including the I.I. Rabi Award from the International Society for Magnetic Resonance in Medicine and the Research Merit Award from the Royal Society.

