G U E S T   E D I T O R I A L

Digital Document Imaging

Robert P. Loce
Xerox Corporation
Imaging Science Laboratory
MS147-15B
800 Phillips Road
Webster, New York 14580
E-mail: loce.wbst147@xerox.com

Robert M. Haralick
University of Washington
Applied Physics Laboratory
FT-10, Electrical Engineering Department
Seattle, Washington 98115
E-mail: harlick@ptah.ee.washington.edu

Luc M. Vincent
Xerox Corporation
Desktop Document Systems
3400 Hillview Avenue,
Building 3
Palo Alto, California 94304
E-mail: lucv@adoc.xerox.com

Publishers note: This guest editorial for the special section on digital document imaging was omitted from the January 1996 issue of the Journal of Electronic Imaging. We regret this error, and we thank the guest editors for their contributions.

The special section on digital document imaging includes papers on document image decoding, optical character recognition (OCR), document structure analysis, resolution conversion, image enhancement, and image restoration. Most of the papers were originally presented at the Document Recognition II session of IS&T/SPIE's Symposium on Electronic Imaging: Science and Technology held in San Jose, California, in February 1995.

The first paper is on a Markov source model for printed music decoding, by Kopec, Chou, and Maltz. In the paper, a Markov source model is described for a subset of music notation that was developed as an extended example of the document image decoding (DID) approach to document image analysis. The model is based on the Adobe Sonata music symbol set and a finite-state language of textual music messages. The authors define the music message language and describe several important aspects of message imaging. Aspects of music notation that appear problematic for a finite-state representation are identified. An example of music image decoding and resynthesis using the model is presented.

The next paper by Gader, Mohamed, and Keller applies the Choquet fuzzy integral to handwritten word recognition. The word recognition system that they describe assigns a recognition confidence value to each string in a lexicon of candidate strings. A dynamic programming matcher finds a segmentation of the word image for each string in the lexicon. The traditional match score between a segmentation and a string is an average. In their paper, fuzzy integrals are used instead of an average.

Chen, Haralick, and Phillips contributed a paper on extraction of text words in document images based on a statistical characterization. The document layout analysis algorithm that they describe utilizes probabilistic knowledge to optimize its performance. They first describe a method for generating layout ground truth data from LaTeX files, where the bounding boxes for the characters, words, text lines, and text blocks are represented in hierarchies. These ground truth data enable the construction of statistical models that characterize the various layout structures in the document images. They demonstrate this concept through the development of a word segmentation algorithm, which employs the recursive morphological transform to model word shapes in document images. They also describe experiments that evaluate the performance of their algorithm.

In the next paper, Chen, Bloomberg, and Wilcox describe a system for detecting and locating user-specified search strings, or phrases, in lines of imaged text. The phrases may be single words or multiple words, and may contain a partially specified word. The imaged text can be composed of a number of different fonts and graphics.

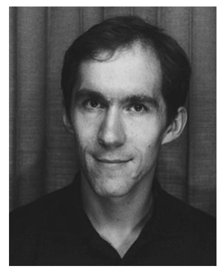
Textlines in a deskewed image are hypothesized using multiresolution morphology. One hidden Markov model is created for each use-specified phrase and another represents all text and graphics other than the user-specified phrases. Phrases are identified using Viterbi decoding on a spotting network created from the models. The operating point of the system can be varied to trade off the percentage of words correctly spotted and the percentage of false alarms.

The paper by Wolberg and Loce first characterizes the degradation of images scanned in the presence of mechanical vibrations. Wolberg and Loce then propose a restoration algorithm that inverts the degradation process and produces a digital image that is an estimate of that same image scanned under ideal conditions. The image restoration algorithm described in their paper makes use of the instantaneous velocity of the scanner array to reconstruct an underlying piecewise constant or piecewise linear model of the image irradiance profile falling onto the sensor array. That reconstructed image is then resampled according to a scanner model operating under ideal conditions. The algorithm is demonstrated on simulated imagery with typical operating parameters.

The final paper in the special section was contributed by Dougherty and Loce. Their paper presents an optimization methodology for binary differencing filters, which are commonly employed in various digital document processing operations. Rather than employing structuring templates that build an entire image, as is done with direct hit-or-miss representation, differencing filters only employ templates that flip values within an image (black to white, white to black). Besides presenting a design methodology, their paper analyzes logic cost and estimation precision. It is also shown how optimal differencing filters may be applied in the digital document setting for image restoration and resolution conversion.

Robert P. Loce is an imaging scientist in the Corporate Research and Technology Division of Xerox Corporation. He joined Xerox in 1981 with an associate degree in optical technology. While working in optical and imaging technology departments at Xerox, he received a BS in photographic science from the Rochester Institute of Technology (RIT) in 1985, an MS in optical engineering from the University of Rochester in 1987, and PhD in imaging science at RIT in 1993. His current work involves simulation of digital reprographic systems and development of image processing algorithms for electronic printing.

Robert M. Haralick is the Boeing Clairmont Egtvedt Professor in Electrical Engineering at the University of Washington. His recent work is in shape analysis and extraction using the techniques of mathematical morphology, robust pose estimation, techniques for making geometric inferences from perspective projection information, propagation of random perturbations through image analysis algorithms, and document analysis. Dr. Haralick served on the faculty of the Electrical Engineering Department at the University of Kansas as a professor from 1975 to 1978. In 1979 Dr. Haralick joined the Electrical Engineering Department at Virginia Polytechnic Institute and State University where he was a professor and director of the Spatial Data Analysis Laboratory. From 1984 to 1986, Dr. Haralick served as vice president of research at Machine Vision International, Ann Arbor, Michigan. He is a fellow of IEEE for his contributions in computer vision and image processing. He serves on the editorial board of *IEEE Transactions on Pattern Analysis and Machine Intelligence*, is an associate editor for the *IEEE Transactions on Image Processing*, and is an associate editor for *Pattern Recognition*. He received a BA in mathematics from the University of Kansas in 1964, a BS in electrical engineering in 1966, and a MS in electrical engineering in 1967. He completed his PhD at the University of Kansas in 1969.



Luc M. Vincent is the director of software development at Xerox Desktop Document Systems Division in Palo Alto, California, where he heads the Image Processing Core Technology group. He received the engineering degree from Ecole Polytechnique, France, in 1986, and the doctorate in mathematical morphology from the Ecole des Mines de Paris in 1990, where he studied under Prof. J. Serra. Dr. Vincent then worked at Harvard University as a postdoctoral fellow, and joined Xerox in 1991. Over the past five years, he has worked on various aspects of optical character recognition and document image analysis, with emphasis on preprocessing and segmentation issues. The page segmentation system he contributed to the development of ranked highest at the 1995 International OCR Accuracy Contest organized by the University of Nevada at Las Vegas. Dr. Vincent has published more than 50 papers in the areas of image processing, morphology, and document analysis. He is currently an associate editor of the *Journal of Electronic Imaging*. Since 1993, he has been chair of the SPIE Document Recognition conference.