Preface

INTRODUCTION

Computer Vision (CV) and Image Processing (IP) are distinct, but at the same time closely related fields. These are being treated as independent fields of studies among the community worldwide. Due to vital importance, effectiveness, and usefulness, extensive advances and discoveries are taking place everywhere. One can find a huge community worldwide contributing to these fields. There is plenty of literature available in the form of books, journals, conference proceedings, Websites, and so forth. One can find many events being held worldwide, numerous laboratories in function, and plenty of workgroups working in these fields. This book is specifically dedicated to the advances in CV and IP.

COMPUTER VISION

Computer vision is a field that includes methods for acquiring, processing, analyzing, and understanding images and, in general, high-dimensional data (Medioni & Kang, 2004; Hartley & Zisserman, 2003; Trucco & Verri, 1998) from the real world in order to produce numerical or symbolic information (Wikipedia, 2012; Shapiro & Stockman, 2001; Morris, 2004). A theme in the development of this field has been to duplicate the abilities of human vision by electronically perceiving and understanding an image (Jahne & Haubecker, 2000). Computer vision is a scientific as well as technological field (Wikipedia, 2012; Shapiro & Stockman, 2001; Morris, 2004; Jahne & Haubecker, 2000; Sonka, Hlavac, & Boyle, 2008). As a scientific field, it is concerned with the theory behind artificial systems that extract information from images. The image data can take many forms, such as video sequences, views from multiple cameras, or multi-dimensional data from a medical scanner. As a technological discipline, computer vision seeks to apply its theories and models to the construction of computer vision systems. Examples of applications of computer vision include systems like detecting events for visual surveillance, navigation of a robot, and automatic inspection in manufacturing applications.

IMAGE PROCESSING

Digital image processing is the use of computer algorithms to perform image processing on digital images. As a subcategory or field of digital signal processing, digital image processing has many advantages over analog image processing. It allows a much wider range of algorithms to be applied to the input data and can avoid problems such as the build-up of noise and signal distortion during processing. Since images
are defined over two dimensions (perhaps more), digital image processing may be modeled in the form of multidimensional systems.

Digital image processing allows the use of much more complex algorithms, and hence, can offer both more sophisticated performance in simple tasks and the implementation of methods that would be impossible by analog means. In particular, digital image processing is the only practical technology for classification, feature extraction, pattern recognition, projection, and multi-scale signal analysis. Some techniques that are used in digital image processing include pixilation, linear filtering, principal components analysis, independent component analysis, hidden Markov models, anisotropic diffusion, partial differential equations, self-organizing maps, neural networks, and wavelets.

LATEST ADVANCES

The chapters in this comprehensive reference explore the latest developments, methods, and approaches to computer vision and image processing in a wide variety of fields and endeavors, providing researchers, academicians, and readers of all backgrounds with an in-depth discussion of the latest advances in imaging tools and technologies.

A. Radgui et al. begin the book with a discussion of their “Adapted Approach for Omnidirectional Egomotion Estimation.” Egomotion estimation is based principally on the estimation of the optical flow in the image. Recent research has shown that the use of omnidirectional systems with large fields of view allow overcoming the limitation presented in planar-projection imagery in order to address the problem of motion analysis. For omnidirectional images, the 2D motion is often estimated using methods developed for perspective images. This chapter adapts motion field calculated using the adapted method that takes into account the distortions existing in the omnidirectional image. This 2D motion field is then used as input to the egomotion estimation process using spherical representation of the motion equation. Experimental results are shown and a comparison of error measures are given to confirm that succeeded estimation of camera motion will be obtained when using an adapted method to estimate optical flow.

This is followed by “Adaptive Kansei Search Method Using User’s Subjective Criterion Deviation” by Yoshitaka Sakurai et al. Sensibility-vectors (kansei-vectors) are useful for retrieving objects like pictures, music, perfumes, and apparel on the Internet. The sensibility-vector is an array of values, each indicating a degree of feeling or impression represented as a sensibility word or kansei word. However, even such an approach leaves a gap between user’s subjective sensibility (image, feeling) value and the corresponding one stored in the database. This chapter proposes a search method to automatically and adaptively decrease such gaps by estimating a Subjective Criterion Deviation (SCD) of the user’s search histories and fuzzy modeling. Conventional methods need tests and questionnaires beforehand to infer a user’s individual sensibility to his or her instinct or impression. The proposed method automatically decreases such gaps without users’ burdens caused by such conventional methods as requiring questionnaires. Moreover, this method reflects the dynamic changes in a user’s preferences. Namely, this method does not need to know a user’s preferences beforehand with questionnaires. An experiment was conducted by building and using a perfume search system. Experimental data results showed that the proposed method is effective.

Motivated by recent results in Joint Source/Channel (JSC) coding and decoding, chapter 3, “Chase-Like Decoding of Arithmetic Codes with Applications” by Amin Zribi et al., addresses the problem of soft input decoding of Arithmetic Codes (AC). A new length-constrained scheme for JSC decoding of
these codes is proposed based on the Maximum a posteriori (MAP) sequence estimation criterion. The new decoder, called Chase-like arithmetic decoder is supposed to know the source symbol sequence and the compressed bit-stream lengths. First, Packet Error Rates (PER) in the case of transmission on an Additive White Gaussian Noise (AWGN) channel are investigated. Compared to classical arithmetic decoding, the Chase-like decoder shows significant improvements. Results are provided for Chase-like decoding for image compression and transmission on an AWGN channel. Both lossy and lossless image compression schemes were studied. As a final application, the serial concatenation of an AC with a convolutional code was considered. Iterative decoding performed between the two decoders showed substantial performance improvement through iterations.

“Automatic Facial Expression Recognition by Facial Parts Location with Boosted-LBP” by Yi Ji and Khalid Idrissi proposes an automatic facial expression recognition system, which uses new methods in both face detection and feature extraction. In this system, considering that facial expressions are related to a small set of muscles and limited ranges of motions, the facial expressions are recognized by these changes in video sequences. First, the differences between neutral and emotional states are detected. Faces can be automatically located from changing facial organs. Then, LBP features are applied and AdaBoost is used to find the most important features for each expression on essential facial parts. At last, SVM with polynomial kernel is used to classify expressions. The method is evaluated on JAFFE and MMI databases. The performances are better than other automatic or manual annotated systems.

Next, “An Evaluation Framework and a Benchmark for Multi/Hyperspectral Image Compression” by Jonathan Delcourt et al. benchmarks three multi/hyperspectral image compression approaches: the classic Multi-2D compression approach and two different implementations of the 3D approach (Full 3D and Hybrid). All approaches are combined with a spectral PCA decorrelation stage to optimize performance. These three compression approaches are compared within a larger comparison framework than the conventionally used PSNR, which includes eight metrics divided into three families. The comparison is carried out with regard to variations in bitrates, spatial, and spectral dimensions variations of images. The time and memory consumption difference between the three approaches is also discussed. Results of this comparison show the weaknesses and strengths of each approach.

In the following chapter, Maria Kulikova et al. describe “A Marked Point Process Model Including Strong Prior Shape Information Applied to Multiple Object Extraction from Images.” Object extraction from images is one of the most important tasks in remote sensing image analysis. For accurate extraction from Very High Resolution (VHR) images, object geometry needs to be taken into account. A method for incorporating strong yet flexible prior shape information into a marked point process model for the extraction of multiple objects of complex shape is presented. To control the computational complexity, the objects considered are defined using the image data and the prior shape information. To estimate the optimal configuration of objects, the process is sampled using a Markov chain based on a stochastic birth-and-death process on the space of multiple objects. The authors present several experimental results on the extraction of tree crowns from VHR aerial images.

Saoussen Ben Jabra and Ezzeddine Zagrouba, in their chapter “A Robust Embedding Scheme and an Efficient Evaluation Protocol for 3D Meshes Watermarking,” propose two main contributions. In the first one, a 3D mesh watermarking using Maximally Stable Meshes detection and multi-signatures embedding is presented. The originality of this scheme is to detect the attack type applied on marked mesh. In plus, it is robust against numerous attacks, blind and invisible. The proposed scheme uses the Maximally Stable Meshes (MSMs) to insert a signature. After MSMs detection using an extension of Maximally Stable Efficient Regions, three MSMs are selected to be marked. Then, three different
signatures are embedded using three different watermarking schemes. This embedding allows knowing the type of the applied attack by detecting which of the signatures resisted. In more, it maximizes robustness by profiting from advantages of every scheme. The second contribution is a new evaluation protocol for 3D watermarking which allows generating a performance score for 3D mesh watermarking schemes. This protocol is based on six criteria having different weights in performance score computing. Finally, this protocol is used to evaluate the proposed watermarking scheme and to compare it with other algorithms. The obtained results verified the good performances of the proposed algorithm, which presents the highest score.

Chapter 8, “Direct 3D Information Determination in an Uncalibrated Stereovision System by Using Evolutionary Algorithms” by Alain Koch and Albert Dipanda, proposes a 3D panoramic shape reconstruction method based on an Uncalibrated Stereovision System (USS) composed of five cameras circularly located around the object to be analysed. First, some interesting points are detected from markers placed on the object such that they are visible by two successive cameras of the USS. These points are then matched on both images acquired by a couple of successive cameras. This process is repeated for all the couples of cameras. Second, by using an evolutionary algorithm, the depth values of the different interesting points are calculated. A comparison with a traditional method based on calibrated cameras validates the accuracy of 3D information provided by the proposed method. Finally, by combining all the interesting points, a panoramic view of the object is obtained.

In “Construction of 3D Triangles on Dupin Cyclides,” Bertrand Belbis et al. consider the conversion of the parametric Bézier surfaces, classically used in CAD-CAM, into patches of a class of non-spherical degree 4 algebraic surfaces called Dupin cyclides, and the definition of 3D triangle with circular edges on Dupin cyclides. Dupin cyclides were discovered by the French mathematician Pierre-Charles Dupin at the beginning of the 19th century. A Dupin cyclide has one parametric equation, two implicit equations, and a set of circular lines of curvature. The authors use the properties of these surfaces to prove that three families of circles (meridian arcs, parallel arcs, and Villarceau circles) can be computed on every Dupin cyclide. A geometric algorithm to compute these circles so that they define the edges of a 3D triangle on the Dupin cyclide is presented. Examples of conversions and 3D triangles are also presented to illustrate the proposed algorithms.

The next chapter, “Evaluation Approach of Arabic Character Recognition” by Hanan Aljuaid et al. proposes and contributes towards designing a complete system for off-line Arabic character recognition. The proposed system is specifically meant for Arabic handwriting recognition, but it equally works for the typed character recognition. It has various phases including preprocessing and segmentation. It also includes the thinning phase and finds vertical and horizontal projection profiles. The recognition phase is managed by a genetic algorithm. The genetic algorithm stands on a feature extraction algorithm that defines six features for each segment. The algorithm for Arabic handwriting recognition obtained a 90.46 recognition rate. The proposed system has been compared with other systems in the literature. It has achieved the second best recognition rate.

In the edge detection, the classical operators based on the derivation are sensitive to noise that causes detection errors. It is even more erroneous in the case of omnidirectional images, due to geometric distortions caused by the used sensors. “Edge Detection by Maximum Entropy: Application to Omnidirectional and Perspective Images” by Ibrahim Guelzim et al. proposes a statistical method of edge detection invariant to image resolution applied to omnidirectional images without preliminary treatments. It is based on the entropy measure. The authors compared its behavior with existing methods on omnidirectional images and perspectives images. The criteria of comparisons are the parameters of
Fram and Deutsch. For omnidirectional images, the authors used two types of neighborhoods: fixed and adapted to the parameters of the sensor. The authors compared the results of detection visually. The tests are performed on grayscale images.

Next, Jyoti Malik et al. explore “Sliding Window Based Fast Corner Matching Palmprint Authentication.” Authentication time is the main and important part of the authentication system. Normally the response time should be fast but as the number of persons in the database increases, there is probability of more response time taken for authentication. The need for fast authentication system arises so that authentication time (matching time) is much less. This chapter proposes a sliding window approach to make a fast authentication system. The highlight of the sliding window method is constant matching time, fast, and can match translated images also. Several palmprint matching methods like match by correlation, and others, are dependent upon the number of corners detected and so is the matching time. In the sliding window method, matching time is constant as the numbers of matching operations are limited and the matching time is independent of the number of corners detected. The palmprint corner features are extracted using two approaches: Phase Congruency Corner Detector and Harris Corner Detector, which are binarized so that only useful information (features) is matched. The two approaches of Phase Congruency Corner Detector and Harris Corner Detector, when matched with hamming distance using sliding window, can achieve the recognition rate of 97.7% and 97.5%, respectively.

Muhammad Ali Norozi then presents a review of “Faster Ranking Using Extrapolation Techniques.” Extrapolations are techniques in linear algebra that require little additional infrastructure that must be incorporated in the existing query-dependent Link Analysis Ranking (LAR) algorithms. Extrapolations in LAR settings rely on the prior knowledge of the (iterative) process that created the existing data points (iterates) to compute the new (improved) data point, which leads to the desired solution faster than the original method. In this chapter, the author presents novel approaches using extrapolation techniques to speed-up the convergence of query-dependent iterative methods, link analysis-based ranking methods, where hyperlink structures are used to determine relative importance of a document in the network of inter-connections. The author uses the framework defined in HITS and SALSA and proposes the use of different Extrapolation techniques for faster ranking. The chapter improves algorithms like HITS and SALSA using Extrapolation techniques. With the proposed approaches, it is possible to accelerate the iterative ranking algorithms in terms of reducing the number of iterations and increasing the rate of convergence.

In “A Semi-Supervised Metric Learning for Content-Based Image Retrieval,” I. Daoudi and K. Idrissi propose a kernel-based approach to improve the retrieval performances of CBIR systems by learning a distance metric based on class probability distributions. Unlike other metric learning methods, which are based on local or global constraints, the proposed method learns for each class a nonlinear kernel that transforms the original feature space to a more effective one. The distances between query and database images are then measured in the new space. Experimental results show that the kernel-based approach not only improves the retrieval performances of kernel distance without learning, but also outperforms other kernel metric learning methods.

Next, a novel Adaptive Lossy Image Compression (ALIC) technique is proposed to achieve high compression ratio by reducing the number of source symbols through the application of an efficient technique. “Image Compression Technique for Low Bit Rate Transmission” by Shaimaa A. El-said et al. proposes an algorithm based on processing the Discrete Cosine Transform (DCT) of the image to extract the highest energy coefficients in addition to applying one of the novel quantization schemes proposed in the present work. This method is straightforward and simple. It does not need complicated
calculation; therefore, the hardware implementation is easy to attach. Experimental comparisons are carried out to compare the performance of the proposed technique with those of other standard techniques such as the JPEG. The experimental results show that the proposed compression technique achieves high compression ratio with higher peak signal to noise ratio than that of JPEG at low bit rate without the visual degradation that appears in case of JPEG.

Shervan Fekri Ershad writes about a “Defect Detection Approach Based on Combination of Histogram Segmentation and Probabilistic Estimation Technique.” Defect detection is a problem that has received much attention from image processing scientists. If defect detection is done correctly and accurately, it can be used in visual inspection systems. Consequently, many different methods have been proposed to solve this problem. Most of these methods have computation complexity, and they do not guarantee for accurate results in every application. In this chapter, an approach is proposed for defect detection based on a main box including two stages. In the first stage, the image is divided into the two groups of pixels based on histogram segmentation, and in the second stage, the defects in the defected group are highlighted based on probability estimation algorithm by using a train stage. This approach by using a train set guarantees the accurate result for nearly all of the applications of defect detection. To prove the quality of proposed approach, the algorithm is applied on stone images to detect the porosity.

The exponential growth of digital image data has created a great demand for effective and efficient schemes and tools for browsing, indexing, and retrieving images from a collection of large image databases. To address such a demand, S. Sathiya Devi and R. Krishnamoorthi propose a new content based image retrieval technique with the orthogonal polynomials model in their chapter, “Rotation Invariant Texture Image Retrieval with Orthogonal Polynomials Model.” The proposed model extracts texture features that represent the dominant directions, gray level variations, and frequency spectrum of the image under analysis and the resultant texture feature vector becomes rotation and scale invariant. A new distance measure in the frequency domain called Deansat is proposed as a similarity measure that uses the proposed feature vector for efficient image retrieval. The efficiency of the proposed retrieval technique is experimented with the standard Brodatz, USC-SIPI, and VisTex databases, and is compared with Discrete Cosine Transform (DCT), Tree Structured Wavelet Transform (TWT), and Gabor filter-based retrieval schemes. The experimental results reveal that the proposed method performs well with less computational cost.

In “Character Segmentation Scheme for OCR System: For Myanmar Printed Documents,” Htwe Pa Pa Win et al. describe automatic machine-printed Optical Characters or texts Recognizers (OCR), which are highly desirable for a multitude of modern IT applications, including Digital Library software. However, the state of the art OCR systems cannot do for Myanmar scripts as the language poses many challenges for document understanding. Therefore, the authors design an Optical Character Recognition System for the Myanmar Printed Document (OCRMPD), with several proposed techniques that can automatically recognize Myanmar printed text from document images. In order to get a more accurate system, the authors propose the method for isolation of the character image by using not only the projection methods but also structural analysis for wrongly segmented characters. To reveal the effectiveness of the segmentation technique, the authors follow a new hybrid feature extraction method and choose the SVM classifier for recognition of the character image. The proposed algorithms have been tested on a variety of Myanmar printed documents and the results of the experiments indicate that the methods can increase the segmentation accuracy as well as recognition rates.

Finally, in “Detecting Corners for 3D Objects” by Misbah Irshad et al., a corner detection algorithm for 3D objects is presented. This algorithm is an extension of the corner detection scheme for planar objects (Beus & Tiu, 1987). This algorithm finds corners and other high curvature points for 3D objects.
CONCLUSION

In all, the chapters in this book present readers with a full view of some of the most up-to-date discoveries in the fields of computer vision and image processing. The mix of practical applications and theoretical research included in this reference volume effectively illustrates the best practices alongside novel techniques, encouraging practitioners and educators to join in creating the next generation of machine and robotic vision applications.

Muhammad Sarfraz
Kuwait University, Kuwait

REFERENCES


ENDNOTES

