C.A. Kulikowski

Department of Computer Science,
Rutgers University,
New Brunswick NJ, USA

Synopsis

Image and Signal Processing

Introduction

Medical image and signal analysis is a broad informatics field that is becoming progressively more specialized as an increasingly wider variety of advanced imaging modalities is developed to help visualize the structures and recently some of the functions of the human body. Imaging integrates signals from adjacent regions of the body, usually focusing on organs or tissues suspected of being involved in pathology, though whole-body imaging is frequently carried out tomographically as well.

With technological advances, it is also more frequent for multiple signal and imaging modality studies to be carried out on the same patient, a process that not only assists in diagnosis, but is often critical to guiding the choice of treatment. In the case of radio- and lasertherapy, accurate and detailed imaging is essential in planning the positioning of radiation beams, so they might destroy targeted pathological tissue with minimal damage to adjacent healthy tissue. Under these circumstances, it is hardly surprising that one of the most important aspects of image and signal processing is ensuring that repeated observations of the target anatomical regions be as carefully located and superimposed (or registered) as possible.

The referees of the IMIA Yearbook have selected four papers in the general field of signal and image analysis for this 1999 issue, two describing new algorithms for image registration integrated within complex radiation treatment systems; one applying a classical pattern-recognition approach to the localization of anatomical landmarks, and the last demonstrating continuing progress in one of the earliest automated signal interpretation problems of medical informatics: the analysis of ECG waveforms.

The first paper, "Image processing algorithms for retinal montage, synthesis, mapping, and real-time location determination" [1] demonstrates the power of image processing methods in producing a clinically useful montage of video retinal angiogram data for guiding laser retinal surgery. It exemplifies the increasing maturity of the field, drawing from a wide variety of image processing techniques in acquisition, object detection, recognition, matching and montage, as well as validation, and refinement of image registration transformations to produce a wide-area analysis. It uses a judicious mix of hardware- and software-based methods to ensure speed when it is needed (when the laser is turned on for surgery) with accuracy checked against a carefully updated wide-area montage of the retina, validated cummulatively throughout the procedure.

The integration of techniques is novel and addresses an important medical problem: choroidal neovascularization (CNV), for which existing procedures have been less than adequate, failing in about 50% of cases. The problem is significant because it is associated with macular degeneration in aging, various post-traumatic and post-inflammatory conditions, degenerative myopia and others, making it a fairly common cause of blindness. Laser cauterization of the neovascularized vessels is the current treatment. The surgeon must register the memorized monochrome image on which the treatment region is mapped with the corresponding region on the patient's retina, seen in color through a slit-lamp microscope, and then press a pedal to activate the laser beam. Complete and accurate localization of all CNV vessels is difficult under these circumstances, but it is needed to avoid recurrent treatments and a prognosis that may be worse than that associated with no treatment at all. Automated registration and montage of multiple retinal images captured by video is essential to provide a wide-enough field of view to ensure completeness of coverage, and serve as input to computer-controlled instruments that deliver the appropriate dose of laser energy to the retinal regions affected by CNV. Such an instrument is currently being developed by the authors. This paper describes the automated imaging techniques and their testing in prototype form.

While computer-based retinal imaging has been the subject of active clinical and bioengineering studies for about 30 years, it is only in the past decade that more general computer vision algorithms have been developed for integrated mapping, montage, and recognition. The present authors have combined a suite of methods, enabling control of tradeoffs between speed of operation and the accuracy of localization and robusteness under high image variability due to ocular movements and illumination changes. Central to their work is the recognition that accurate real-time location determination is only essential during laser activation. An initial, highly accurate wide-area reference montage of the patient's retina is used to guide tracking with fast, suboptimal techniques, which are continuously monitored and corrected when possible errors in tracking are detected. The laser is only turned on when there is very high confidence in the recognition of vascular landmarks and the robustenss of their tracking over successive frames, and it can be turned off at the slightest hint of questionable results, such as inconsistencies in updated montage registration.

From an informatics perspective, the main algorithms involved are for:

1. rapid detection and characterization of vascular landmarks (branching and cross-over points), using a simple edge-dispersion measure around suspected landmarks, avoiding more time-consuming segmentation, thinning and skeletonization of vessels. Edge direction histograms are used to obtain unique signatures for each landmark point for subsequent tracking;
2. rapid matching of sets of vascular landmark points, by scoring possible transformations between pairs of points, using a method that rapidly rejects bad correspondences and then saves only the top five most similar transformations between sets of point clusters. using a least squares matching method;
3. sequential similarity detection based on absolute differences between normalized intensities, evaluating deviations between successive frames and helping eliminate poorly matched transformations arising from approximations in the earlier steps through rescaling operations;
4. wide-area montage and map synthesis built cummulatively from an acceptable initial image of the retina;
5. real-time localization and tracking using a suboptimal 1-D local template-based method, which assumes small movements and fixed magnification, but is over-ruled and corrected by the more accurate point-based methods when a highly sensitive confidence level on the template matching fails to be met.

The laser can be turned off within 3.7 ms when this condition is detected, a speed the authors report is an order of magnitude faster than required for real-time operation. The tracking algorithm was evaluated against six manually located landmark points, yielding a mean squared difference of 1.9 pixels, which is close to the 1.35 pixels reported as the inherent error of the manually determined points.

In "An automatic six-degree-of-freedom image registration algorithm for guiding frameless stereotaxic radiosurgery" [2] the author presents an algorithm for carrying out image-guided frameless stereotaxic neurosurgery and has tested it through simulations that correlate well with phantom studies in a robot-based radiosurgery system. Completely automating image registration in frameless stereotaxic surgery is essential for resolving the conflicting needs of accurately locating targeted anatomical regions of interest (traditionally done with fiducial points on a frame attached to the patient's head) and the surgeon's requirement of easy and open access to as much of the cranium as might be required in the course of a procedure. Additional advantages of the frameless system are in facilitating finer control of radiation dosing (fractionation) and surgery outside the cranium, as well as in relaxing the same-day requirement for imaging, planning, and treatment, enabling them to be carried out on an ambulatory basis.

Complete automation requires extremely reliable robust alignment and registration to correct for ambiguities and significant errors in position determination. Without the external fiducials of a frame to match planning coordinates to those obtained during treatment, the system is more subject to errors from changes in a patient's position during the course of radiosurgery. Working with a specific frameless system (Cyberknife), the author tests and evaluates three different approaches to registration of digitally reconstructed radiographs (DRRs) during treatment, with the CT study or simulator image from the earlier phase of treatment planning. The selected method consists of an iterative reprojection of pairs of DRRs from a set of perturbed patient poses derived from the CT planning data, until a DRR matches the actual treatment radiograph. This is performed in real time, followed by the extraction of a set of heuristic features corresponding to points along the outer contour of the skull, with a chi-square statistic used for comparison evaluation.

High speed in registration is achieved by identifying a number of small regions of interest encompassing the skull edges, and computing only those pixels in the DRRs. Magnified and demagnified versions of the skull DRR silhouette are subtracted prior to treatment to determine these regions of interest so that only 5 to 10% of the total image area needs processing. A Sobel edge-detector filter is used within each region of interest to determine skull contour edges along rays projected from the image center. The first moment of pixel intensity values along each ray is used to improve the precision of registration based on the skull contour coordinates. Modeling rotations of the skull by rotating camera geometry with reference to the stationary CT study (in an Eulerian convention) during DRR calculation was found to be computationally faster than rotating CT voxel data against a camera-determined coordinate system. The DRR ray-tracing algorithm resolves outer edges of the skull by adjusting integration steplength according to the local gradient of the Hounsfeld numbers.

Besides compensating for differences in contrast between DRRs and the acquired radiographs, the algorithm applies an antialiasing filter that also smoothes the stepwise contour of the skull. Best fit by chi-square minimization for the six free parameters of translational and angular position is carried out by a nonlinear hybrid method of gradient search, combined with a Taylor-series approximation to the fitting function near the minimum. A lookup table of DRR feature vectors is computed prior to patient treatment over the whole range of values of the six free parameters to empirically estimate gradients by finite differences. This algorithm typically reached a good minimum in four iterations, requiring five pairs of DRRs when starting from a fixed initial position of zero translation and rotation. Registration for the phantom skull was achieved on a scale of one second, with a translational uncertainty of 0.5 to 1.0 mm, and a rotational uncertainty of 0.6 to 1.3 degrees. This compared favorably to other alternative techniques for registration reported in the literature.

The paper on "Automatic radiographic identification of cephalometric landmarks" [3] reports on a traditional Bayesian posterior probability approach to the identification of cephalometric landmarks for orthodontics and dentofacial orthopedics.

The authors report on the results of this automated method, which they call spatial spectroscopy (SS), based on preliminary tests for detecting 15 craniofacial landmarks from 14 radiographs taken on different X-ray units with a wide range of exposure parameters and techniques.

Features are extracted from the images by convolution with a Gaussian filter and computing the image intensity distribution gradient in the x and y directions. Seventy-five features are obtained this way for characterizing each pixel, and normal densities are fitted to the empirical distributions at known landmark locations. These comprised four cranial base, two maxillary, five mandibular, and four dental landmarks. Each of the 15 landmarks was manually identified by one observer on a computer monitor on five separate days on the set of fourteen 64x64 cephalometric images. The average of these five attempts was used as a reference value (gold standard) for comparison with the automated technique.

The overall results showed no statistically significant differences in mean landmark identification errors between the manual and automatic identification methods, the total mean error for both being less than one pixel (3.1 mm). This is greater than the landmarking error of 1.26 mm reported for manual identification of cephalometric radiographs obtained at high resolution. However, the present study was carried out at minimum resolution (0.16 cm^2/pixel) to limit computational time and memory requirements. Only one landmark (the sella) showed differences between the methods.

Furthermore, landmarks reported in the literature, identified manually with large errors, were found to have comparably large errors on this study, while those with small errors from previous studies were likewise found to have small errors on the automated study. The authors conclude with a discussion of the possible causes of these spatial inaccuracies and the design of future studies.

The last paper pertains to the "analysis of beat-to-beat frequency variability of ECGs using 2-D fourier transforms" [4]. There are considerable difficulties in detecting very faint, fragmented, and delayed high frequency ECG signal components (late potentials) in patients with sustained ventricular tachychardia. The authors [4] have developed an analytical approach called Single-Beat Spectral Variance (SBSV) to characterize the beat-to-beat variability of the frequency components of these ECG segments, while taking advantage of signal averaging to improve the signal-to-noise ratio (SNR) of the measurements. The technique is based on two-dimensional Fourier transforms of 80 ms segments (the first dimension) of 128 consecutive beats each (the second dimension) at the end of the QRS-complex, with the J-point defined as the peak of the autocorrelation function of the spatial velocity. The study reports on an index function derived from the analysis which helps differentiate post-MI patients with and without ventricular arrhythmias from each other and from healthy volunteers.

Signal input is from three bipolar orthogonal ECG leads, recorded simultaneously with a special low-noise and high-gain amplifier, sampled at 1,000 Hz and digitized with 16 bits accuracy. The 2-D Fourier transform is computed after discarding ectopic beats, calculating the J-points and extracting the 80 ms segments of interest, filtering to suppress artificial frequency components introduced by the temporal discontinuity of the sampling, and augmentation by zero-padding of the time segments to 256 ms. Haberl's method of spectro-temporal mapping is then applied, with the beat-to-beat periodicity of the frequency content of the signal characterized by its cycles-per-beat (cpb) plotted against the frequency spectrum of the signal. By not using transformations along the beat axis, the 2-D transform exactly reproduces 1-D transform information after signal averaging by the 2-D transform at zero cpb, while beat-to-beat variability is contained in the non-zero cpb variations of the 2-D spectrum. The authors carried out simulations illustrating the problem of detecting late potentials with incremental, rather than periodic, time delays relative to the QRS-complex, showing how the former (termed Wenckenbach-like conduction patterns) are not detectable after signal averaging.

A late-potential index function (IF) is defined to discriminate those late potentials that are detectable after signal averaging (the constant ones), from those that are not (those with marked Wenckenbach-like conduction patterns). The effectiveness of this index in determining pathology, based on the SBSV data, was tested on 95 subjects divided into three groups. The first included 35 patients in the post-MI phase with well-documented sustained ventricular tachychardia, the second consisted of 50 patients without ventricular arrhythmias in two years post-MI, and the third group comprised 10 healthy volunteers.

Results of the study indicated marked differences in the spectral map and the IF for both individuals and groups of patients in the three categories. In the first group of most at-risk individuals, the SBSV method classified 83% as pathologic vs. 62% by the conventional Simson method. In the second group, 90% were classified as normal vs. 80% for the Simsom method. None of the healthy volunteers had late potentials detected by either method. These results show the promise of this new method of detecting ECG late potentials by the variation in their beat-to-beat spectral characteristics without having to resort to sensitive recording techniques.

Conclusion

Signal and image procesing methods described in these papers are quite varied, depending on the specific type of medical application involved. The image recognition and registration algorithms used for radiation treatment emphasize how careful problem specification and decomposition can establish the steps where simple, rapid, but not necessarily most accurate processing, is essential, as long as it is followed by other steps that filter out false negatives, and that incorporate enough redundancy, feedback, and instrumentation fail-safe control mechanisms to ensure that the final results are highly robust against various sources of noise, and patient and environmental variability.

The first two papers also reinforce a commmon theme that is becoming increasingly important for the practical, automated application of image processing in patient care: the registration of multiple image modalities in real time. It is a significant advance that both these contributions describe prototype implementation of the imaging algorithms within complex, automated instrumentation and control systems. However, systematic investigation of how the algorithms perform in clinical settings when fully integrated with the instrumentation still lies in the future. Reported results from current tests are difficult to consistently compare with those from other instrument systems due to differences in instrumentation, their approach to task specification, and the assumptions behind the techniques and their application. This introduces an interesting challenge for medical informatics. As fully automated imaging and signal processing technologies mature, we can anticipate an increasing need to extend information interchange formats (such as DICOM) currently developed for routinely used clinical imaging modalities to the much wider variety of specialty measurement systems. This presents especially difficult problems in balancing the needs of expressivity in describing the nuances of complex systems and their data, with those of the simplicity and transparency best suited to standardizing any required specifications and data models. The longer-term experience of clinical laboratory systems in these matters may prove useful for study.

Meanwhile, the problem-specific nature of the algorithmic analyses contained in the papers of this section contrasts the continued application of basic pattern recognition for preliminary and straightforward classification problems as opposed to the domain- and problem-specific integration issues that arise in system-embedded imaging and signal processing.

References

1. Becker DE, Can A, Turner JN, Tanembaum HL, Roysam B. Image processing algorithms for retinal montage synthesis, mapping, and real-time location determination. IEEE T Bio-Med Eng 1998;45:105-18.
2. Murphy MJ. An automatic six-degree-of-freedom image registration algorithm for image-guided frameless stereotaxic radiosurgery. Med Phys 1997;24:857-66.
3. Rudolph DJ, Sinclair PM, Coggins JM. Automatic computerized radiographic identification of cephalometric landmarks. Am J Orthod Dentofac Orthop 1998;173-9.
4. Spiegl A, Steinbigler P, Schmucking I, Knez A, and Haberl R. Analysis of beat-to-beat variability of frequency contents in the electrocardiogram using two-dimensional fourier transforms. IEEE T Bio-Med Eng 1998;45:235-41.

Address of the author:
C. A. Kulikowski,
Department of Computer Science,
Rutgers University,
New Brunswick, New Jersey,
USA
E-mail: kulikows@cs.rutgers.edu