

# Sensor, Signal, and Imaging Informatics: Big Data and Smart Health Technologies

S. Voros<sup>1</sup>, A. Moreau-Gaudry<sup>1,2</sup>, Section Editors for the IMIA Yearbook Section on Sensor, Signal and Imaging Informatics

<sup>1</sup> UJF-Grenoble 1 / CNRS / INSERM, TIMC-IMAG UMR 5525, Grenoble, F-38041, France

<sup>2</sup> UJF-Grenoble 1 / CHU / INSERM CIT803, Grenoble, F-38041, France

## Summary

**Objectives:** This synopsis presents a selection for the IMIA (International Medical Informatics Association) Yearbook 2014 of excellent research in the broad field of Sensor, Signal, and Imaging Informatics published in the year 2013, with a focus on Big Data and Smart Health Technologies

**Methods:** We performed a systematic initial selection and a double blind peer review process to find the best papers in this domain published in 2013, from the PubMed and Web of Science databases. A set of MeSH keywords provided by experts was used.

**Results:** Big Data are collections of large and complex datasets which have the potential to capture the whole variability of a study population. More and more innovative sensors are emerging, allowing to enrich these big databases. However they become more and more challenging to process (i.e. capture, store, search, share, transfer, exploit) because traditional tools are not adapted anymore.

**Conclusions:** This review shows that it is necessary not only to develop new tools specifically designed for Big Data, but also to evaluate their performance on such large datasets.

## Keywords

Big data, smart health technologies

Yearb Med Inform 2014;150-3

<http://dx.doi.org/10.15265/IY-2014-0035>

Published online August 15, 2014

## Introduction

The selection is composed of papers addressing innovative approaches in the field of sensor, signal, and imaging informatics with a special focus on Big Data and Smart Health Technologies. Our initial automatic selection returned 1361 articles. We pre-selected 21 articles as candidate best paper in the double blind review process. 15 articles were selected as candidate best papers. The four best papers were selected thanks to an external review process.

The selection first highlights a few examples of technologies allowing to capture patient information. We then present a selection of tools adapted to the exploitation of big data. Finally, we show some of the challenges that still need to be addressed.

## Smart Health Technologies: Towards the Capture of Large Heterogeneous Quantities of Patient Information

The emergence of patient specific sensing tools widens the arsenal of the clinician to acquire relevant data for diagnosis or treatment of his/her patient. These heterogeneous data range from physiological and environmental information for the smart health monitoring of elderly people, as illustrated by the review [1], to medical images such as thermogram images for the detection of breast cancer [2] or forces applied to tissues during needle insertion thanks to innovative robotic assistants [3].

Acquired data can be combined with biomedical models to synthesize patient specific information that will guide diagnosis or treatment. For instance, [4] proposes a fluids dynamics analysis to guide surgeons during aortic replacement in the aim of reducing the risk of aneurysm formation or aneurysmal dilatation. These few examples illustrate that highly multimodal information of different degrees of abstraction can be used to assist the clinician in the care of his patient, resulting in challenges in its acquisition, storage and exploitation.

## Tools for the Exploitation of Big Data in Health

In this synopsis, we will focus on the question of big data exploitation, which we found prevalent during our review process, but the question of data security [1] is also of great importance in the health domain. Three types of tasks involved in big data exploitation were identified: classification, anomaly detection and clustering.

**Classification** consists in determining to which category an observation belongs, based on a training set of observations of known category. In [7] the authors thoroughly compare four supervised learning methods to classify Alzheimer Disease patients and controls, and to predict conversion of mild cognitive impairment to Alzheimer disease on a large database of MRI images (345 participants from the AddNeuroMed cohort). Several characteristics from each patient (age, education, APOE genotype) and from

each patient's MRI were extracted to provide a set of patient features. The authors then constructed 17 models using four supervised learning methods: orthogonal partial least-squares to latent structures, decision trees, artificial neural networks and support vector machines. They performed the classifications using MRI data only or MRI data combined with demographic information or MRI data combined with APOE genotype. The results shows that the performances of the classifiers were similar (~81% accuracy) and that the addition of demographic information of APOE genotype did not improve the classification rates. The classifiers did not perform as well in predicting conversion. [5] reviews predictive data mining approaches on monitoring data from Intensive Care Units (ICU), and shows that most of the approaches for short or medium term predictions are based on classification approaches, treating the time-series data as time-independent. The authors insist on the challenges of reducing false alarm rates and of ensuring good quality data. They stress that the presented approaches lack large-scale validation to be used in clinical settings.

**Anomaly detection** consists in detecting items or events that do not conform to a dataset. Rather than using classification-based approaches for automatic prediction from an ICU monitoring systems like [5], the authors of [6] treat the problem as a novelty detection task, where they model a "normal" patient and identify deviations from this "normality". Several approaches for novelty detection are compared (inference with kernel estimates, modified kernels and one-class support vector machines) on a real large-scale clinical setup (10.000 patients). The limits of common metrics used for the evaluation of monitoring systems, that assume time-independent data, are addressed. [8] also addresses image segmentation as an anomaly detection problem. Based on the observation that multi-atlas segmentation approaches are usually limited to structures represented by the atlases and anatomically consistent among the studied subjects, the authors propose to compute an out-of-atlas likelihood from a multiple-atlas (15 labeled MRI images from the OASIS dataset). This out-of-atlas likelihood is used for the automatic segmentation of malignant glioma

**Table 1** Best paper selection of articles for the IMIA Yearbook of Medical Informatics 2014 in the section 'Sensor, Signal and Imaging Informatics'. The articles are listed in alphabetical order of the first author's surname.

Section
Sensor, Signal and Imaging Informatics
<ul style="list-style-type: none"> <li>▪ Asman AJ, Chambless LB, Thompson RC, Landman BA. Out-of-atlas likelihood estimation using multi-atlas segmentation. <i>Med Phys</i> 2013 Apr; 40(4):043702.</li> <li>▪ Clifton DA, Wong D, Clifton L, Wilson S, Way R, Pullinger R, Tarassenko L. A Large-Scale Clinical Validation of an Integrated Monitoring System in the Emergency Department. <i>IEEE J Biomed Health Inform</i> 2013;17(4):835-42.</li> <li>▪ De Lorenzo D, Koseki Y, De Momi E, Chinzei K, Okamura AM. Coaxial needle insertion assistant with enhanced force feedback. <i>IEEE Trans Biomed Eng</i> 2013 Feb;60(2):379-89.</li> <li>▪ Xie Y, Ho J, Vemuri BC. Multiple Atlas construction from a heterogeneous brain MR image collection. <i>IEEE Trans Med Imaging</i> 2013 Mar;32(3):628-35.</li> </ul>

("abnormalities" compared to a "normal" MRI scan) and for quality control of Diffusion Tensor Images (DTI).

Finally, **clustering** consists in discovering groups or structures of common characteristics in a dataset. Multi-atlas construction consists in automatically generating a set of template images that are representative of a population group and that can then be used for image segmentation, as representatives of the studied population. In [9], the authors propose an innovative approach for multi-atlas construction that provide images that respect rotational invariance and are sharper and more informative compared to existing approaches. The authors argue that the obtained templates are more representative of medical images and thus more useful for medical applications. They also use the MRI OASIS dataset to validate their framework.

## Challenges in the Exploitation of Big Data in Health

Although a lot of challenges were identified in the selected papers, the following especially grabbed our attention. Through a review of automatic delineation methods for radiotherapy [10], the authors argue that current automatic segmentations method are not yet able to match the accuracy of the expert clinician. Approaches incorporating the clinician's expertise and clinically meaningful information are now required, as well as new metrics of image quality characterizing fitness to purpose. As highlighted in [1] and

[5], false positive rates and lack of large-scale evaluations are still limitations for systems to be clinically used. A preliminary step before large-scale validations could be to provide frameworks for common quantitative evaluation methodologies enabling a comparison of methods developed by different groups. Such an approach is proposed by [11], with the development of patient-based synthetic quality assurance CT phantoms for the evaluation of registration methods for radiotherapy planning.

## Conclusion

Smart Health Technologies involving new sensors, new signals (and new signal processing approaches), and new technics of imaging informatics, at the service of Big Data are nowadays a real challenge to improve the care of targeted population of patients. Furthermore, the evaluation of such technologies, particularly on Big Data Sets, is mandatory to hope to be able to use them in daily clinical situation.

## References

1. Baig M, Gholamhosseini H. Smart Health Monitoring Systems: An Overview of Design and Modeling. *J Med Syst* 2013;37(2):1-14.
2. Francis SV, Sasikala M. Automatic detection of abnormal breast thermograms using asymmetry analysis of texture features. *J Med Eng Technol* 2013;37(1):17-21.
3. De Lorenzo D, Koseki Y, De Momi E, Chinzei K, Okamura AM. Coaxial Needle Insertion Assistant With Enhanced Force Feedback. *IEEE Trans*

- Biomed Eng 2013 Feb;60(2):379-89.
4. Heim L, Poole RJ, Warwick R, Poullis M. The concept of aortic replacement based on computational fluid dynamic analysis: patient-directed aortic replacement. *Interact Cardiovasc Thorac Surg* 2013;16(5):583-7.
  5. Güiza F, Van Eyck J, Meyfroidt G. Predictive data mining on monitoring data from the intensive care unit. *J Clin Monit Comput* 2013;27(4):449-53.
  6. Clifton DA, Wong D, Clifton L, Wilson S, Way R, Pullinger R, et al. A Large-Scale Clinical Validation of an Integrated Monitoring System in the Emergency Department. *IEEE J Biomed Health Informatics* 2013;17(4):835-42.
  7. Aguilar C, Westman E, Muehlboeck J-S, Mecocci P, Vellas B, Tsolaki Met, al. Different multivariate techniques for automated classification of MRI data in Alzheimer's disease and mild cognitive impairment. *Psychiatry Res* 2013;212(2):89-98.
  8. Asman AJ, Chambless LB, Thompson RC, Landman BA. Out-of-atlas likelihood estimation using multi-atlas segmentation. *Med. Phys* 2013;40(4):043702.
  9. Xie Y, Ho J, Vemuri BC. Multiple Atlas construction from a heterogeneous brain MR image collection. *IEEE Trans Med Imaging* 2013;32(3):628-35.
  10. Whitfield GA, Price P, Price GJ, Moore CJ. Automated delineation of radiotherapy volumes: are we going in the right direction? *Br J Radiol* 2013;86(1021):20110718.
  11. Nie K, Chuang C, Kirby N, Braunstein S, Pouliot J. Site-specific deformable imaging registration algorithm selection using patient-based simulated deformations. *Med Phys* 2013;40(4):041911.

## Correspondence to:

Sandrine Voros  
 Laboratoire TIMC-IMAG, équipe GIMCAO  
 IN3S, pavillon Taillefer  
 Faculté de Médecine  
 38706 La Tronche Cedex, France  
 Tel: +33 4 56 52 00 09  
 Fax +33 4 56 52 00 55  
 E-mail: Sandrine.Voros@imag.fr

## Appendix: Content Summaries of the Selected Best Papers for the IMIA Yearbook 2014, Section Sensors, Signal and Imaging Informatics

**Asman AJ, Chambless LB, Thompson RC, Landman BA**

**Out-of-atlas likelihood estimation using multi-atlas segmentation**

**Med Phys 2013 Apr; 40(4):043702**

Multi-atlas segmentation consists in combining the registration of multiple expert-segmented example images (atlases), with a target image. The conflicts between the atlases are resolved using different strategies for label fusion. This approach has proven to be robust across a large range of applications, however their scope is limited to structures that are represented by the atlases, and to structures that are anatomically consistent across potential target subjects. The authors propose to estimate the out-of-atlas likelihood, i.e. the likelihood that a multi-atlas segmentation estimate is not representative of the atlas. This estimation can be used to determine and localize an abnormality in a target image. Two applications exploiting this out-of-atlas likelihood are presented. The first application is the detection of malignant gliomas in the human brain: 15 T1-weighted whole brain MRI that have been labeled constitute the atlas. 30 T1-weighted brain MRI scans with malignant gliomas are registered with the atlas, and the out-of-atlas likelihood estimator can consistently and reliably declare image voxels to be malignant. The second application is the detection of large-scale imaging artifacts in diffusion tensor imaging datasets for image “quality control”, for which the authors present encouraging qualitative results.

**Clifton DA, Wong D, Clifton L, Wilson S, Way R, Pullinger R, Tarassenko L**

**A Large-Scale Clinical Validation of an Integrated Monitoring System in the Emergency Department**

**IEEE J Biomed Health Inform**  
**2013;17(4):835-42**

The authors present the large-scale clinical validation of an integrated monitoring system (10 000 patients) that combines electronic patient records with high-rate acquisition of patient physiological data. After having presented the existing standard of care (based on manual “early warning score”- EWS - systems), the authors present the infrastructure that they have deployed to create their monitoring system: a peer-to-peer network of bed-side monitors, hand-held PDAs and wall-mounted touchscreens. The bed-side monitors allow for

the acquisition of vital-sign data using ECG electrodes, pulse oximeters and sphygmomanometers. The hand-held PDAs allow for the input of vital-sign data from manual patient review. Based on the measurements gathered through this infrastructure, the authors propose to identify deviations from normality using a novelty detection approach: a model of “normal” patient physiology is built and test data are compared to this model to raise alerts if they differ significantly from it. Several novelty detection approaches are developed: inference with kernel estimates, modified kernels and one-class support vector machines. The question of the evaluation of these approaches is then discussed, with a highlight on the limits of the existing evaluation metrics when applied to such large-scale, time series data, and an evaluation strategy is proposed. Finally the authors compare their novelty detection approaches with manual EWS systems. They show that the support vector machine is the superior classifier in terms of identifying patient deterioration.

**De Lorenzo D, Koseki Y, De Momi E, Chinzei K, Okamura AM**

**Coaxial needle insertion assistant with enhanced force feedback**

**IEEE Trans Biomed Eng** 2013  
**Feb;60(2):379-89**

Many medical procedures involve the insertion of needles in soft tissues, some of which do not require visual guidance. In such procedures, offering the capacity to sense the efforts the needle exerts on tissues could be very useful (for an estimation of the needle's depth of insertion and the sensing of structures that are penetrated). Sensing the efforts applied at the tip of a needle is very challenging because they are masked by the efforts resulting from the friction between the needle shaft and the surrounding tissues. The authors present a new robotic system for needle insertion that incorporates haptic feedback and is able to enhance the operator's sensing of the forces applied at the needle tip. The proposed system comprises a hollow outer needle in which the coaxial inner needle is inserted. With this construction, the inner

needle is not subjected to the friction of the tissues along its shaft. Several force sensors measure the tip force, the shaft force, the operator force and the actuator force. A control scheme of the robotic assistant is proposed, that amplifies the forces between the needle and tissue, and two force-feedback modes are possible (tip forces or tip and shaft forces). The authors evaluate their system through experiments on a phantom mimicking brain vessels of different thicknesses and placed at a different depth inside an artificial soft tissue. 11 volunteers were asked to perform 36 biopsies each with the feedback mode randomly selected and varying conditions (membrane thickness, membrane depth and amplification factor for the feedback). The volunteers always sensed membrane perforation better when they were only provided with tip force information, and their sensing decreased with the depth of the membrane inside the artificial soft tissue.

**Xie Y, Ho J, Vemuri BC**

**Multiple Atlas construction from a heterogeneous brain MR image collection**

**IEEE Trans Med Imaging 2013**

**Mar;32(3):628-35**

Atlases are informative representatives of a population of images. Computing an atlas from a collection of images is a fundamental problem in medical imaging, which is often formulated as a mean estimation problem. For instance, a common way to compute a multiple atlas is to fit a Gaussian Mixture Model to the input images using the Expectation-Maximization algorithm. A common characteristic of atlases constructed using a “mean” approach is their blurriness and absence of sharp anatomical details. As a result, these atlases do not mimic real images, which could be critical for some applications. The authors give the example of image registration where it could be difficult to find

image correspondences between real sharp images and blurry atlas images. After having identified the causes for blurry images in the common approaches for atlas construction, the authors present a novel algorithm that produces rotationally invariant atlas images retaining clear important anatomical structures common to the subpopulation of images they represent. Their approach is based on manifold learning methods and produces rotationally invariant atlas images. It can also automatically or semi-automatically partition the image collection into a more homogenous sub-collection of images for multi-atlas construction. The approach is validated on MR images from the free OASIS dataset for single atlas and multiple atlas construction and the obtained atlases capture image sharpness and structures better with the proposed approach than with state-of-the-art approaches.