Article:

Frangi, AF orcid.org/0000-0002-2675-528X, Tsaftaris, SA and Prince, JL (2018) Simulation and Synthesis in Medical Imaging. IEEE Transactions on Medical Imaging, 37 (3). pp. 673-679. ISSN 0278-0062

https://doi.org/10.1109/TMI.2018.2800298

Simulation and Synthesis in Medical Imaging

Alejandro F. Frangi, *Fellow, IEEE,* Sotirios A. Tsaftaris, *Member, IEEE,* and Jerry L. Prince, *Fellow, IEEE*

(Invited Editorial)

***Abstract*—This editorial introduces the Special Issue on Simulation and Synthesis in Medical Imaging. In this editorial, we provide working definitions to so-far ambiguous terms of simulation and synthesis in medical imaging. We also briefly discuss the synergistic importance of mechanistic and phenomenological models of medical image generation. Finally, we provide an overview of the twelve papers that were accepted covering both mechanistic (5) and phenomenological (7) medical image generation. This rich selection of papers covers applications in cardiology, retinopathy, histopathology, neurosciences, and oncology. It also covers all mainstream diagnostic medical imaging modalities.**

***Index Terms*—Simulation, Synthesis, Modelling, Imaging.**

I. Introduction

THE medical image community has always been fascinated by the possibility to create simulated or synthetic data upon which to understand, develop, assess, and validate image analysis and reconstruction algorithms. From very basic digital phantoms all the way to very realistic *in silico* models of medical imaging and physiology, our community has progressed enormously in the available techniques and their applications. For instance, mechanistic models (imaging simulations) emulating the geometrical and physical aspects of the acquisition process have been used now for a long time. Advances on computational anatomy and physiology have further enhanced the potential of such simulation platforms by incorporating structural and functional realism to the simulations that can now account for complex spatio-temporal dynamics due to changes in anatomy, physiology, disease progression, patient and organ motion, etc. More recently, developments in machine learning together with the growing availability of ever larger scale databases have provided the theoretical underpinning and the practical data access to develop phenomenological models (image synthesis) that learn models directly from data associations across subjects, time, modalities, resolutions, etc. These techniques may provide ways to address challenging tasks in medical image analysis like cross-cohort normalization, image imputation in the presence of missing or corrupted data, transfer of knowledge across imaging modalities, views or domains. To this date, however, these two main research avenues (simulation and synthesis) remain independent efforts despite sharing common challenges. This special issue provides a birds' eye overview the state-of-the-art in methods and algorithms at the bleeding edge of synthesis and simulation in/for medical imaging research. We hope this collection will stimulate new ideas leading to theoretical links, practical synergies, and best practices in evaluation and assessment common to these two research directions. We solicited contributions from cross-disciplinary teams with expertise, among others, on machine learning, statistical modelling, information theory, computational mechanics, computational physics, computer graphics, applied mathematics, etc.

II. Context and Definitions

It is helpful at this point to define the concepts of *simulation* and *synthesis* in the context of this special issue, that is, in medical imaging. We noted that while the concept of simulation is, in general, very ample and unspecific, on the other side, there was virtually no formal definition for image synthesis.

The concepts of image simulation and synthesis can be ambiguous (or even exchangeable) if one attends to the definitions of these terms by authoritative dictionaries like Oxford (OED) or Merriam-Webster (MWD):

> **Simulation** [OED] *n* • 3. The technique of imitating the behaviour of some situation or process (whether economic, military, mechanical, etc.) by means of a suitably analogous situation or apparatus, esp. for the purpose of study or personnel training.
> **Simulation** [MWD] *n* • 3a: the imitative representation of the functioning of one system or process by means of the functioning of another a computer simulation of an industrial process; b: examination of a problem often not subject to direct experimentation by means of a simulating device.
> **Synthesis** [OED] *n* • 1. Logic, Philos., etc.: a. The action of proceeding in thought from causes to effects, or from laws or principles to their consequences. (Opposed to analysis n. 3).
> **Synthesis** [MWD] *n* • 1 a : the composition or combination of parts or elements so as to form a whole.

The concept of synthesis currently in use in computer vision and medical image analysis contrasts strikingly as almost opposite to that traditionally used in philosophy or science[1]. In computer graphics, *realistic image synthesis* "is the process of creating images that are, in some way, accurate representations

[1]The Oxford English Dictionary provides contextual quotes that illustrate this contrast. For instance, from T. Hobbes in *Elements Philos.* iii. xx. 230, 1656: "Synthesis is Ratiocination from the first causes of the Construction, continued through all the middle causes till we come to the thing it selfe which is constructed or generated.", and from I. Newton in *Opticks* (ed. 2) iii. i. 380, 1718: "The Synthesis consists in assuming the Causes discover'd, and establish'd as Principles, and by them explaining the Phnomena proceeding from them." Source: http://www.oed.com/view/Entry/196574.

A.F. Frangi is with the CISTIB Centre for Computational Imaging & Simulation Technologies in Biomedicine, Electronic & Electrical Engineering Department, University of Sheffield, Sheffield, UK. e-mail: a.frangi@sheffield.ac.uk.

S.A. Tsaftaris is with the Institute of Digital Communications School of Engineering, University of Edinburgh, Edinburgh, UK.

J.L. Prince is with the Image Analysis & Communications Lab, Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, USA.

of a real scene. Often, but not always, the images are meant to be viewed by a human observer. Therefore, the accuracy is with respect to the human visual system. Sometimes the image needs to be predictive, guaranteeing that the viewer would have the same visual experience if they were actually in the scene. In some cases the image need only be plausible, the viewer is convinced that the scene could actually be real." [2] While medical image computing is interested in visually plausible results, one is usually also interested in the quantitative assessment of the synthesised images or, at least, in figures of merit that are meaningful for the intended task (e.g. diagnostics, planning, prognosis, et.c). In the sequel, we attempt to provide some distinction and definition between the concepts of *image synthesis* and *image simulations* based on the literature and praxis of our medical imaging community.

At one level, in using the concepts of simulation and synthesis, our community usually makes a fundamental ontological distinction best described by referring to mechanistic and phenomenological models, respectively. By simulation, we usually start of from first principles while in synthesis we start off with abundant data. We also usually assume behind these concepts a natural information processing direction: from data to models, in the case of synthesis; and from models to data, in the case of simulation. Simulation implies the existence of an abstraction of the knowledge we possess, usually in the form of first principles, that is used to derive instances of that knowledge in an scenario that is fully controlled by the selection of simulation parameters. Synthesis, on the contrary, implies the ability to extract or summarise (synthesise) knowledge from a collection of representative examples from a wider population or phenomenon. This is usually accomplished through statistical or phenomenological models unless a mechanistic model is available in which case on is able to perform data assimilation or parameter identification resulting in a mechanistic model. Conversely, one is able to simulate new image (or shape) examples from an image (or shape) synthesis method but we tend to talk then of generative models and these are usually phenomenological in nature.

We offer the following two definitions:

(Image) **Synthesis** [ours] *n* • The generation of visually realistic and quantitatively accurate images through learning-based generative models of phenomenological mature with application to the problems of interpolation, super resolution, normalisation, modality propagation, etc.

(Image) **Simulation** [ours] *n* • The application of mechanistic first principles from imaging physics, organ physiology, and/or their interaction to produce virtual images that are visually realistic and physically/clinically plausible, and generated under controlled imaging protocols.

III. Mechanistic or Phenomenological?

It is beyond the possibilities of this editorial to review the considerable progress made over the past decades in both physical models of image formation as well as in machine learning techniques for image synthesis. This special issue is a modern and exciting excerpt of the most recent developments. We would like, however, to put these two approaches underpinning these special issue in the wider context of current trends in science and data science.

There are opportunities and limitations in approaching image generation from a mechanistic or a phenomenological standpoint, some of epistemological reach. Some people argue that with increasing availability of big data, computational resources, and breakthroughs in artificial intelligence, data-driven phenomenological models will eventually supersede the need of mechanistic theories [3], while others seriously content this viewpoint [4]. The complexity of image generation process, the need to model detailed and accurately the geometry and physics of imaging as well as the variability and uncertainty associated with anatomical and physiological factors, all seem to favour those challenging the need or feasibility of generating truly accurate medical images from first principles. In Chapter 12 of his book, Helbing [5] presents an interesting cautionary argument that contrasts with Anderson's vision of Big Data (assuming that we no longer will need theory and science). Fig III shows Helbing's model for digital growth in computational resources doubling about every 18 months (Moore's law), and data resources doubling about every 12 months (soon every 12 hours!). While these two resources follow an exponential growth, the complexity of the processes that these resources help to elucidate or decide on (e.g. parametric complexity of the computational methods, ontological complexity of health data) follow a factorial growth as they are based on combinatorial combinations and system networks, respectively. The above implies the problem of "dark data", i.e. the share of data we will not be able to process is increasing with time. As a consequence, we need to know what data to process and how, which requires science and understanding of the underlying mechanisms that relate data and phenomena so that algorithmic complexity can be tractable.

IV. Special Issue Statistics

Twenty-four manuscripts were received for this special issue. Two were immediately rejected while another ten were rejected after a revision round. Twelve papers were final accepted after peer-review covering both mechanistic (5) and phenomenological (7) modelling and data generation. This rich selection of papers covers applications in cardiology, retinopathy, histopathology, neurosciences, and oncology. It also covers all mainstream diagnostic medical imaging modalities. Two manuscripts were dealt with by Associate Editors Mehrdad Gangeh and Hayit Greenspan to avoid potential conflicts of interest. Each paper was reviewed, at least, by three expert reviewers.

[2] Smits B, https://www.cs.utah.edu/~bes/graphics/overview

[3] Anderson C. "The end of theory: the data deluge makes the scientific method obsolete, *Wired*, http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory, Jul 23, 2008

[4] Mazzocchi F. "Could Big Data be the end of theory in science? A few remarks on the epistemology of data-driven science". *EMBO Rep*. 2015 Oct;16(10):1250-5.

[5] Helbing D, Thinking Ahead-Essays on Big Data, Digital Revolution, and Participatory Market Society, Springer, 2015.

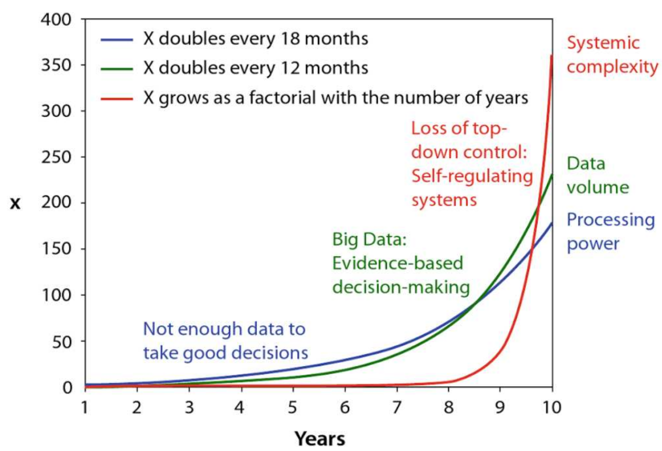

Fig. 1. Helbing's model for digital growth where systemic complexity (e.g. algorithmic parametric complexity and complexity of health data) grows at a factorial rate compared to the exponential rate of data and computing resources. Courtesy of D Helbing. Reprinted with permission.

V. Special Issue Overview

The special issue comprises 12 papers covering both simulation and synthesis. Simulation papers focus on either generating computational phantoms of anatomy or physiology in health and disease, or aim at developing computational phantoms of image formation. In the first category of simulation papers, Seagars *et al.* start off by reviewing what is arguably one of the most widespread digital phantoms in computational human anatomy and physiology of the human thorax. The authors overview the four dimensional (4D) eXtended CArdiac-Torso (XCAT) series of phantoms, which cover a vast population of phantoms of varying ages from newborn to adult, each including parametrised models for the cardiac and respiratory motions. This paper illustrates how these phantoms found great use in radiation dosimetry, radiation therapy, medical device design, and even the security and defence industry. Abadi *et al.* extend upon the capabilities of the XCAT series of computational phantoms, and propose a detailed lung architecture including airways and pulmonary vasculature. Eleven XCAT phantoms of varying anatomy were used to characterize the lung architecture. The XCAT phantoms were utilized to simulate CT images for validation against true clinical data. As the number of organs described as numerical phantoms as XCAT models increases, the potential use of such models as a tool to virtually evaluate the current and emerging medical imaging technologies increases. The paper by García *et al.*, the authors consider the challenging task of evaluating the correlation of the parenchymal patterns (i.e. local breast density) as provided by mammography with MRI volume information. Differences in distributions (MRI versus x-ray) and radical deformation present (due to how the breast is imaged during mammography and MR) render this problem also relevant from a registration perspective. The authors in tackling this challenge, employ a subject-specific biomechanical model of the breast to assist the MRI volumes to X-ray mammograms. When converged, a direct projection of the MR-derived glandular tissue permits the comparison to the corresponding mammogram. Roque *et al.* a reaction-diffusion model of tumour growth. The predicting tumour growth (based on models) and particularly its response to therapy is a critical aspect of cancer care and a challenge in cancer research. In this work, the authors derive an image-driven reaction-diffusion model of avascular tumour growth, that permits proliferation, death and spread of tumour cells, and accounts for nutrient distribution and hypoxia. The model parameters are learned (and evaluated) based on longitudinal time series of DCE-MRI images. Rodrigo *et al.* study the influence of anatomical inaccuracy in the reconstruction of Electrocardiographic Images (ECGI) in non-invasive diagnosis of cardiac arrhythmias. The precise position of the heart inside the body is important for accurate reconstructions but often not accurately known. They explored the curvature of L-curve from the Tikhonov regularization approach, which is one methodology used to solved the inverse problem, and dicovered that optimization of the maximum curvature minimizes inaccuracies in the atrial position an orientation. Such automatic method to remove inaccuracies in atrial position improves the results of ECGI. Moreover, it allows to apply ECGI technology also where the electric recording, usually done via Body Surface Potential Mapping (BSPM) and the anatomical CT/MRI images are not recorded one after another, which could lead to potential expand of ECGI use to larger group of patients. Polycarpou *et al.* propose a digital phantom to synthesise 3D+t PET data using a fast analytic method. The proposed method derives models of cardiac respiration and motion based on real respiratory signals derived from PET-CT images are combined with MRI-derived motion modelling and high resolution MRI images. In addition, this study incorporates changes in lung attenuation at different respiratory cycle positions. The proposed methodology and derived simulated datasets can be useful in the development and benchmarking of motion-compensated PET reconstruction algorithms by providing associated ground-truth of various controlled imaging scenarios.

This issue also comprises several papers using phenomenological or data-driven methods for image synthesis or generating annotated reference datasets. Some methods are hybrid combining both generative with mechanistic approaches. Zhou *et al.*, for instance, undertake to generate realistic synthetic cardiac images, of both ultrasound (US), and cine and tagged Magnetic Resonance Imaging (MRI), corresponding to the same virtual patient. This method develops a synthesis by registration approach where an initial dataset is segmented, transformed and warped -as needed- to generate a motion- and deformation-informed set of both cMRI, tMRI and US. Only the motion model in this method is derived from an actual physical model while the image intensity is created through mapping reference values from literature. In a related paper, Duchateau *et al.* also focus on the automatic generation of a large database of annotated cardiac MRI image sequences. Their approach, like the one of Zhou *et al.*, combines both mechanistic motion models of cardiac electro-mechanics with anatomical augmentation via data-driven non-rigid deformations. The proposed method requires the existence of a small database of cine CMR sequences that serve as seed to augment the anatomical variability by creating simulations of cardiac electro-mechanics under diverse conditions. Augmented data is created by warping image intensities in the original sequence through the electromechanical simulation. This method ensures the material point correspondence between frames complies with a mechanistic electromechanical model yet image appearance is not altered compared to that of the original dataset used. The authors apply this approach to generate a database of subjects myocardial infarction under controlled conditions in infarct location and size. Finally, Mattausch and Goksel's paper focuses on how to reconstruct the distribution of ultrasound image scatterers of tissue samples non-invasively. The recovered scatterer map will inform a realistic ultrasound image simulation under different viewing angles or transducer profiles. The robustness of this technique relies on obtaining images from multiple view points to accurately assess scatterer distribution, without which the forward problem is not accurately solved. Besides an inversion strategy, the authors contribute a novel beam-steering technique to insonate the tissue rapidly and conveniently acquiring multiple images of the same tissue. The authors also demonstrates that the scatterer map offers a new tissue representation more convenient to edit the tissue definition to create controlled variations.

Several papers focus on machine learning for image synthesis to tackle problems as diverse as generating benchmark data, image normalisation, super resolution or cross-modality synthesis, to name just a few. One topic attracting several submissions is adversarial learning. For instance, Costa *et al.* propose a combination of adversarial networks and adversarial auto-encoders to develop synthetic retinal colour images. Adversarial auto-encoders are use to learn a latent representation of retinal vascular trees and generate corresponding retinal vascular tree masks. Adversarial learning, in turn, is use to map these vascular masks into colour retinographies. The authors present a learning approach that jointly learns the parameters of the adversarial network and auto-encoder. The authors extensively validated of the quality of their synthetic images. The data produce can help generating valuable labelled ground-truth data for testing or training retinal image analysis methods. Ben Taieb and Hamarneh also use adversarial learning to address the problem of histopathology normalisation. Recognizing the large variability between staining processes in different histopathology laboratories, the authors propose a method that aims to emulate stain characteristics from one laboratory to the other. Treated as a style transfer problem (to adopt the term from computer vision literature) the authors proposed a deep neural network that learns to map input images to output images that best match the distribution characteristics of a reference set of data, thus achieving stain normalization. A combination of generative, discriminative and task specific networks jointly optimized achieve the desired objective of finding stain normalizations suitable for segmentation or classification tasks. Chartsias *et al.* propose an approach to MRI synthesis that is both multi-input and multi-output and uses fully convolutional neural networks. The model has two interesting properties: it is robust to handle missing data, and, while it benefits from, does not require, additional input modalities. The model was evaluated on the ISLES and BRATS datasets and demonstrate statistically significant improvements over state-of-the-art methods for single input tasks. Using dictionary learning, Huang *et al.* present a method that can synthesize data across modalities using paired and unpaired data. Relying on the power of cross modal dictionaries they establish matching functions that can discover cross-modal sparse embeddings even when unpaired and unregistered data are available. Furthermore, considering that across modalities different distributions may be present, a manifold geometry formulation term is considered. They extensively evaluate their method on two publicly available brain MRI datasets.

VI. Acknowledgements

We thank Associate Editors Gangeh and Greenspan and the numerous reviewers for their contribution to the selection process of the manuscripts in this special issue. We also thank Prof D Helbing from ETH Zürich for pointing us out to interesting literature and sharing his views with AFF.