A virtual imaging platform for multi-modality medical image simulation.

Tristan Glatard, Carole Lartizien, Bernard Gibaud, Rafael Ferreira da Silva, Germain Forestier, Frederic Cervenansky, Martino Alessandrini, Hugues Benoit-Cattin, Olivier Bernard, Sorina Camarasu-Pop, et al.

A Virtual Imaging Platform for multi-modality medical image simulation

Tristan Glatard[1], Carole Lartizien[1], Bernard Gibaud[2], Rafael Ferreira da Silva[1], Germain Forestier[2], Frédéric Cervenansky[1], Martino Alessandrini[1], Hugues Benoit-Cattin[1], Olivier Bernard[1], Sorina Camarasu-Pop[1], Nadia Cerezo[3], Patrick Clarysse[1], Alban Gaignard[3], Patrick Hugonnard[4], Hervé Liebgott[1], Simon Marache, Adrien Marion[1], Johan Montagnat[3], Joachim Tabary[4], and Denis Friboulet[1]

***Abstract*—This paper presents the Virtual Imaging Platform (VIP), a platform accessible at http://vip.creatis.insa-lyon.fr to facilitate the sharing of object models and medical image simulators, and to provide access to distributed computing and storage resources. A complete overview is presented, describing the ontologies designed to share models in a common repository, the workflow template used to integrate simulators, and the tools and strategies used to exploit computing and storage resources. Simulation results obtained in 4 image modalities and with different models show that VIP is versatile and robust enough to support large simulations. The platform currently has 200 registered users who consumed 33 years of CPU time in 2011.**

***Index Terms*—Medical image simulation, ontology, workflows, distributed computing infrastructures.**

I. Introduction

Medical images can be simulated from digital models of the human body for a variety of applications in research and industry, including fast prototyping of new devices and the evaluation of image analysis algorithms [1], [2], [3]. Several image modalities are commonly simulated, among which Magnetic Resonance Imaging (MRI), ultrasound imaging (US), Positron Emission Tomography (PET), and Computed Tomography (CT).

However, image simulation remains mastered only by a few, due to the variety, complexity and heaviness of simulation pipelines. Both the simulation code and the physical model of the imaged object can be very elaborate and specific to the imaging modality. They are usually not designed by the same person or group, and need to be shared. Sharing models requires consistent descriptions so that they can be reused in simulations of different modalities. This is challenging because simulators need different types of information, for instance relaxation times for MR, or radiopharmaceutical activity for PET. Moreover, models contain information not only about anatomy but also about pathologies and external entities present in the body during image acquisition.

Additionally, computing times and the volume of produced data also limit the level of realism and the size of the simulation scene. Although distributed computing infrastructures (DCI) can help in supporting the computing load and storing the generated data, using them should not become an additional burden to simulation users. High-level interfaces should allow their transparent exploitation, together with customized deployments and optimization.

A platform making simulations less ominous, especially for newcomers, is therefore needed. This platform should facilitate access to models and physical parameters, to simulators of different modalities, and to appropriate computing power and storage. This paper describes VIP, an openly-accessible online platform sharing models and simulators, and supporting the execution of heavy simulations. It extends the summary presented in [4] with a more detailed platform description and additional simulations illustrating different image modalities.

The manuscript is organized as follows. After a review of related work in Section II, we describe in Section III how ontologies were created to describe models in VIP. In Section IV we show how a workflow-based template was used to integrate simulators of 4 different imaging modalities and deploy them on a DCI. Methods and techniques employed to support simulation execution are described in Section V, highlighting computation management, data provenance recording, storage of models and simulated data, and interface. Finally, Section VI exemplifies VIP on PET, US, MR and CT simulations of various models. Usage statistics are also reported.

II. Related work

Several databases of simulated images are available for whole-body or brain PET imaging [5], [6], [7], [8], and for brain MRI, e.g. Brainweb [9], offering an online simulation service for MRI. The platform in [10] provides simulation-based interactive tutorials on imaging modalities. Simulators are usually dedicated to a particular modality, for instance SPECT [11], CT [12], mammography [13], MRI [14], or US [15], [16], [17], with the notable exception of GATE [18], which was validated for PET, SPECT, CT, and even radiation therapy. VIP targets a multi-modality online simulation service relying on existing codes.

The sharing of object models is important for initiatives related to the Virtual Physiological Human (VPH) program. FieldML [19] was developed for this program as a standard to represent information contained in models. It can structure repositories and databases such as those available in euHeart [20], focusing on the modeling of the heart. Several

[1]Université de Lyon, CREATIS; CNRS UMR5220; Inserm U1044; INSA-Lyon; Univ. Lyon 1, France
[2]INSERM / INRIA / CNRS / Univ. Rennes 1, VISAGES U746, Rennes, France
[3]CNRS / UNS, I3S laboratory, MODALIS team, Sophia Antipolis, France
[4]CEA-LETI-MINATEC, Recherche technologique, 17 Rue des martyrs, 38054 Grenoble Cedex 09, France

ontologies are available to describe human anatomy [21], pathology [22], qualities [23], and radiological terms [24], but none can completely describe the exhaustive list of physical and biological parameters targeted for VIP models. A specific development is therefore needed.

One approach to medical image simulation is to simulate images from geometrical models parametrized with distributions of physical properties such as magnetic properties, echogenicity, radioactivity, or chemical composition. This approach is flexible since geometrical models can be deformed to a particular individual [25], [26]. However, obtaining realistic simulations is challenging because parameter distributions are not easily estimated and human morphology is difficult to model from geometrical objects alone. An alternate [27] is to create a digital object from a real acquisition by extracting object geometry and physical parameters from signal properties. Resulting images usually look very realistic, but this method requires an *in-vivo* acquisition for each simulation, which constrains investigation. Both model-based and image-based approaches are targeted in VIP.

Some online platforms, commonly called scientific gateways, integrate medical imaging software with access to computing and storage resources. neuGRID [28], CBRAIN[1], and LONI pipeline [29] (Laboratory Of Neuro Imaging) target neuroimaging data sharing and analysis using DCIs. In neuGRID, users can start remote desktop sessions on machines where image analysis tools and clients are pre-installed to access resources of the European Grid Initiative (EGI)[2]. CBRAIN is a web portal where neuroscientists can launch data processings such as segmentation tools on several clusters. The location of files and executions is controlled from the interface. In the LONI pipeline, users describe their own image processing pipelines executed on local resources. Neurolog [30] focuses on the sharing and reuse of heterogeneous data and tools produced by distributed sites. The OntoNeuroLOG ontology describing involved datasets and entities is used to harmonize database schemas, and annotate tools. e-bioinfra [31] is a web portal for executing image processing tools on distributed computing resources, e.g. Freesurfer and the Functional MRI of the Brain Software Library (FSL[3]). Some scientific gateways such as EUMEDGrid[4], DECIDE[5], and GISELA [32] are built from a set of reusable portlets providing basic functions such as authentication and data transfer.

Compared with the reviewed alternatives, VIP is distinguished by (i) an interface dedicated to medical image simulation, including simulators of 4 image modalities and a repository to store physical and biological models, (ii) a workflow-based methodology that allows fast porting of new simulators with minor adaptations, (iii) a web interface that totally relieves users from resource management decisions.

To objectively compare the activity of the described platforms, the consumed cumulative CPU time metric gives an indication on their usage, therefore usability. In 2011, the accounting system of the EGI[6] reported that neuGRID consumed 10.3 CPU years, EUMEDGrid consumed 3 CPU years, and e-bioinfra consumed 46.3 CPU years. The consumption of VIP is reported in Section VI.

[1]http://cbrain.mcgill.ca
[2]http://www.egi.eu
[3]http://www.fmrib.ox.ac.uk/fsl
[4]http://applications.eumedgrid.eu
[5]http://applications.eu-decide.eu
[6]http://accounting.egi.eu

III. Model Integration

We chose ontologies to share model semantics in the platform. Apart from structuring the model repository, two additional roles of ontologies are interesting [33]: (1) they provide a standard vocabulary to refer to conceptual entities, and (2) they model conceptual entities formally, using axioms expressed in a logical language. In VIP, model files describing geometrical objects (represented as voxel maps or meshes) are annotated with concepts defined in our application ontology called OntoVIP (described below). Annotated models are stored in the repository shown on Fig. 1.

A. Ontology design

OntoVIP integrates components describing representational objects (model layers, model layer parts, etc.) and their associated real-world entities (anatomical/pathological structures, foreign bodies, etc.). To increase the potential sharing and interoperability of our semantic model with other systems, we reused existing ontologies or extracted relevant parts from them. Anatomical terms were extracted from the Foundational Model of Anatomy (FMA) [21], pathological structures were extracted from the Mammalian pathology ontology (MPATH [22]), qualities from the phenotype and trait ontology (PATO [23]), and foreign bodies and external agents from RadLex [24]. Although extractions are easily performed to extend the list of terms, they have to be performed in moderation to avoid slowing down the reasoning engine by confusing annotation with too many available entities.

The above components were imported as modules into OntoVIP following the integration approach used in [34]. This integration relies on the foundational ontology of Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE [35]) and on other core ontologies guaranteeing overall consistency. Models are described using two basic entities: (1) `values-layers` describe data files associating each point of a 3D map with a particular value, e.g. the activity of a radiopharmaceutical; (2) `object-layers` describe data files containing labels associated with `object-layer-parts` linked to anatomical, pathological, foreign-body or external-agent objects. Each object can be associated with a `mathematical-distribution`, to specify the range of an associated physical parameter, e.g., T1 values. Layers can be associated to different time references in order to model time-varying phenomena, e.g., movement of an organ, metabolism of a substance, or growth of a tumor. Two different time scales are considered: (1) `instants` denote variations during an image acquisition and (2) `time-points` denote changes between different imaging procedures. A detailed

description of this ontology, specifying the various representational entities and how they refer to real-world entities is provided in [36].

As discussed above, the motivation for ontological modeling is two-fold. First, it provides an organized vocabulary to describe information. This is useful to query the repository using search criteria corresponding to model categories (e.g. static vs dynamic) or model content, such as anatomical structures or contrast agents. Queries may use terms extracted from a relational database, but the ontology brings flexibility with the desired specificity. For example, a lesion may be annotated using the generic term `neoplasm` or the more specific term `glioma`. Thanks to relations in the ontology, searching for models containing `neoplasm` returns models annotated with `neoplasm` and `glioma`. Other inferences could be made from `part-whole` relationships, such as FMA's `constitutionalPart` / `constitutionalPartOf`, to enable selecting entities that are part of a structure of interest.

The second motivation to ontological modeling is their reasoning capabilities. For instance, rules were implemented to check the compatibility of models with modalities and simulators (see top bar on Fig.1). A model is compatible with a simulator if it has all the required physical parameters. These rules are important to prevent users from launching unsuccessful simulations on incomplete models.

B. Semantic model repository

Models can be uploaded to VIP in the form of data files bundled with OntoVIP annotations expressed as Resource Description Framework (RDF[7]) triplets. For instance, in the `mymodel rdf:type ontovip:static-model` triplet, `ontovip:static-model` denotes the ontology class modeling static models. During model import, the consistency of assertions is checked before being added to the repository.

A standalone software currently assists users with model annotation. Knowledge stored in OntoVIP is leveraged to enrich annotations with limited manual effort from the user. For example, when a user annotates a data file with the term `Brain`, the software automatically detects that it is an `anatomical-object` using subsumption information in OntoVIP. This annotation software is progressively integrated in the online platform.

Models uploaded to the repository can be browsed according to their model parts, time information (static or dynamic), and presence of pathology, foreign body or external agent. Beyond that, rules are applied to check if a model can be used in a simulation of a particular modality, i.e., all the required physical parameters are defined. This is one of the added values of the ontology-based approach for manipulating simulation models. Fig. 1 shows a cardiac model in the model repository. Models stored in the repository can also be visualized in a 3D interface where simulation scenes can be defined, as seen on Fig. 2.

[7]http://www.w3.org/RDF

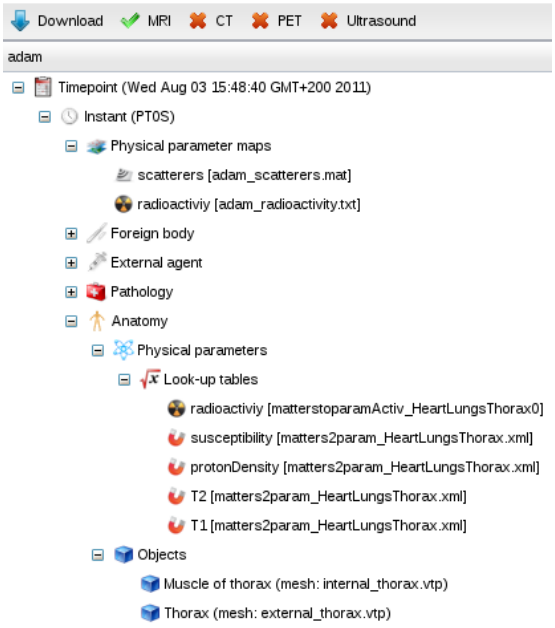

Fig. 1. VIP model repository showing a cardiac model ready for MRI simulation, where physical parameters of other modalities are missing (for readability purposes, the object list in the anatomical layer was truncated).

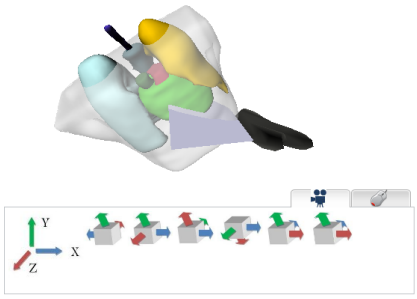

Fig. 2. Screenshot of the simulation scene interface. A thorax model was selected from the semantic model repository and an US transducer is positioned to simulate an apical echocardiography (black object on the right with sectorial imaging plane in gray).

IV. Simulation Workflows

VIP integrates simulators based on their workflow description, and without modifying their code so that their validation remains under the responsibility of their developers. The workflow representation was chosen because it provides structured, graph-based representations of applications, which is useful for several reasons. Firstly, graph representations are parallel languages, which facilitates workflow deployment on distributed computing infrastructures. Workflows also facilitate data annotation because dependencies between data files, parameters, programs and results (a.k.a data provenance) are clearly expressed, which structures output data. Finally, workflows foster reusability of software components among simulators. Ultimately, it is envisioned that the platform could facilitate simulator integration by suggesting workflow components based on existing tools (e.g. format converters).

Integrating image simulators in VIP with minimal software refactoring is difficult due to the variety of their interfaces, parameters and characteristics related to (i) the format and in-

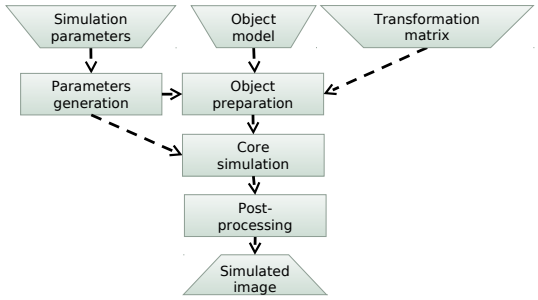

Fig. 3. VIP template represented with the Conceptual Workflow formalism.

formation required in the models; (ii) the acquisition protocols and imaging sequence and (iii) the computing model and its parallelization. To facilitate simulator integration, we proposed in [37] a workflow template composed of the components presented in Fig. 3. *Object preparation* converts a repository-annotated model into the native format suitable for the simulator. It combines the different model layers, defines the scene geometry from the geometrical transformations produced from the 3D interface, splits dynamic models into multi-static models, and adjusts physical parameters if needed. *Parameter generation* generates simulator parameter files from parameter values. Depending on the type of users, some parameters may be pre-set and hidden. *Core simulation* invokes the simulation code. It splits the simulation into independent chunks of data on which the code is iterated. Finally, *post-processing* performs image reconstruction and format conversion.

To facilitate the integration of simulators, the *Conceptual Workflows* formalism was developed [38]. Conceptual Workflows represent a family of workflows complying with a given template as illustrated on Fig. 3 for VIP. This formalism not only provides a frame for the design of new workflows, but it also assists users in workflow creation and validation. Following this formalism, image simulators of 4 modalities were integrated using the Gwendia (Grid Workflow Efficient Enactment for Data Intensive Applications) workflow language [39]: *SIMRI* [40] for MRI, *FIELD-II* [41] for US imaging, *SORTEO* [6] for PET, and *SINDBAD* [42] for CT.

Object preparation for SIMRI applies the scene transformation (produced by the interface on Fig. 2), format conversion, layer combination, and time splitting to the model. *Core simulation* is based on the concurrent computation of the magnetization of spin vectors and uses the Message Passing Interface (MPI[8]) since the simulator was already parallelized using this framework. Image reconstruction by inverse Fourier transform of the k-space is included in SIMRI.

Object preparation for SORTEO prepares the model for both emission and transmission simulations. Radioactivity is extracted from the model description and added to the description of the acquisition protocol. Photo-electric and Compton cross-sections of the materials are obtained from specific look-up tables. For *core simulation*, SORTEO uses a two-step Monte-Carlo simulation. First, single events are simulated to evaluate the probability of a photon to be detected on a detector element when emitted in a region of the model. Second, the simulator generates the true, scattered and random coincidence events using a deadtime model which incorporates these singles rates. As Monte-Carlo simulations, these steps can be split in any number of jobs lower than the number of independent random events. SORTEO produces sinograms or list-mode data. Although data correction and reconstruction are not included in VIP, the documentation refers to the STIR software[9] for both analytic and iterative reconstructions. Reconstruction examples will also be available in the portal.

Object preparation for FIELD-II either samples scatterer positions and amplitudes from distributions of the model, or applies the scene transformation to existing scatterers. *Core simulation* concurrently simulates the radio-frequency (RF) lines involved in the simulation. It supports both 1D (linear and sectorial) and 2D transducers. Once all lines are simulated, an RF matrix is assembled and the final image is reconstructed using envelope detection and Cartesian reconstruction.

Object preparation for SINDBAD generates cross-sections from material properties (chemical composition), and writes the scene transformation in SINDBAD format. *Core simulation* was implemented as a two-level splitting: each projection of the 3D scan is computed concurrently, and in addition, Monte-Carlo simulations in each projection are split depending on the number of involved photons. The reconstruction of 3D images from projections is not provided, but the documentation refers to the image reconstruction toolbox of Fessler[10] that can be used for that purpose.

Different types of scanners can be used in VIP, depending on the capabilities of the simulators. For SORTEO and SINDBAD, scanners can be defined in parameter files; for FIELD-II, the definition of the transducer is one of the simulation parameters; for SIMRI, different values of B0 and antenna profiles can be set. VIP is also extensible to other simulators, for instance to include another modality, or to compare simulators of a same modality. A new simulator can be integrated as a Gwendia workflow, following the template on Fig. 3. Due to technical constraints, workflow development still requires a developer's assistance. More image processing tools can also be integrated in VIP. Around 10 tools are currently available, including the GATE simulator, and Freesurfer for neuroimaging[11].

V. Simulation Execution

A. Deployment on distributed computing resources

Simulation parallelization relies on data parallelism to avoid refactoring simulator codes. Simulators are iterated on subsets of the simulation scene or of input parameter sets, and the resulting partial results are eventually merged by another process. Simulations are launched on the *biomed* virtual organization of the European Grid Infrastructure (EGI)[12]. EGI has access to more than 100 computing clusters worldwide, however these are not dedicated and performance can

[8]http://www.mcs.anl.gov/research/projects/mpi

[9]http://stir.sourceforge.net

[10]http://web.eecs.umich.edu/~fessler/code

[11]http://surfer.nmr.mgh.harvard.edu

[12]http://www.egi.eu

be significantly hampered by the activity of other users, or resource downtimes. To address this, VIP accesses resources by means of the Distributed Infrastructure with Remote Agent Control (DIRAC [43]) pilot-job system: computing jobs are not directly queued to the infrastructure, but they are kept on the VIP server until agents actually reach resources and fetch them. This technique is well used in large distributed systems as it improves fault tolerance and overall performance.

As described in [44], the pilot-job system was extended so that VIP users can also use their own clusters to complement grid resources. Results in [44] show that introducing even a small fraction of dedicated computing resources has an important impact on the performance of the simulation.

B. Data provenance

Simulations produce large amounts of diverse data, difficult to manipulate by non-expert users. To address this, provenance metadata is recorded during execution.

Provenance is registered on-the-fly in an OPM-compliant RDF repository (Open Provenance Model [45]) tracking simulation component invocations, and the related consumed and produced data. The relations among files and parameters produced during a simulation are recorded as provenance graphs.

Such a provenance model enables inference of new statements about the simulated data. For instance, rules can be applied to provenance paths to produce annotations describing simulation results and build a shared, structured database of simulated data from user activity.

The scalability of semantic repositories is challenging since they contain information about all the data generated by the platform. For instance, a single execution of the SORTEO core simulation workflow produces 15,000 triplets. To tackle this issue, we distinguish provenance metadata from the annotations of simulated data. Provenance metadata is generic, technical, and volatile whereas simulated data annotations represent only a few statements with a long-term added-value. Consequently, a dedicated short-term semantic repository is deployed to store provenance metadata of each simulation execution. Once generated, simulated data annotations are pushed to a long-term semantic repository. If necessary, the short-term provenance repository can then be cleaned-up.

C. Storage

VIP stores files on the EGI data management system. Files are distributed on storage resources at the different sites supporting the *biomed* virtual organization. These resources expose a homogeneous interface through which files can be listed and transferred. A central logical file catalog provides a common indexing space for these files. 3.5 PB are available for storage, among which 2 PB are currently used.

Although storage availability is not ensured (availability usually ranges from 80% to 95%), replication is possible by linking several physical files to the same logical name. Availability of critical files (e.g. application workflows) is thus ensured. The platform integrates its own local data manager, where critical files are replicated and accessed when EGI storage is down. Files are also cached by jobs on the execution nodes to further decrease transfer error rate.

D. Interface

VIP is available as a web portal where users can access the model repository (browsing and importing), define a simulation scene through the 3D interface, launch new simulations, monitor performance, and transfer input/output files. Documentation and a messaging system to contact the support team are available in the portal[13].

Performing a simulation consists of (i) selecting a model from the repository (see Fig. 1), (ii) defining the simulation scene (see Fig. 2), (iii) adding simulation parameters, (iv) launching the simulation, monitoring jobs, and downloading results from file transfer interface. In case the user is already familiar with the simulator parameters, then the first three steps are reduced to only parametrizing the core simulator.

Authentication is done with login and password. Users are mapped to a robot certificate used for all grid authentications. For security purposes, the portal keeps track of all user operations, in particular file transfers and simulation execution. The VIP open access policy fosters experimentation: accounts are created based on valid email address only. Different user levels are distinguished: beginners can launch only one simulation at a time, and cannot write in shared folders. Advanced users have extended rights, but they must briefly describe their activities, and have a grid certificate registered in a virtual organization. To avoid technical problems, the certificate is not used to log in to the portal but only for administrators to check user identity. Users are organized in groups defining access rights to applications and data.

A specific interface component enables file transfer between local user machines and grid storage. The upload process consists in (i) uploading the file from the local machine to the portal and (ii) transferring the file to the grid through an asynchronous pool of transfers processed sequentially. Download is performed similarly, in the opposite direction. This two-step process avoids connectivity issues between user machines and distributed grid hosts. The transfer pool manages the load of concurrent transfers performed by the server to avoid network clogging on the portal machine. It also replicates files to ensure availability, and it caches a local copy of critical files such as application workflows. An optimized file browser is available to interact with the transfer service. It caches grid directory content and uses native catalog commands to ensure fast browsing. File permissions are also enforced by the browser: users have a private folder readable only by themselves, and each group has a shared folder.

VI. Results

A. Integrated Models

A heart-thorax model (ADAM) and a full-body model (Zubal) are currently integrated. ADAM [46] contains geometrical descriptions for lungs, thorax, aorta, myocardium, spine, atria, and ventricles, with associated physical parameters for

[13] http://vip.creatis.insa-lyon.fr/documentation

T1, T2, T2*, proton density, echogenicity, and radioactivity. It also has a tumor and a needle. The Zubal phantom (full body with arms)[14] contains 126 labels with an emphasis on brain structures. Simulations can also be conducted with external models (e.g. the 4D NURBS-based Cardiac-Torso, XCAT [47]), but the 3D simulation scene interface is then unavailable and the model has to be manually converted to the file formats required by the simulators.

B. Image Simulations

Whole-body and cardiac images simulated with VIP in 4 modalities are described below along with the acquisition parameters, models and resulting images. Reported cumulative CPU times give an estimation of the simulation time on a single CPU core. Simulation times obtained with VIP are only indicative since they depend on the parallelization parameters and uncontrolled grid load during simulation execution.

1) Whole-body imaging: a whole-body PET/CT acquisition was simulated using SORTEO for PET and SINDBAD for X-ray CT scan. The anatomical and activity distribution models were constructed from the XCAT phantom which was fitted to a specific patient anatomy. In addition to the main organs, XCAT includes the vessels and the airway tree which are necessary for high resolution simulations such as CT.

The PET acquisition protocol consisted of a static 224 s acquisition described in [48]. The scanner geometry was that of the PET/CT Philips GEMINI system (Philips Healthcare, Cleveland, OH). The 18F-FDG (fluorodeoxyglucose) activity distribution was obtained from [5]. Three bed fields were used to cover thorax and abdomen. Each bed field had the dimension of the axial field of view (FoV) of the scanner, i.e, 47 cm. The overlap between beds was set to 50% of the axial FoV to compensate for the sensitivity loss on both extremities of the axial FoV. List-mode data were reconstructed with one-pass list-mode expectation minimization [49] using 5 iterations and 8 subsets, resulting in 128x128x81 images with 4 mm isotropic voxels.

The CT simulation was exclusively analytical although SINDBAD also has a Monte-Carlo mode. The scanner model was a simplified version of a Philips Scanner (Philips MX8000) which is the CT component of the PET/CT GEMINI scanner. We assumed a point source located at 0.75 m from the center of the thorax (also considered as the rotation axis of the CT system) and at 1.2 m from the plane detector. We used a cone-beam geometry and a standard X-ray energy spectrum with a tube voltage of 110 kVp and an aluminium filter of 2.5 mm. The typical values for intensities and scanner rotation length corresponded to a very low-dose acquisition (resp. 1 mA and 0.5 s). According to the magnification factor scale in this scanner geometry (from 1.2 to 2.4 for the phantom), the voxel size of the CT model was chosen to be half the pixel size of the detector. Precisely, the XCAT phantom was converted into a 500x500x689 volume with an isotropic voxel size of 1 mm. A set of 480 projections were reconstructed with a Feldkamp cone-beam reconstruction algorithm resulting in a 450x450x450 matrix with an isotropic pixel size of 1mm.

[14]http://noodle.med.yale.edu/zubal/

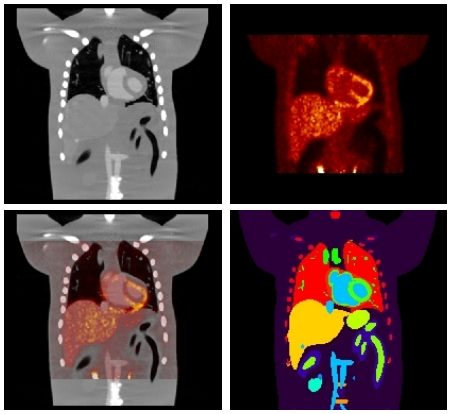

Fig. 4. Whole-body FDG-PET and CT simulations (coronal slices). From left to right and top to bottom: simulated 0.5 s CT acquisition, simulated static 224 s FDG-PET acquisition, overlaid CT and FDG-PET simulations, XCAT-based voxel model.

Fig. 4 shows the XCAT model and coronal slices extracted from the CT and PET 18F-FDG simulations. Validation of the first and second order simulated PET image statistics against clinical ones has been performed in [5]. For CT, the airway tree of the XCAT model was well rendered in the simulation, in spite of the low resolution of the simplified scanner model (2x2 mm detector size) and the simple analytical model. Adding the noise from the scattered radiation and non-ideal detector would allow to reproduce accurate count rates and noise texture in the reconstructed image. Such validation is out of scope here. The CT simulation represented a cumulative CPU time of 12.8 hours, computed in 1.9 hours on VIP.

2) Heart imaging: A 2D MR balanced steady state free precession (bSSFP) sequence at 1.5 T was simulated on a cardiac cycle (14 instants) extracted from the ADAM model. Scan parameters similar to the simulation conducted by Tobon *et al.* in [25] were chosen: repetition/echo time = 2.9/1.2 ms and flip angle = 45°, matrix size 256x256, number of signal average: 1. Magnetic parameters for blood, lung and myocardium were also extracted from [25] while those of fat, muscle, spine, and spinal cord were obtained from [50], [51] and [52]. Fig. 5 (top-left) shows one instant of the simulated sequence. The whole sequence can be seen from the VIP gallery[15]. Despite the limited number of anatomical structures compared to real images, contrasts in the blood pool and left-ventricle are coherent with the image obtained in [25] (see Fig. 5, bottom-left), with a blood hypersignal in the cavities as opposed to myocardium. More information about the validation of SIMRI is found in [40]. Each instant represented a cumulative CPU time of 31±4 min, computed in 5±2 min on VIP.

A healthy and a pathological FDG-PET acquisition were simulated with ADAM using SORTEO. The scanner geometry of ECAT EXACT HR+ (CTI/Siemens Knoxville) was used, and 18F-FDG was the radiotracer to study glucose metabolism. The simulated activities were respectively 45.1 MBq for the healthy case and 43.8 MBq for the pathological case as used in [46]. The raw data was corrected for attenuation and reconstructed using a standard 3D filtered back-projection

[15]http://vip.creatis.insa-lyon.fr/gallery

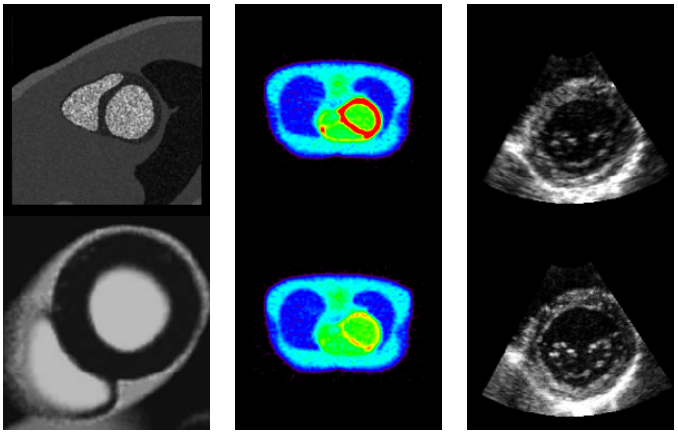

Fig. 5. Left-top: cardiac short-axis MRI simulated with VIP from ADAM. Left-bottom: simulation obtained by Tobon *et al.* in [25] with XCAT. Center: simulated cardiac FDG-PET at end diastole, transverse view (top: healthy; bottom: pathological). Right: simulated (top) and real (bottom) images from a 2D+t echocardiography.

algorithm resulting in a 128x128x63 3D image with a voxel size of 5.15x5.15x2.42 mm. Fig. 5 (center) shows the resulting simulated FDG-PET end diastolic instant for both acquisitions. It clearly shows a high homogeneous FDG uptake in the healthy myocardium and lower fixation for the pathological case. The simulation of one instant of the cardiac cycle represented 91 CPU hours, obtained in 39 hours on VIP.

A 2D+t echocardiographic sequence was also simulated with an image-based approach. It was obtained by deforming a scattering map with a known motion model. The acquisition was simulated with FIELD-II using a 64-element sectorial probe at 3.75 MHz. A sampling frequency of 40 MHz was set. The view angle was 66° and the pitch was set to half of the wavelength to avoid grating lobes effects. One frame of the obtained simulated sequence is shown on Fig. 5 (top-right) with the corresponding real image (bottom-right). For a better dynamic perception, the full sequence can be found online[16] along with the associated benchmark velocity field. As detailed in [53], both visual appearance and motion are very realistic due to the use of a real sequence as a template. As the true motion for the synthetic sequences is known, this sequence can be used as benchmark for the evaluation of myocardium motion estimation algorithms. The cumulative CPU time of the simulation was 42 hours, obtained in 4 hours on VIP.

C. Platform usage

To date, 230 users are registered in VIP, among which 64 logged in during the last month. In 2011, the average monthly CPU consumption was 2.75 years and the yearly cumulative was 33 years. This level of activity is comparable to the main grid platforms integrating medical imaging applications (see values reported in Section II). VIP is significantly used, which is an indicator of its availability and usability.

VII. Conclusion

VIP is a versatile, open-access platform for multi-modality medical image simulation. This article provides a complete overview of VIP and describes design issues and solutions selected to enable sharing of object models, the integration of simulators, and the execution of simulations on distributed computing resources. Results show that VIP can run simulations of 4 modalities and different organs. Usage statistics show that it has a good potential to go beyond a proof-of-concept level. The development of VIP started in 2009 and mobilized 3 developers for 2 years. Operations require about 2 hours daily.

Some limitations remain, partly due to the adoption of a web infrastructure. A centralized web portal simplifies usability and user support, but it is a network bottleneck when large files are transferred. To address this, files could be directly transferred from user hosts to storage sites. Performance of small simulations is another issue due to the sharing of computing resources among several users. Although pilot jobs are used to reduce latency, jobs still have to queue from 1 minute to 1 hour. A few dedicated resources could be used to address this problem. The model repository also has to be exemplified on more models. For instance, some terms will probably have to be added to the ontology when specific medical applications are targeted. Finally, simulators and tools still cannot be integrated directly from the web interface because it requires substantial knowledge about the infrastructure.

VIP operates as an image simulation and image processing service for the academic community. Maintenance costs remain reasonable thanks to the use of external, open-source software, and the exploitation of public distributed computing and storage resources.

[16]http://www.creatis.insa-lyon.fr/us-tagging/news

Acknowledgment

This work is funded by the French National Research Agency (ANR) under grant ANR-09-COSI-03. We thank the European Grid Initiative and “France-Grilles”.

References

[1] A. D. Gilliam and S. T. Acton, “Echocardiographic simulation for validation of automated segmentation methods,” in *IEEE ICIP*, 2007, pp. 529–532.

[2] M. Prastawa, E. Bullitt, and G. Gerig, “Simulation of brain tumors in MR images for evaluation of segmentation efficacy,” *Medical Image Analysis*, vol. 13, no. 2, pp. 297–311, 2009.

[3] G. Wagenknecht, H.-J. Kaiser, T. Obladen, O. Sabri, and U. Buell, “Simulation of 3D MRI brain images for quantitative evaluation of image segmentation algorithms,” in *SPIE Medical Imaging*, vol. 3979, 2000, pp. 1074–1085.

[4] T. Glatard, A. Marion, H. Benoit-Cattin, S. Camarasu-Pop, P. Clarysse, R. Ferreira da Silva, G. Forestier, B. Gibaud, C. Lartizien, H. Liebgott, K. Moulin, and D. Friboulet, “Multi-modality image simulation with the virtual imaging platform: illustration on cardiac MRI and echography,” in *IEEE ISBI*, Barcelona, 2012.

[5] S. Tomei, A. Reilhac, D. Visvikis, N. Boussion, C. Odet, F. Giammarile, and C. Lartizien, “OncoPET_DB: a freely distributed database of realistic simulated whole body [18f]FDG PET images for oncology,” *IEEE Trans on Nuclear Science*, vol. 57, no. 1, pp. 246–255, 2010.

[6] A. Reilhac, G. Batan, C. Michel, C. Grova, J. Tohka, D. L. Collins, N. Costes, and A. C. Evans, “PET-SORTEO: Validation and Development of Database of Simulated PET Volumes,” *IEEE Transactions on Nuclear Science*, vol. 52, pp. 1321–1328, 2005.

[7] I. Castiglioni, I. Buvat, G. Rizzo, M. Gilardi, J. Feuardent, and F. Fazio, “A publicly accessible Monte-Carlo database for validation purpose in emission tomography,” *Eur. J. Nuc. Med. Mol. Imag.*, vol. 32, pp. 1234–39, 2005.

[8] S. Stute, P. Tylski, N. Grotus, and I. Buvat, "Lucas: Efficient monte carlo simulations of highly realistic pet tumor images," in *IEEE NSS-MIC*, Dresden, 2008, pp. 4010–12.

[9] R. K. Kwan, A. C. Evans, and G. B. Pike, "MRI simulation-based evaluation of image-processing and classification methods." *IEEE Transactions on Medical Imaging*, vol. 18, no. 11, pp. 1085–1097, 1999.

[10] A. Dikshit, D. Wu, C. Wu, and W. Zhao, "An online interactive simulation system for medical imaging education," *Computerized Medical Imaging and Graphics*, vol. 29, no. 6, pp. 395–404, 2005.

[11] F. Beekman and M. Viergever, "Fast SPECT simulation including object shape dependant scatter," *IEEE TMI*, vol. 14, pp. 271–282, 1995.

[12] D. Lazos, K. Bliznakova, Z. Kolitsi, and N. Pallikarakis, "An integrated research tool for X-ray imaging simulation," *Computer Methods and Programs in Biomedicine*, vol. 70, no. 3, pp. 241–251, 2003.

[13] D. E. Peplow and K. Verghese, "Digital mammography image simulation using Monte Carlo," *Medical Physics*, vol. 27, pp. 568–579, 2000.

[14] I. Drobnjak, D. Gavaghan, E. Sli, J. Pitt-Francis, and M. Jenkinson, "Development of a functional magnetic resonance imaging simulator for modeling realistic rigid-body motion artifacts," *Magnetic Resonance in Medicine*, vol. 56, no. 2, pp. 364–380, 2006.

[15] W. Wein, A. Khamene, D.-A. Clevert, O. Kutter, and N. Navab, "Simulation and fully automatic multimodal registration of medical ultrasound," in *MICCAI'07*, 2007, vol. 4791, pp. 136–143.

[16] D. Aiger and D. Cohen-Or, "Real-time ultrasound imaging simulation," *Real-Time Imaging*, vol. 4, pp. 263–274, 1998.

[17] R. Shams, R. Hartley, and N. Navab, "Real-time simulation of medical ultrasound from CT images," in *MICCAI'08*, 2008, pp. 734–741.

[18] S. Jan, D. Benoit, E. Becheva, T. Carlier, F. Cassol, P. Descourt, T. Frisson, L. Grevillot, L. Guigues, L. Maigne, C. Morel, Y. Perrot, N. Rehfeld, D. Sarrut, D. R. Schaart, S. Stute, U. Pietrzyk, D. Visvikis, N. Zahra, and I. Buvat, "Gate v6: a major enhancement of the gate simulation platform enabling modelling of ct and radiotherapy," *Physics in medicine and biology*, vol. 56, no. 4, pp. 881–901, 2011.

[19] G. R. Christie, P. M. F. Nielsen, S. A. Blackett, C. P. Bradley, and P. J. Hunter, "FieldML: concepts and implementation," *Roy Soc of London Philosophical Transactions Series A*, vol. 367, pp. 1869–1884, 2009.

[20] N. Smith, A. De Vecchi, M. McCormick, D. Nordsletten, O. Camara, A. Frangi, H. Delingette, M. Sermesant, J. Relan, N. Ayache, M. Krueger, W. Schulze, R. Hose, I. Valverde, P. Beerbaum, C. Staicu, M. Siebes, J. Spaan, P. Hunter, J. Weese, H. Lehmann, D. Chapelle, and R. Razavi, "euHeart: Personalized and integrated cardiac care using patient-specific cardiovascular modelling," *Journal of the Royal Society Interface Focus*, vol. 1, no. 3, pp. 349–364, 2011.

[21] C. Rosse and J. Mejino, "The foundational model of anatomy ontology," *Anatomy Ontologies for Bioinformatics*, pp. 59–117, 2008.

[22] P. Schofield, G. Gkoutos, M. Gruenberger, J. Sundberg, and J. Hancock, "Phenotype ontologies for mouse and man: bridging the semantic gap," *Disease models & mechanisms*, vol. 3, no. 5-6, pp. 281–289, 2010.

[23] G. Gkoutos, E. Green, A.-M. Mallon, J. Hancock, and D. Davidson, "Using ontologies to describe mouse phenotypes," *Genome Biology*, vol. 6, pp. 1–10, 2005.

[24] C. Langlotz, "Radlex: a new method for indexing online educational materials," *RadioGraphics*, no. 26, pp. 1595–1597, 2006.

[25] C. Tobon-Gomez, F. M. Sukno, B. H. Bijnens, M. Huguet, and A. F. Frangi, "Realistic simulation of cardiac magnetic resonance studies modeling anatomical variability, trabeculae, and papillary muscles," *Magnetic Resonance in Medicine*, vol. 65, no. 1, pp. 280–288, 2011.

[26] A. Lemaitre, P. Segars, S. Marache, A. Reilhac, M. Hatt, S. Tomei, C. Lartizien, and D. Visvikis, "Incorporating patient specific variability in the simulation of realistic whole body 18F-FDG distributions for oncology applications," *Proceedings of the IEEE*, vol. 97, no. 12, pp. 2026–2038, 2009.

[27] O. Kutter, R. Shams, and N. Navab, "Visualization and GPU-accelerated simulation of medical ultrasound from CT images," *Computer Methods and Programs in Biomedicine*, vol. 94, no. 3, pp. 250 – 266, 2009.

[28] G. Frisoni, A. Redolfi, D. Manset, M. Rousseau, A. Toga, and A. Evans, "Virtual imaging laboratories for marker discovery in neurodegenerative diseases." *Nat Rev Neurol.*, vol. 7, no. 8, pp. 429–38, 2011.

[29] D. E. Rex, J. Q. Ma, and A. W. Toga, "The LONI pipeline processing environment," *NeuroImage*, vol. 19, no. 3, pp. 1033 – 1048, 2003.

[30] J. Montagnat, A. Gaignard, D. Lingrand, J. Rojas Balderrama, P. Collet, and P. Lahire, "NeuroLOG: a community-driven middleware design," in *HealthGrid*, Chicago, USA, June 2008, pp. 49–58.

[31] S. Shahand, M. Santcroos, Y. Mohammed, V. Korkhov, A. C. Luyf, A. van Kampen, and S. D. Olabarriaga, "Front-ends to Biomedical Data Analysis on Grids," in *HealthGrid*, Bristol, UK, june 2011.

[32] R. Barbera, F. Brasileiro, R. Bruno, L. Ciuffo, and D. Scardaci, "Supporting e-science applications on e-infrastructures: Some use cases from latin america," in *Grid Computing*, 2011, pp. 33–55.

[33] T. Gruber, "Toward principles for the design of ontologies used for knowledge sharing," *International journal of human computer studies*, vol. 43, no. 5, pp. 907–928, 1995.

[34] L. Temal, M. Dojat, G. Kassel, and B. Gibaud, "Towards an ontology for sharing medical images and regions of interest in neuroimaging," *Journal of Biomedical Informatics*, vol. 41, no. 5, pp. 766–778, 2008.

[35] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, and A. Oltramari, "WonderWeb Deliverable D18. The WonderWeb Library of Foundational Ontologies and the DOLCE ontology," 2003.

[36] G. Forestier, A. Marion, H. Benoit-Cattin, P. Clarysse, D. Friboulet, T. Glatard, P. Hugonnard, C. Lartizien, H. Liebgott, J. Tabary, and B. Gibaud, "Sharing object models for multi-modality medical image simulation: a semantic approach," in *IEEE CBMS'2011*, Bristol, 2011.

[37] A. Marion, G. Forestier, H. Benoit-Cattin, S. Camarasu-Pop, P. Clarysse, R. F. da Silva, B. Gibaud, T. Glatard, P. Hugonnard, C. Lartizien, H. Liebgott, J. Tabary, S. Valette, and D. Friboulet, "Multi-modality medical image simulation of biological models with the Virtual Imaging Platform (VIP)," in *IEEE CBMS*, Bristol, UK, 2011, pp. 1–6.

[38] N. Cerezo and J. Montagnat, "Scientific workflows reuse through conceptual workflows," in *WORKS'11*, Seattle, 2011, pp. 12–18.

[39] J. Montagnat, B. Isnard, T. Glatard, K. Maheshwari, and M. Blay-Fornarino, "A data-driven workflow language for grids based on array programming principles," in *WORKS'09*, Portland, 2009, pp. 1–10.

[40] H. Benoit-Cattin, G. Collewet, B. Belaroussi, H. Saint-Jalmes, and C. Odet, "The SIMRI project : a versatile and interactive MRI simulator," *J of Magnetic Resonance Imaging*, vol. 173, no. 1, pp. 97–115, 2005.

[41] J. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 39, no. 2, pp. 262–267, 1992.

[42] J. Tabary, S. Marache, S. Valette, W. Segars, and C. Lartizien, "Realistic X-Ray CT simulation of the XCAT phantom with SINDBAD," in *IEEE NSS and MIC Conference*, Orlando, USA, October 2009, pp. 3980–3983.

[43] A. Tsaregorodtsev, N. Brook, A. C. Ramo, P. Charpentier, J. Closier, G. Cowan, R. G. Diaz, E. Lanciotti, Z. Mathe, R. Nandakumar, S. Paterson, V. Romanovsky, R. Santinelli, M. Sapunov, A. C. Smith, M. S. Miguelez, and A. Zhelezov, "DIRAC3 . The New Generation of the LHCb Grid Software," *Journal of Physics: Conference Series*, vol. 219 062029, no. 6, 2009.

[44] R. Ferreira da Silva, S. Camarasu-Pop, B. Grenier, V. Hamar, D. Manset, J. Montagnat, J. Revillard, J. R. Balderrama, A. Tsaregorodtsev, and T. Glatard, "Multi-Infrastructure Workflow Execution for Medical Simulation in the Virtual Imaging Platform," in *HealthGrid*, Bristol, 2011.

[45] L. Moreau, J. Freire, J. Futrelle, R. McGrath, J. Myers, and P. Paulson, "The Open Provenance Model: an overview," in *Provenance and Annotation of Data and Processes*, 2008, vol. 5272, pp. 323–326.

[46] R. Haddad, "Un modéle anthropomorphique et dynamique du thorax respirant et du coeur battant," Ph.D. dissertation, INSA Lyon, 2007.

[47] W. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui, "4D XCAT phantom for multimodality imaging research," *Med. Phys.*, vol. 37, no. 9, pp. 4902–4915, 2010.

[48] S. Marache, R. Prost, J. M. Rouet, and C. Lartizien, "Incorporating patient-specific variability in the OncoPET_DB database," in *IEEE NSS-MIC*, Valencia, 2011.

[49] A. J. Reader, S. Ally, F. Bakatselos, R. Manavaki, R. J. Walledge, A. P. Jeavons, P. J. Julyan, S. Zhao, D. L. Hastings, and J. Zweit, "One-pass list-mode EM algorithm for high-resolution 3-D PET image reconstruction into large arrays," *IEEE Trans on Nuclear Science*, vol. 49, pp. 693–699, 2002.

[50] P. Schmitt, M. A. Griswold, P. M. Jakob, M. Kotas, V. Gulani, M. Flentje, and A. Haase, "Inversion recovery TrueFISP: quantification of T1, T2, and spin density," *Magn Reson Med*, vol. 51, no. 4, pp. 661–7, 2004.

[51] G. J. Stanisz, E. E. Odrobina, J. Pun, M. Escaravage, S. J. Graham, M. J. Bronskill, and R. M. Henkelman, "T1, T2 relaxation and magnetization transfer in tissue at 3T," *Mag Res Med*, vol. 54, no. 3, pp. 507–12, 2005.

[52] E. R. Melhem, D. A. Israel, S. Eustace, and H. Jara, "MR of the spine with a fast T1-weighted fluid-attenuated inversion recovery sequence," *AJNR Am J Neuroradiol*, vol. 18, no. 3, pp. 447–54, 1997.

[53] M. Alessandrini, H. Liebgott, D. Friboulet, and O. Bernard, "Highly realistic simulation of echocardiographic image sequences for ground-truth validation of motion estimation." in *IEEE ICIP'12*, Orlando, 2012.