HyperProbe

Project title: Transforming brain surgery by advancing functional-guided neuronavigational imaging

Project acronym: HyperProbe

Grant Agreement: 101071040

Call identifier: HORIZON-EIC-2021-PATHFINDERCHALLENGES-01

D3.1 Digital phantom delivery

Lead partner: UCL

Author(s): Frédéric Lange (UCL), Charly Caredda (UCBL), Luca Giannoni (UNIFI), Ilias Tachtsidis (UCL)

Work Package: WP3

Due date: Month 12

Actual delivery date: 29/09/2023

Type: DATA

Dissemination level: PU

Funded by the European Union

European Innovation Council

Tables of contents

Abbreviations

DIS	Digital instrument simulator
FOV	Field of view
FWHM	Full width at half maximum
HIS	Hyperspectral imaging
NIRS	Near-infrared spectroscopy
MC	Monte-Carlo
NA	Numerical aperture
WD	Working distance
WP	Work package

Disclaimer

The opinions stated in this report reflect the opinions of the author(s) and not the opinion of the European Commission.

Introduction

This deliverable presents the development of a digital phantom that will act/ be used as an instrument simulator (DIS). The main objective of the deliverable was to develop a computational tool able to simulate brain tissue optical parameters and incorporate realistic instrument specifications, in order to (i) assist the development of the HyperProbe2 and 2.1 prototypes, and (ii) generate synthetic data to help the development of the machine learning algorithm of work package (WP) 4.

Indeed, in the recent year, simulation tools have been used to model entire instruments[1]. This possibility can help to refine systems development by testing various components or implementation strategies before physical system construction, reducing the development costs. We thus want to develop a digital phantom simulator that can incorporate the main parameters of our HyperProbe instruments, in order to support its instrumentation development. Indeed, optimizing parameters such as the number of wavelengths used in HyperProbe2.1, to reduce the complexity of the system without compromising on its accuracy, can be time consuming if implemented physically. Optimizing these parameters at first in a DIS would help to reduce the development time and cost.

In this report, we will describe the main blocks and parameters of our DIS and demonstrate the digital phantom simulator. The simulator relies on Monte-Carlo (MC) simulations of the propagation of the light in tissues, to generate a numerical/digital phantom, followed by the incorporation of the actual instrument parameters, acting like a digital twin of the instrument. It worth noting that the code of this simulator will be publicly released in a GitHub repository upon publication. Moreover, a full dataset of hyperspectral images generated with this framework can be found here: https://doi.org/10.5281/zenodo.8386156.

1. Instrument simulator framework

The DIS is based on a 2-stage approach depicted in figure 1. Indeed, the simulator needs to emulate 2 distinctive aspects of an actual measurement: the object to image (the digital phantom) and the instrument. The first step is thus to simulate the propagation of the light through tissue, i.e., the object to image. To do so, the virtual objects need to emulate the propagation of the light inside it. This is most often done by using a MC approach, which will be describe in section 1.1. The second step takes the raw output of the MC simulations and incorporates the real instrument parameters in order to produce realistic images. This step will be described in section 2.2.

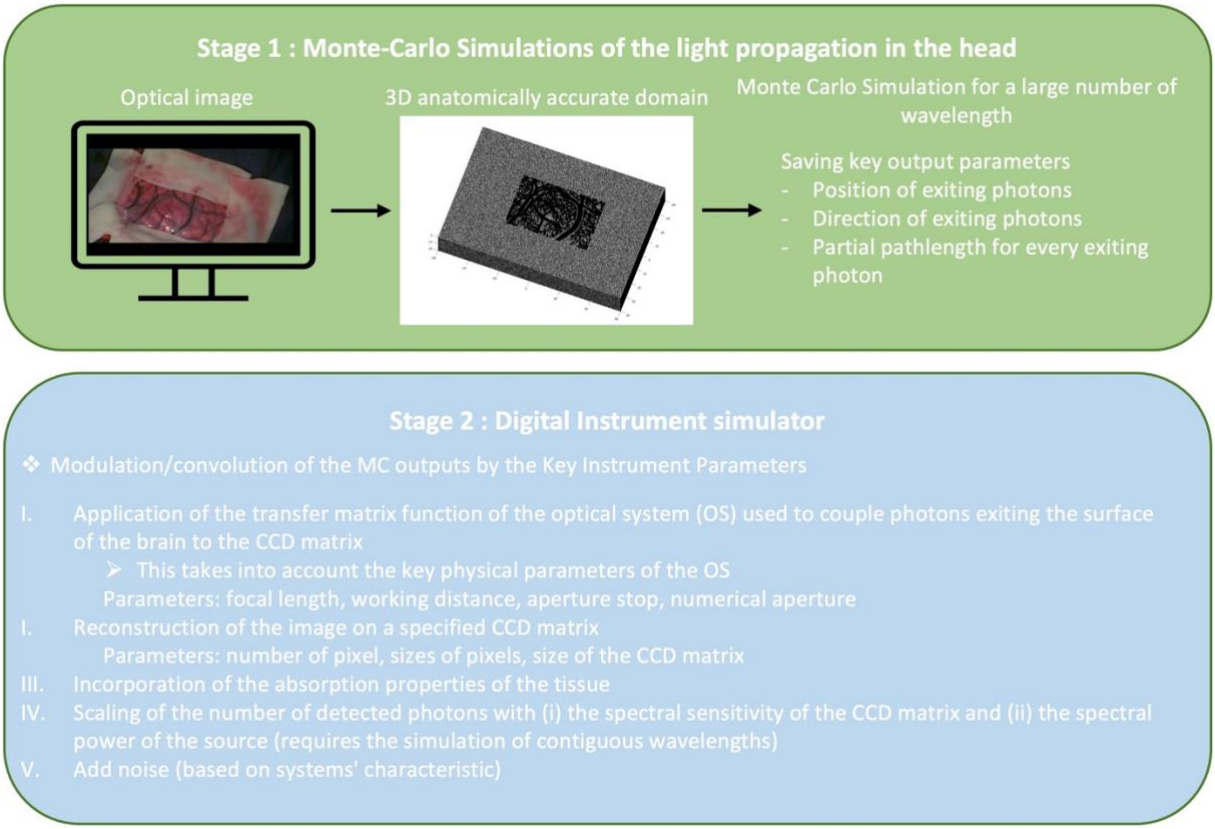

Figure 1.Summary of the main steps to design the digital instrument simulator (DIS).

1.1. Monte-Carlo Methods

MC methods to simulate the propagation of photons in tissues are widely used in diffuse optics technics[2]. They are considered as the gold standard to describe accurately the propagation of light in biological tissues. This technique is based on the random walks that packets of photons make as they travel through an object, in our case tissues, which are chosen by statistically sampling the probability distributions for step size and angular deflection per scattering event. When the propagation of enough photons has been simulated, the net distribution of all the photon paths yields an accurate approximation to reality. This technic has been used for more than 30 years, as described in the recent special issue of Journal Biomedical Optics[2]; it worth noting that there are several codes available to run these simulations. Two of the most popular ones are the Mesh-based Monte Carlo (MMC)[3] and Monte Carlo eXtreme (MCX)[4] codes[5]. These codes allow to simulate the propagation of photons in an arbitrary 3D-volume. The main difference between MMC and MCX is the way the domain is discretized. For MMC, the domain is composed of tetrahedron, producing a mesh volume, while MCX uses voxels. The rest of the framework of these tools are almost identical, with only a few specific additional options in each of the codes, but the general principal remains the same. The main benefit of using the meshed based approach is the ability to more accurately render complicated shapes, such as ones with a lot of curvature. However, this comes at the cost of a computational burden, as many tetrahedron are required to render such domains. The interested reader can refer to these papers for more details[3].

As these codes (MMC and MCX) can produce the same output files, they both can be used in our simulator. The first step is to produce a domain, either as a 3D-mesh if MMC is used, or as a voxel-based volume if MCX is used. As we want to model the light propagation in the exposed cortex, we want to be able to model a domain taking into account the blood vessels that will be very absorbent, and have different optical properties compared to the gray matter. Therefore, in order to produce realistic domains/phantoms for our simulations, we used real

images of exposed cortex (provided by the consortium partners, UCL and UCBL) to generate our model/digital phantom. We later in the report show examples of how to produce the spatial domain based on real high-definition cortical images to resolve an accurate representation of the brain for both the meshed and voxelized volumes.

1.1.1. Geometry definition

a) Example with mesh-based domain

We report here an example of a 3D-mesh generation from a grayscale image. This work had been previously reported in Giannoni *et al* [6].

Here, we used a grayscale image of the exposed cortex, showing a 1.2 × 1.2 mm field of view (FOV) of the surface of the brain of a mouse and composed of 400 × 400 pixels. The image is firstly manually segmented to obtain a binary mask that differentiates between blood vessels and the surrounding brain tissue. A 3-D binary volume of the pial vasculature (1.2 × 1.2 × 0.1 mm) is then generated by expanding the mask along the vertical direction while symmetrically eroding the sections of the vessels from the central plane. This is done to replicate the curvature of the vascular geometry. The 3-D binary volume of the pial vasculature is then converted into a meshed volume using iso2mesh[7], a tool widely used in the community. This constitutes the first medium of the final domain. The pial vasculature volume is then encased in a 2.4 × 2.4 × 1 mm slab reproducing the surrounding mouse subpial gray matter. The extra layers added to the 1.2 × 1.2 mm FOV have the purpose of minimising boundary effects during the MC simulations, i.e., distortion of the light propagation due to a finite domain. The workflow diagram describing this process is illustrated in Fig. 2.

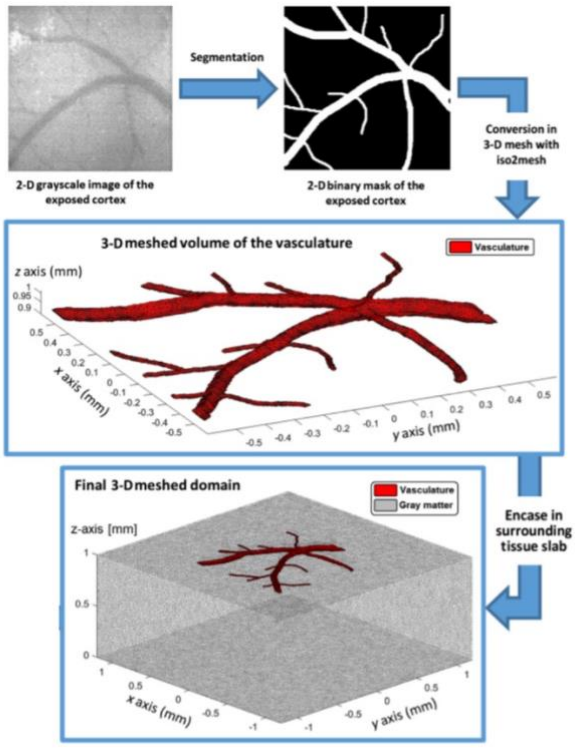

Figure 2. Workflow diagram of the methodology used to create a 3-D meshed domain of the exposed cortex: from an in vivo 2-D image (in grayscale), a binary mask is first created (in black and white) identifying the two media; then a 3-D mesh of the pial vasculature (in red) is generated, as well as a slab of subpial gray matter (in gray) encasing it. Modified from Giannoni et al.[6].

b) Example with voxel-based domain

We report here an example of the generation of a voxel-based domain from an RBG image (resolution 75μm) of the human exposed cortex provided by the UCBL partner. In this case, the RGB was segmented with a semi-automated procedure into three classes: gray matter, large blood vessel, and capillaries. Pixels were clustered into ten clusters using the K-means algorithm from the python library OpenCV (v4.8.0). The components of each cluster were manually sorted and attributed to the three classes. Three functional regions were defined as 1cm disk based on electrical brain stimulation findings, which lead to three other classes: activated grey matter, activated large blood vessel and activated capillaries.

Once the image was segmented into six classes, we modelled the brain volume. The binary segmentation masks were replicated along the z axis on 2cm to avoid any photon loss. Then we modelled the blood vasculature, using morphological erosion. The structuring element used for the erosion was set to 0 (in pixels) for z=0 (in pixels) and was increased of 1 pixel while increasing z axis. The binary volumes of the six classes were finally merged together with a final isotropic resolution of 75μm.

The RGB image, the result of the segmentation and the final image if the 6 classes are presented in Fig. 3.

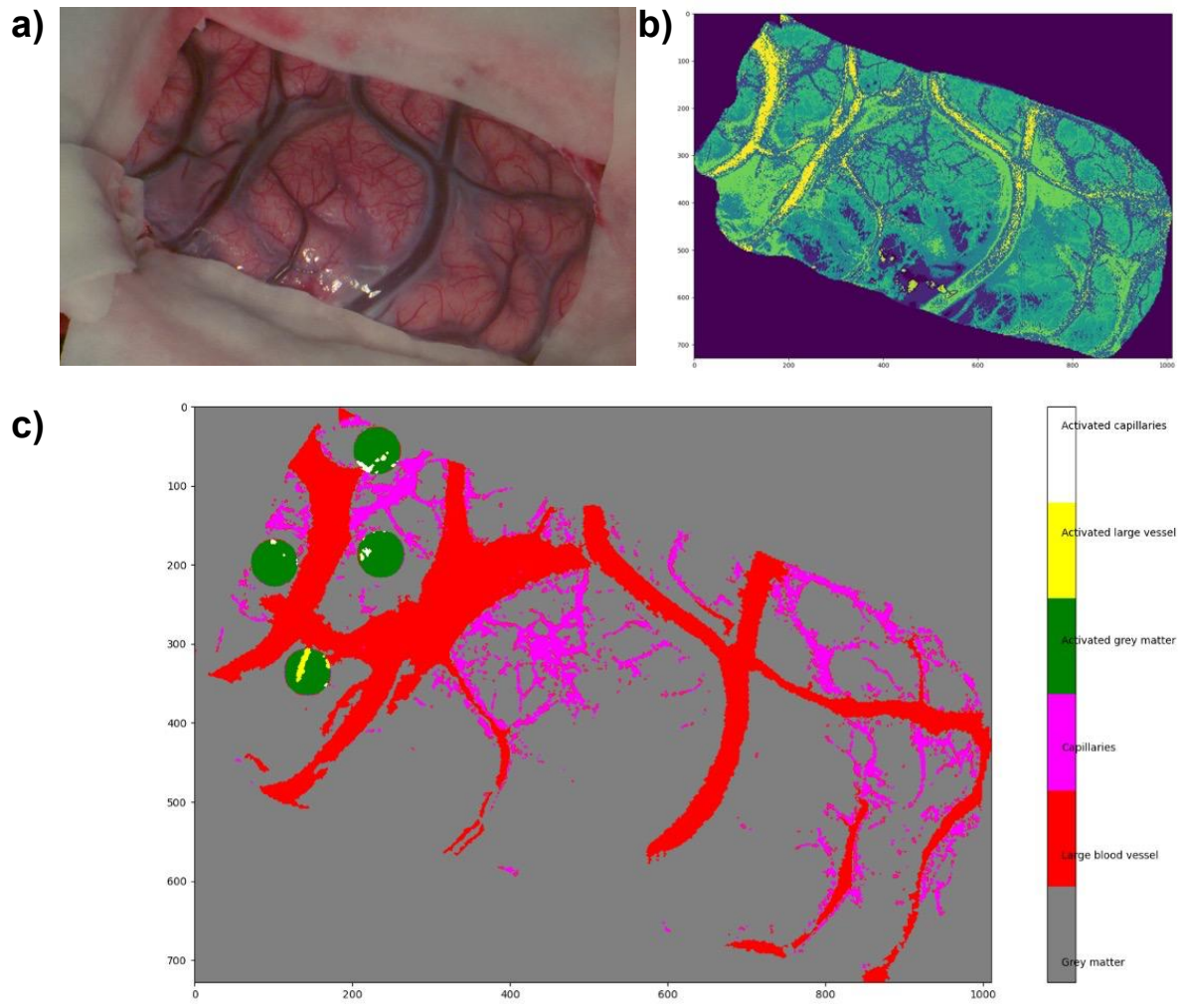

Figure 3. Example of the segmentation process. a) raw RGB image on the exposed cortex. b) Map of the segmented vessels. c) Final Map of the different classes taken into account.

Funded by the European Union
European Innovation Council

1.1.2. Monte-Carlo framework

Once the domain defined, the 2 simulators require the same input parameters in order to run the simulation. The first step is to assign the optical properties (absorption coefficient μ_a, scattering coefficient μ_s, anisotropy g, and refractive index n), based on their tissue type (for more details, see Giannoni *et al* [6]), to each element of the volume. Once the domain has been defined (i.e., both geometry and optical properties), the final step is to position a source (with its numerical aperture (NA) and the number of photons launched) and a single large detector on top of the medium, covering the entire field of view (FOV) of interest, in order to detect the reflected photons. Once all of these parameters have been set, the simulation can be run.

Multiple parameters can be saved for each simulation but the ones that we considered are the exit position and direction of each detected photon, together with their partial pathlength (PPL), i.e. the length that each photon has spent in each classes of the domain. It worth noting that that these parameters are not affected by the absorption and that these simulations are thus called white MC[8]. The advantage of this approach is that the absorption can then be considered *a posteriori*, by using the Beer-Lambert law[9]. Thus, when changing the absorption parameters of the mesh (i.e., to simulate a brain activation for example), the simulation does not need to run again, speeding up the processing. In this case only 1 simulation per wavelength is required, considering that the scattering properties of the medium do not change. A scattering change can be considered but a new simulation with a new set of μ_s would need to be rerun. We summarize all the parameters of the MC simulation in figure 4 below.

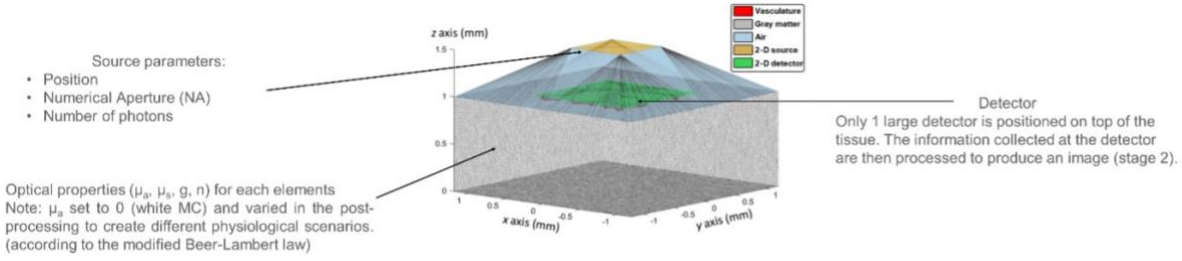

Figure 4.Summary of all parameters of the simulation. In this case for the MMC code. A specificity of MMC is that arbitrary source and detector sizes and positions can be defined. In the case of MCX, the entire superior face is considered as the detector, and the source is also of the same size as the superior surface.

1.2. Stage 2 – Digital instrument simulator (DIS)

The second part of the simulator consist of generating realistic data that considers the physical parameters of an optical system (OS) (both light source and light detection). This type of approach has already been introduced previously. We can cite for example the work of Sudakou *et al*.[10]. In this paper, time domain functional near infrared spectroscopy (TD-fNIRS) measurement were simulated using a MC approach and the responsivity of the detectors of a particular system was introduced to simulate realistic data acquired by this system. The authors showed that the consideration of the responsivity of the detectors was shifting the optimal wavelengths to use for their application, highlighting the importance of taking into account the physical limits of the instrument in simulations. We want to take a similar approach here and consider all the key parameters of both the source and detection scheme, to generate a complete digital instrument simulator (DIS) to complete the digital phantom.

1.2.1. Transfer matrix

The first step is to calculate the position of each photon exiting the tissue onto the detector (for example a CCD camera). This is done by taking the initial exit photon position and direction and applying the transfer matrix of the OS that couples the light to the detector, and which considers all its key elements (i.e., focal length, working distance, size of the optics, numerical

aperture). From a computational point of view, the ABCD matrix analysis is used in order to couple the photon exiting the brain to the image plane[11]. Moreover, the exit angle of the photon it used to reject any photon that would be exit the brain outside of the acceptance angle of the lens. This allows us to reject all the photons that would not be detected by the system. The final position of the photons on the image plan are then sorted and binned together in a matrix according to the principal characteristics of the CCD matrix (i.e., number and sized of pixels, and size of the matrix).

An example of the effect of a lens on the image of the diffuse reflectance is presented in figure 5. On the left the reflectance, calculated as in Yao *et al.*[12], at the surface (perfect image) is plotted. Then the effect on the image produced by 2 lenses with different focal lens (f=30mm and f=31 mm) positioned at the same working distance (WD=400 mm) on a 5 x 6 mm sensor (100x80 pixels). One can see the blurring effect of a non-ideal optics on the image quality. It is worth noting that we can also see the effect of the lenses on the FOV, as the image formed by the lens on a real sensor is cropped compared to the entire surface of the brain.

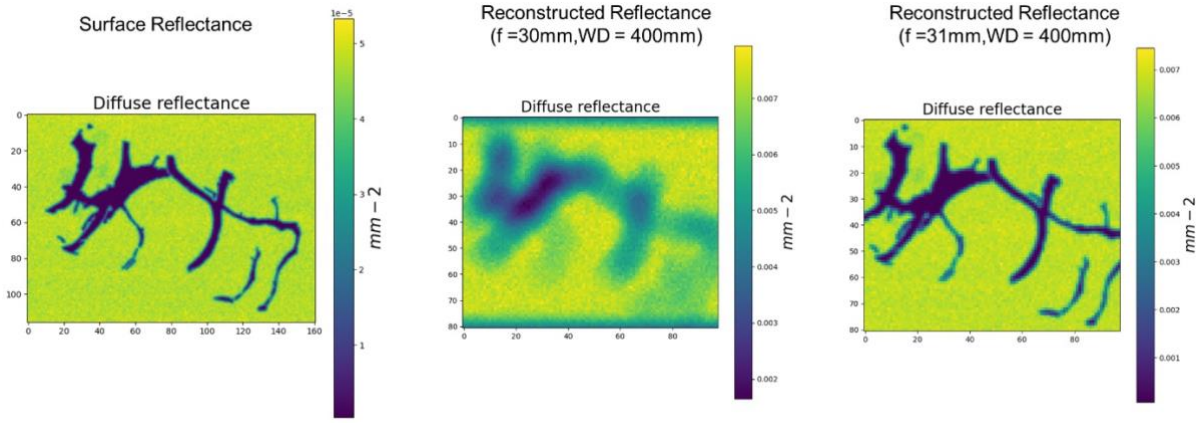

Figure 5. Effect of a 30mm and 31mm focal lenses at 400mm from the brain on the image quality of the diffuse reflectance. The left panel is the reflectance at the surface of the brain (i.e., perfect image), the central panel is the f=30mm lens (blurred image) and the right panel is the f=31mm lens (in focus).

1.2.2. Absorption scaling

As mentioned earlier, the simulation is run as a white MC, meaning that the absorption of the tissue is set to 0. Therefore, to produce a sensible image, one need to incorporate the absorption of the tissue to produce a realistic image. This is done by using the PPL information. Using the Beer-Lambert law, the weight of each photon is adjusted considering the absorption of each tissue type. The absorption is set based on the tissue composition and its physiological state, as $\mu_a(\lambda) = \sum_i \varepsilon_i * C_i$, with ε_i the extinction coefficient of the ith chromophore and C_i the concentration of that chromophore[9]. Here, C_i will be a function of the physiological state of the tissue. This state can include dynamic physiological processes such as brain/cortical activation that thus lead to changes in the optical parameters[13]. For example, in the image generated above, the tissue composition of gray matter and blood vessels were generated according to the tissue composition taken from the litterature[14], reported in table 1.

Table 1. Tissue constituent used to generate absorption of the tissues.

Tissue Constituents	Grey matter	Blood vessel
Water (%)	70	0
Fat (%)	10	0
[HbO_2] (µMol)	65	2375
[HHb] (µMol)	22.1	125
[oxCCO] (µMol)	5	0
[redCCO] (µMol)	1	0

1.2.3. Sensitivity scaling

Finally, the number of photons detected is adjusted according to the overall light spectral sensitivity of the instrument. The idea here is that typical source cannot be considered as Dirac delta function and even a source specified at a specific wavelength is spread over a few nm. Moreover, the sensor has a specific light detection efficiency and spectral resolution, which will scale the response differently at different wavelength. An example of the spectral characteristic of different sources and of the sensitivity of the detector is presented in figure 6. On figure 6(a), the spectral shape of a supercontinuum laser source filtered with acousto-optics tunable filters (from the HyperProbe-1 instrument) is presented. We note that the light source response of each wavelength is spread over a few nm and follows a gaussian distribution. On figure 6(b), the sensitivity of the detector a CMOS sensor (CS135MU - Thorlabs) is presented. We can see the sensitivity of the sensor varies widely as function of wavelength. Taking into account this information is important given the large bandwidth of our instrument.

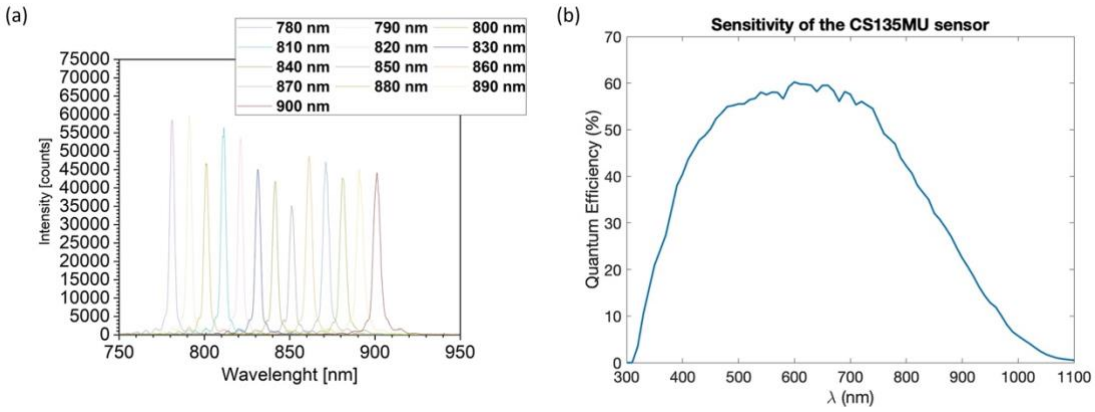

Figure 6. (a) - Example of the spectral shape of a supercontinuum laser source filtered with acousto-optics tuneable filters (from the HyperProbe-1 instrument). (b) - Example of responsivity from a CMOS sensor (CS135MU - Thorlabs).

Therefore, to simulate the recording of an image at a specific wavelength, and considering actual instrument parameters, simulation of continuous wavelength needs to be performed in order to integrate the results over the bandwidth of the source. Figure 7 gives an example of this.

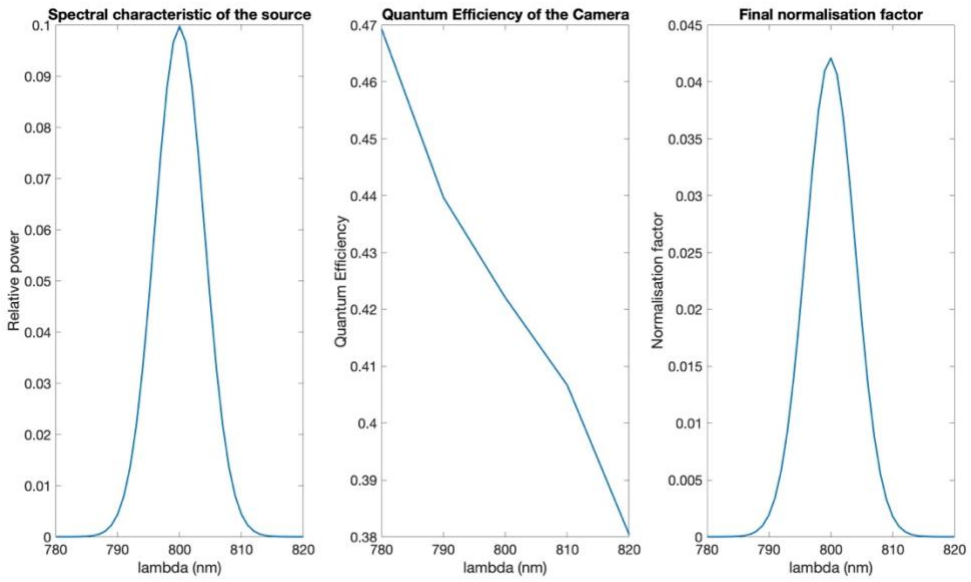

Figure 7. Left - Spectral characteristic of a light source with a gaussian distribution at 800 nm with a FWHM of 10nm. Middle – Responsivity of the CS135MU sensor in the same wavelength range of the source. Right – Final normalisation vector to be taken into account to model the effect of the sensor and the spectral distribution of a source centred at 800nm with a gaussian distribution of 10 nm of FWHM.

The Left panel present the spectral distribution of a source centered at 800nm with a FWHM of 10 nm. In order to be able to simulate a recording using this source, one needs to simulate a high number a wavelength around the central one, in order to integrate the results. Here we used 41 wavelengths (800 +/- 20nm with a step of 1nm). On top of the spectral distribution of the source, the sensitivity of the sensor in that range (middle panel) is taken into account for each wavelength. This results in a final normalization factor for each recorded wavelength (right panel). The normalized intensity for each wavelength is then integrated over this spectral range in order to produce the final realistic image, presented in figure 8.

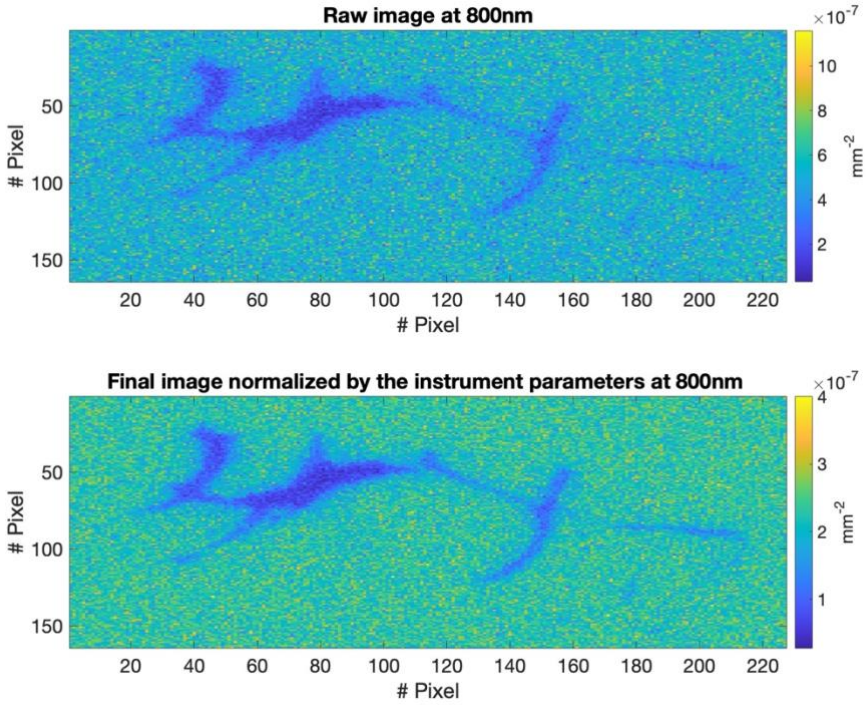

Figure 8. Example of the production of an image considering the responsivity of the sensor and the spectral shape of the source.

1.2.4. Incorporation of noise

The final step of the process is to add realistic noise to the images produce. It is a simple addition of noise pattern to the final images produced. We will use the noise level characterized in our real instrument and incorporate it in our final images.

2. Example of simulated Instrument parameter: The numerical aperture

Here, we present an example of the use of our DIS on a specific parameter. Indeed, this simulator can be used to model the entire hardware, but one of the benefits of numerical approaches is to be able to isolate one parameter to study its effect independently. Here we report a quick study of the effect of the NA on the number of detected photons. To do so, we have used the raw output of one of the simulations presented in Giannoni *et al.*[6], with a mesh presented in figure 9.a. By knowing the directions of every exiting photon comprised in the aperture of the system, we can select the photons that are not exceeding the maximum angle of the NA of the system, and can be detected. For example, by taking an arbitrary OS using a lens with a F-number of 1.7 (half-angle 17.11°), we see that only 8.1% of the total number of exiting photons can be collected by the system. An illustration of the exiting photons together with their angular distribution is presented in figure 9b and 9c.

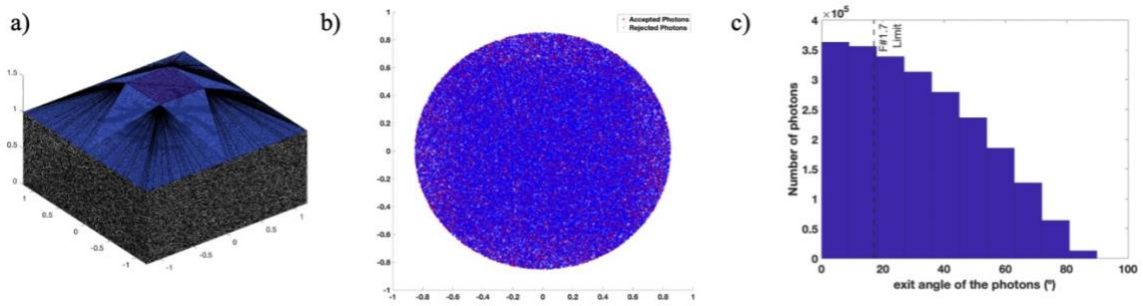

Figure 9. a) Illustration of the 3D-Mesh used for this simulation, comprising the source, the air and the brain tissues. More information can be found in reference x. b) Example of an image of the exiting photons passing though the aperture of the OS. The red dots represent the photons able to be detected by the system (here F#1.7). c). Histogram of the angular distribution of the exiting photons passing through the aperture of the system. The limit angle of the F#1.7 is also marked.

This example shows how we can use our DIS to study a specific parameter. This could be used to optimize the development of our system in order to maximize its light harvesting by looking at the benefits of using OS with various NA. It can notably help us to gain time in the development by quantifying the potential gain of using different optics and focusing on the most significant parameters. For example, using an OS with a F#1.4 would collect 4% more light compared to the F#1.7. This information can be used in combination with other system's parameter in order to decide if such improvement would be significant for the data quality of our images, and thus whether or not it should be implemented physically.

This examples also highlight one of the benefits of using MC simulator. As the information used by the instrument simulator only require basic information about the output of the simulation, one can use previously simulated data to incorporate its instrument parameter. It is worth noting that even if we have used MMC and MCX, any MC code could be used, as soon as they also output the position and direction of the exiting photons, together with their PPL.

However, one can note that the number of photons captured by the simulated instrument here is greatly reduced compared to the initial number of photons detected in the simulation presented. Therefore, one needs to be careful about the number of photons simulated in order to get a sufficient SNR in the final images.

3. Notes on computational cost of the simulator

The initial development of the software was done using MMC and MATLAB. However, the specificity of the HSI of the human brain made us adopt different strategies. The first issue that we encountered during the testing phase using the image of the human exposed cortex was the difficulty to run meshed-based simulation. Indeed, the FOV being large (more than 5 cm^2) the mesh generated was extremely large, resulting in memory issue during the simulations. We thus used the voxel-based version on the image which is less intensive from a memory point of view.

The second issue that we faced was the number of detected photons. Indeed, since the FOV is very large, a large number of photons needs to be detected in order to produce an image with a sufficient SNR. This led to issues with the processing time of the output files of the simulation in MATLAB. Indeed, the reconstruction of a single image would take more than 1 day. We therefore transitioned to C++ to implement the DIS to reduce processing time drastically (reconstruction of an image in 2 min).

Therefore, we managed to produce a tool that is more suitable for large FOV system. As a reference, a set of 41 simulations (1/wavelength) were run on the IN2P3 computation grid

Funded by the European Union
European Innovation Council

(https://www.in2p3.cnrs.fr/en) using Nvidia A100 GPUs. The run time per simulation was 10min for a 3D volume of 820 x 1137 x 265 voxels (resolution 75 µm) and 5e7 photons launched. Regarding the postprocessing, using the C++ implementation of the code, the finale images could be processed with a runtime of 2 min/image.

Conclusions

We have reported here the development of a digital phantom and its use to produce an optical DIS designed to optimize the development of a novel hyperspectral system for application in brain/cortex imaging. This digital phantom is based on MC simulations of the light propagation of the photons in tissues in order to model as realistically as possible the photon trajectories. Then the raw outputs of the MC simulations are integrated with the key instrument parameters in order to produce realistic images.

As the initial code of the DIS has been finalized, it can now be used to evaluate a variety of instrument parameters. We will now use the basic phantoms developed in WP3 to evaluate the data produced by the DIS to check the agreement between the instrument and its digital twin. This will allow us to validate the accuracy of our numerical tool before using it to optimize key parameters of the optical system, like the number of wavelengths used, without the need to physical implement all the solutions. Once this evaluation done, and the ultimate characteristic of the system agreed on, this simulator will be able to be used in order to generate large amount of synthetic data to help the training of ML algorithms as part of the WP4 work.

Bibliography

[1] Carles, G., Zammit, P. and Harvey, A. R., “Holistic Monte-Carlo optical modelling of biological imaging,” Sci. Rep. **9**(1), 15832 (2019).

[2] Fang, Q., Martelli, F. and Lilge, L., “Special Section Guest Editorial: Introduction to the Special Section Celebrating 30 years of Open Source Monte Carlo Codes in Biomedical Optics,” J. Biomed. Opt. **27**(08), 2022–2024 (2022).

[3] Fang, Q., “Mesh-based Monte Carlo method using fast ray-tracing in Plücker coordinates,” Biomed. Opt. Express **1**(1), 165 (2010).

[4] Cai, F. and He, S., “Using graphics processing units to accelerate perturbation Monte Carlo simulation in a turbid medium.,” J. Biomed. Opt. **17**(4), 040502 (2012).

[5] “http://mcx.space.”, .

[6] Giannoni, L., Lange, F. and Tachtsidis, I., “Investigation of the quantification of hemoglobin and cytochrome-c-oxidase in the exposed cortex with near-infrared hyperspectral imaging: A simulation study,” J. Biomed. Opt. **25**(4) (2020).

[7] Fang, Q. and Boas, D. A., “Tetrahedral mesh generation from volumetric binary and grayscale images,” Proc. - 2009 IEEE Int. Symp. Biomed. Imaging From Nano to Macro, ISBI 2009(Isbi), 1142–1145 (2009).

[8] Pifferi, A., Taroni, P., Valentini, G. and Andersson-Engels, S., “Real-time method for fitting time-resolved reflectance and transmittance measurements with a Monte Carlo model,” Appl. Opt. **37**(13), 2774 (1998).

[9] Kocsis, L., Herman, P. and Eke, A., “The modified Beer–Lambert law revisited,” Phys. Med. Biol. **51**(5), N91–N98 (2006).

[10] Sudakou, A., Wojtkiewicz, S., Lange, F., Gerega, A., Sawosz, P., Tachtsidis, I. and Liebert, A., “Depth-resolved assessment of changes in concentration of chromophores using time-resolved near-infrared spectroscopy: Estimation of cytochrome-c-oxidase

uncertainty by Monte Carlo simulations," Biomed. Opt. Express **10**(9) (2019).

[11] Thetford, A., "Introduction to Matrix Methods in Optics," Opt. Acta Int. J. Opt. **23**(3), 255–256 (1976).

[12] Yao, R., Intes, X. and Fang, Q., "Direct approach to compute Jacobians for diffuse optical tomography using perturbation Monte Carlo-based photon 'replay,'" Biomed. Opt. Express **9**(10), 4588 (2018).

[13] Caredda, C., Van Reeth, E., Mahieu-Williame, L., Sablong, R., Sdika, M., Schneider, F. C., Picart, T., Guyotat, J. and Montcel, B., "Intraoperative identification of functional brain areas with RGB imaging using statistical parametric mapping: Simulation and clinical studies," Neuroimage **278**(July), 120286 (2023).

[14] Charly Caredda. Imagerie optique spectrale : applications cliniques et précliniques. Imagerie médicale. Université de Lyon, 2020. Français. NNT : 2020LYSE1231 . tel-03602297 Caredda, C., "Imagerie optique spectrale : applications cliniques et précliniques" (2020).