Reconstruction of 4D spatio-temporal phenomena
Final report (2023)

1. Objectives of the project

The main focus of this research has been the solution of inverse problems where the parameters to be reconstructed are functions of both space and time. Among the possible practical applications, we primarily targeted medical imaging, i.e. *dynamic tomography*. In dynamic Positron Emission Tomography (PET) a time dependent concentration function should be estimated for each 3D grid point of the field of view. The number of grid points can easily be in the range of several hundred million.

Iterative inverse problem solvers simulate the estimated model, compare its output to the measured output, and modify the model estimation accordingly. As simulation is repeated many times and the data to be reconstructed as well as the measured data may have billions of elements, the simulation efficiency and performance are key factors in the solution.

There are four main considered tools for efficiency: non-analog simulation approaches that allow partial analytic solution of the original analytically non-tractable problem, importance sampling in Monte Carlo simulation, the exploitation of the massive parallel architecture of the GPU, and proper representation of the continuous data in high dimensions.

2. Research results

For simulation, we developed gathering type methods for partial differential equations and integral equations for efficient GPU implementation of particle/energy transport computation. We aimed at gathering type algorithms that can be implemented on parallel computers without causing write collisions. Gathering type approaches often correspond to adjoint transport operators, for which the governing equations must be established. In particular, we focused on the light transfer in surface models and participating media, and positron and gamma photon transport in volumetric models describing living organs or the body.

2.1. Dynamic Positron Emission Tomography

With dynamic positron emission tomography (PET), we study the metabolic processes by monitoring the uptake of a radioactive tracer. The goal is to reconstruct the time-activity curves (TACs) of the voxels, from which the relevant kinetic parameters can be obtained. These curves are assumed to have the algebraic form of a pre-determined kinetic model. Plausible algebraic forms have a limited number of free parameters and thus can be used even for low-statistic measurements. The algebraic form derived from the analysis of compartmental models depends nonlinearly on the nonnegative parameters to be determined. Because of the iterative nature of the maximum likelihood–expectation maximization (ML–EM) reconstruction, the fitting result of the previous step can serve as a good starting point in the current step; thus, after the first iteration we have a guess that is not far from the solution, which allows the use of gradient-based local optimization methods. However, finding good initial guesses for the first ML–EM iteration is a critical problem since gradient-based local optimization algorithms do not guarantee convergence to the global optimum if they are started at an inappropriate location.

The kinetic model typically involves the amount of radiotracer in the blood, which should be determined either by direct measurement or by extracting it from the PET data using image-derived or model-based methods. However, the direct measurement of the blood concentration is complicated, and the results of the image-derived and the model-based methods are also not reliable because of the partial volume effect and the unknown fraction of blood in the voxels. Moreover, in direct dynamic tomography, the kinetic model is fit from the very beginning of the reconstruction, thus the blood input function is needed even before the image-derived or model-based approaches can provide it. Our contributions to the solution of this problem are the following:

- We examined the robust solution of the fitting problem both in the initial phase and during the ML–EM iteration[1].

[1] L. Szirmay-Kalos, Á. Kacsó, M. Magdics, and B Tóth: *Robust compartmental model fitting in direct emission tomography reconstruction.* **The Visual Computer** (2021)

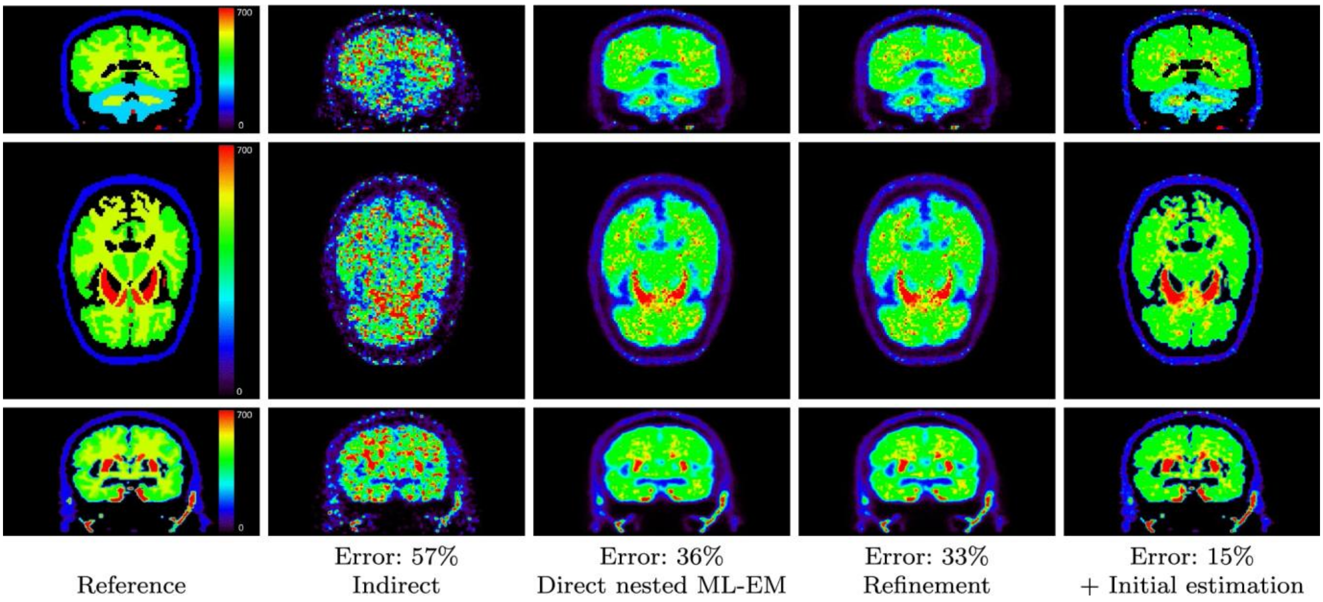

Figure 1. *Comparison of separated spatial and temporal fitting (Indirect), the state-of-the-art (Direct nested ML-EM) to the Refinement and Initial estimation strategies proposed by us.*

- We developed a method that is based on compartmental modeling and the Feng blood input function model defined by a fourth-order exponential equation with a pair of repeated eigenvalues, but does not require the blood input function and the fraction of blood parameters explicitly[2]. Thus, the model can be used in cases when we wish to have the robustness and advantages of compartmental modeling, but no blood input function measurement is available. The test results show that the added error in the TACs of not knowing the blood input is reduced in comparison with the spline-based TAC representation. The method can be applied in reference tissue analysis and in image-derived input function approaches.
- We reviewed the possibilities of various post-processing filters to enhance the resolution and the noise characteristics of the reconstruction and proposed improvements to incorporate information from multiple modalities[3].
- We proposed the incorporation of the modeling of inter-crystal scattering into DOI-based PET reconstructions[4]. The transfer probabilities caused by inter-crystal scattering are determined off-line with Monte Carlo simulation and are built into the system matrix as a factored component.
- We elaborated a class of methods to take into account and to compensate patient movements during PET reconstruction[5].

We applied deep learning techniques for medical imaging[6]. Originally, we planned the reconstruction of the hidden blood input function with deep learning. The implemented approach is not bad, but did not meet or original expectations. However, the gained experience helped to successfully solve a medical diagnosis and segmentation problem for Bladder Tumors.

---

[2] L. Szirmay-Kalos, M. Magdics, D. Varnyú: *Direct dynamic tomographic reconstruction without explicit blood input function*. **Biomedical Signal Processing and Control**, Volume 80, Part 1, 104313, 2023.

[3] D. Varnyú, L. Szirmay-Kalos: *Comparison of Non-Linear Filtering Methods for Positron Emission Tomography*. **Infocommunication Journal** 12 (2), 63-70

[4] L. Szirmay-Kalos, D. Varnyú, M. Magdics, and B. Tóth: *Modeling of Depth of Interaction with Inter-crystal Scattering for PET Reconstruction*. Molecular Imaging Conference (MIC), Paper 1459, Boston, 2020

[5] L. Szirmay-Kalos, D. Varnyú, and M. Magdics: *Motion Compensation for Dynamic PET with Continuous Motion Blur*. Molecular Imaging Conference (MIC), Paper: 1358, Manchester, 2019.

[6] Varnyú, Dóra and Szirmay-Kalos, László: *A Comparative Study of Deep Neural Networks for Real-Time Semantic Segmentation during the Transurethral Resection of Bladder Tumors*. **Diagnostics**. pages 2849, 2022.

2.2. Monte Carlo particle transport

Monte Carlo particle transport simulation in inhomogeneous participating media is important in algorithmic physics, rendering, and medical imaging. Monte Carlo methods require a lot of volume samples since the extinction coefficient needs to be integrated along light paths. Ray marching makes small steps, which is time consuming and leads to biased algorithms. Woodcock-like approaches use analytic sampling and a random rejection scheme guaranteeing that the expectations will be the same as in the original model. These models and the application of control variates for the extinction have been successful to compute transmittance and single scattering but were not fully exploited in multiple scattering simulation.

During the project, we have improved this method using a non-analog approach[7]. Our method attacks the multiple scattering problem in heterogeneous media and modifies the light–medium interaction model to allow the use of simple analytic formulae while preserving the correct expected values. The model transformation reduces the variance of the estimates with the help of Rao-Blackwellization and control variates applied both for the extinction coefficient and the incident radiance. Based on the transformed model, efficient Monte Carlo rendering algorithms are obtained. We have demonstrated the performance of the new method in materials of Rayleigh and Henyey-Greenstein scattering.

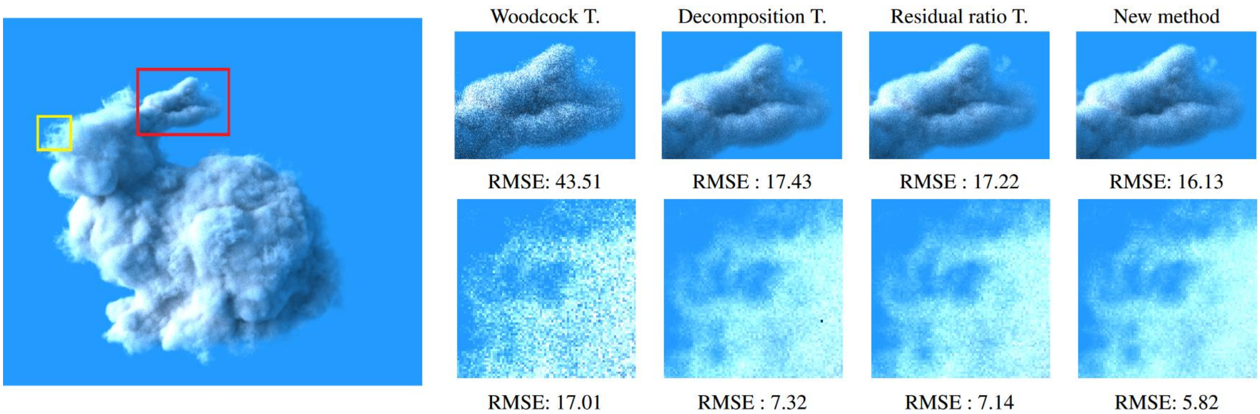

Figure 2. Comparison of the Blackweelization and control variate based method to state-of-the-art techniques.

2.3. Adaptive multiple importance sampling

Monte Carlo ray tracing is a fundamental algorithm to solve Fredholm type integral equations describing light transfer. At each light-material interaction, the light path is continued into a random direction, thus paths will be random samples of a high, in fact infinite dimensional space. The reason of using random samples rather than deterministic regular grids is that random samples fill high dimensional spaces with lower discrepancy, i.e. estimate high-dimensional integrals with lower error. In Monte Carlo simulation, the distribution from which the samples are drawn is a key design decision. The integration error depends on how well the applied Probability Density Function (PDF) mimics the integrand. However, the integrand is not available analytically, it becomes known just after the sampling, so mimicking is only partially possible, the unknown part should be added with heuristics or learned from previous samples. Multiple Importance Sampling (MIS) mixes different heuristics keeping their advantages. Our contribution was the analysis of and the development of methods for the optimal combination taking advantage of the statistics of samples obtained so far.

- We analyzed the variance of different combination techniques and concluded that the equal sample count proposed by Veach can be improved far beyond he suggested[8],[9].

---

[7] L. Szirmay-Kalos, M. Magdics, M. Sbert: *Multiple Scattering in Inhomogeneous Participating Media Using Rao-Blackwellization and Control Variates*. **Computer Graphics Forum**, Vol 37, No 2, 2018.

[8] M. Sbert, V. Havran, L. Szirmay-Kalos, Víctor Elvira: *Multiple importance sampling characterization by weighted mean invariance*, **The Visual Computer**, June 2018, Volume 34, Issue 6–8, pp 843–852

[9] M. Sbert, V. Havran, L. Szirmay-Kalos: *Multiple importance sampling revisited: breaking the bounds*. **EURASIP Journal on Advances in Signal Processing** 2018 (1), 15

- We have shown that the Importance Sampling can also be considered using the theory of statistical divergences[10], and reviewed the application of different divergences, including Kullbach-Leibler, chi-square, Hellinger, Tsallis, etc.
- The weights of optimal combinations can be obtained by solving a non-linear equation where parameters are statistical estimates thus their variances affect the obtained weights. Thus, there are two problems: the statistical uncertainty of previous estimates as well as the robustness and added computational cost of non-linear equation solving. Therefore, we have shown that complex non-linear equations can be approximated by simple linear equations that have robust solution[11].
- Importance sampling methods have been applied in both physically based[12] and non-photorealistic rendering[13].

Figure 3. *Importance sampling in Metropolis path generation.*

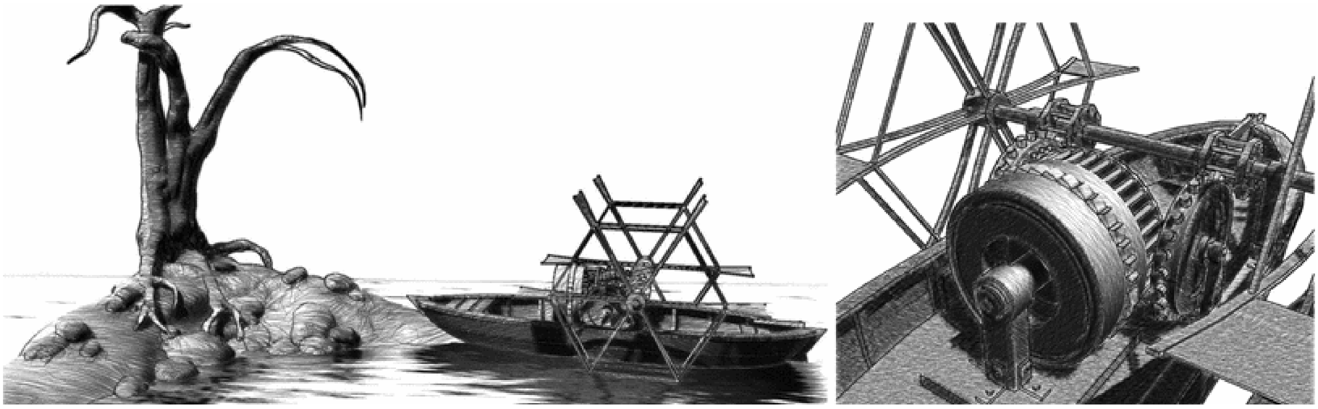

Figure 4. *Stroke placement with importance sampling.*

[10] M. Sbert, L. Szirmay-Kalos: *Robust Multiple Importance Sampling with Tsallis φ-Divergences*. **Entropy**, Volume 24, Issue 9, 10.3390/e24091240, 2022.

[11] M. Sbert, L. Szirmay-Kalos: *A linear heuristic for multiple importance sampling*. **EURASIP Journal on Advances in Signal Processing** (1687-6172 1687-6180): 2023 1 Paper 31. (2023)

[12] L. Szirmay-Kalos and L. Szécsi: *Improved Stratification for Metropolis Light Transport*. **Computers and Graphics** 68C. 2017. pp. 11-20.

[13] T. Umenhoffer, L. Szirmay-Kalos, L. Szécsi, Z. Lengyel, and G. Marinov: *An image-based method for animated stroke rendering*. **The Visual Computer**, June 2018, Volume 34, Issue 6–8, pp 817–827.

2.4. Sampling and reconstruction of three-dimensional scalar fields

3D scalar fields are usually represented by a Cartesian grid and values between the grid points are reconstructed by using trilinear interpolation while gradients are estimated by calculating central differences on the fly. In a GPU implementation, this requires seven trilinear texture samples: one for the function reconstruction, and six for the gradient estimation. In a series of papers we stepped beyond the state of the art in this problem:

- We showed that the six additional samples can be used not just for gradient estimation, but for significantly improving the quality of the function reconstruction as well[14]. As the additional arithmetic operations can be performed in the shadow of the texture fetches, we can achieve this quality improvement for free without reducing the rendering performance at all. Therefore, our method can completely replace the standard trilinear interpolation in the practice of GPU-accelerated volume rendering.
- Furthermore, we also showed that, using only eight trilinear texture fetches, high-quality filters such as the triquadratic B-spline or the Mitchell-Netravali notch filter can be efficiently evaluated together with their analytic derivatives, which are necessary for the shading computations[15].
- For visualization models that do not require gradient estimation, such as translucent volume rendering or Maximum Intensity Projection (MIP), we proposed a fast GPU-accelerated evaluation of the classical Catmull-Rom spline interpolation that requires only eight trilinear texture fetches[17].
- As the Face-Centered Cubic (FCC) lattice has been recently proposed for sampling scalar fields below the Nyquist limit, we adapted our Cosine-Weighted B-spline (CWB) reconstruction technique for the FCC lattice. Note that, previously, we proposed CWB reconstruction for the Body-Centered Cubic (BCC) lattice, which is optimal for sampling isotropically band-limited scalar fields above the Nyquist limit[16].

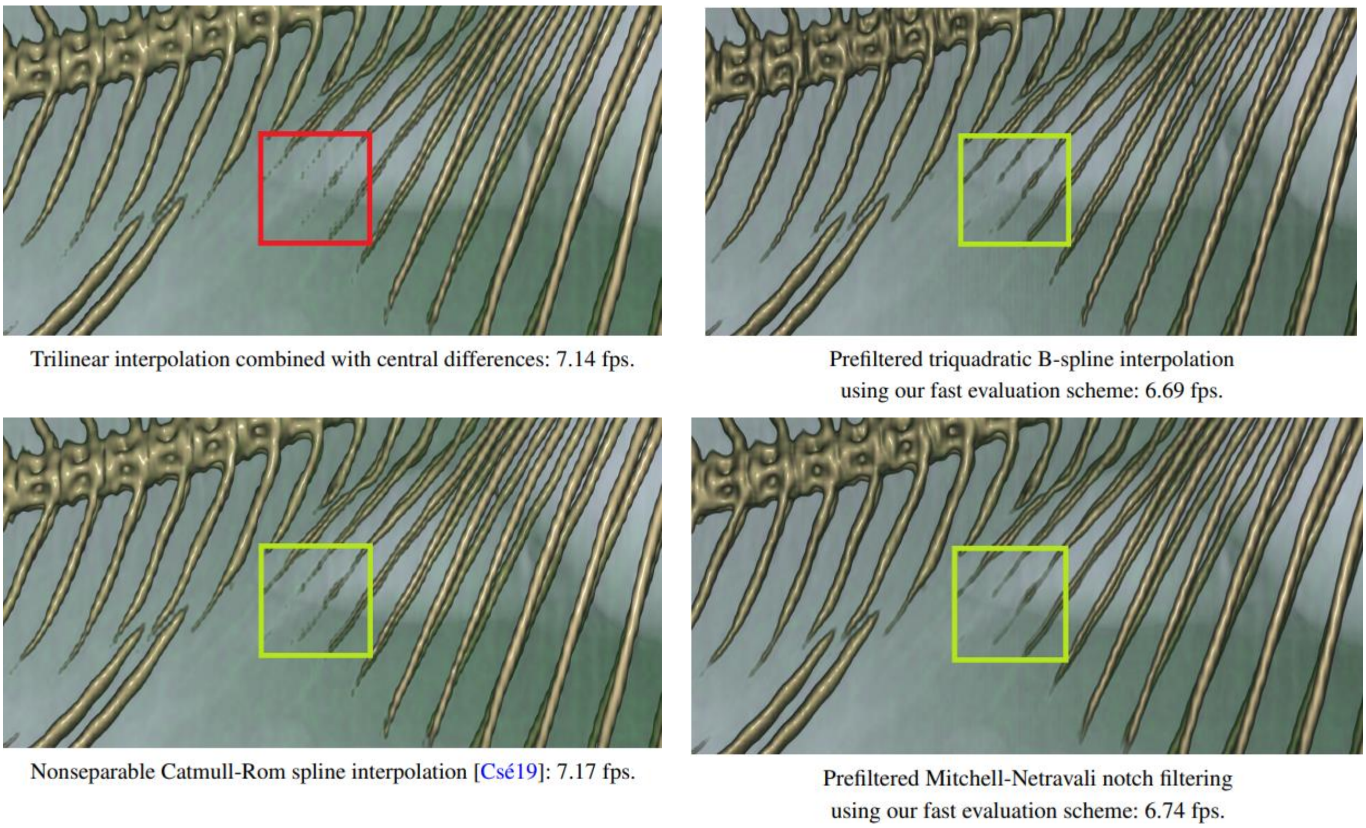

Figure 5. *Comparison of the new method with state-of-the art.*

2.5. Visualization of Non-Euclidean worlds

Curved spaces are very un-intuitive to our eyes trained on Euclidean geometry. Games provide an interesting way to explore these strange worlds. Games are written with the help of modeling tools and game engines based on Euclidean geometry. We have addressed the problem of adapting 3D game engines to the rules of curved spaces, including elliptic and hyperbolic geometries[14]. We considered the conversion of Euclidean objects, geometric calculations, transformation pipeline, lighting and physical simulation.

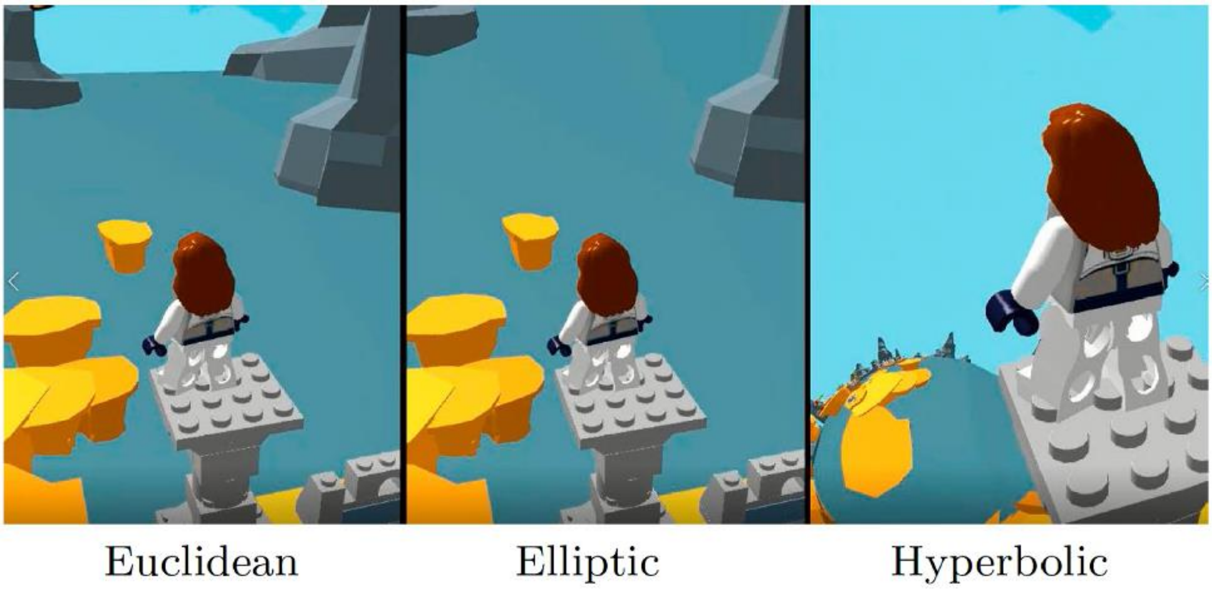

Figure 6. *Comparison of using the axioms of Euclidean, Elliptic and Hyperbolic geometries, respectively.*

2.6. Techniques in numerical optimization

Many methods used in this project are related to numerical optimization, where the gradient and Hessian computations are important steps. We also had results in these fields:

- We have proposed the filtering and convexification of the target function during optimization to make sure that the process does not stuck in local minima[15].
- We have generalized the concept of dual-numbers for the computation of higher-order automatic derivation[16].

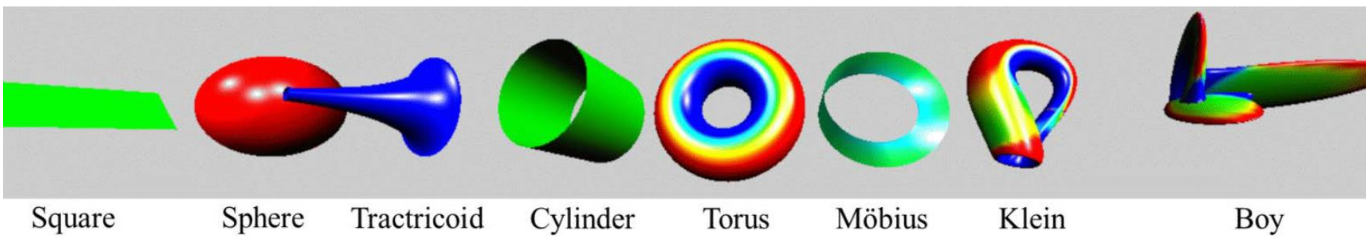

Figure 7. *Gaussian curvature of classical parametric surfaces calculated with higher order dual numbers.*

[14] L. Szirmay-Kalos, M. Magdics: *Adapting Game Engines to Curved Spaces*. **The Visual Computer**, 2021.

[15] László Szirmay-Kalos: Kinetic Model Fitting with Forced Convexification. Molecular Imaging Conference (MIC), Paper: 1357, Manchester, 2019.

[16] L. Szirmay-Kalos: *Higher Order Automatic Differentiation with Dual Numbers*. **Periodica Polytechnica Electrical Engineering and Computer Science**, 2020. https://doi.org/10.3311/PPee.16341

3. Utilization

The theoretical results have been utilized in practical applications. Our PET related results have been built into the Mediso's Tera-Tomo software system developed for PET reconstruction. The results on the non-linear PET filters were exploited in the Claritas' iPet software, which has received Federal Drug Administration (FDA) clearance in the USA. While investigating the application of deep learning techniques in medical imaging, we implemented several neural networks for medical image segmentation. One of them is about to turn to a commercial application, called CystoSmart, and its FDA examination will start soon.

During the project, three participants, Kacsó Ágota, Tóth Márton, and Rácz Gergely received their PhD degrees based mainly on the results achieved by them in this project.