Computed Tomographic Capture of Cultural Objects: Best Practices for the Use of CT Technology in GLAM Institutions

Matthias Göggerle

Deutsches Museum

Translated by Daniel Dumas

29.09.23 / 03.11.23 (transl.)

Content

1. Introduction

As part of the 3D-Cipher project funded by the German Federal Ministry of Education and Research (BMBF), the Deutsches Museum carried out extensive 3D-CT indexing of its cryptological collection from October 2020 to September 2023. In line with the objectives of the eHeritage funding programme, this enabled the provision of digital indexing data for research purposes. Since many of the cryptological devices and machines are based on complex technology that cannot be seen from the outside, industrial computer tomography was chosen as the measurement method.[1] Extensive 2D digitisation and occasional 3D surface scans were also carried out.

The findings on the feasibility of CT projects in cultural institutions gained over the course of this project are compiled here as an example of best practices and do not claim to be exhaustive. In addition to the comprehensive technical recommendations on the scanning modalities of the MUSICES project at the Germanisches Nationalmuseum,[2] the focus is primarily on the feasibility from the perspective of cultural institutions.

The structure of the "digital guide" is chronologically oriented to an ideal project sequence. At the outset, there are project-preparatory considerations and processes that necessarily precede an extensive CT measurement initiative. In the project itself, the focus is on organisational processes and the actual measurements. This is followed by a look at how to work efficiently and as profitably as possible with the measurement data obtained. The final sections deal with the subsequent use and provision of the research data and results.

The considerations recorded here have been developed with a view to technically complex multi-material objects that require different scanning equipment than, for example, biological specimens.[3] In natural history fields, extensive CT measurements have long been part of analysis and indexing, but due to different prerequisites and questions, they can only serve as a blueprint for technical objects to a limited extent.[4]

2. Preliminary Considerations

2.1. How to start?

Computed tomographic measurements are often preceded by concrete questions that can largely be divided into the following categories:

- Conservation issues
- Object-specific questions
- Didactical questions
- Digital applications

While the first two points concern research, didactical questions and the generation of data for further processing in other digital applications are primarily addressed to external actors. All of these points are valid reasons for using computed tomography but should be weighed against less expensive and more resource-efficient alternatives.

---

[1] CT technology as such will not be discussed here. A good overview of industrial X-ray computed tomography can be found in Hanke 2010.

[2] See Bear 2018.

[3] For the use of industrial CT scanners in the micrometre range see Du Plessis 2016.

[4] The Museum für Naturkunde should be mentioned here as an example of one of the leading museums for CT measurements in Germany: https://www.museumfuernaturkunde.berlin/en/science/micro-ct-laboratory. The following applies to all links: date of retrieval 27.09.2023.

The next step is to check which institutions have suitable CT equipment and what costs can be expected here.[5] Simple, smaller measurements are possible for less than €1,000 but for more complex and larger objects the costs can go into the high five-figure range. This makes it all the more important to have a clear cost estimate at this stage.

Realistic time planning is also advisable at the beginning of the project if, for example, the results for an exhibit must be available by a certain date. The cooperation partners or service providers may only be able to offer certain time frames, for example, if the costs are lower in the context of a scientific collaboration than in the case of industrial contracts. If necessary, a tendering procedure must be carried out in public institutions, which in turn should be taken into account in the time planning. Likewise, the preparation and especially the analysis of the data takes time.

The desired results and objectives should be communicated prior to the scanning work so that they are taken into account during the measurements. If, for example, only a partial section is relevant for a specific research question, other parameters can be used than when the entire object is scanned evenly. Here, too, the expectations should be adjusted according to the conditions.

2.2. Planned use of the data

In addition to planning the actual CT measurements, comprehensive planning for the use and further processing of the data to be collected should be carried out in advance. All measures should be scrutinised with regard to financial feasibility, existing structures (personnel, hierarchical, technical), time expenditure, and benefits. This becomes all the more relevant the more one goes beyond the purely internal and project-specific use of the data for research and conservation questions to the subsequent use in larger research and mediation contexts.

If the CT data are intended purely for internal purposes (i.e., to serve as a basis for research work or concrete restoration measures), only derivatives in the form of reports and screenshots may be needed, for example, if the analyses are to be carried out by third parties. If this work is to be carried out in-house, it makes sense to additionally store the reconstructed CT data (see Section 4). This ensures that the data can be reproduced and validated at a later time or used for new analyses. Care should be taken to ensure that the CT data are not exclusively in a proprietary format but in a common open-access format. This avoids the need to use a specific, and sometimes expensive, CT processing programme.

As soon as the data is intended for publication in any form, be it the reconstructed image stacks or a derivative thereof, considerations must flow into the data formats and the usage scenarios (more on data formats in Section 5.1). Ideally, all data needed to generate the relevant final results (analyses, videos, surface models) are stored. This ensures traceability as well as re-implementation if desired in the future.

2.3. Logistics

The issue of logistics is particularly important when either the scanning facility is far away or the objects are fragile or difficult to transport. Basically, the usual considerations for art transport apply, with regard to transport, insurance, etc. Especially in the case of more complex measurements or measurements encompassing several objects, it should be clarified how long the objects will be off-site and, above all, whether the desired storage conditions can be met in terms of security and climatic conditions.

[5] It is assumed at this point that an industrial CT scanner is not available nor is the institution planning on acquiring one.

Even though computed tomography is non-invasive, object-specific requirements should be communicated and agreed upon in advance. For example, it often makes sense from a metrological point of view to wedge or partially disassemble an object in order to improve the measurement results.[6]

3. CT Measurements

The actual CT measurements were comprehensively documented by the MUSICES recommendations for the specific case of musical instruments. These recommendations are in many cases applicable to other cultural objects.[7] At this point, therefore, only the deviations and the special case of metal-containing and thus dense materials, which are often accompanied by measurement artefacts, will be discussed.

Measurement artefacts occur very frequently in computed tomographic measurements and have various causes.[8] Many of the artefacts can be avoided directly during the CT measurement by calibration, repeated measurements or by post-processing the CT reconstructions. However, some effects can only be controlled to a limited extent and should therefore be considered in advance during planning. Beam hardening in particular can lead to disturbing image errors. Beam hardening occurs in particularly dense materials such as metal when penetrated by X-rays, as the photons scatter to different degrees.[9] The result is visible “stripes” in the CT images. These occur in particular when the density of two materials lying directly on top of each other differs significantly. In the case of strong beam hardening effects, materials can be so extremely superimposed that certain areas can hardly or not at all be recognised.

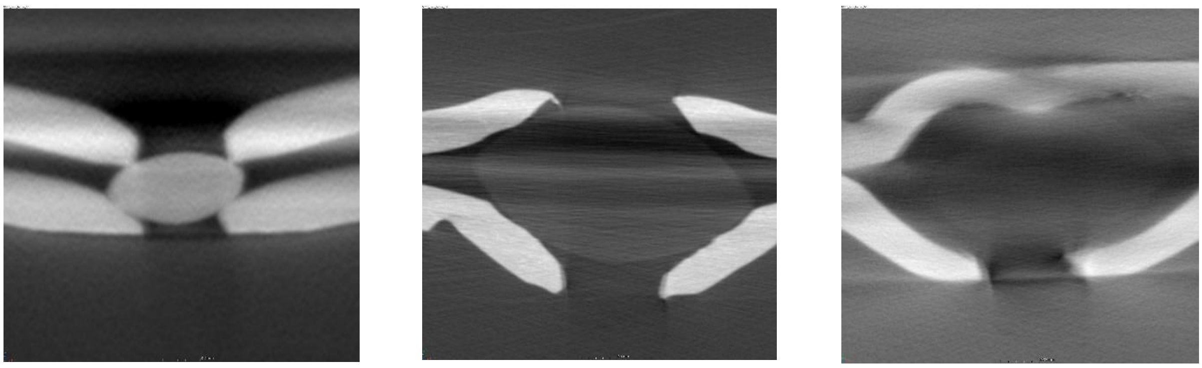


Figure 1. Differently pronounced beam hardening in three different single-lens microscopes. The lens in the left CT scan can be seen very well, while it can only be guessed at in the right image.
CT scans: Klaus Achterholz, TU Munich.

[6] Bear 2018, pp. 56–62.
[7] Ibid, pp. 16–19.
[8] Boas 2012.
[9] Lifton 2013, p. 1.

There are numerous approaches to minimise measurement artefacts such as beam hardening.[10] Even if it is not necessary for most users to deal intensively with these technical specifications, it is nevertheless useful to be aware of the possible consequences or the technical limitations of computed tomography.

4. Data Preparation

4.1. Software

Specialised software is necessary for working with CT image data. There is an abundance of free and premium software, which differ mainly in terms of usability, complexity, and range of functions.[11] When choosing software, it should ideally be clarified in advance which tasks the cultural institution itself will take on and which will be carried out by the scanning institution. Depending on how extensive the planned CT data to be used is and the degree to which the institution is contributing to a project, it becomes increasingly important to use a software that offers a wider range of functions.

For smaller CT analyses with very specific questions from the point of view of object history or conservation, the acquisition of one's own CT processing software is not absolutely necessary, as an evaluation can be carried out by the scanning service provider. The results can then be transmitted in the form of analysis reports, screenshots, and, if necessary, videos.

For more extensive projects, the acquisition of one's own software and familiarisation with it can prove useful and sensible. This is the case, for example, if repeated analyses are to be carried out with the measurements that exceed the cooperation agreement with the service provider. Similarly, in-house data processing may be useful for further processing of the CT data for mediation purposes in various forms. Here, further processing software may be needed, for example, for transformation into surface models or for simulation purposes.

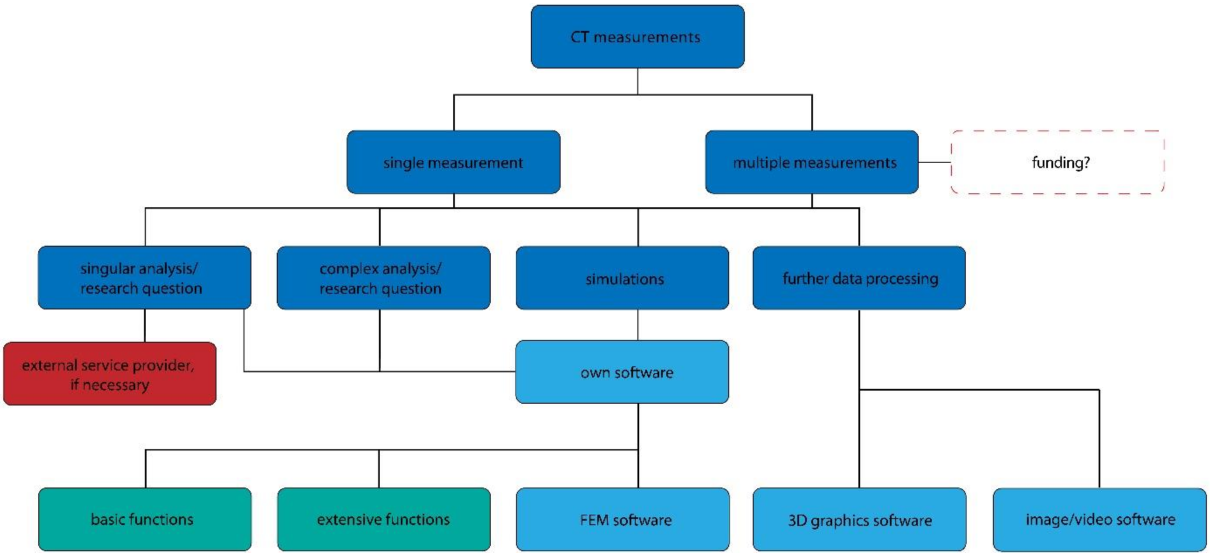


Figure 2. Graphic by author.

[10] See, among others, Lifton 2013.
[11] At this point, no recommendation for or against individual programmes should be made. A good, albeit not entirely up-to-date, overview of free programmes is provided here: https://peterfalkingham.com/2019/02/18/free-software-for-ct-segmentation-2019/.

The functional scope of different programmes differs considerably in some cases. The obligatory step of CT reconstruction is usually carried out directly after the CT measurements by the experts and should only be part of the functional scope for institutions with their own CT systems. The viewing of the individual sectional images in x, y, z axes and a 3D rendering are part of the minimum scope of most CT programmes. Many programmes differ from each other, especially in terms of quality and options for 3D rendering. Other features include geometric measurements and analysis tools such as porosity, deformation, and deviation analyses or target/actual comparisons. In the case of premium products, there are sometimes additional modules that have to be paid for and purchased as required. The possible input and export formats for the various data (CT, analyses, meshes, etc.) should also be considered. For example, some proprietary formats can only be read by your own software. Usually, however, common open-source formats are supported by most programmes. This is important, among other things, if further work is to be done with the data in FEM or 3D surface programs.[12]

Since many first-time users will be working with CT software, the accessibility, ease of use, and learning curve of the programmes should not be underestimated. In the case of freeware, there are often dedicated communities and extensive online tutorials, but these often require basic knowledge. Programming skills are also required in some cases, which make the programmes customisable but also more complex to use. Programmes that have to be paid for often offer training, but this is usually accompanied by additional costs in the three- to four-digit range.

In summary, the decision for or against CT processing software should be made on the basis of the following criteria:

- Type and scope of processing
- Financial means
- Human resources
- Time

4.2. Hardware

When considering suitable PC hardware, a distinction must again be made between how and to what extent the CT data is to be worked with.

A standard computer is sufficient for viewing the reconstructed image stacks alone. If 3D rendering or reslicing (realignment of the image stacks) is also to be carried out, the requirements increase. A powerful graphics card and sufficient RAM are the most important factors. For example, if the reconstructed volume is loaded completely into the working memory. With large volumes in high-gigabyte ranges, usual laptops, but also desktop PCs, reach their limits. Many programmes offer the possibility to load only a part of the data set or to skip individual slice images in order to reduce the size. However, this reduces the resolution.

For many steps of data preparation, such as surface determination, surface and volume meshing and animations, a powerful computer is also an advantage. For example, creating a polygon mesh of a large surface can take 30 to 60 minutes of computing time, even with high-end hardware.

Precise recommendations are difficult at this point and depend on the individual case as well as on constantly changing technologies. Basically, it is advisable to comply with the hardware requirements of the software providers and to rely on the expertise of the experts for the computed tomographic measurements. However, a dedicated mid-range graphics card and rather more RAM are generally

[12] The individual steps of data preparation are described in Section 4.3.

sensible investments. There is no need to invest a disproportionate amount of money in this area, but one should be aware of the possible limitations in terms of feasibility and computing time.

4.3. Steps in data preparation

The preparation of the CT measurements can be divided into different steps, which have to be carried out partially or as a whole, depending on the question and the desired result. The most important steps are listed and explained below.

4.3.1. CT reconstruction

Directly after the CT measurements, the CT reconstruction takes place (i.e., the calculation of the 3D volume data sets). This step is usually carried out directly by the scanning cooperation partner. The two-dimensional images of the computed tomographic measurement are converted into 3D volumes by interpolation. In this process, the grey values lying between two images are averaged.[13] In this highly automated step, unwanted image artefacts (beam hardening, ring artefacts, etc.) can be partially reduced.

4.3.2. Surface determination

After the CT reconstruction, further work can be done with the sectional images in suitable CT processing programmes. This is usually the first time that researchers come into contact with the CT data themselves.

The property that different materials absorb X-rays to different degrees is visualised as grey values in CT programs. This also means that, unlike surface scan methods, there is no clearly defined surface. Surface determination is needed for many other analyses, including all geometric measurements such as distances, angles, lengths, shape or position tolerances, nominal/actual comparisons or inclusion analyses, as well as for the creation of surface models. It is one of the most important steps in working with CT data.[14] This makes clean surface determination all the more important. For objects with different materials, basically every material can be defined as a surface.

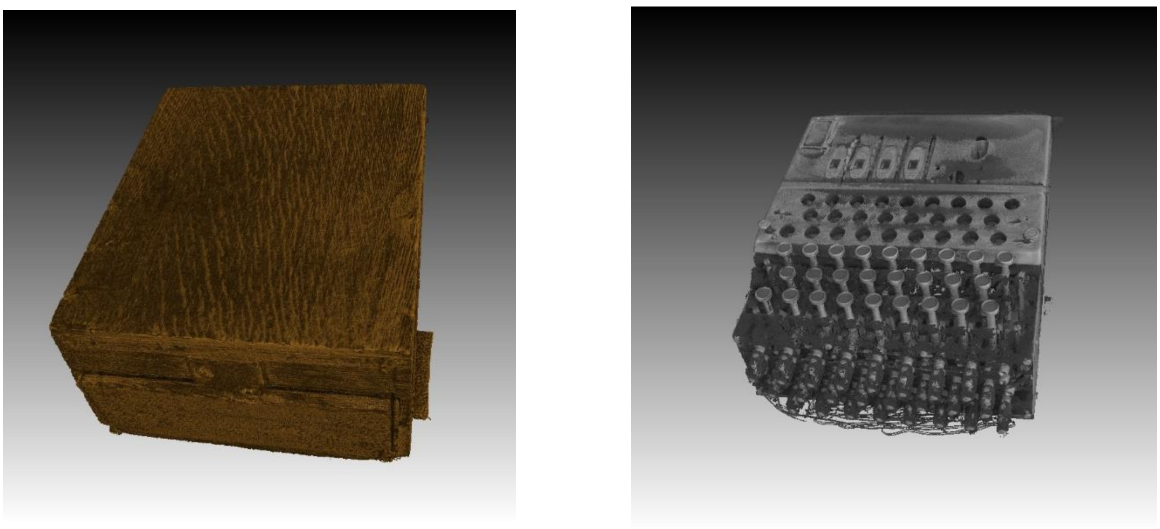


Figure 3. Various surface determinations of an Enigma M4, partially coloured.

[13] Lucchichenti 2005, p. 2147.
[14] For a comprehensive description of the surface determination of CT data see Kroll 2013.

The determination itself is carried out semi-automatically by the CT software using various threshold value methods.[15] The more similar different materials are in terms of their density—or their grey values in the CT measurement—the more difficult a clean surface determination becomes. This becomes clear when looking at a 2D sectional image without interpolated representation (Figure 4). The individual voxels (three-dimensional grid points) are assigned different grey values, which in sum appear like a coherent surface, but cannot be defined as an exact surface. The algorithms can compensate for this to a certain extent, but there may be deviations of greater or lesser magnitude. Especially with very exact analyses, this can sometimes lead to distorted measurement results. Therefore, it may make sense to repeat this step iteratively until a sufficiently exact accuracy has been achieved. Alternatively, this can be done based on manual segmentations, on which a cleaned surface determination can then be carried out (see Section 4.3.4).

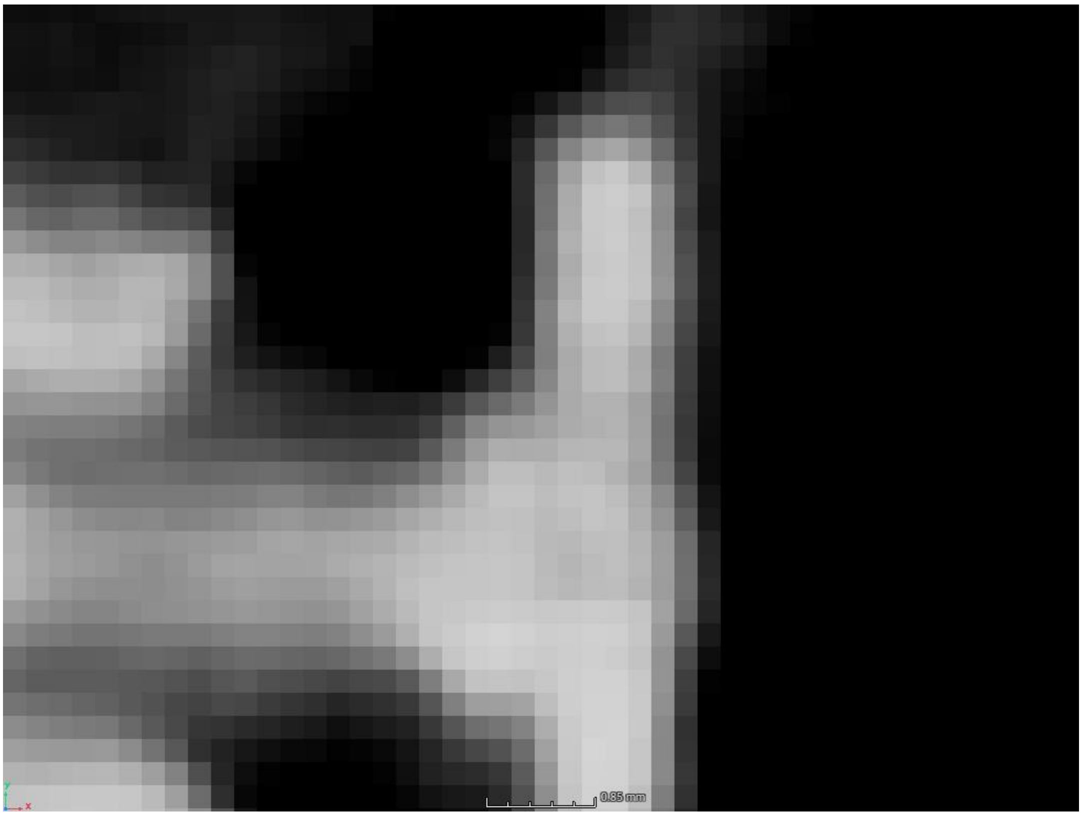


Figure 4. 2D CT cross-sectional image without interpolarised representation.

[15] Yang 2022 provides an overview of various algorithms.

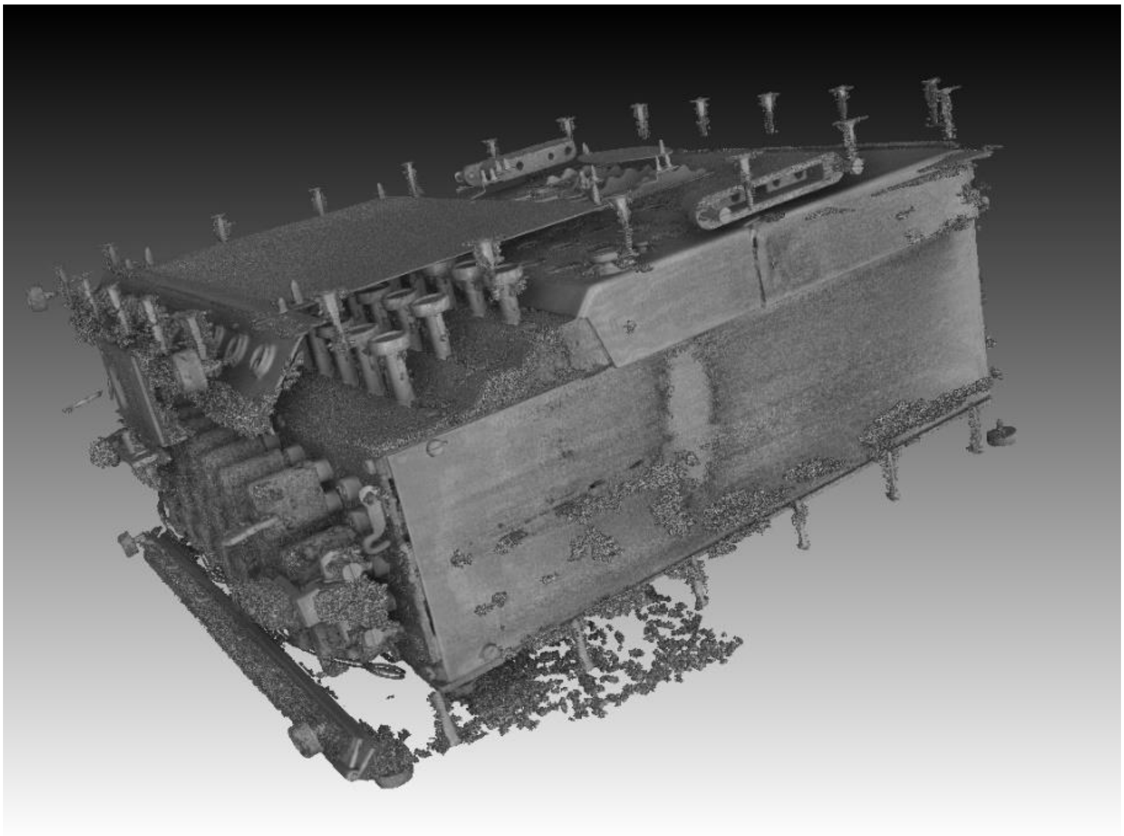


Figure 5. Unclean surface determination of an Enigma M4 with artefacts.

4.3.3. Metrology and analyses

After the CT data set has been defined with a surface, further analyses and measurements can be carried out. For this purpose, it often makes sense to abstract parts of the data set and describe them with the help of geometric shapes. For example, points, lines, circles, and surfaces, but also three-dimensional shapes such as spheres or cylinders are adapted to the object. A cleanly performed surface determination is essential here in order to achieve results that are as exact as possible. With the help of ideal shapes, deviations and deformations of the objects can then be detected, visualised, and numerically quantified.

Simple but very powerful tools are measurement analyses for distances and angles. This allows precise measurements of hard-to-reach components, which provide information about constructional details or can serve as a template for detailed replicas. The possibility of reslicing the sectional images, i.e., shifting the viewing plane, which is inherent in computed tomography, enables comprehensive analyses of the scanned objects.

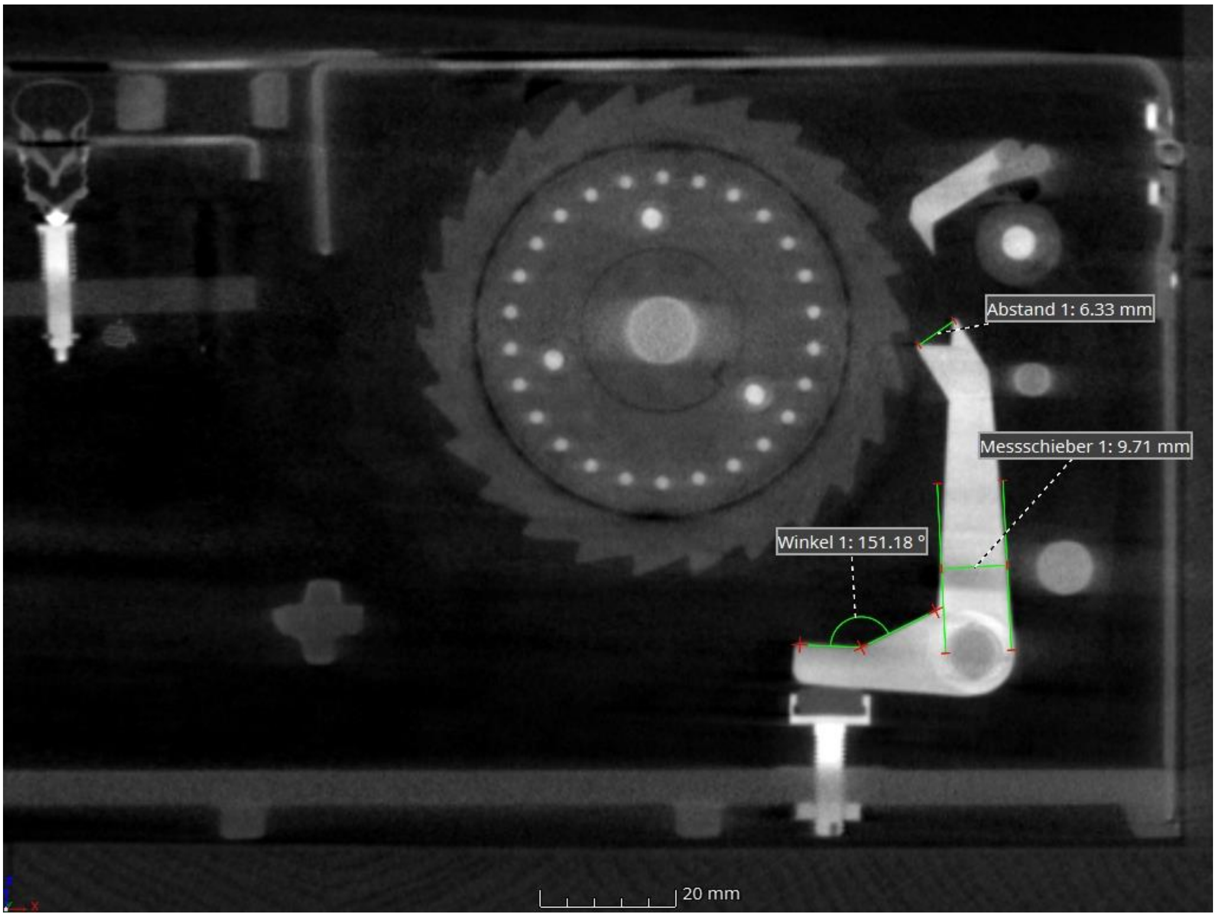


Figure 6. Distance and angle analyses on components of an Enigma M4.

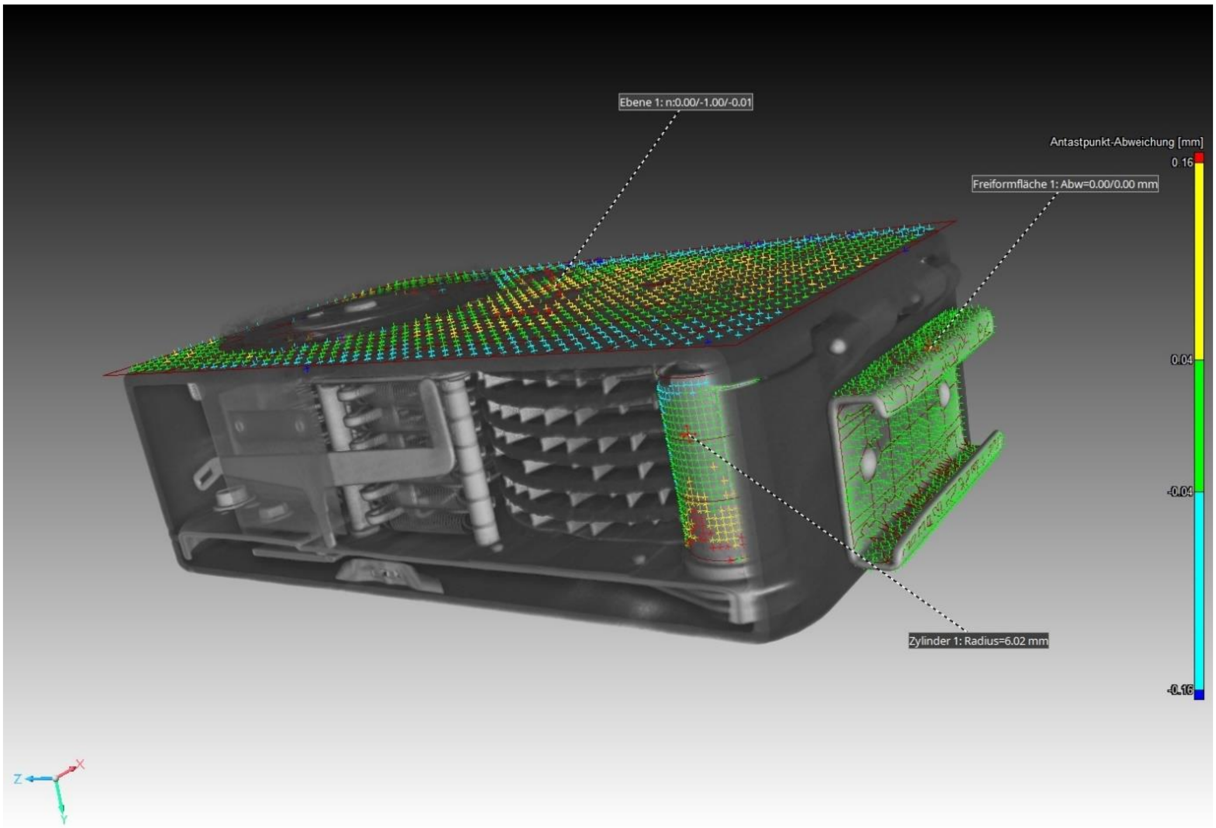


Figure 7. Geometric shapes adapted to a CT data set and the real deviations.

Even more complex analyses can be realised with the help of geometric shapes, such as angular distances, parallelism, sphericity, etc. The metrological methods are mainly used in material and component analysis to detect deviations from a given standard and to define tolerance ranges. By combining CAD models with blueprints of ideal geometric models, the smallest deviations can be detected and documented. However, CAD models are not usually available for use in cultural institutions. Nevertheless, comparative analyses can be carried out on the basis of the CT data, for example by scanning and comparing two objects of identical construction. To do this, both CT data sets must be loaded and digitally overlaid. Here, again, a surface determination is necessary, based on the characteristics of which the software can carry out the alignment. If necessary, the definition of geometric shapes is useful to improve the accuracy of the overlay. Because just as with the determination of the surface, the quality of the superimposition has a direct influence on the significance of the results. Otherwise, there is a danger of drawing conclusions about the physical objects that are in fact merely the result of the software setup. With careful preparation, this target/actual comparison finally provides a deviation diagram for all voxels, indicated in distance to the reference object (in nm, µm or mm). Tolerances, search distances, etc., can refine the analysis.

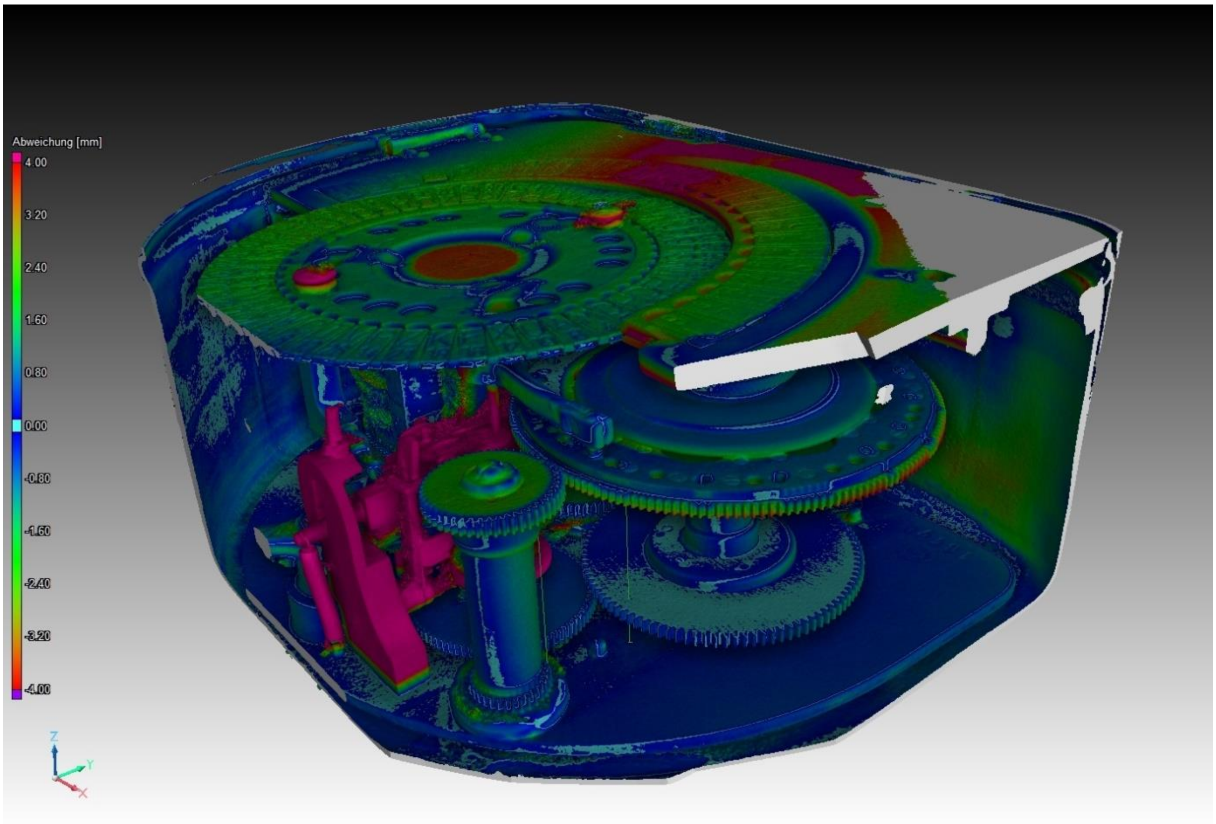


Figure 8. Digital overlay of two Kryha standard machines with coloured deviation diagram. The pink-coloured mechanics at the bottom left of the picture are not present in the comparison object.

Especially from a conservation perspective, deformation and porosity analyses can provide helpful knowledge about the condition of objects. With the help of digital volume correlation, the change in the object's condition between two CT scans can be visualised. This is particularly useful for documenting material degradation or particularly stressed components.[16] Inclusion or porosity analyses can be used for foreign body analysis, for example. Here, the algorithm uses threshold values to calculate the proportion of foreign material within a defined area. For the CT data, this differentiation is based once

[16] CT scans can complement the extensive analytical methods (spectroscopy, XRF, etc.) used in restoration research.

again on the assigned grey values and, analogous to surface determination, is more exact if the materials being compared significantly differ from one another. At the Deutsches Museum, for example, this analysis method made it possible to quantitatively measure the decomposition of a wooden strut on the basis of the proportion of air it contained. Based on these results, further conservation measures could then be planned in a targeted manner.

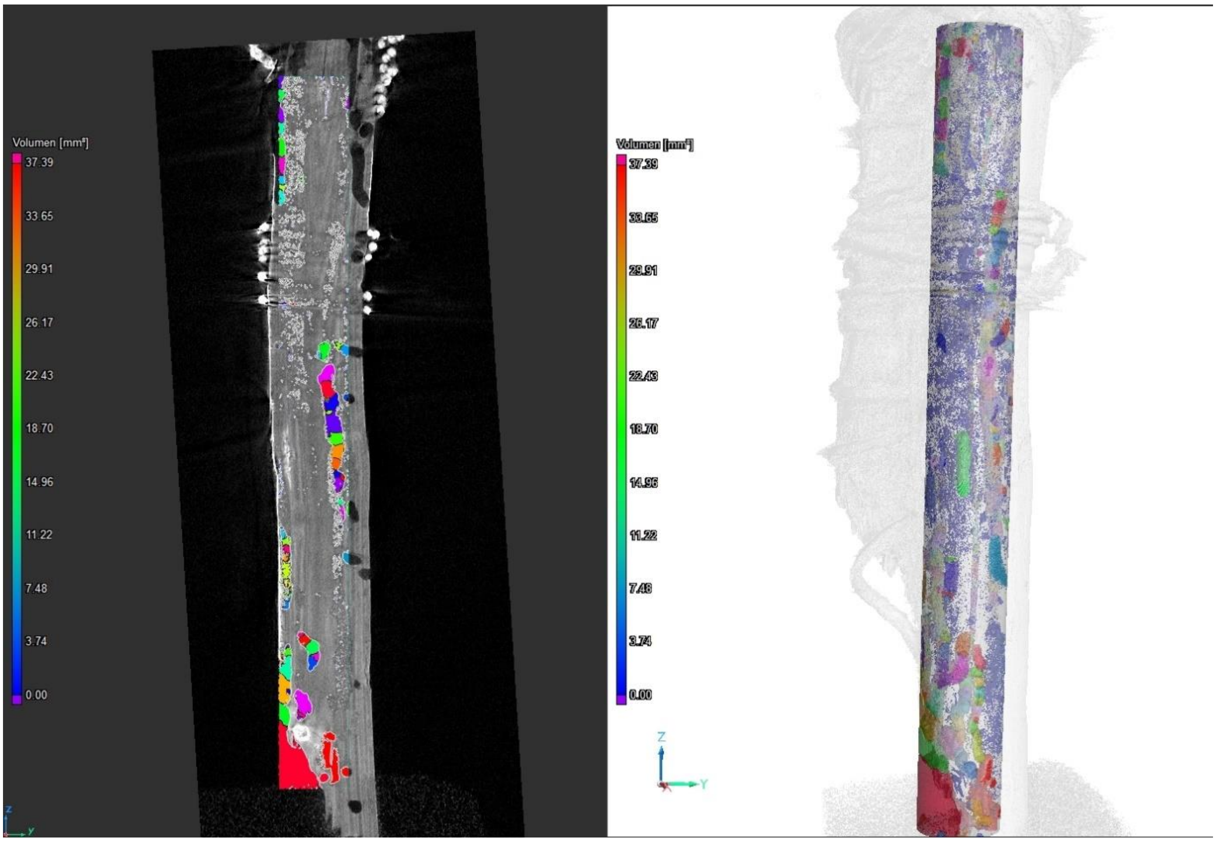


Figure 9. Porosity analysis of a strut of a glider by Otto Lilienthal in 2D and 3D view. The coloured areas depict air pockets that have been categorised and defined according to pore size.

4.3.4. Segmentations

Together with surface determination, segmentation is one of the most important tools for processing CT data. Here, the CT images are divided into individual regions of interest (ROI) and sharply separated from each other. In this way, for example, different materials or mechanical components can be exposed. Segmented components can then serve as the basis for animations and further visualisations that are intended to explain individual elements of an object in more detail. In addition, surface models can be created specifically on their basis and in part, the problem of different materials can be circumvented when determining the surface.

The segmentation of a CT data set can be done in different ways and depends on the material composition of the objects, the quality of the CT scans, the technical complexity, and the chosen objective. In addition, different CT programmes offer various segmentation options, some of which are based on their own algorithms.

Basically, three segmentation techniques can be distinguished: threshold-based, manual geometric, and automated segmentations. In recent years, there have been increasing attempts to automate the segmentation process with the help of machine learning and artificial intelligence.[17] Although some

[17] Examples of AI-based segmentation can be found, for example, in Borrelli 2021, Chen 2022, and Lang 2022.

software programmes already offer AI-based segmentation, at the moment there is often no way around manual segmentation. Threshold-based segmentation is based on the grey values of the voxel data and is closely related to surface determination (see Section 4.3.2). Here, too, the following applies: The more different the materials and thus their assigned grey values are, the easier they can be separated in the CT data set using threshold-based segmentation. A common example of this is the separation of wood and metal parts.

The segmentation process becomes more complicated with material mixes whose densities are very similar or when splitting components of the same material type. In the case of technical artefacts, this will often be the case when mechanics are to be broken down into their components. Since CT measurements are always an approximation depending on the chosen resolution, among other things, some components may “merge” in the digital CT dataset that are actually structurally separate. Segmenting these parts is therefore only possible manually, requiring considerable time.

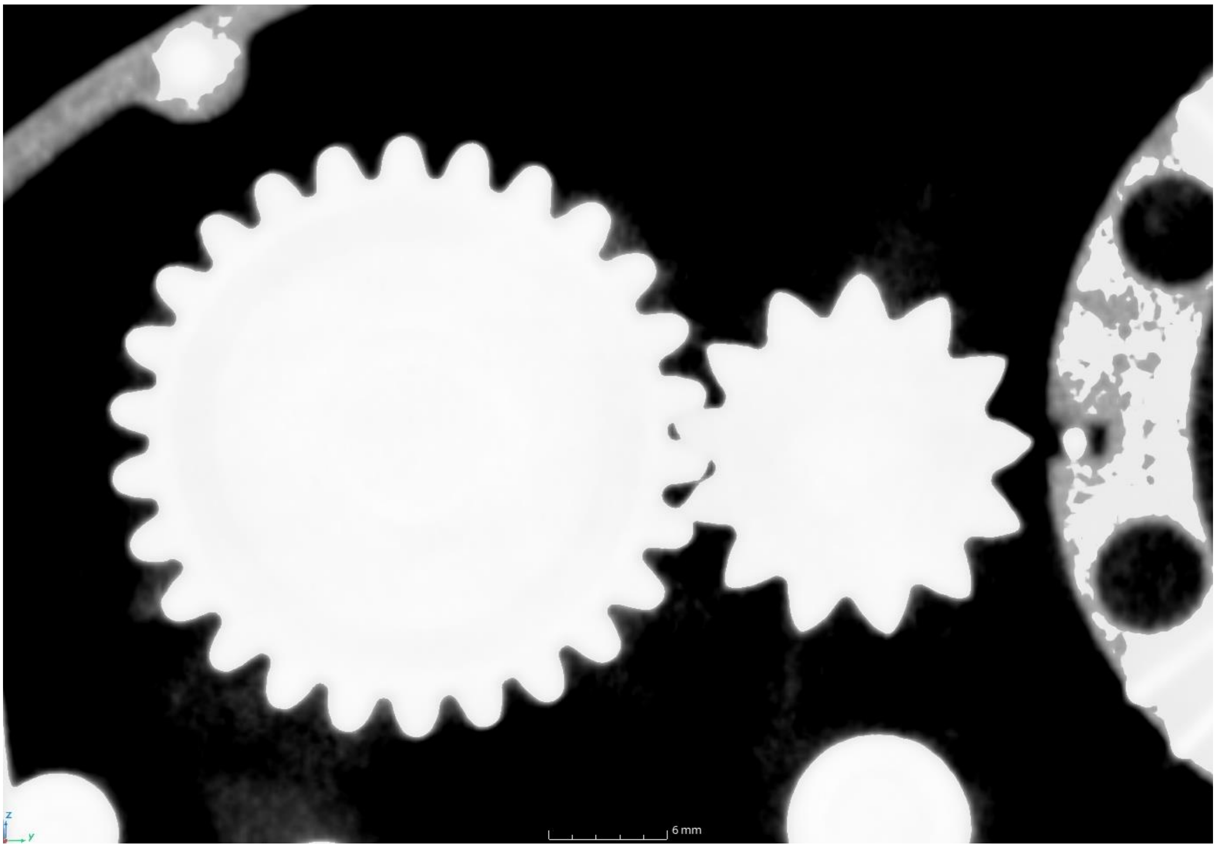


Figure 10. 2D-CT image of two apparently fused gears.

Manual segmentation works by selecting and defining the individual ROIs by drawing them in the 2D or 3D view. Here, too, various geometric shapes such as rectangles, circles, etc., are often available as tools with which to optimise this laborious process. Experience shows that the best results can be achieved with an iterative process of different steps, which, depending on the individual case, consist of geometric- and threshold-based segmentation, the use of geometric shapes (see Section 4.3.3), and ROIs based on surface determination.

The segmentation of the components of a Hagelin M-209 cipher machine is described here as an example. The mechanical system consists of several interlocking components, some of which are made of the same or similar material in a confined space. Neither a purely threshold-based segmentation nor a fast geometric segmentation led to acceptable results.

In the first step, a surface determination was carried out that corresponds to the grey values of the relevant parts of the mechanics.[18] From this, an ROI of the volume can be generated on the basis of the grey values, which can serve as a reference with regard to the grey values of all future ROIs. This means that all ROIs created afterwards are intersected with the volume ROI and form a subset of it. Subsequently, a manual segmentation was carried out for the components that correspond to "simple" geometric shapes. In this example, this worked particularly well with the cylindrical cage at the rear end of the cipher machine. Details were then reworked at individual points until the result met the quality requirements. To remove the grey values that do not belong to the selected component, such as the trapped air, the ROI was cut with the volume ROI as described above. This way, this method also works for components such as the printer axle, consisting of individual gears, as misleading surrounding air can be subtracted. This step was repeated analogously for all geometrically simple shapes.

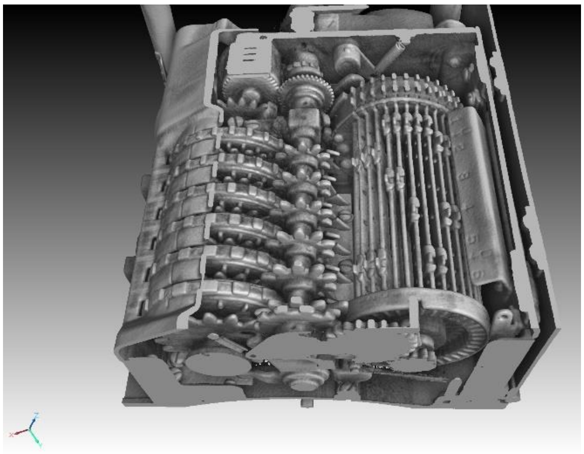
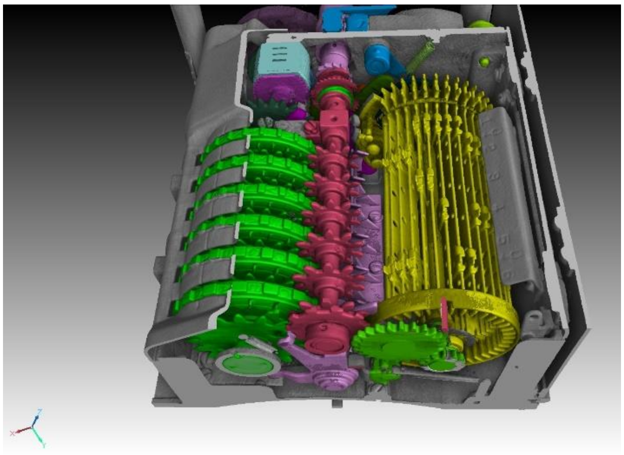
For complicated shapes such as the levers for transferring the cage to the rotors, an intermediate step was necessary. Thanks to the surface determination, a geometric free-form surface can be defined, from which an ROI can be created. Here, too, rework was necessary, such as extending the ROI by the search distance X or smoothing the ROI. The sometimes time-consuming process leads to significantly more accurate results than manual segmentation. Finally, the ROIs created in this way were cut again with the volume ROI.

ROIs can not only be intersected with another ROI as a subset but can also be subtracted from each other. This is particularly helpful because, in the case of parts that are close together, manual detail processing is only required once and this ROI can then be subtracted from the second one. Therefore, it can be suggested that creating a greater number of ROIs leads to a faster segmentation of the remaining parts.

All these processes require a certain training period to achieve good results. Fortunately, the processes can usually be directly reviewed, reworked, and corrected. Finally, it should be noted that full segmentation is often not necessary and the scope should be based on the desired outcome.

Figure 11. Hagelin M-209 cipher machine. Photo: Konrad Rainer, Deutsches Museum, CC-BY-SA 4.0.

[18] This step can be repeated as desired with other grey values if, for example, another material is to be the focus.

Figures 12 and 13. Unsegmented and segmented mechanics of the Hagelin M-209.

4.3.5. Surface models

If further 3D models are to be created with the CT data, a conversion into 3D surface models (surface or polygon meshes) can be useful. Many CT programmes offer this function and export in common formats. Surface definition is once again the decisive step for the quality of the mesh, as it defines the surface on which the polygons are mapped, which ultimately creates the 3D model. Here it is important to note: The tighter the mesh and the smaller the search distance, the longer it takes to calculate the model and the larger the surface model will be later. Since CT data naturally do not have any textures, unevenness is much more noticeable than in surface models created by 3D surface scans, which can conceal inaccuracies with textures. As a result, 3D models from CT data often consist of significantly more polygons and are correspondingly larger, which in turn can have a negative effect on the display in a web viewer or the performance in a 3D scene. If possible, surface models should therefore preferably be created from designated methods such as photogrammetry, laser scans, etc., instead of CT data. However, since CT data can be used to create surfaces that cannot be captured by other scanners (such as the mechanics inside a closed enclosure), they can still be a useful addition, especially if the CT measurements are already available.

Figure 14. An unclean surface model with incorrectly recognised material as well as gaps due to unclean surface determination of the CT data. Hagelin CD-57 cipher machine.

Surface models have multiple application purposes such as online display in a variety of 3D viewers, as a basis for animations and interactive scenes or as a template for 3D prints. There are approaches to combine computed tomographic 3D CT data with surface scan data and display it in one data set. The advantage is obvious, one obtains an accurately textured surface paired with the depth-defining information of computed tomography. Technically, the undertaking is not trivial, as different data models have to be merged. The result is a point cloud that is partly composed of the surface data and partly of the CT data.[19]

4.3.6. Volume meshing

Volume meshes can be created from the CT data, which, unlike polygon meshes, also depict the interior of the objects as a (tetrahedral) mesh. The very specific application is limited to finite element simulations for structural analyses such as strength and deformation investigations, fluid mechanics simulations, and special rendering applications. For most applications in GLAM institutions, however, these methods will not play a role.[20]

[19] The Gyrolog project has implemented this using historical gyroscopic instruments as examples: https://portal.wissenschaftliche-sammlungen.de/CollectionActivity/182412. For a description of the methodology, see Zhan 2020 and Fritsch 2021.

[20] For more on the use of volume meshing, see Varotsis 2021 and Kuhlmann 2022.

4.3.7. Animations

Renderings can be created indirectly via the diversions of the surface models or directly based on the CT models. Many CT programmes offer extensive rendering settings and animation options. For example, the animation can be exported in common video formats for further processing in video programs. For interactive animation applications such as in the gaming sector, conversion to surface models usually has to be resorted to.

The advantage of direct animation of the CT data is the possibility of seamless representation of different surfaces. In addition, a time-consuming conversion of the CT data into surface models is not necessary, as very good results can be achieved with the software's own settings. This allows videos to be created that can impressively convey the power of the CT data of "looking inside".[21]

5. Data Export

Data export is of great importance when working with CT (and also other) data. As long as you are only working within one programme, the question of data formats and exports does not arise. But at the latest when the data have to be exported and converted for other applications, there is no way around data export. It is therefore advisable to deal systematically with the possibilities and limitations of different formats at an early stage.

Which data or derivatives are to be exported again depends on what is to be achieved with the CT data. For researchers, CT reconstructions are of particular interest in order to carry out their own analyses without loss of information. Analysis reports are particularly interesting if they are to serve as a basis for further work such as publications. From a mediation point of view, images, animations, and surface models are relevant results. For archiving, online presentation, and possible Linked Data approaches, metadata also play an important role.[22]

5.1. Data formats

When choosing the data formats, it is again necessary to differentiate between the individual work steps. For example, derivatives from the CT data as surface models or volume meshes are ideally saved in common formats (obj, glb, stl, ply or inp, pat, bdf) for further processing.

The CT measurements themselves are often stored in proprietary formats of the processing software. As long as the data are used exclusively for singular research questions and are thus only needed in a closed ecosystem, these formats can be worked without any problems. As soon as the CT data are passed on to third parties, be it researchers or the general public, it makes sense to convert them to an open format. The CT data can be saved as image stacks. These are composed of the individual sectional images in one direction, whereby the number of images corresponds to the number of voxels in this direction. CT programmes then reconstruct the volume data from the image stacks. The sectional images can be saved in formats such as jpg or tiff and viewed through common image processing programmes. The DICOM format, which originated in the medical field, has established itself as the standard for CT, MRI, and ultrasound images and offers numerous metadata options, which are, however, optimised for medical purposes (patient data, etc.).

Analyses and measurements can usually only be saved within a project instance of the CT software. An export in an image stack for further processing in other programmes is usually not possible. Exports in report form (pdf, excel, etc.) for analyses nevertheless offer a possibility to export corresponding information.

[21] See an animation of the 3D cipher project of a NEMA cipher machine: https://youtu.be/FCbzTbclfBw.

[22] See Section 6.

Screenshots and animations can be exported in standard image formats (png, jpg, tiff) or video formats (avi, mp4 etc.) and then edited in appropriate programmes. The animation itself cannot be changed after export.

Annotations, i.e., labels attached to the CT slice images consisting of a free text and a position in x, y, z direction, can be exported in table formats if necessary. A difficulty arises from the reference system, i.e., the global or local coordinate system, which cannot be exported. Likewise, the third-party programme must allow the import of the annotations in this format, which is not necessarily given. Ideally, it is considered in advance for what annotations are needed and whether they can be added at a later stage (e.g., in online 3D viewers or in animations) instead of directly in the CT processing software.

5.2. Export of ROIs

ROIs have a special significance as a segmented subset of a CT dataset. In the narrower sense, no new research or analysis data are generated and all information is already available in the CT reconstructions. However, due to the importance of the segmentations for many further work steps and the effort involved, data export and backup of the ROIs make perfect sense.

ROIs can usually only be stored and retrieved within a software ecosystem.[23] One way of saving is to export them as 3D surface models. However, this involves the loss of the volume information of the CT data. Another option is to export in tabular form with the most important information of grey values, number of voxels, position, and standard deviation.[24] Here a graphical view of the ROIs is not given by all CT viewers.

Another option is to export as a CT dataset, i.e., as a subset of the original dataset. The advantage is that segmented volumes can be loaded and displayed by all CT viewers, provided they are saved in a compatible format. The disadvantage is that it is a very large amount of data, as all CT slice images of the corresponding sections are stored again. In addition, the ROIs generated in this way cannot necessarily be correctly aligned with each other again, as the orientation and position within the global coordinate system are not saved. A workaround is therefore required if the individual ROIs are to be reassembled afterwards. Various CT programmes offer semi-automated processes here, so this method can be considered the most promising unless one wants to remain within the original programme.

6. Data Link

6.1. Further digitisation measures

Often, digitisation actions are not limited to the use of computed tomography, for example, when a museum collection is to be made accessible. The digital copies produced from this are a dematerialised, digital representation of physical objects enriched with a deeper level of information.[25] Classically, a basic digital capture of meta- and research data is carried out (e.g., in databases).[26] The more value is placed on machine-readable data input in this process, the better it can be linked later (Linked Open Data). 2D object photography is also one of the common digitisation methods and forms the most important recognition feature of objects, both in internal databases and in online collections.[27]

[23] There are some workarounds, for example the open-source application Osirix. However, it is not yet possible to view previously defined segmentations with a CT viewer of one's choice.
[24] See Zwingenberger 2008.
[25] In distinction to Digital Born Objects, which are not created on the basis of a physical model. Regarding this, see Carius 2022, pp. 19–22.
[26] On basic digital recording and standards, see Deutscher Museumsbund 2022.
[27] Ibid, pp. 39–41.

In actual 3D digitisation, computed tomography is at best a marginal phenomenon. The lion's share is realised by 3D surface scanning methods. The most important methods used here are photogrammetry, laser scanners, and structured light methods. In photogrammetry, several images are taken with a camera, which are subsequently converted into a three-dimensional point cloud by software. In the structured light method, a light pattern is projected onto the object and captured by a camera. By distorting the pattern on uneven surfaces, the software can use triangulation to calculate the coordinates of the individual measurement points in three-dimensional space and thus generate the 3D model. Laser scanners work on a similar principle: by projecting, deflecting, and reflecting a laser onto the object. The laser beams are then captured by a detector camera and a 3D point cloud is also created using triangulation methods based on software.[28]

All methods have their advantages and disadvantages. What they have in common is that the result is a three-dimensional point cloud that can generate a polygon mesh by means of triangular meshing. Analogous to mesh creation from CT data, there is also a loss of information here. The finer the mesh, the better the result. On the other hand, the computing times and the data sizes increase. The methods mentioned here all only record the surface and, unlike CT measurements, do not provide information about areas hidden from the camera. Instead, the texture of the objects is also recorded, which is not possible with the X-rays of computed tomography.

6.2. Metadata

Metadata is important additional information about an object that is intended to structure and describe it in more detail.[29] Ideally, the metadata created should also be referenceable. Object-specific metadata includes, for example, type of material, manufacturer, dating, etc., but also classifications according to object type. There are established standards for different subject areas and metadata such as DublinCore, CIDOC CRM or LIDO.[30]

Digitisation processes also generate new, technical metadata. In 2D photography, this includes information about the camera model used, the name of the photographer, resolution, etc. Image rights and licences also fall under metadata if, for example, the digitised images are published online. In 3D digitisation processes, other metadata are generated.

In CT digitisation, a distinction must again be made between the degree of processing of the data in relation to metadata. For the actual CT measurements, there are standards for recording scanner modalities, etc.[31] This results in the minimum requirements of the metadata for reconstructed CT images:

- Unique ID
- File size
- Media type
- File format
- Target grid size
- Voxel spacing
- Unit for the voxel distance
- Colour depth

The unique ID as well as the media type (e.g., CT image stack) may only be of interest during later processing. The file format can be, for example, a raw file or an image stack (DICOM, jpg). The target

[28] A detailed description is provided by Henkensiefken 2022, pp. 7–15.
[29] Dörfler 2020, p. 39.
[30] Bullin 2020, p. 11.
[31] Bear 2018, pp. 33–55.

grid size describes the size of the scanned and reconstructed area in x, y, z direction and depends on the voxel spacing—the resolution so to speak. Here the distance of one voxel centre to the next is measured, also in x, y, z direction. If the file size is reduced by skipping individual voxels, not only does the voxel spacing change, but also the target grid size. The voxel spacing is usually specified in mm or µm. The colour depth can be, for example, 8 or 16 bits.

This information is essential for the correct display of the data in CT processing programmes and must be specified when loading. In certain cases, the information is stored in the metadata of the image stacks themselves (e.g., DICOM). However, raw files do not support this. It therefore makes sense to store metadata separately.

The exchange format for metadata LIDO, which is common in the museum sector, offers only insufficient metadata elements for the special case of CT measurements and therefore cannot capture all relevant metadata.[32] A workaround for storing and presenting the metadata is therefore necessary, for example in the form of separate text files.

In addition to the descriptive metadata, it may be useful to document and store the process-based paradata. This refers to the descriptive data of data creation and data interpretation.[33] On the basis of well-documented paradata, it is ideally possible to later trace the individual work steps from an original state to the present final state. For computed tomographic data, this can mean, for example, the number of measurements including the changed parameters or the parameters of the post-processing (surface determination, measurement analyses, generation of surface models, etc.).

6.3. Rights of use

Terms of use of digitised material must primarily be negotiated and defined within the institution.[34] A distinction must be made between the rights of use of metadata and the actual digitised material. Usually, the descriptive metadata are provided with a maximum free licence, as they have neither artistic value nor research-specific relevance. Moreover, metadata are an essential part of digitised work and should be easily accessible.

In principle, the digitised material itself can be provided with very open or very restrictive licences. Current recommendations are clearly in favour of the most comprehensive and open re-use options possible, which is sometimes also made a condition for funding.[35]

Where applicable, third-party copyrights must be taken into account, which can represent a restriction for open licences. In the case of computed tomography, this concerns the scanning itself as well as any data processing. Here, the desired usage scenarios should be agreed and specified in advance.

6.4. Linking the Data

Linked data has great potential in terms of richness of information and discoverability if data has been processed accordingly. While there are already established recommendations and standards for many digitisation methods, this is not yet the case with computed tomography. Linking it with other object-describing data and metadata is therefore anything but trivial.

At this point, the procedure used in the 3D Cipher project will be described. Due to the given structures and technical peculiarities, it can at best be regarded as an intermediate step towards an ideal solution.

---

[32] Ibidem, p. 55.

[33] Dörfler 2020, p. 44f.

[34] The granting of Creative Commons licences, for example, is well established: https://creativecommons.org/.

[35] This is the case for the DFG, for example. See the DFG Code of Practice, pp. 42–44 and Deutscher Museumsbund 2022, p. 39.

The project aimed to link the following digital data of the cryptological collection of the Deutsches Museum:

- 2D object data
- 3D CT datasets (web viewer)
- 3D CT datasets (web repository)
- Research data (as download)
- Annotations
- 3D surface models
- Animations

The original approach of integrating the 3D CT viewer developed for the project directly into the online database Deutsches Museum Digital (DMD) on the object pages could not be realised during the project period. Since the data available in LIDO-XML format is generated from the internal museum database, the CT data would have to be fed in afterwards. In addition, the web viewer in its current form is not suitable for embedding, partly because of the size of the data sets to be loaded.[36] As a solution, the web viewer is hosted in a separate instance on the DMD's servers.[37] The individual datasets were enriched with annotations that can be downloaded in json format. In addition, there is a link to Zenodo[38] to download the uncompressed CT image stacks, the segmentations and, if available, the surface models.

The web repository MorphoSource is also used, as it is an established platform for storing, displaying, and exchanging CT datasets.[39] Here, too, a link back to the datasets in the DMD and to the 3D Cipher project page was created. This serves as a landing page and links to all the above-mentioned pages and vice versa. Some animations are also stored here, which also cannot be elegantly linked to the other CT data and their derivatives on one page.

The procedure presented can at best be seen as a compromise solution in terms of technical feasibility. The 3D Cipher project and the experience gained from it serve as a first-use case for NFDI4Memory, which will be launched in 2023, with regard to data quality. Ideally, solutions for the problems mentioned here regarding data structure and data linkage will be developed in the near future.[40]

7. Long-term Archiving

Digital long-term archiving is an important topic, especially in museum digitisation, as preservation is one of the core tasks of a museum. For CT data, storage in an open and established standard, such as the already mentioned DICOM or tiff formats, is therefore a good choice. For metadata, formats such as txt, pdf or XML are suitable.[41] When providing data on one's own servers, it should be guaranteed that the data are migrated to new versions and digital infrastructures, and continue to be available. In the 3D Cipher project, the very large CT datasets and the associated metadata were therefore stored

[36] Following the project, however, work will continue on integrating 3D and also 3D-CT models into the Deutsches Museum Digital database.
[37] See https://ctview.dmd.zone/.
[38] Zenodo is CERN's online storage service for scientific research data: https://zenodo.org/.
[39] MorphoSource is mainly used by natural history collections to upload biological specimens: https://www.morphosource.org/.
[40] https://4memory.de/.
[41] https://www.forschungsdaten.info/themen/veroeffentlichen-und-archivieren/formate-erhalten/.

on the Zenodo web platform and are available for download using persistent identifiers.[42] As a renowned institution, CERN can guarantee long-term availability. In addition, the reconstructed image stacks are stored on the servers of the Deutsches Museum.

There is also the question of how to proceed with derivatives that have been created from the CT data. For the important and time-consuming steps of segmentation and the generation of surface models, data sets were also deposited with Zenodo in order to guarantee long-term availability. Other derivatives, such as illustrative videos, were deliberately not archived for the long term because, as snapshots, they do not contain any generically new information. This was a conscious decision within the project and may be handled differently in future projects according to the objectives.

8. Conclusion

The empirical values presented here are intended to list the peculiarities of computed tomographic measures and serve as initial information for planned research projects in cultural institutions. As is so often the case, the most important factors are time and money. Therefore, it should be considered very well in advance which goals are to be achieved with the use of CT measurements and whether, if necessary, other measures are more target-oriented and economical.

Based on a work plan, CT measurements can be planned and carried out accordingly, usually by an external institution. If the data processing is to be carried out by the cultural institution itself, the costs for hardware, software, and personnel must be taken into account. In addition, the time required should be realistically estimated and any training periods should be taken into account. It is advisable to avoid proprietary formats wherever possible and to use open standards, especially with regard to long-term archiving. Linking reconstructed CT data with other digitised data is desirable, but represents a major challenge.

With correspondingly careful planning, computed tomographic measurements are a great enrichment for the digitisation of cultural objects. This is true both for conservation and object-historical questions as well as from a mediation perspective. Therefore, it is to be hoped that the importance of CT analyses will increase and that formal and technical developments will emerge that make the use of the technology more cost-effective, standardised, and thus, more attractive.

9. Bibliography

Bär, Frank et al. Recommendations for the Three-Dimensional Computed Tomography of Musical Instruments and Other Cultural Artefacts. Nuremberg/Fürth (2018). URL: https://www.gnm.de/fileadmin/redakteure/Forschung/pdf/2018-10-11_MUSICES_Recommendations.pdf.

Boas, F. Edward; Fleischmann, Dominik: CT artifacts: Causes and reduction techniques. In: Imaging in Medicine (2012), Volume 4, Issue 2, pp. 229–240.

Borrelli, Pablo et al.: Artificial intelligence-aided CT segmentation for body composition analysis: a validation study. In: European Radiology Experimental (2021), Volume 5, 11, pp. 1–6. URL: https://doi.org/10.1186/s41747-021-00210-8.

Bullin, Martin; Henrich, Andreas: Die inhaltsbasierte Bildsuche und Bilderschließung: Ansätze und Problemfelder (*Content-Based Image Search and Image Exploration: Approaches and Problem Areas).* In: Hastik, Canan; Hegel, Philipp (eds.): Bilddaten in den digitalen Geisteswissenschaften. Wiesbaden (2020), pp. 11–33. https://www.doi.org/10.13173/9783447114608.

[42] For the project's datasets on Zenodo, see https://doi.org/10.5281/zenodo.8113300.

Carius, Hendrikije; Fackler, Guido: Ausstellungen digital kuratieren. Formen und Diskurse, Herausforderungen und Chancen (*Curating Digital Exhibitions. Forms and Discourses, Challenges and Opportunities).* In: Carius, Hendrikije; Fackler, Guido (eds.): Exponat—Raum—Interaktion. Perspektiven für das Kuratieren digitaler Ausstellungen. Göttingen (2022), pp. 15–38. https://doi.org/10.14220/9783737012584.

Chen, Xiuyuan: AI-based chest CT semantic segmentation algorithm enables semi-automated lung cancer surgery planning by recognizing anatomical variants of pulmonary vessels. In: Frontiers in Oncology (2022). https://doi.org/10.3389/fonc.2022.1021084.

Deutscher Museumsbund e.V. (ed.): Handreichung Digitale Grunderfassung: 10 Grundsätze (*Basic Digital Recording Handout*). Berlin (2022). https://www.museumsbund.de/praktische-handreichung-unterstuetzt-museen-bei-der-digitalisierung-ihrer-sammlungen/.

DFG Code of Practice. Digitalisierung (*Digitisation*). https://www.dfg.de/download/pdf/foerderung/programme/lis/12_151_v1216_de.pdf.

Dörfler, Franz Georg: Ein Blick in die Punktewolke. Herausforderungen und Möglichkeiten im Umgang mit Daten von photogrammetrisch erzeugten 3D Modellen (*A look into the point cloud. Challenges and opportunities in dealing with data from photogrammetrically generated 3D models*). Graz 2020.

Du Plessis, Anton et al: Laboratory X-ray micro-computed tomography: a user guideline for biological samples. In: Gigascience (2017), Volume 6(6), pp. 1–11. https://doi.org/10.1093%2Fgigascience%2Fgix027.

Fritsch, Dieter et al.: Making Historical Gyroscopes Alive-2D and 3D Preservations by Sensor Fusion and Open Data Access. In: Sensors (2021), Volume 21(3), 957, pp. 1–31. https://doi.org/10.3390/s21030957.

Hanke, Randolf: Computertomographie in der Materialprüfung. Stand der Technik und aktuelle Entwicklungen (*Computed tomography in materials setting. State of the art and current developments)*. German Society of NDT. Erfurt (2010). https://www.ndt.net/?id=10133.

Henkensiefken, Claus: 3D-Digitalisierung im Museum. Erläuterungen zu Scanverfahren und deren Anwendung im musealen Bereich. Erfahrungsbericht. (*3D-Digitalisation im Museum. Explanations of scanning processes and their application in the museum sector*. *Field report*). (2022). https://www.museum4punkt0.de/ergebnis/3d-digitalisierung-im-museum/.

Kroll, Julia: Aufgabenangepasste, kontrollierte Oberflächenextraktion aus 3D-Computertomographiedaten (*Task-adapted, controlled surface extraction from 3D computed tomography data*). Stuttgart (2014). http://dx.doi.org/10.18419/opus-6843.

Kuhlmann et al.: From µCT data to CFD: an open-source workflow for engineering applications. In: Engineering Applications of Computational Fluid Mechanics (2022) 16:1, pp. 1706–1723. https://doi.org/10.1080/19942060.2022.2109758.

Lang, Thomas; Sauer, Tomas: AI-Supported Segmentation of Industrial CT Data. 11th Conference on Industrial Computed Tomography, Wels (2022), pp. 1–7. https://doi.org/10.58286/26602.

Lifton, Joseph J.; Malcolm, Andrew A.; McBride, John W.: The Application of Beam Hardening Correction for Industrial X-ray Computed Tomography. 5th International Symposium on NDT in Aerospace, 13–15 November 2013, Singapore. e-Journal of Nondestructive Testing Vol. 18(12). https://www.ndt.net/article/aero2013/content/papers/14_Lifton_Rev2.pdf.

Luccichenti et al.: 3D reconstruction techniques made easy: Know-how and pictures. In: European Radiology (2005), Volume 15, pp. 2147–2156.

Varotsis, Alkaios Bournias: Meshing in FEA, CFD and manufacturing. Webblog (2021). https://www.ntop.com/resources/blog/meshing-in-fea-cfd-manufacturing/.

Yang, Xiuyuan; Sun, Wenjuan; Giusca, Claudia L.: An automated surface determination approach for computed tomography. In: NDT and E International (2022), Volume 131, pp. 1–17.

Zhan et al: Computed Tomography Data Colouring Based on Photogrammetric Images. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information XLIII-B2-2020 (2020), pp. 361–368. https://doi.org/10.5194/isprs-archives-XLIII-B2-2020-361-2020.

Zwingenberger, Allison: Exporting regions of interest in Osirix. In: Veterinary Radiology (2008). https://www.veterinaryradiology.net/486/exporting-regions-of-interest-in-osirix/.