UNCERTAINTY VISUALIZATION OF COMPUTED TOMOGRAPHY DATASETS FROM COMPLEX COMPONENTS USING STATISTICAL ANALYSIS

Martin Allerstorfer [a], Christoph HEINZL [a], Johann KASTNER[a], Eduard Gröller[b]

[a]Upper Austrian University of Applied Sciences, Wels Campus, Austria
[b]Vienna University of Technology, Institute of Computer Graphics and Algorithms, Austria

Abstract. This paper describes an approach for statistical analysis of industrial 3D X-ray computed tomography (3DCT) data and different means of visualization for the statistical information. The main purpose of a statistical analysis is, that the information concerning spatial uncertainty of industrial 3DCT data is often neglected for typical data evaluations. However, the data inherent uncertainty information mainly contributes to the result of each data evaluation. To address this issue, a data processing pipeline is proposed, which consists of three major steps: anisotropic diffusion for prefiltering, K-means clustering for an automatic definition of the specified classes in the dataset and finally the calculation of probability volumes applying the Bayes Theorem. The first part is a prefiltering step, which aims to reduce noise without blurring edges and so to increase the quality of classification. The second step is the statistical analysis: First, a K-means algorithm is used to compute probability information based on the provided data. The results are used to define probability density functions, which describe the probability of a certain data value to belong to a specific class or material. Secondly the Bayes Theorem is applied to calculate a multichannel probability volume consisting of posterior probabilities for each channel (material) and each spatial position. These probabilities are finally used for uncertainty visualization. Concerning visualization we present two methods to visualize the calculated probability information: direct volume rendering (DVR) of the probability applying a transfer function for color coding and opacity mapping and secondly the extraction of isolines which interpolate a certain probability value. Both visualization techniques are tested on test specimens as well as real world industrial components.

1 Introduction and Motivation

It is well known that uncertainties are often the result of error-prone measurements or incomplete sampling and appear in applications processing a large amount of data. One such example is industrial 3D X-ray computed tomography (3DCT) which is used for non destructive testing and metrology applications. 3DCT is based on the following principle: X-rays are shot from the X-ray source through the specimen to the detector, which measures the incident radiation. The specimen performs a full 360 degree rotation. For each orientation of the specimen the X-ray attenuation is recorded in penetration images which are mathematically reconstructed into a 3D grid of grey values. [7] One crucial disadvantage of this technique is the occurrence of artefacts. Some of the most common artefact types are noise-induced streaks, aliasing, beam hardening, partial volume and scattered radiation effects. Figure 1b shows an artefact prone measurement of an industrial component where grey values are modified so that the application of a transfer function introduces holes and additional material in the 3D rendering. If we compare the rendered dataset with a photograph of the original work piece (Figure 1a) these structures and holes cannot be detected in reality.

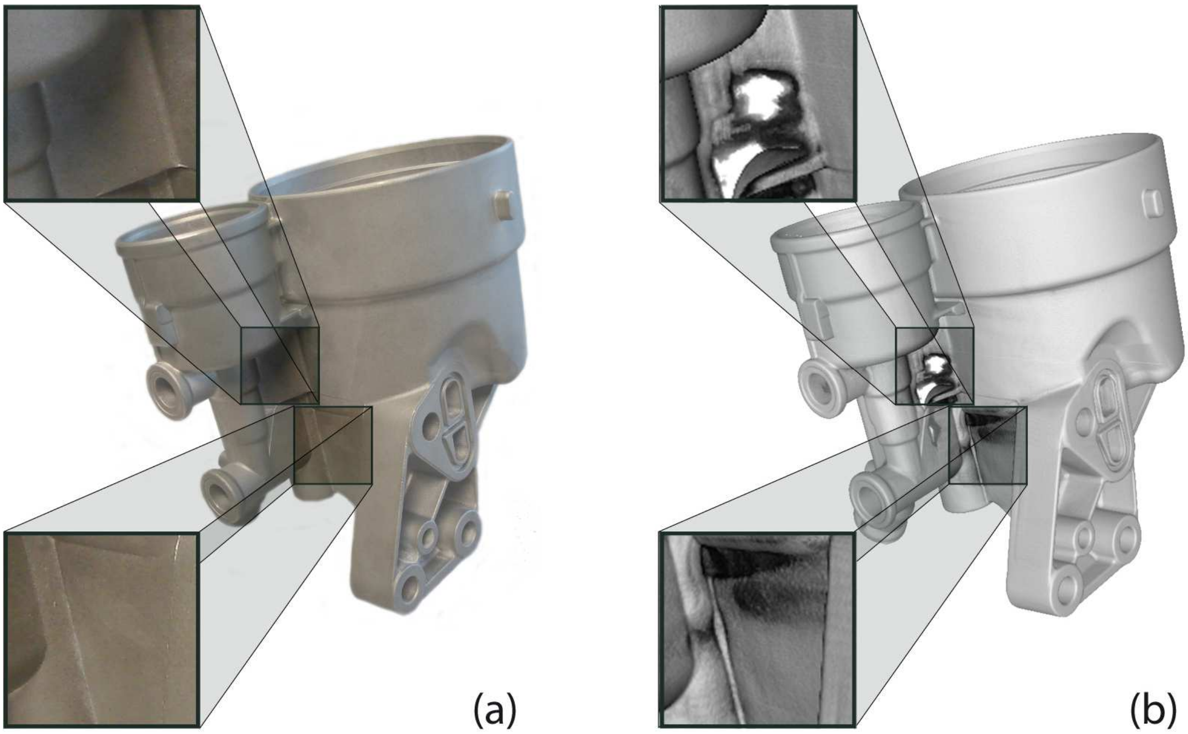

Figure 1: Industrial oil filter housing specimen with complex geometry. (a) Photograph of the specimen. No holes exist and the surface is smooth. (b) 3D rendering of the CT dataset showing rendering artefacts. Holes and additional material appear in the rendering.

The characteristics and strength of such artefacts strongly depend on the material, geometry, position and orientation of the specimen in the ray as well as measurement parameters. In non destructive testing artefacts can have a huge impact on the quality and precision of an evaluation, e.g. they affect the precision of a dimensional measurement because it is hard to find appropriate global thresholds without additional information. Usually only one grey value is used as global threshold for the distinction of the interface between material and air, e.g. to extract an isosurface. In most cases however a single threshold is not enough, especially when artefacts interfere with transitions between material and air. Therefore it is difficult to distinguish between correct and incorrect values. The introduction of uncertainty information makes it possible to estimate the error of measurement at a certain spatial position in the dataset and thus use it to increase the interface quality between different materials.

Besides dimensional measurement, also simple visualization is prone to errors without considering uncertainty information. Visualization of these errors helps to identify possible artefacts and prevents the user from deducing incorrect results.

2 Related Work

Due to the fact, that all recording devices tend to be susceptible to noise, several techniques have been developed to improve image quality. Perona and Malik [3] introduced the method of Anisotropic Diffusion or also known as Perona-Malik diffusion. This method allows the reduction of noise like Gauss filtering but preserves edges and fine details. In their work they use it for edge detection. Anisotropic Diffusion is applied in the presented work in order to improve the quality of classification.

A large variety of uncertainty visualization techniques have already been proposed so far like glyphs, geometry and attribute modification, addition of geometry [1] as well as point based probability surfaces [2] and visualization based on isosurfaces [6].

Kniss et al. [4] use the Bayes Theorem to let the user explore uncertainty, risk, and probabilistic decision of surface boundaries interactively. Furthermore they present a method of how to use uncertainty and risk information for visualization and a data dimensionality reduction scheme to reduce memory usage. Additional information and introduction to the topic of statistical classification and segmentation can be found in a comprehensive book by Duda et al. [5]. It also includes a description about the K-means clustering algorithm which is used in our work for classification and extraction of Gaussian probability functions. Lundström et al. [8] calculate local histograms to estimate probabilities. These histograms are generated by Gaussian curve fitting for Partial Range Histograms extracted from the main histogram. This method has been developed for the medical area and therefore it is only partly applicable for industrial components. They also present an optimization method called Blockwise Spatial Coherence Refinement which uses spatial coherence of material to adapt their ranges interactively. Pang et al. [1] give a considerable overview about uncertainty in general, visualization methods and combination of methods respectively. Additionally they categorize the methods by application (Radiosity, Animation, Interpolation and Flow) and approach (glyphs, adding geometry, modifying geometry, animation, etc.) because not all methods are applicable in each area. Grigoryan and Rheingans [2] present a surface based method to show surface uncertainty. Their approach is to render a displacement of surface points instead of the surface itself. The strength of displacement depends on the uncertainty at a certain location and takes place along the surface normal. Malik et al. [11] present a method which allows direct comparison of surface models and volumetric datasets. Uncertainties are visualized through color coding and Boxplots. In our work we apply isolines and present the methods of pseudo colors on three work pieces.

3 Pipeline

The major part of our approach is the calculation of probability information which is the basis for uncertainty visualization (Figure 2). The pipeline is implemented as follows: First, we extract uncertainty information by using anisotropic diffusion and the K-means algorithm. Then Bayes Theorem is applied to cluster material and calculate posterior probability. Based on this information we present two different techniques of uncertainty visualization.

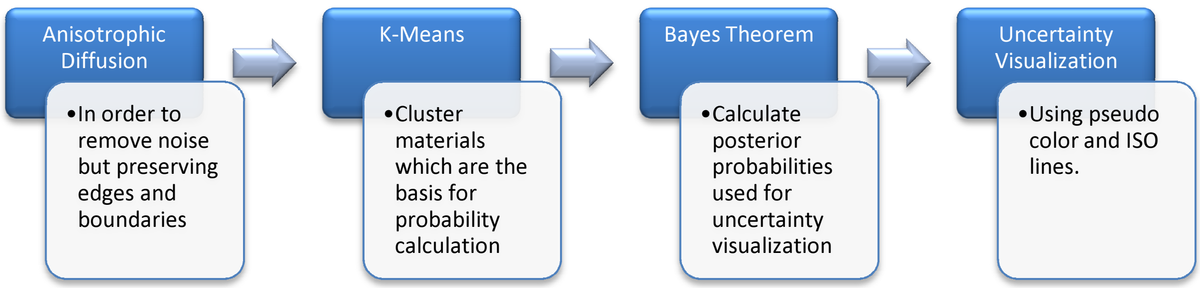

Figure 2: Simple representation of the pipeline for statistical analysis of 3DCT data consisting of anisotropic diffusion, the K-means algorithm, the application of the Bayes Theorem and the final step of visualization.

3.1 Anisotropic diffusion

Noise reduction is a common means for image quality improvement mostly achieved by convolving the original image with application specific kernels, such as a Gaussian kernel. The major drawback of kernel based filtering is however that it blurs the whole image without consideration of edges or fine details. What we require is a method that smoothes the interior of each region while it preserves material boundaries. Anisotropic diffusion by Perona and Malik [3] fulfills this criterion. Their approach is to filter the data by following the partial differential anisotropic diffusion equation:

$$I_t = div(c(x,y,t)\nabla I) = c(x,y,t)\Delta I + \nabla c * \nabla I \tag{1}$$
$$c(x,y,t) = g(||\nabla I(x,y,t)||) \tag{2}$$

div is the divergence operator, *c* represents the conduction coefficient of smoothing, ∇ and Δ are the gradient and Laplacian operators respectively. The coordinates *x* and *y* define the image location and I_t is the changing image over time *t*. In order to achieve the goal of preserving edges and boundaries the function *c(x,y,t)* has to be nonnegative and monotonically decreasing with *g(0) = 1* whereas high values are returned for interior regions and low values for boundaries. Therefore using gradient information for the parameter of function *g* fulfils our requirements.

3.2 K-means

A K-means algorithm [5] is a very simple and fast method used for clustering of similar objects to a known number of *k* clusters. For 3DCT scans of single material work pieces K-means proved to be an efficient way to distinguish materials. The basic idea is an iterative procedure as follows:

1.) Define *k* random cluster mean values for the first iteration step.
2.) Assign each grey value to the cluster with the closest distance to the mean value.
3.) Recalculate the mean value of each cluster.
4.) Continue with step 2 until the cluster mean values do not move anymore or a predefined maximum iteration depth is reached.

The Euclidian distance is used as a distance function and the Euclidean sum of squares is used as a mean value. Based on the assumption that homogenous materials in 3DCT scans are continuously distributed around a mean value, we use Gaussian probability functions as classifiers. The resulting classes of the K-means algorithm allow to calculate necessary mean μ and standard deviation σ values. In order to gain smooth transitions between low and high probabilities we increased σ to 3σ for the presented work pieces which was found empirically.

3.3 Bayes Theorem

After applying the K-means algorithm on the dataset we set up a probability density function for each class or material. Using this probability density function, the probabilistic likelihood $P(\vec{x}|\omega_i)$ as it is called in [4] is calculated. We however want to evaluate the posterior distribution $P(\omega_i|\vec{x})$ which normalizes the likelihood and weights against prior probability. Therefore we make use of the Bayes Theorem:

$$P(\omega_i|\vec{x}) = \frac{P(\vec{x}|\omega_i)P(\omega_i)}{\sum_{i=1}^{C} P(\vec{x}|\omega_i)P(\omega_i)} \tag{3}$$

C denotes the number of classes and $P(\omega_i)$ defines the apriori probability. As the priors should not influence the posterior probabilities the prior probability is $1/C$ for all classes. The divisor of formula 3 serves as a scaling factor to guarantee that all posteriors of each class sum to one. The feature vector $\vec{x}$ for CT datasets could consist of the spatial grey values, gradient magnitudes and similar properties. Due to performance reasons and artefact influences we decided to use only a one dimensional feature vector consisting of the spatial grey values.

The result of statistical analysis is a multi channel dataset of channel size *k*. Each channel represents the posterior probability for each material. Tested work pieces consisting of air and material are visualized by always selecting the material component.

4 Results and Discussion

Testing and evaluation of the described pipeline was done by scanning three work pieces. These scans were performed in a HWM Rayscan 250E 3DCT device with a 225 keV microfocus and a 450kV macrofocus tube. We used the libraries ITK [9] and VTK [10] to implement the pipeline in C++.

Work piece one (Figure 1a) is an aluminum oil filter housing with complex geometry. Especially central areas of the specimen are strongly affected by artefacts and smooth surfaces become rough (lower left corner of Figure 1b). The work piece was scanned using: 810 projection images, 210 kV, 830 µA, 500 ms integration time and 1 mm Cu prefiltering.

In Figure 3 we see the result after applying the pipeline on the oil filter housing. Strong artefacts, which are manifested as holes and thickening (Figure 1b), are also shown in the uncertainty visualization as red and orange areas. Interpreting this information it becomes clear at first sight, that the rendered holes and additional material are artefacts. Less obvious artefacts are shown in the lower right corner of Figure 3 which are caused by the drill holes.

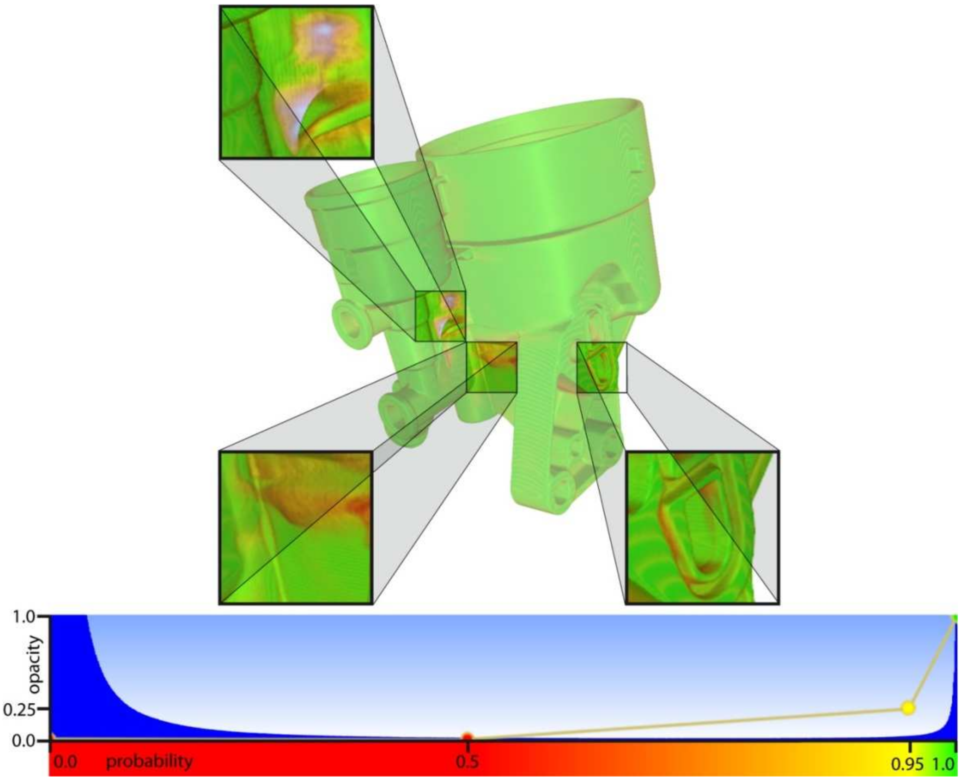

Figure 3: Uncertainty visualisation. The zoom-ins show areas severely affected by artefacts. The transfer function and probability distribution is shown below the rendered image.

Work piece two (Figure 4a) is an aluminum cylinder with a drill hole along the longitudinal axis. On the top there is a concentric ring with two vertical and one horizontal drill hole (Figure 4b), in the middle a bigger horizontal hole and on the bottom four vertical ones. The work piece was scanned using: 720 projection images, 210 kV, 315 µA, 500 ms integration time and 1 mm Al prefiltering.

This work piece was scanned in order to present another method for uncertainty visualization: isolines. Each isoline represents a certain probability value. This kind of visualization is a good way to identify probabilities of inner structures. In Figure 4 we used 10 isolines between the probabilities 0.3 (yellow) and 0.9 (blue).

We get the highest uncertainties on the top and especially in the bottom area of the cylinder. The first conclusion we can infer by analyzing Figure 4 is that the drill holes in the concentric ring are less artefact prone than the longitudinal one as there are no light green and yellow isolines. In the area of the horizontal drill hole blue isolines indicate high probabilities near the outside of work piece two and decrease with the distance to the center. Furthermore variances in the distance between isolines are an indicator for artefacts as shown in Figure 4c. On the lower right side of this depiction isolines are close together whereas large distances occur on the upper left side. This means that uncertainties are not equally distributed along the drill hole.

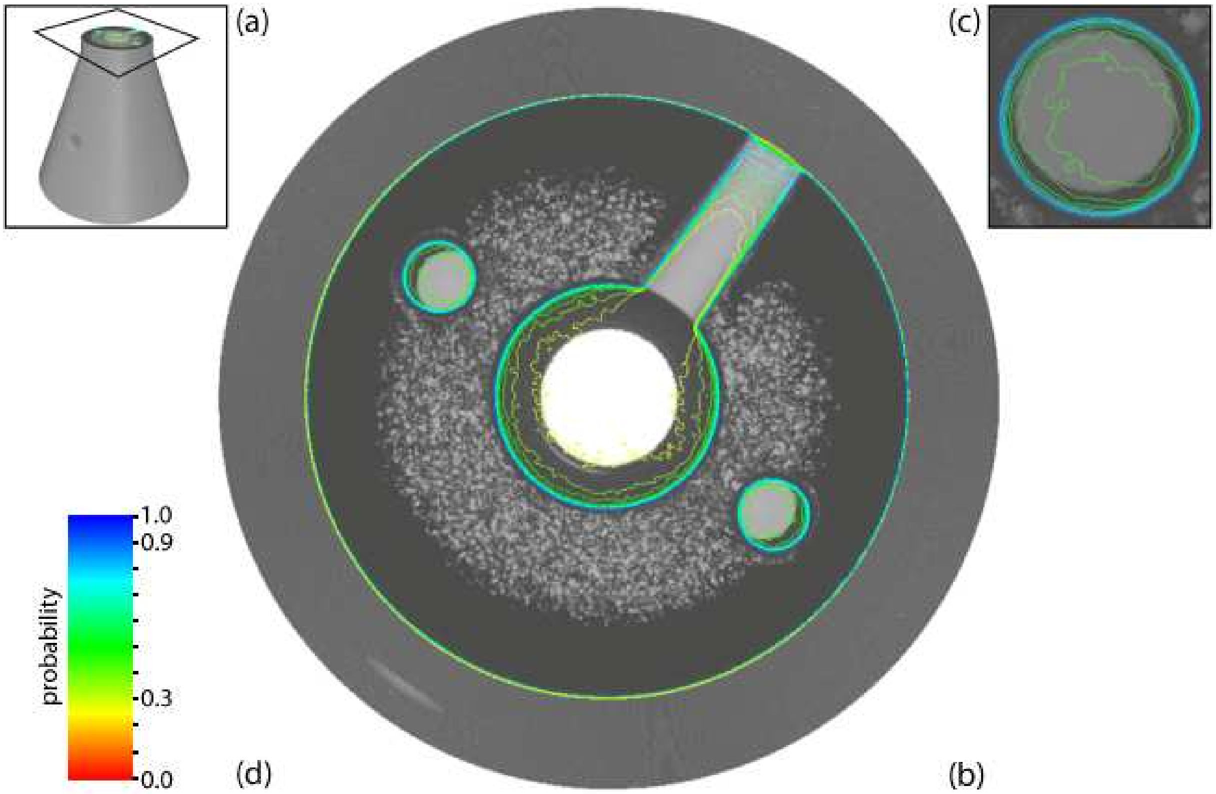

Figure 4: Uncertainty visualisation of work piece two using isolines. Blue lines represent high, green lines middle and yellow lines low probabilities. In this example 10 isolines are drawn in the range from 0.3 to 0.9. The slice depicted in (a) is shown in (b). The enlargement of the upper left hole (c) is rendered with reduced lighting for better visualization. (d) The used color coding.

Work piece three is an aluminum engine piston of a motorcycle (Figure 5). Due to high penetration length and scattered radiation the main artefacts appear on the cylinder cap and in the piston rings. The piston is scanned using: 630 projection images, 200 kV, 500 µA, 500 ms integration time and 0.5 mm Cu prefiltering.

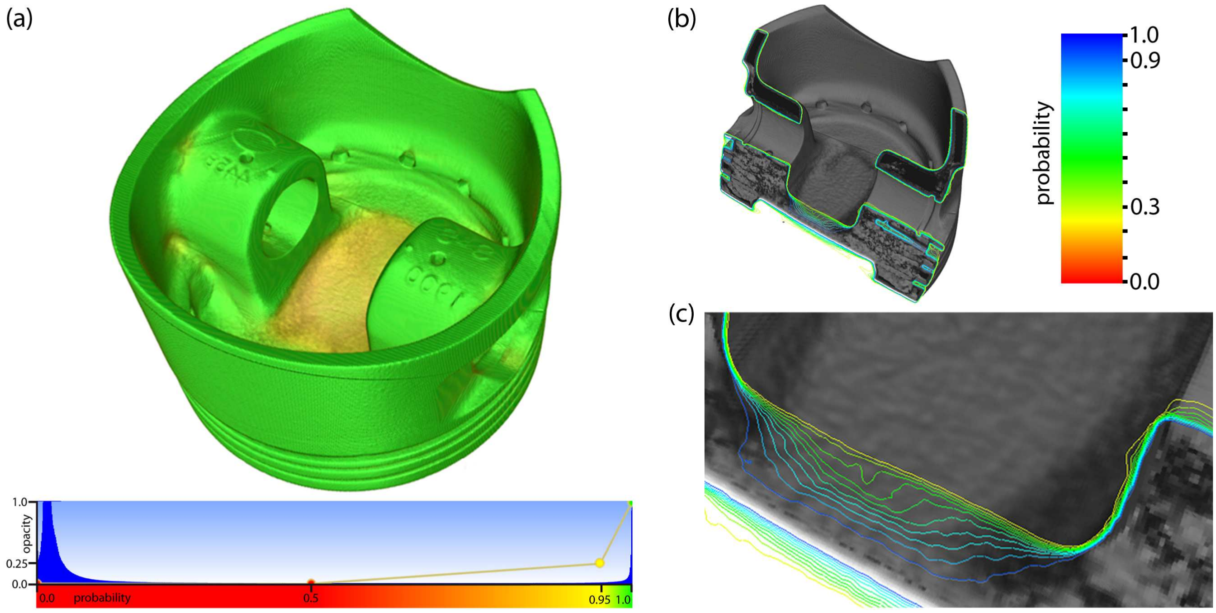

Figure 5: Pseudo color (a) and isoline visualization (b) of a motorcycle piston. (c) Detailed view of the most interesting area. The used transfer function and the used colors for isolines are shown below and next to the images. For the isoline visualization 10 lines between the probabilities 0.3 and 0.9 are drawn.

Pseudo color visualization shows the artefacts as expected (yellow areas of Figure 5a). Probabilities of inner structures are depicted by 10 isolines between the probabilities 0.3 and 0.9 (Figure 5b and 5c). In most areas the isolines are close together except in the area of the detailed view (Figure 5c). There the space between the same isolines increases. This indicates that the area of high uncertainty is larger and therefore artefacts can be identified in this region.

5 Summary and Conclusion

The presented calculation of uncertainty information as well as the two visualization techniques are robust and lead to expected results. The pipeline is a combination of CT scanning, Anisotropic Diffusion, the K-means algorithm and Bayesian decision theory. Tests on industrial work pieces as well as homogeneous test parts showed that probability data allows to identify artefact affected areas precisely also without knowledge of the real specimen. Pseudo colors are mainly applied on objects for uncertainty visualization of surfaces. Isolines are more convenient to view uncertainty information about inner structures.

In the future a more sophisticated probability calculation by using multi dimensional feature vectors including gradient magnitude is possible. This would lead to an even more precise distribution of materials and therefore an improved accuracy in uncertainty visualization.

Acknowledgements

This work was supported by the bridge project "SmartCT – Genaue Geometriebestimmung und Interfacecharakterisierung von Mulitmaterialbauteilen mittels Kelgestrahl-CT" of the Austrian Research Promotion Agency (FFG), Grant No. 818108 as well as the K-project for Non Destructive Testing, Grant No. 820492.

References

[1] Alex T. Pang, Craig M. Wittenbring, and Suresh K. Lodha, Approaches to Uncertainty Visualization, The Visual Computer, Volume 13, Issue 8, Pages: 370 – 390

[2] G. Grigoryan, and P. Rheingans, Point-based probabilitstic surfaces to show surface uncertainty, Visualization and Computer Graphics, IEEE Transactions on Volume 10, Issue 5, Sept.-Oct. 2004 Pages: 564 – 573

[3] P. Perona, and J. Malik, Scale-Space and Edge Detection Using Anisotropic Diffusion, Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume 12, Issue 7, Jul 1990 Pages: 629 – 639

[4] J. M. Kniss, R. VanUitert, A. Stephens, G.-S. Li, T. Tasdizen, C. Hansen, Statistically Quantitative Volume Visualization, Visualization Conference 2005, Pages: 287 – 294

[5] R. O. Duda, P. E. Hart, and D. G. Stork. Pattern Classification (2nd Edition). Wiley-Interscience, 2000.

[6] P. Rhodes, R. Laramee, R. Bergeron, and T. Sparr, Uncertainty Visualization Methods in Isosurfcae Rendering, Eurographics 2003, Pages: 83 – 88.

[7] J. Kastner, E. Schlotthauer, P. Burgholzer, and D. Stifter. Comparison of x-ray computed tomography and optical coherence tomography for characterisation of glass-fibre polymer matrix composites. In Proceedings of World Conference on Non Destructive Testing, pages 71 – 79, 2004.

[8] C. Lundström, P. Ljung, A. Ynnerman, Local Histograms for Design of Transfer Functions in Direct Volume Rendering, Visualization and Computer Graphics, IEEE Transaction on Volume 12, Issue 5, Nov. – Dec. 2006, Pages: 1570 – 1579

[9] L. Ibáñez, W. Schroeder, L. Ng, J. Cates, The ITK Software Guide, Kitware, Inc., 2005

[10] W. J. Schroeder, K. M. Martin, L. S. Avila, C. C. Law, The VTK User’s Guide, Kitware, Inc., 2006

[11] M.M.Malik, C.Heinzl, E.Gröller, Computation and Visualization of Fabrication Artifacts, Journal of WSCG, Vol. 17, No. 1, 2009, Pages: 17 – 24