Project title: Implications of Medical Low Dose Radiation Exposure

Grant Agreement Number: 755523

Call identifier: NFRP-2016-2017

Topic: NFRP-9

Deliverable D2.12

Correlation analysis between objective and subjective image quality in chest CT patient parameter and indication dependent

Lead partner: OvGU

Author(s): Christoph Hoeschen

Work Package: 2

Estimated delivery: 31 May 2020

Actual delivery: 16 July 2020 (v1), 27 August 2020 (v2)

Type: Report

Dissemination level: Public

This project has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 755523.

Table of contents

List of figures

Abbreviations

MTF – modulation transfer function

NPS – noise power spectrum

1. Introduction

The overall objective of task 2.1. of the MEDIRAD project is to develop an optimisation strategy for CT imaging of the chest region. This optimisation needs to take into account the exposure of the patient as well as the achievable image quality. As it is very difficult to predict the achievable diagnostic performance based on the image quality depending on the patient characteristics, the image acquisition parameters and the reconstruction set-up, it is of uttermost importance for the optimisation process to evaluate the image quality of such clinical investigations. The first question which needs to be addressed in this context is, how image quality should be defined for this purpose. Regarding the optimisation that is intended as an outcome of MEDIRAD, the exposure for the patient should be reduced as much as possible while the diagnostic accuracy has to be maintained. The diagnostic accuracy depends on the visibility of pathological structures. However, such pathological structures are very different / depend on clinical indications and are not existing in all patients. Therefore, it is necessary to find structures of normal anatomy which can represent the pathological structures in a way that if they are clearly visible the pathologies would most probably be visible as well. The next task that comes up at this point is the following aspect. The evaluation if the defined structures are clearly visible and the overall image quality regarding resolution, contrast, noise and artifacts is suitable needs a subjective evaluation by a large group of readers. Each of which would have to evaluate many image examples for each technique to get reliable results for the effect of each optimisation approach.

The other approach that is available is to determine image quality parameters using objective physics-based image quality descriptors. Such image quality descriptors are typically derived by evaluating images acquired from physical test phantoms. These evaluations can be pretty easily performed, once the methodology is established for a given imaging task. When evaluating the literature, however, in many cases there is not or hardly a good correlation between the physical image quality determined on images of physical test phantoms and the subjective image quality performance-based evaluation. Especially, with respect to the iterative reconstruction methodologies getting more and more prominent in CT imaging and their interaction with the used acquisition parameters it is very difficult to use images based on physical test phantoms. The reason for this is, that for objects that are imaged which do have only a limited amount of structures the iterative reconstruction can really reduce noise without deteriorating fine structures. That is hardly possible in real clinical imaging situations on real patients due to the large amount of structures inside the human body. This missing correlation would limit the applicability of an optimisation based on phantom based physical image quality descriptions. Therefore, it had been the idea of task 2.1.1 in the MEDIRAD project to develop a methodology which allows the correlation of objective physics-based image quality parameters and subjective image quality evaluations. The main idea was to

- Identify relevant clinical indications, especially those for chest CT imaging not using contrast media
- Identify the relevant structures within such CT images which allow a suitable subjective image quality evaluation
- Develop physics-based image quality descriptors and corresponding evaluation methods within the clinical images, preferably even on the same structures
- Compare the subjective and the objective image quality parameters as determined on clinical chest CT images

2. Correlation analysis

The work of subtask 2.1.1 which is the basis of this deliverable has been performed according to the DoA.

2.1 Decision on clinical indications

As it can be seen in the report of Milestone M26 which had been provided as a report in time the group of experts necessary for the proposed procedure had been established within the first year. This group decided on the clinical indications that should be used for the optimisation approach. This decision was achieved using a Delphi-process to come to a common proposal within the group of experts.

2.2 Decision on subjective image quality criteria for chest CT imaging

As the next step the same group of experts decided in a face-to-face meeting, again based on the procedure defined as the Delphi process, on the dedicated structures to evaluate the image quality subjectively as well as about general image quality descriptors to be evaluated like noise impression, overall image quality and artifacts. As described in the book of abstracts for the European radiation protection week 2018 (ERPW18, (1)), this was the proposed structures and evaluation criteria:

"

…they defined 5 anatomical structures to be assessed, using a 5 level Likert scale for defining is these structures could be sharply visualized, as follows:

1- confident that the criterion is not fulfilled,
2- somehow confident that the criterion is fulfilled,
3- I do not know if the criterion is fulfilled,
4- somehow confident that the criterion is not fulfilled,
5- confident that the criterion is not fulfilled.

The 5 selected structures were chosen because they can be precisely defined and represent different degree of anatomical precision and underlying spatial resolution."

It was also stated, that all assessments have to be performed using the lung window on axial transverse images reconstructed with an algorithm promoting fine details. All structures except the major fissure should be analyzed on the right lung, since the major fissure is incomplete on the right lung in many normal subjects.

The structures are: "

1. Major Fissure of the left lung
2. B1 (apical bronchus of right upper lobe): 3 divisions on the axial plane
3. B6: 3 divisions on the axial plane
4. Right inferior pulmonary vein: 3 divisions on the axial plane
5. A6: 4 divisions on the axial plane

"

And further more for the general image quality aspects it was mentioned, that

In addition to the evaluation of the anatomical details stated above, there will be an overall assessment of as well the noise impression and the appearance of artifacts, using a 3-point scale, which was defined as follows: "

1- absent,
2- present but not disturbing,
3- present and disturbing.

"

Finally the radiologists should provide an overall assessment regarding the diagnostic acceptability, again making use of a 3-point scale: "

1- fully acceptable,
2- probably acceptable,
3- unacceptable

".

2.3 Preparation of subjective image quality evaluation study

Once these definitions had been taken, ethical approval has been gained at all participating centers for collecting the image data for the subjective image quality evaluation study. Afterwards, more than 400 clinical cases had been collected representing the three identified clinical indications. The data had been anonymized and stored on a secured server. Now, they are ready to be used in the subjective image evaluation trial (final trial will be on 300 cases) and on corresponding evaluation of the physical image descriptors on the same patient images.

An existing software tool developed in earlier EC-funded projects for subjective image quality analysis (ViewDEX) had been adapted to the necessities of the subjective image quality evaluation in MEDIRAD. The original version of ViewDEX was used for evaluation of projectional radiography, while here CT-image data sets have to be evaluated, so additional functionality for showing all images of a study and corresponding documentation had to be implemented. This has been achieved and the new version has been presented with some examples to a large number of radiologists from the group of experts. They all agreed that the tool can be used in a very convenient way for the evaluation of the image data sets.

2.4 Development of a methodology to determine physics-based image quality parameters in clinical images

During the preparation of the project it was foreseen to use methodologies for evaluating resolution related parameters like the modulation transfer function (MTF) from patient images based on methodologies as they were to our knowledge at that time developed by a group at Duke university. However, during the project it turned out that this was not suitable for our purpose since the group at Duke used the borders of the patient (edge between skin and air) for the evaluation of the derived spatial resolution. Due to aspects like higher noise, less contrast, more scatter radiation contribution and more overlaying structures inside the human body where all relevant structures are located it needs to be assumed that such determination of results is not really patient related and would not result in a good correlation to the subjective image quality evaluation. In addition, we found out, that a large number of parameters such as patient anatomy, potential implants, patient motion, variations due to the pathology or a CT system with a non-suitable setup may affect the image resolution inside the body of the patient, which would not be determined by the method proposed by Duke university (3). But in general it can be stated, that the MTF can be computed by analyzing boundaries between two clearly separated objects. (4). Thus, in MEDIRAD, an algorithm for MTF determination on the clinical structures of interest was developed as described in the SPIE presentation (2):

Actually, the MTF was computed on boundaries for different objects in each thoracic CT image. As a first step, a ROI with 16 by 16 pixels containing a boundary typically between bone and background was chosen manually. An algorithm using active contours was the applied to the ROI for detecting the boundary between the desired segments. (5) The image gradient was computed along the detected edge in order to obtain an edge spread function (ESF) perpendicular to the object boundary. Always taking into account the direction of the image gradient, many separate ESFs were derived. By determine the first derivative of the ESFs, corresponding Line spread functions (LSF) were computed. The MTF was then calculated by Fourier transformation of the averaged LSF. The general procedure for the MTF determination is summarized in Fig. 1. To develop a first approach for a quantitative comparison to the subjective evaluation of different image qualities, we propose to compute the integral under the MTF curve. As it can be easily understood, a smaller area under the curve corresponds to a poorer detail representation in the image.

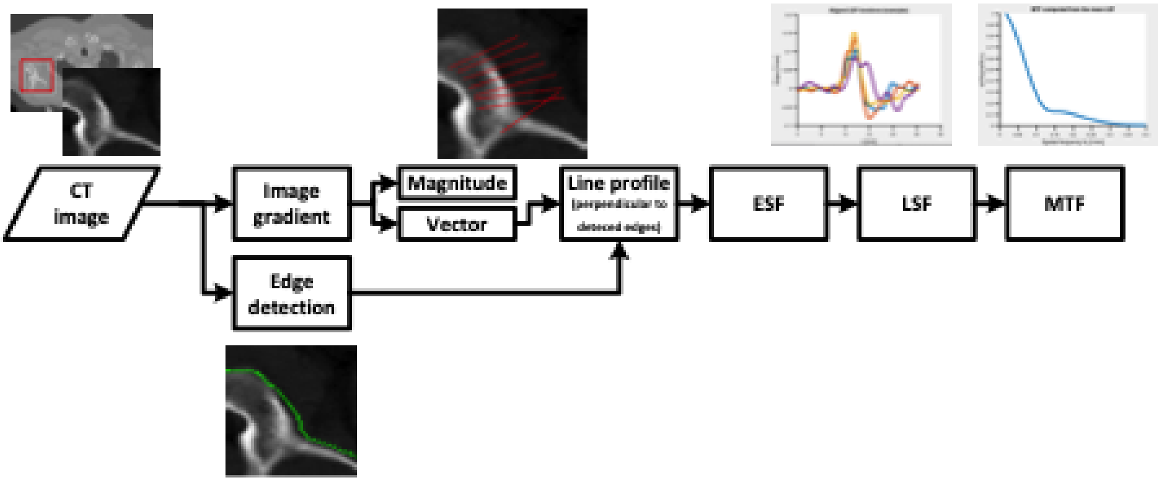


Figure 1: The general workflow of the MTF computation for a CT image as developed. (from reference 2)

As it is well known, in fourier-based approaches, the noise is typically characterized by the noise power spectrum (NPS). This also holds for CT imaging. Assuming the assumption that the signal is spatially invariant and the detector would be of unlimited size (which is not really correct, but can be assumed for certain signals to be valid),the NPS provides information about the noise amplitude, but also the frequency distribution of the noise in the image. The noise in CT images is among other reasons caused by the quantum noise, by the electronic noise of the detection system and also depending on the used reconstruction technique. (2) Other image processing, but also image acquisition schemes (like slice thickness) will influence the noise as well. All this has to be taken into account. This is not reflected in the approach presented by Duke university since it tries to get rid of the structures by evaluating many slices in connection which will result in different measurement results depending on slice thickness as we found when evaluating their proposed method. So also for the noise evaluation, MEDIRAD had to develop its own methodology.

Calculating a NPS that is meaningful in terms of describing image noise requires a homogenous background i. e. without any (anatomical) structures. This requirement is really challenging for determine NPS in real clinical CT images. As a first approach, the NPS was therefore calculated from the phantom images described in Section 2.5 (2). As for the evaluation of the MTF, we did choose a 16 by 16 pixel region-of-interest (ROI), which was chosen in a way that it was homogenous meaning without any background structures. (2) Meanwhile, we developed a methodology to automatically determine the regions suitable for NPS evaluation in the close vicinity of the relevant structures. The next step was, to derive the NPS from the CT images according to the recommendations of the ICRU Report 87 (6). For detrending, a two-dimensional polynomial was fitted to this ROI, the polynomial was then used for correcting the ROI values. That allowed to remove the very-low frequency background. Afterwards, the NPS can be computed by using a 2D-Fourier transformation to the values of the corrected ROI R_i:

$$NPS(f_i) = \frac{v_x v_y}{N_x N_y} |\mathrm{DFT}\{R_i - \overline{R_i}\}^2|$$

Where v_X and v_Y are the pixel dimensions in 6illimetres and N_X and N_Y are the number of pixels in the *x* and *y* dimension (here 16 and 16). As described in the above mentioned ICRU report 87 (6), the 1D-NPS can be obtained by radially averaging the 2D-NPS. Larger values of the NPS correspond to stronger noise in the CT image. The general computational procedure is summarized in Fig. 2. (2)

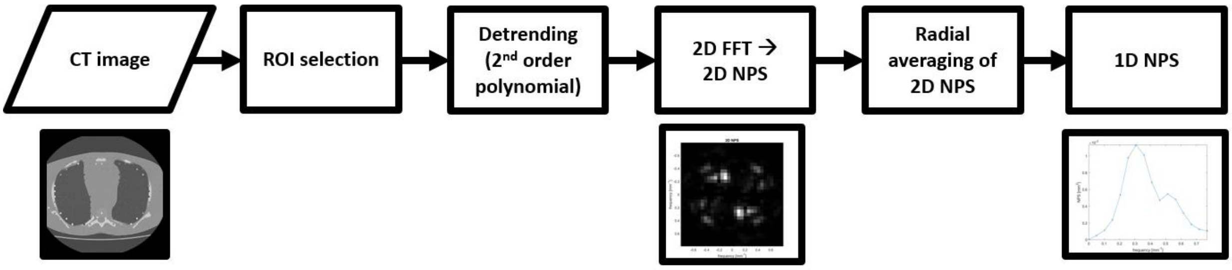


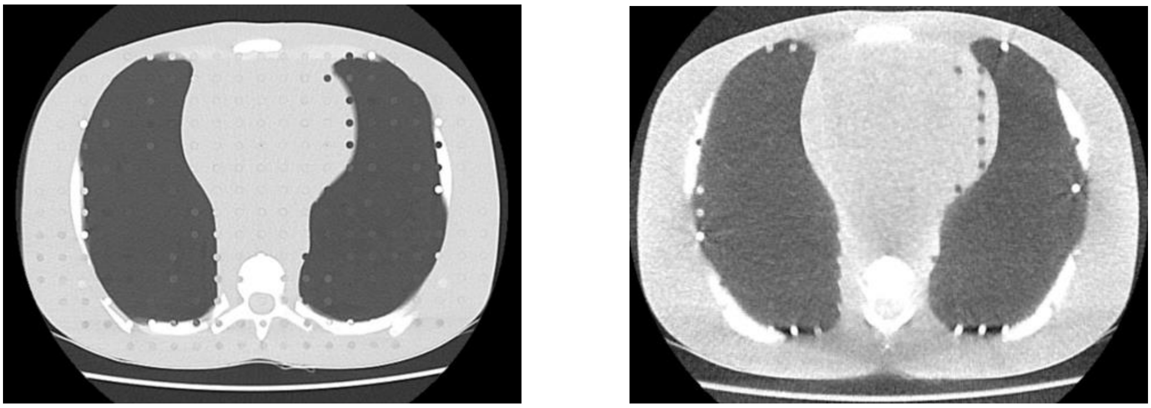
Figure 2: General workflow of the NPS computation for a chest CT image within the Medirad project. (from reference 2)

2.5 Test image data sets for the evaluation

As stated in the proceedings for the presentation at SPIE 2020 (2), phantom and real clinical patient images were used for the evaluation of the developed methodology:

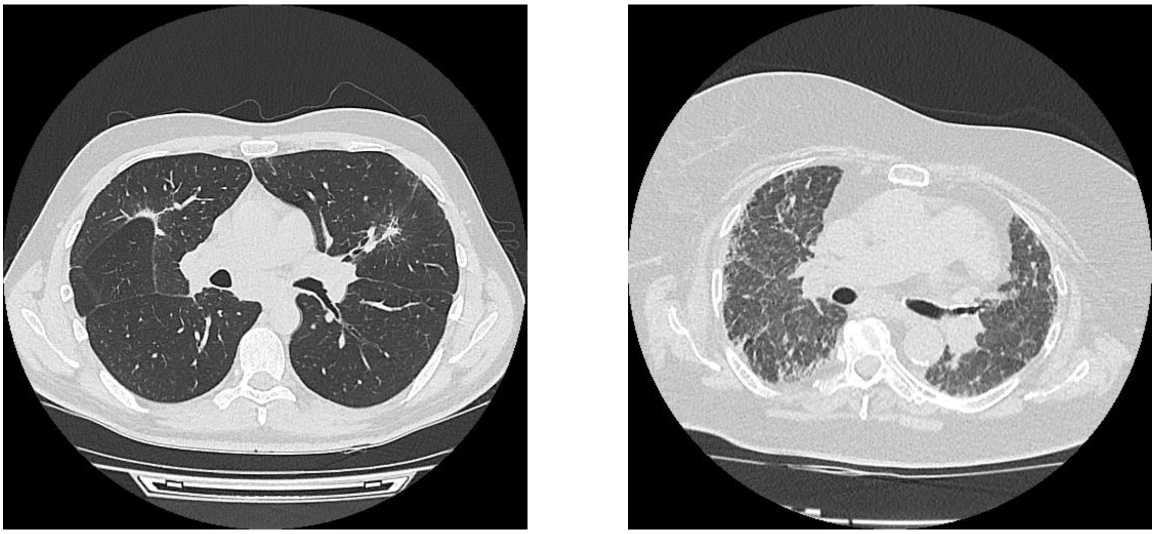
"A CT phantom (Dosimetry phantom, Cirs Inc. USA) was scanned with varying parameters using a clinical CT system (Aquilion Prime, Toshiba). The reconstruction was done by an iterative reconstruction technique. The quality of the acquired phantom images was altered by changing the tube voltage and tube current. Three different beam qualities were used, e. g. 120 kV/80 mA, 100 kV/50 mA and 80 kV/10 mA. Furthermore, two different the slice thicknesses were considered, i. e. 1 mm and 5 mm. The in-plane pixel size was 0.65 mm x 0.65 mm." In figure 3 two phantom images exemplarily chosen are shown. They have been acquired using different tube settings, both are reconstructed with a slice thickness of 5 mm. As such images have been gathered by imaging the phantom, the background shown in the images is nearly homogeneous. This allows to test the algorithm described before.

The patient images were taken from the data provided by the Otto-von-Guericke university hospital for the clinical study using the same CT system and reconstruction technique as described in the paragraph before. For our purposes here we used the major fissure of the left lung and the right inferior pulmonary vein as two of the areas considered in this study. The subjective image quality was categorised by radiologists from the hospital, that provided the images. The classification stated either acceptable or not-acceptable for diagnostic purposes. To judge the images and its acceptability the radiologists did use the structures and approaches as described in the previous subsections. Two exemplary chest CT images are depicted in Fig. 4. (2).



(a) 120kV, 80mA (b) 80kV, 10mA

Figure 3: Exemplary CT image of the investigated torso phantom. (from reference 2)



(a) Thorax CT rated as “acceptable”. (b) Thorax CT rated as “not-acceptable”.

Figure 4: Exemplary thorax CT images from patients with different quality. (from reference 2)

2.6 Results of the objective image quality analysis and the correlation approach

In Figure 5 1D-NPS determined from different phantom CT images are shown. It is easy to identify the increased low-frequency noise in the images acquired with low tube current (10 mA or 50 mA). One can also see that for images reconstructed with 1 mm slice thickness the higher frequency components of the NPS show larger magnitudes compared to the values for those images reconstructed with 5 mm slice thickness. (2) These findings could be expected because

- Thicker slices allow the averaging over many more detected rays for the reconstruction of one voxel in a slice, thus the quantum noise and especially the high-frequency components will be reduced.
- Lower mAs result in less photons per irradiated volume and thus also on an increased quantum noise. However, in the case of iterative reconstruction if there is a certain sufficient number of photons (meaning above a certain exposure threshold) and if there are not too many structures like in the phantoms high level noise will be suppressed so that there are no strong differences in the higher frequency region.

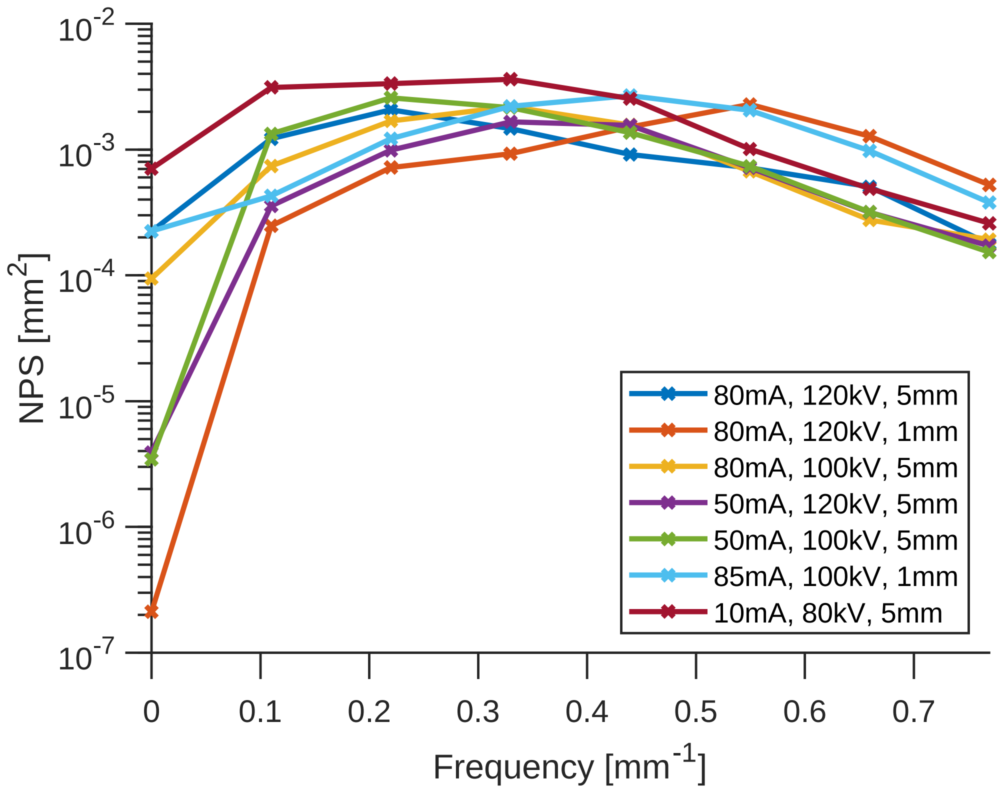
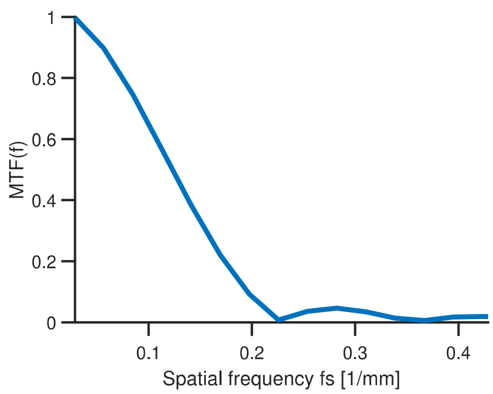
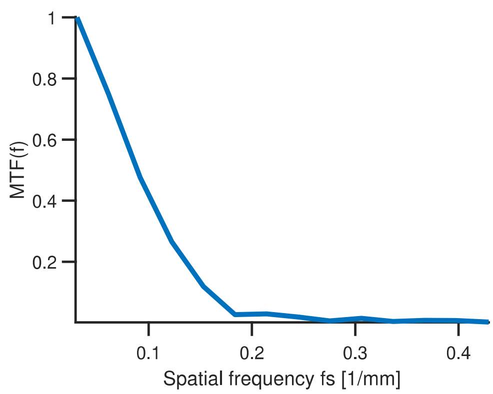


Figure 5: NPS spectra of the acquired phantom images using different tube settings and slice thicknesses. (from reference 2)

In figure 6 two exemplary MTF curves derived as described above are presented. The MTFs were calculated from two different chest CT images (compare Fig. 4). One of the images was marked as "acceptable" while the other was classified as "not-acceptable". The MTF of the image annotated as "acceptable" authenticates that higher frequency components presented stronger in the anatomical structures. We found that the area under the MTF shown in Fig. 6a ("acceptable") was 0.11 while the value derived for the area under the MTF curve in Fig. 6b ("not-acceptable") was 0.07. (2)



(a) MTF from the thorax CT shown in Fig. 2a



(b) MTF from the thorax CT shown in Fig. 2b

Figure 6: MTF results from the CT thorax images shown in Fig 4. (from reference 2)

As mentioned before, in the meantime a methodology had also been developed to determine "structure- free" areas in the clinical images to use the developed NPS determination method in real patient images. It is based on a region segmentation approach, which in its main components is shown in figure 7:

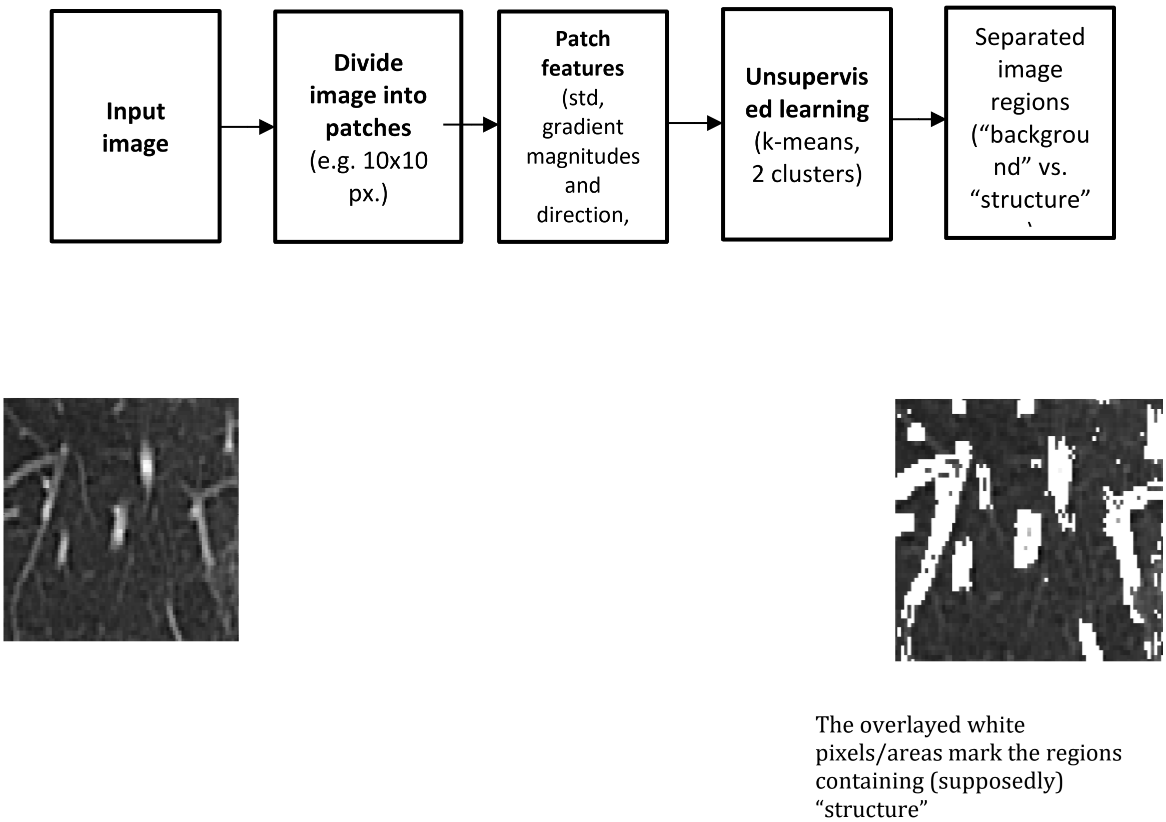


Figure 7: Workflow for region segmentation for evaluating suitable regions for NPS measurement.

One exemplarily solution is shown in figure 8:

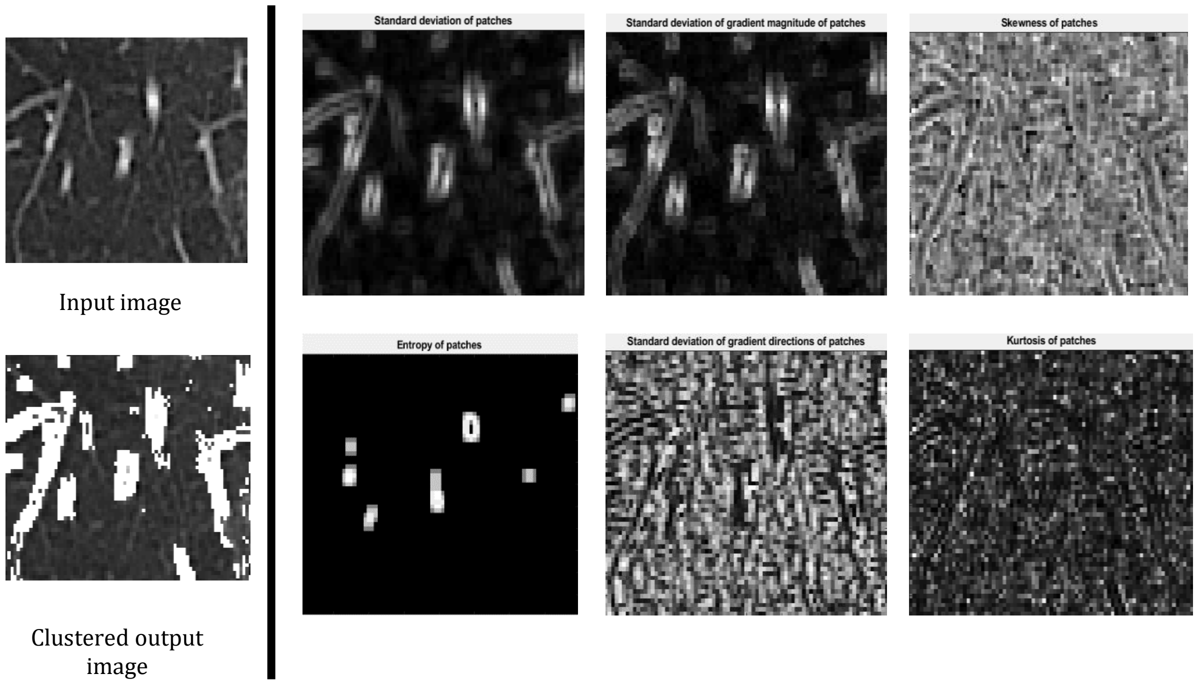


Figure 8: An exemplary region segmentation evaluation for determination of structure-free region in clinical image.

The next step was to define a suitable combination of criteria. Also, this is performed and implemented.

The correlation analysis so far shows that the determined resolution as well as the noise properties in the physics-based evaluation correlate with the acceptable or non-acceptable subjective image quality ratings and it corresponds to the predicted outcome of changes.

3. Conclusion

Within this work,

- A dedicated study was initiated for determine subjective image quality evaluations on chest CT for three relevant clinical indications. All necessary preparations are done, the study is going to be performed and finalized during the MEDIRAD project.
- A framework for objective image quality analysis based on NPS and MTF in real clinical patient images is proposed. The algorithms are developed and evaluated using phantom images or real thorax CT images annotated by different radiologists.
- The comparison between subjective image quality evaluation and physical measures show good correlation. In addition, the results obtained from the physical measurements are consistent with the imaging parameters.
- As soon as the clinical study has been performed for subjective image quality analysis and the same images have in total been evaluated by the physics-based methods developed, the optimisation can be performed as promised in the DoA with respect to the image quality.
- As described in another deliverable based on the work of subtask 2.1.2. the dose determination for a corresponding optimisation has also been finalized (see deliverable D 2.13). Thus, the development of the optimisation software can be done as promised based on very promising scientific data.
- As far as it can be judged, the shown approaches allow for the first time a complete semi-automatic evaluation of physics-based image quality parameters in real clinical patient images which are correlated to subjective image quality evaluation. There are peer-reviewed papers about this in preparation.

4. References

[1] Marie-Pierre Revel, Maciej Pech, Zahra Passand, Amelia Estevao, Graciano Paulo, Joana Santos, Apostolos Karantanas, Angelica Svalkvist, Ase Johnsson, Jenny Vikgren, Christoph Hoeschen: SUBJECTIVE QUALITY ASSESSMENT OF THORACIC CT IMAGES: ERPW18, Rovinj, https://erpw2018.com/wp-content/uploads/2018/09/ERPW-abstracts-book_web_pages.pdf

[2] Passand, Z. and Hoeschen, C.: Image quality assessment of real patient thorax CT images using modulation transfer function and noise power spectrum, SPIE Proc. https://doi.org/10.1117/12.2550073 (2020).

[3] Sanders, Jeremiah & Hurwitz, Lynne & Samei, Ehsan. (2016). Patient-specific quantification of image quality: An automated method for measuring spatial resolution in clinical CT images. Medical Physics. 43. 5330-5338. 10.1118/1.4961984.

[4] Rolstadaas, L. and Wasbø, E., "Variations in mtf and nps between ct scanners with two different ir algorithms and detectors," *Biomedical Physics & Engineering Express* **4**(2), 025009 (2018).

[5] Chan, T. F. and Vese, L. A., "Active contours without edges," *IEEE Transactions on image processing* **10**(2), 266–277 (2001).

[6] Boone, J. M., Brink, J., Edyvean, S., Huda, W., Leitz, W., McCollough, C. H., McNitt-Gray, M., Dawson, P., Deluca, P. L., Seltzer, S., et al., "Radiation dose and image-quality assessment in computed tomography," *Journal of the ICRU* **12**(1), 9–149 (2012).