

CHALMERS



How does image-enhancing software affect the detection performance in computed tomography?

Master of Science Thesis

DANIEL OLSSON

Department of Signals and Systems
Division of Biomedical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden, 2010
Report No. EX004/2010

Abstract

This thesis investigates the influences of applying an image-enhancing software for computed tomography produced images, whose purpose is to reduce the radiation dose. This is mainly accomplished with a detection task aimed at human observers. Synthetic created low-contrast object serves as stimuli in different noisy environments, received by a water phantom study. To assess observer performance the methodology of the "two-alternative forced choice (2AFC)" was used for images unprocessed and processed, with the image-enhancing software for the comparison. The sensitivity for discrimination is measured as the detectability index for the proportion of correct scores in the 2AFC, based on the assumption of two Gaussian distributions with same variances for unbiased observers. The experiment includes three different sizes for the synthetic object, represented by simulated nodules, for one contrast. Result from five observers shows an overall higher average of correct scores for the processed images. But above all, a remarkable higher and uniform discrimination amongst the observers is received. The result also indicates a possible reduction of radiation dose, for all three object sizes.

Sammanfattning

Detta examensarbete utreder påverkan av ett bildbehandlingsprogram för bilder skapade med datortomografi, vars syfte är att reducera strålningsdosen. Detta görs i huvudsak genom en detektionsuppgift ämnade åt mänskliga observatörer. Syntetiska lågkontrast objekt har skapats och agerar stimuli i olika brusmiljöer som erhållits genom en vattenfantomsstudie. För att bedömma observatörernas förmåga används en framtvingad tvåvals metod för oprocessade och processade bilder för en jämförelse. Känsligheten för urskillningen mäts som detekterings index för antal rätt i $2AFC_n$, baserat på antagandet om två Gaussiska distributioner med samma varians för obiaserade observatörer. Experimentet inkluderar tre olika storlekar för det syntetiska objektet, som representeras av simulerade noder, för en kontrast. Erhållna resultat från fem observatörer visar ett generellt högre medelvärde, för de processade bilderna. Men framför allt, ett anmärkningsvärt högre, och enhetlig urskillning mellan samtliga observatörer. Resultaten indikerar också på en möjlig reduktion av strålningsdos, för samtliga objekt storlekar.

Acknowledgement

First of all I want to thank Markus Håkansson at Södra Älvsborg hospital and Magnus Båth at Sahlgrenska University hospital for me letting me write my thesis within the field of medical imaging at Södra Älvsborgs hospital. Thanks are also given to all the friendly and helpful employees of the hospital, partly for participating in the evaluation study, but also for all the support. I also want to thank my supervisor Artur Chodorowski at Chalmers University for the technical support and guidance along the path of this thesis work.

Secondly I want to thank Sapheneia for letting me use their product. Special thanks are given to Marie Samuelsson and Kajsa Tibell for all the support and help.

Thanks are also given to my girlfriend Malin for all the patience, support and understanding during my ups and downs, my daughter Abra for holding both my feet on the ground, my mother in law Christina for all the support, all the time.

Finally I want to thank Jesper Nordén for correcting my English.

Contents

1	Introduction	1
1.1	Purpose and task description	1
1.2	Background	1
1.3	Delimitation	2
2	Overview of computed tomography imaging	3
2.1	Introduction	3
2.2	The computed tomography chain	4
2.2.1	Principles of generating X-rays	4
2.2.2	Attenuation	5
2.2.3	Computed tomography detectors	6
2.2.4	CT number and Hounsfield unit	7
2.2.5	Projection, reconstruction and reconstruction algorithm	7
2.3	Dose regulations and image quality	9
3	Approach the image	10
3.1	Image properties	10
3.1.1	Image model degradation	10
3.1.2	Spatial resolution	11
3.1.3	Image contrast	11
3.1.4	Noise	12
3.2	Signal-to-Noise ratio	13
4	Signal detection	14
4.1	Introduction	14
4.2	Receiver operating characteristic method	14
4.3	Two alternative forced choice method	15
4.3.1	The assumption of two distributions	15
4.3.2	Sensitivity	16
5	Material and methods	18
5.1	Data acquisition	18
5.2	The DICOM file	19
5.3	Experiment method and implementation	20
5.3.1	Image selection	20
5.3.2	Synthetic nodule simulation	22
5.3.3	Selection of mask radius	23
5.3.4	Nodule insertion	25
5.3.5	Nodule localization	27
5.3.6	2AFC image compound	29
5.4	Pilot validation	30
5.5	Image enhancement tool	30
5.6	Evaluation of unenhanced and enhanced images	30
5.7	Mathematical image analysis	31

6	Results	32
6.1	Pilot validation	32
6.2	Image analysis and measurements	34
6.2.1	Mathematical image analysis	37
6.3	Evaluation of unenhanced and enhanced images	38
7	Discussion	41
7.1	The work of this thesis	41
7.2	Validity of selected images	41
7.3	Nodule size and contrast	42
7.4	Implementation accuracy	43
7.4.1	Insertion procedure	43
7.4.2	Evaluation of the 2AFC implementation	44
8	Conclusion	45
8.1	Summation	45
8.2	Recommendations and future work	45

Abbreviations

ALARA	As Low As Reasonably Achievable
BKE	Background Known Exactly
CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine
FD	Fourier Domain
FOV	Field Of View
FT	Fourier Transform
FW5M	Full Width at Fifth Maximum
HU	Hounsfield Unit
keV	kiloelectronVolt
kV	kilo Voltage
kVp	peak kilovoltage
mAs	milliAmpere second
MTF	Modulation Transfer Function
OTF	Optical Transfer Function
PACS	Picture Archiving and Communication System
ROC	Receiver Operating Characteristic
ROI	Region Of Interest
PSF	Point Spread Function
SD	Spatial Domain
SDT	Signal Detection Theory
SKE	Signal Known Exactly
2AFC	two-Alternative Forced Choice

1 Introduction

This chapter will introduce the reader to the thesis by giving a description of the task followed by a background.

1.1 Purpose and task description

The purpose of this thesis is to investigate the influences of using a post-processing enhancement software, aimed at CT reconstructed images. Since the lack of evidence are rather limited, Södra Älvsborgs hospital are interested in what information there is to extract by using the image-enhancing software for detection tasks. The assignment is to create images, which will be subjectively and mathematically observed for the cases, unprocessed and processed images by using image-enhancing software. This assignment has been divided into following groups:

- Data acquisition from CT phantom study at Södra Älvsborgs hospital.
- Create images consisted of synthetic low-contrast objects for detection tasks.
- Process the images with image-enhancing software.
- Find suitable method for the evaluation of the images.
- Implement the evaluation using human and/or mathematically observers.
- Data analyze

1.2 Background

In the early 1970s, Godfrey Hounsfield introduced the first computed tomography (CT) scanner for commercial use [1]. Using only a single X-ray tube and detector, the first generation CT was born and a new era of imaging was initiated. Since the release of the first generation CT with scan times up to a few minutes, the technique has been developed with different revolutionized technical solutions. Facing the fifth generation helical CT, the scan times are reduced to a few milliseconds [2].

In comparison to conventional X-ray, the digitized CT image technique shows reconstructed structures built up by tissues through a slice of a particularly anatomy. Examination of diseases in the region of the head or thorax e.g. looking for tumours or illness in pancreas, CT provides such valuable information conventional X-ray could not provide with [3].

There is a huge disadvantage using CT for medical imaging, namely the ionized radiation. The awareness of the ionized radiation CT contributes with, approaches of low-dose solutions trough post-processing algorithms has become a fact. The purpose with the image-enhancing software is to reduce image noise and enhance anatomical structures with maintained image quality, whilst the radiation dose is decreased.

Södra Älvsborgs hospital, where this thesis has been carried out as a part of the Master of Science in the Biomedical Engineering and Electrical Engineering at Chalmers University of Technology, shows skepticism among the low-dose solution and are not that convinced, even though some hospitals around Sweden already implemented this solution for some modalities. Today the knowledge concerning the post-processing software are rather insufficient and the lack of evidence pointing the direction that image-enhancing software could be used to decrease the radiation dose is limited.

1.3 Delimitation

The thesis will only treat post-processing software aimed at CT reconstructed images. The resemblances and differences between anthropomorphic phantoms and human tissues will not be taken into account when dealing with the phantoms during measurements, but discussed in chapter 7. The thesis will neither deal with the impact and influences for different type of CT machines and reconstruction algorithms since only one CT machine has been used for measurements in the work of the thesis. This thesis neglects all other source of noise other than the quantum noise.

2 Overview of computed tomography imaging

This chapter will give an overview of CT and the chain of components included for standard procedures during CT examination. Particularly, a more detailed description about the function of CT and the technique in mapping volume elements are described, since this diverge from conventional radiography.

Within the electromagnetic spectrum, X-rays can be found in the wavelength range of approximately 10^{-8}m to 10^{-12}m , with a photon energy spectra between 0.1keV to 100keV depend on the field of application. X-rays are classified as ionized radiation resulting in a health problem, addressed both the patient and the surrounding population.

2.1 Introduction

CT is a digitized diagnostic imaging method used to produce individual transverse thin slices of anatomical structures, one at the time. One image slice reflects the attenuation of the X-rays passing through the inhomogeneous human body. The attenuation diverges from place to place since the human body includes different type of tissues. By combining adjacent slices, it is possible to display three-dimensional images. This is an advantage in comparison to conventional radiography since the detection of small irregularities increases. With conventional radiography, three-dimensional anatomy are projected onto a two-dimensional plane resulting in the lack of information caused by overlapping anatomy [1]. The importance of CT can be found in examinations and diagnostics within the skull, spinal cord, thorax, and abdomen e.g. looking for cerebral hemorrhage, tumours and diseases of the liver [3].

Even though CT only constitute almost 5% of all X-ray examinations, it contributes with almost 30-40% of the collective radiation [4]. This is because a relatively high radiation dose is given during examination in comparison to conventional radiography. The huge benefits of providing valuable information is of great significance within medical diagnostic imaging and could not be replaced with other imaging techniques with respect to the financial and time aspects.

2.2 The computed tomography chain

General CT systems consists of an X-ray source, attenuation profile, detectors and reconstruction profile. Also, produced data from examinations are stored as Digital Imaging and Communications in Medicine (DICOM) files in the Picture Archiving and Communication System (PACS). The chain of components are presented in Figure 1.

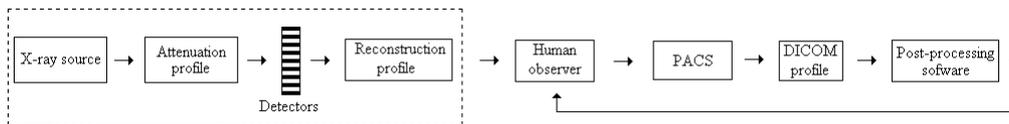


Figure 1: Schematic overview of the CT chain. The "human observer" indicates the subjective response for the analysis of images before and after applying image-enhancing software.

2.2.1 Principles of generating X-rays

The X-ray technology for radiography use a X-ray tube in purpose to produce and emit photons, see Figure 2. The underlying theory and mechanisms behind this production starts of by describing the parts of an X-ray tube.

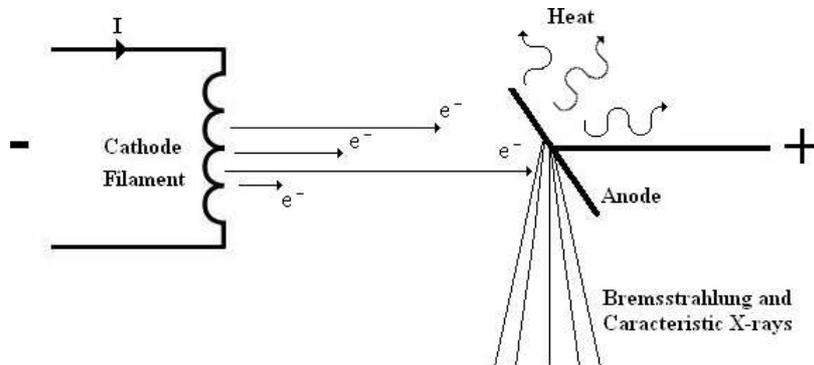


Figure 2: Simplified schematic of an X-ray tube. Excluded from this illustration is a generator with adjustable filament power supply for the tube current control, and a high voltage supply for the peak kilovoltage (kVp). These two adjustable parameters is a central part [1] of producing X-rays and gives the ability to control the exposure of energy through the body.

The purpose of the filament is to generate free charged electrons which are accomplished by the high temperature that arises when the tube current passes through. When the filament reaches a sufficiently large temperature, loosely bounded charges will set free and accumulated about the filament [1, 2].

Due to the arise of a potential difference between the anode and cathode, high kinetic electrons starts to accelerate toward the anode. The kinetic energy that is not transformed into heat constitutes the X-rays. There are two known processes for producing X-rays, characteristic X-rays and bremsstrahlung [1], see Figure 3.

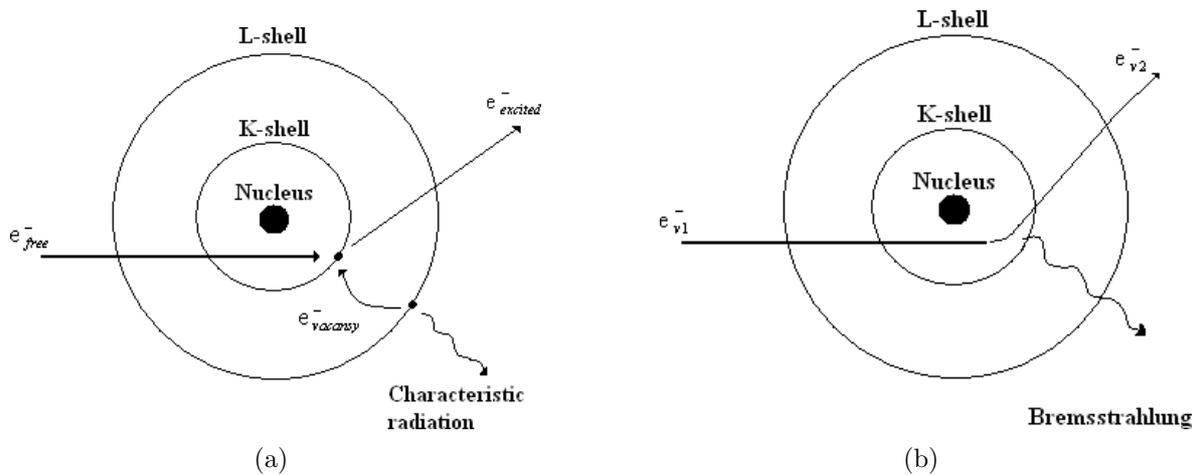


Figure 3: Two significant processes within the production of X-rays where (a) characteristic X-ray and (b) bremsstrahlung.

Characteristic X-rays contribution to a standard diagnostic beam is about 20% [1], and the process in how to generate the photons are different from bremsstrahlung. If the incoming electron, that goes toward the anode, have sufficiently enough energy it will hit and remove an orbital electron from an atom within the anode and thereby create a vacancy throughout excitation. This space will be filled with an outer-orbital electron. When the outer-orbital electron leaves its initial state to fill a vacancy, energy will be emancipated, i.e. an electron leaves from L-shell to K-shell in the atom, and produces characteristic X-rays, also illustrated in Figure 3(a) [5].

Bremsstrahlung is the result when using the motion of electrons and the kinetic energy within it. When the electron moves with a velocity, v_1 , towards the atomic nucleus, Coulomb forces will affect the electron and force it into another path as can be seen in Figure 3(b). Within this change of motion the electron loses velocity. The new velocity will be v_2 , where $v_1 > v_2$, resulting in the release of energy and a bremsstrahlung photon is generated by an arbitrary value between zero and the whole of the kinetic energy of the electron [5].

2.2.2 Attenuation

The basic idea of generating CT images can be described by the mapping of linear attenuation coefficient values μ . The linear attenuation coefficient are the result of different X-ray photon energies and its interaction with different tissues, i.e. different densities and compounds of the material. The X-ray photon interacts with the material (tissue) through two dominant effects known as the photoelectric and Compton effect [1].

For a photoelectric effect, which dominates for photon energies less then 100keV, see the explanation for the production of characteristic X-rays.

The Compton effect, which dominates for photon energies above 100keV, is the result of the interaction between an incoming X-ray photon and an outer-orbital electron, called

the recoil electron. The X-ray photon emits some energy to the recoil electron and changes its course of origin, now with less energy. As the recoil electron leaves a scattered photon is emitted [2, 1].

The X-ray spectrum is divided into primary and secondary X-rays, where the primary X-rays are produced from the anode plate due to characteristic X-rays or bremsstrahlung. The secondary X-rays are produced from Compton effect due to scattering. The consequences of these two effects is a reduction of beam intensity, since the intensity is proportional to number of X-ray photons in it [6].

The beam intensity I is a variable that is related to the thickness x of the object, and the linear attenuation coefficient μ according to the expression

$$I = I_0 e^{-\mu x} \quad (1)$$

where I_0 is the intensity of origin. This equation is only valid under the influence of mono-energetic photons, i.e. all X-ray photons are possessed with the same energies [1]. Unfortunately the X-ray source does not produce pure mono-energetic photons and eq.(1) can only be seen as an approximation. The result of excluding different possessed energy levels for the X-ray photons, i.e. low and high energetic photons, are described as the phenomena beam-hardening. The rise of beam-hardening causes artifacts and distortion in the CT image but can be reduced through different methods [6, 7].

As mentioned earlier the linear attenuation coefficient was mapped during scanning. This means that for each location of the anatomic slice the value of μ is calculated as a variable that depends on the interaction of the X-ray photon energy and the material compound. The meaning of location is here referred to as a volume element or voxel. The cross-sectional area of a voxel are in the order of 1mm^2 and a few millimeter thick [1]. When the values of μ are calculated the receiving is a two-dimensional map of linear attenuation values throughout that specific slice. Calculations of the linear attenuation coefficients includes heavy computational calculations and are together with the image reconstruction determined.

2.2.3 Computed tomography detectors

The detectors receive the photons that have passed throughout the material and contains the information about the attenuation. Often the detectors are developed from scintillators. The scintillators are characterized by the ability of "burst of light" when incident photons arrive and hits the scintillator. The scintillator transmits a visible light which is detected by photomultiplier which generate a pulse of charge. Common materials used in medical scintillators are sodium¹ iodide doped with thallium² [1]

¹Na, Atomic number 11

²Tl, Atomic number 81

2.2.4 CT number and Hounsfield unit

The Hounsfield unit (HU), named after 1979 nobelprize winner Sir Godfrey N Hounsfield, describes the relationship between the linear attenuation coefficient μ , and that of water μ_{H_2O} . The HU are also known as the CT number. Calculation of the CT number, for a tissue at a specific point, is accomplished by the expression

$$CT_{number} = \frac{\mu - \mu_{H_2O}}{\mu_{H_2O}} \times 1000 \quad (2)$$

where water is chosen as a reference material since it constitutes almost 80% - 90% of soft-tissue mass [1]. A value of $CT_{number}=0$ would correspond to $\mu=\mu_{H_2O}$ and a value of $CT_{number}=-1000$ to $\mu=0$. The value of the CT number is often represented as an integer value between -1000 and 3000 where a relatively low value indicates on soft tissue and a high value on compact bone [7].

When the linear attenuation values are displayed as a pixel transverse slice matrix, each value of attenuation will have a corresponding gray-scale value, where the intensity of the gray-scale value depends on the calculated attenuation. Normally each pixel corresponds to a voxel of attenuation [1]. Some common CT numbers for various tissue types are presented in Table 1.

Tissue type	CT number [HU]	Tissue type	CT number [HU]
Bone	1000+	White matter	24-36
Liver	50-80	Water	0
Muscle	44-59	Fat	-20 to -100
Blood	42-58	Lung	-300
Gray matter	32-44	Air	-1000

Table 1: Common CT numbers of various tissue types [6].

2.2.5 Projection, reconstruction and reconstruction algorithm

As mentioned earlier in section 2.2.2 the linear attenuation coefficient was calculated according to eq.(1) for a homogeneous object. For an inhomogeneous object, i.e. different type of tissues, this would result in a sum of different values for μ along the path of x . The beam intensity I is described by the expression

$$I = I_0 e^{\sum_{i=1}^{i=\infty} -(\mu_i x_i)} \quad (3)$$

where I_0 is the intensity of origin, μ the linear attenuation coefficient and x the distance. One projection p , for an angle, say $\theta = 0$, can mathematically be expressed as

$$p_{\theta=0} = -\ln \frac{I}{I_0} = \sum_{i=1}^{i=\infty} \mu_i x_i \quad (4)$$

where the latter expression is the result of taking the natural logarithm of eq.(3). The result is the projection throughout an object along the X-ray path of a particularly ray

for that specific angle. The creation of the first projection profile is thereby obtained.

By letting the angle θ grow sufficiently by rotation through the object, several projection profiles can be obtained in purpose to evaluate the average linear attenuation for each voxel. The projection profiles are the basis for the reconstruction process. The process within the actual reconstruction differ from different modalities. There are several types of reconstruction algorithm methods used today, depend on the anatomy. The two dominant ones, for regular CT examination, is called Fourier reconstruction and filtered back-projection [8]. These two reconstruction algorithms are used since they are much faster then others [1].

To get a clearer view the projection and reconstruction are illustrated in Figure 5 and Figure 6 by using Matlab's **radon** and **iradon**. The input image being projected and reconstructed are shown in Figure 4.

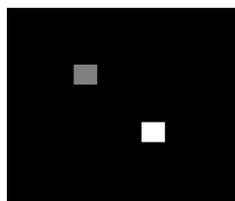


Figure 4: A reference image, 256×256 matrix, containing two squares filled with white and gray respectively.

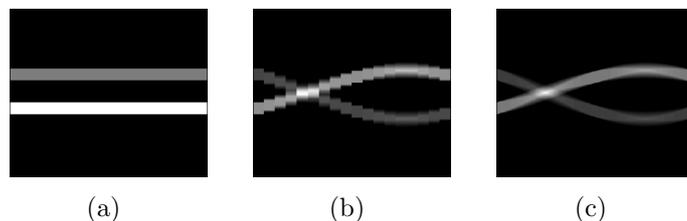


Figure 5: Projection profiles through Matlab's Radon transform for different amount of angles where (a) one projection, (b) $\frac{180}{10}$ projections and (c) 180 projections.

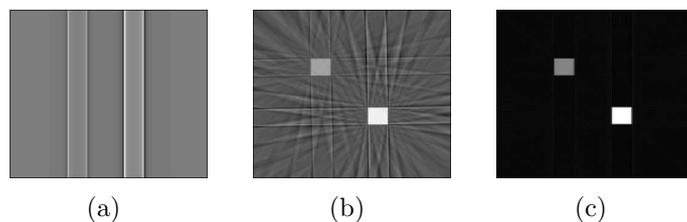


Figure 6: Reconstructed images through filtered back-projection using Matlab's inverse Radon transform where (a) one projection, (b) $\frac{180}{10}$ projections and (c) 180 projections.

The Radon transform (**radon**) calculates the projection of the image intensity along a radial line for a specific angle. By varying the angle, a number of projections of the image

intensity at different angles are obtained. The reconstructed image is calculated as the inverse Radon transform (**iradon**) for a parallel beam of the projection profile data, for the different angles.

2.3 Dose regulations and image quality

There are essentially two major discourses when discussing medical imaging in the quality perspective. A separation is drawn between equipment, such as technical and physical parameters of a system, and interpretation and detection of an image done by the radiologist [9]. Standard procedures among hospitals occasionally occur during CT examinations. In a technical perspective, CT systems differ from different systems. Generally the physical and image parameters that will affect the image quality can be divided into these two groups [2].

- For physical parameters
 - Tube voltage (kV)
 - Tube current (mA)
 - Time of exposure (s)
 - The slice size
 - Pitch (table movement/numbers of rotations)
 - The thickness of the object

- For image parameters
 - Kernel
 - Reconstruction algorithm
 - Field of view (FOV)

No further detailed description in how these parameters affect each other mutually will be given. Neither how they should be adjusted to maintain the best result in the image quality perspective. For further reading see for instance [4]. Standard procedures with common routines and lock-up tables among hospitals are commonly used and differ with the pathology. The ambition among hospitals, for regular CT examinations, is to apply the principle "As Low As Reasonably Achievable" (ALARA) to answer the increasingly proportion of ionized radiation [10]. The common denominator for the parameters mentioned above is that they will affect the image quality in terms of noise, contrast and spatial resolution combined with the technical setup [4, 8]. This will be discussed further on in the next chapter.

3 Approach the image

The meaning of quality can be described in many different ways. The quality of an imaging system, such as CT systems, must fulfil specified characteristics to be diagnostically useful. This chapter will discuss important measurement properties when judging the quality of an digital image, but also possible adjustments that can be received for CT images.

3.1 Image properties

The content of an arbitrary two dimensional image is a spatial pattern of intensities at some location and are therefore rather to be analyzed in one, or both, of two domains. These are known as the Spatial Domain (SD) and Frequency Domain (FD). Approaches of an image in the SD refers to the image plane itself, whilst FD is a representation of the spatial frequencies the image is composed of, by decomposing the image into sums of sine and/or cosine components i.e. Fourier Transform (FT) [11]. The reason for using the two domains is that the information content may be easier to interpret in one or the other.

3.1.1 Image model degradation

For an arbitrary linear, position-invariant degradation process the input and output relationship can be modeled, both in SD and FD, as the convolution in SD and multiplication in FD with the degradation function $h(x, y)$ or $H(u, v)$ according to the expressions

$$g(x, y) = f(x, y) * h(x, y) + \eta(x, y) \quad (5)$$

$$G(u, v) = F(u, v)H(u, v) + N(x, y) \quad (6)$$

where $f(x, y)$ and $F(u, v)$ represents the input images and $g(x, y)$ and $G(u, v)$ the output images. The terms $\eta(x, y)$, $N(u, v)$ denotes the additive random noise for SD and FD respectively [11, 12]. The noise will be discussed further in subsection 3.1.4.

Point Spread Function

The term $h(x, y)$ represent the spatial degradation function for an arbitrary image. In physical optical systems it is possible to measure the attached degradation function. For imaging system this degradation function is called Point Spread Function (PSF). The PSF can be seen as a quality measurement, deciding how much a system will spread or blur a point of light, i.e. the impulse response for a specific imaging system is decided by its PSF. In the FD domain the PSF is transformed via FT to the optical transfer function (OTF). Assuming that $\eta = N = 0$ the output images can be expressed as

$$g(x, y) = f(x, y) * PSF(x, y) \quad (7)$$

$$G(u, v) = F(u, v)OTF(u, v) \quad (8)$$

where $PSF(x, y)$ and $OTF(u, v)$ represents the PSF or OTF of that specific system [6, 11, 12].

Modulation Transfer Function

The Modulation Transfer Function (MTF) for an imaging system can be obtained by taking the magnitude of the OTF expressed as

$$MTF(u, v) = |OTF(u, v)| \quad (9)$$

where the measure of MTF examines the magnitude for each frequency component the image possesses. Thereby it is possible to analyze which of all frequency components being attenuated or passed by the system [6, 12].

3.1.2 Spatial resolution

In common terms the spatial resolution is a measure of describing resolved details in the image e.g. separation of small features such as adjacent lines or dots. The spatial resolution of a digital image is characterized by the systems PSF, or equivalently by the MTF. Another important factor concerning the spatial resolution is the number of pixel used to represent the image. This is decided by the selection of sampling rate [6, 11].

3.1.3 Image contrast

To meet the variety of requirements in digital images these normally undergoes some sort of enhancement to suit specific requirements. Approaches of the image includes numerous of post-processing techniques to visualize relevant data. The medical image is usually built up by different shades of gray where the differences of gray levels are determined by the density of the tissues. Common digitization of CT images occurs with 12-bit, resulting in 4096 different shades of gray for each pixel [6]. To improve the contrast some modalities uses contrast agents which is a chemical substance, normally iodine³.

To improve the pictorial information for human interpretation in CT images, this can be accomplished by adjustments of the contrast through the window width and window level, called window technique. The window level value can be seen as the average value for a specific range of gray values, i.e. window width. When adjusting window width this alter contrast in the image, and by the selection of window level different structure selection can be captured. Figure 7 illustrates the change of window width with a constant window level of the CT number, $N_{CT} = -400$. In Figure 8 the window level has been changed from $N_{CT} = -400$ to $N_{CT} = -824$.

³I, Atomic number 53

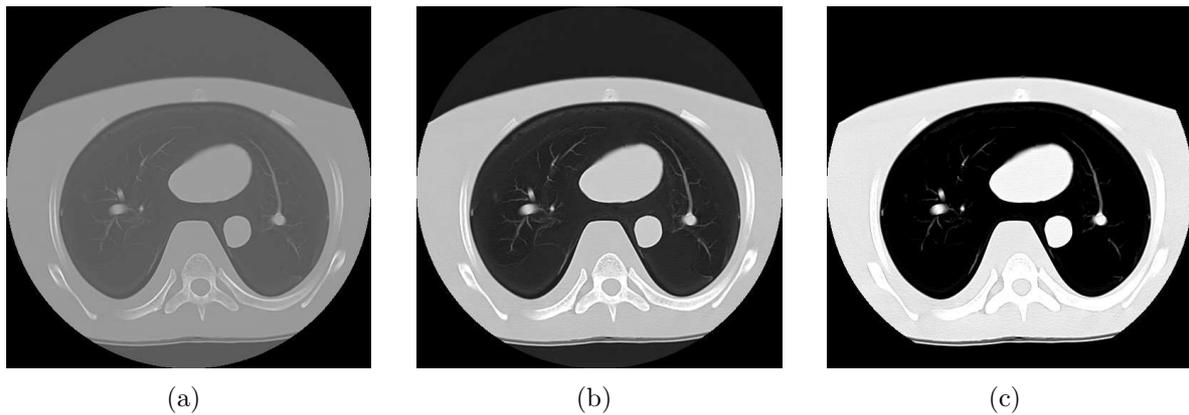


Figure 7: Thorax tomogram of Lucas the phantom object for different size of window width with a constant window level value of $N_{CT} = -400$ where (a) $N_{CT} \pm 976$,(b) $N_{CT} \pm -224$ and (c) $N_{CT} \pm -524$. Image is captured from Södra Älvsborgs hospital.

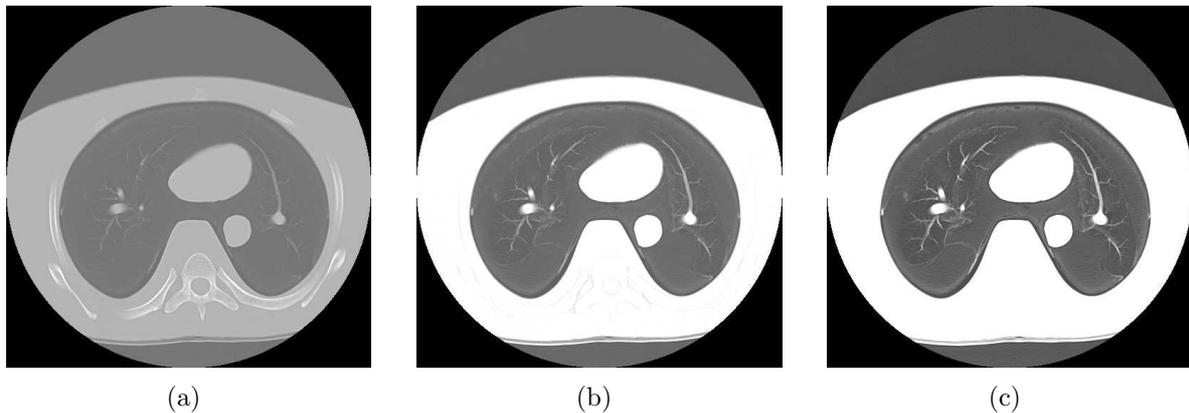


Figure 8: The same thorax tomogram as in Figure 7 but window level is changed from $N_{CT} = -400$ to $N_{CT} = -824$ resulting in a threshold where (a) $N_{CT} \pm 976$,(b) $N_{CT} \pm -224$ and (c) $N_{CT} \pm -524$. Image is captured from Södra Älvsborgs hospital.

A comparison between Figure 7(c) and Figure 8(c) indicates that by changing the window level different structures can be emphasized. The vessels in Figure 8(c) are clearly seen in comparison to Figure 7(c) at the expense of not receiving information about the bone structures.

3.1.4 Noise

The origin of a low quality image are caused by the unavoidable quantum noise which possesses the ability to corrupt the image by a statistically random uncorrelated noise. The quantum noise, or photon noise, is due to the fact that source emitted photons will vary per time unit, but also the photons interaction with the material, i.e. the linear attenuation coefficients at a specific location will differ from time to time. The random quantum noise usually obeys the Poisson distribution to describe independent counting events and is directly related to the radiation dose [6, 4]. To illustrate the impact of the noise components a homogeneous water phantom has been radiated using a CT. From

Table 1 the CT number for water is zero, but since the CT machine is quantum limited a variation of CT numbers will appear described by the Poisson distribution. Figure 9 illustrates the impact of three different radiation dose levels 26mAs, 96mAs and 259mAs respectively for the homogeneous water phantom object.

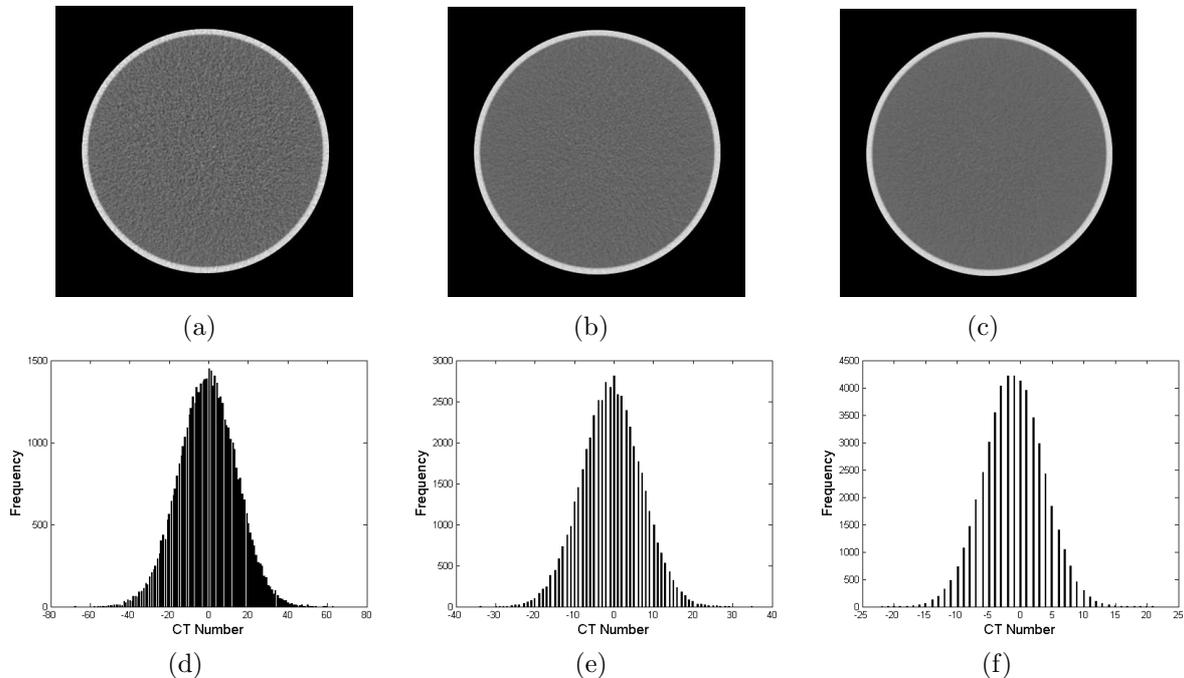


Figure 9: Water phantom object measured for gradual increase of radiation dose where (a) 26mAs, (b) 96mAs and (c) 259mAs and the corresponding histogram shows the fluctuations in CT numbers described by the Poisson distribution where (d) 26mAs, (e) 96mAs and (f) 259mAs.

The histograms from Figure 9(d) - Figure 9(f) show a gradual decrease in fluctuation of CT numbers around zero, i.e. an increased radiation dose will result in a less noisy image since the quanta increases.

3.2 Signal-to-Noise ratio

The relationship between the signal and noise can be received by using the signal-to-noise ratio (SNR) measurement. In a Poisson distribution the standard deviation is numerically equal to square root of N, where N indicates the number of photon carrying the signal [6]. Measuring the ratio can therefore be expressed as

$$SNR = \frac{N}{\sqrt{N}} \quad (10)$$

for N number of photons. An increased photon number will result in less additive noise [6, 4].

4 Signal detection

Since there are numerous ways to interpret an image, depend on background and purpose of the observer, it is well motivated to select a proper and valid method determining the observers performance. This thesis aims at experiment with synthetic low- contrast object inserted in CT produced images for signal detection tasks, using sample images with known background and known signal. The experiment includes unprocessed and processed images with signals of different sizes to be detected in different amount of noise content. By using human observers, the subjective response are measured for the cases unenhanced and enhanced images for the comparison. This chapter gives the reader a general introduction to the psychophysical field with focus lying on the two-alternative forced choice (2AFC) method for signal detection experiments.

4.1 Introduction

The term psychophysics is a discipline within psychology which examines the relationship between a physical stimuli and the subjective response. Psychophysical experiments often aims to distinguish poor auditory or visual signals from noisy environments [13].

In the medical image field there are professionals such as medical physicians and radiologists examining the images. Measuring the observer performance, it can be interpreted and described as the ability of a human observer to subjectively detect relevant information extracted in images. The information can be abnormalities within the image. This often result in an experts opinion. When evaluating images with respect to the quality or information content, the professionals do not always agree. One approach of measuring and quantifying the response performance can theoretically be explained by using the model of signal detection theory (SDT) [13, 14].

4.2 Receiver operating characteristic method

The receiver operating characteristic (ROC) method has been classified as gold standard in medical imaging, for the evaluation and comparison of imaging methods. ROC methodology can be used in a way to graphically generate so-called ROC curves to quantify the performance of giving correct diagnoses for real or simulated cases. To calculate the accuracy of a specific test or model, index A_z often are used. The index A_z represent the area under the ROC curve and take probability values of 0.5 to 1.0. A value of 0.5 indicates fifty-fifty chance of getting a correct answer, and 1.0 no failure. [14, 15].

There are mainly two distinct ways of determining the ROC for the observers performance. The first method, ROC, is determined by asking the observer if a signal is present or absent, in one image at the time. The observer are asked to answer "yes" or "no" depended on if he or she might see the appearance of a signal or not. The second method, which has become the most popular one, ROC is determined by using a graded rating scale. The observer is asked to quantify his or shes certainty, if a signal is present or not. The grading scale design often include grades from one to five, where one represent

”signal definitely present”, and five ”signal definitely not present”. Using ROC methods can be extremely time consuming and demands high carefulness designing it [14, 16].

4.3 Two alternative forced choice method

Measuring the observers performance for visual signal detection tasks, an attractive and rather simple way of evaluating the performance is by the two-alternative forced choice (2AFC) method [16].

In a 2AFC experiment the observer is shown two images at the time with one of them consisting of a stimuli (signal). The observer is asked to decide in which one. The simplest case of 2AFC, to receive information about the detectability of the observer, is known as the signal known exactly/background known exactly (SKE/BKE) case. The SKE/BKE aims at detection of signals with known size, shape and intensity placed in a known position for a uniform background with uncorrelated additive noise. The signal is randomly assigned in the center of one of two noisy background images with equal prior probability. The two images are then displayed for the observer, who are forced to select either one of the two images, to decide which one contains the signal. The observer is presented an arbitrary amount of images. The fraction P , of correct score are then calculated [15, 16] to measure the detectability index d' (sensitivity). The area under an ROC curve A_z , can be shown to be identified as the fraction of correct score P , in a 2AFC experiment [17]. The 2AFC methodology suits experiments where the aim is to discriminate between signal intensities, signal shapes and signal sizes [16].

4.3.1 The assumption of two distributions

Theoretically a binary decision AFC task for discrimination can be described by using the model of SDT, with the assumption of two independent Gaussian distributions, with the same variances. A translation of this assumption is that the two presented images are observed independently [13]. The first distribution represent the image consisted of noise with a signal absent, the other of noise with a signal present. An illustration of the two possible situations are shown in Figure 10. The two Gaussian distributions includes P_L and P_R and represents the proportion of correct hits for a signal present in the left or right interval respectively.

By neglecting the occurrence of biased observers, the assumption that the observer systematically would favor either left over right or vice versa are neglected [16].

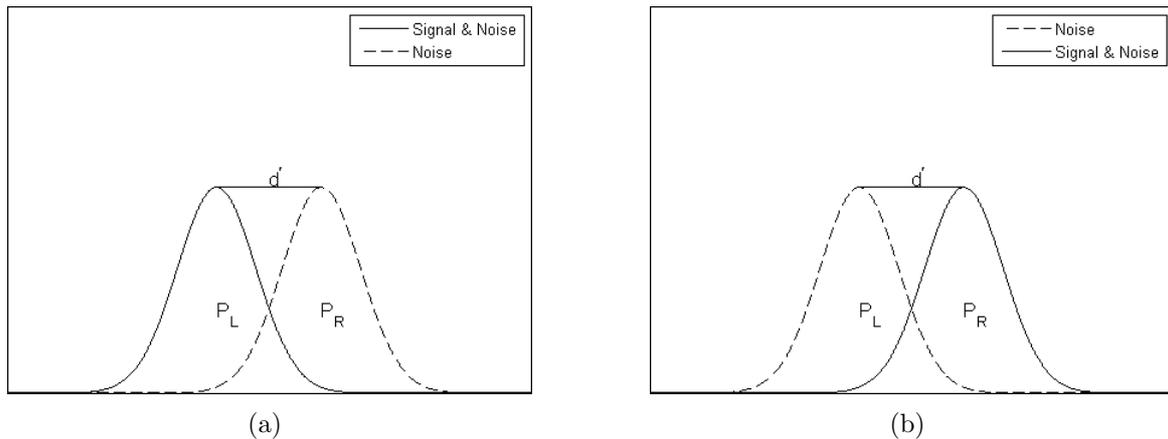


Figure 10: Theoretical representation of a binary AFC task explained by SDT. P_L and P_R are the proportion of correct responses for the observer. The detectability d' indicates the distance between the two distributions.

4.3.2 Sensitivity

The separation between two Gaussian distributions, illustrated in Figure 10, is measured as the sensitivity. The sensitivity or detectability index d' , describes the ability to discriminate between the two distributions. This is illustrated in Figure 11 where the decreasing d' gives an indication of poor discrimination.

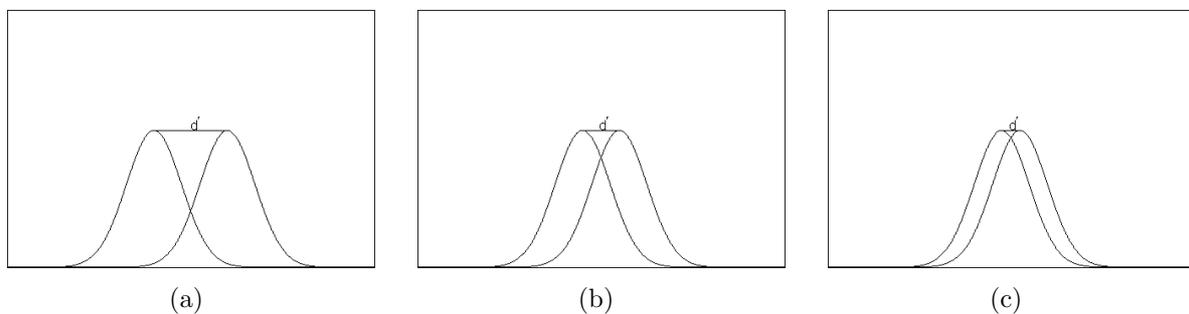


Figure 11: The detectability d' describe the distance of separation between the two distributions and measures the sensitivity for discrimination tasks.

When assuming an unbiased observer, calculation of d' for the fraction of correct score P , is received [16] according to expression

$$d' = \sqrt{2}z(P) = \sqrt{2}\Phi^{-1}(P) \quad (11)$$

where $z(P)$ is the standard score (z-score) which are related to the cumulative normal distribution with the inverse [16]. In a 2AFC experiment, a fraction of correct score $P=0.5$ implies pure guessing, why any value of $P \leq 0.5$ is uninteresting. A value of $P < 0.5$ could indicate a biased observer. The theoretical relationship between the fraction P of correct score obtained in a 2AFC experiment, and the corresponding detectability index d' are shown in Figure 12.

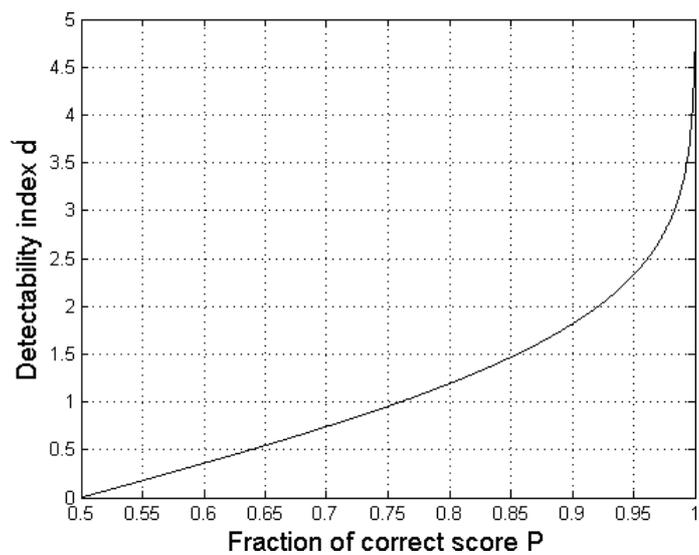


Figure 12: Detectability index d' versus the fraction of correct score above $P \geq 0.5$. A value of $P=0.75$ ($d' \approx 1$) would correspond to 50% detection probability. As P approaches 1, $d' \rightarrow \infty$.

5 Material and methods

With the theory from chapter 4, the methods for the design of the two-alternative forced choice are here described. The basis of the experiment is to investigate the influences of using an image-enhancement software for synthetic low-contrast objects of different sizes, both subjectively and mathematically. The study compares images unprocessed and processed with Clarity™ tuning tool [18]. The materials used for the production of images has been done through Matlab version R2009b [19] and processed with Sapheneia Clarity™ Tuning tool version 2.73. The selection of reference values, kernels and phantom object have been done in consultation with supervisors at Södra Älvsborgs hospital and Sahlgrenska University hospital to receive qualitative results. The evaluation and storage of data has been handled with viewDEX 2.0 [20].

5.1 Data acquisition

The data acquisition were received through a water phantom study carried out at Södra Älvsborgs hospital in Borås. The homogeneous phantom is an instrument used within the hospital for calibration and reference measurements for Siemens CT Somatom Definition. The Somatom Definition is a dual x-ray source helical CT occupied in room 9. The phantom has been used to create data with different amount of noise. Figure 13 shows the working flow for the production of data.

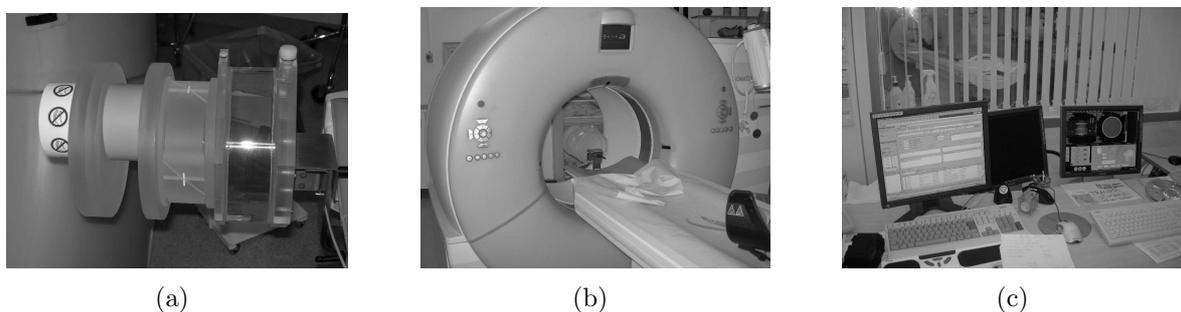


Figure 13: The working flow for data acquisition where (a) water phantom object, (b) gantry of Somatom Definition CT and (c) workstation where images are stored in PACS as DICOM files.

The CT settings for the phantom measurements were assumed to be as for liver measurements, with a reference value of 150mAs. A series of 12 different axial measurements of the water phantom was then collected with an interval of 20% between each measurement (3 measurements above the reference value, and 8 measurements below the reference value). The outcome of the 12 measurements, was 12 different data with respect to the noise.

The Somatom Definition at Södra Älvsborgs hospital comes with a numerous of eligible convolution kernels in purpose to enhance anatomical structures depend on pathology. Data was partly collected with a smooth kernel number "b26f", and partly with kernel number "b30f", at the tube voltage 120kV. The data used in this experiment are the one reconstructed with kernel "b30f", for a slice thickness of 3mm. The differences between

kernel "b26f" and "b30f" is that kernel "b26f" applies an advanced smoothing algorithm in purpose to enhance the anatomical structures [21]. All data was registered in PACS as DICOM-files, and later stored on a CD.

During regular measurements and examinations the mAs value is depending on the thickness of the scanned object, and varies along the object if the size differ. This value is called effective mAs. The CT equipment handles this automatically if not turned off. The water phantom in this study have been illuminated with the same mAs throughout the whole object and no effective mAs was calculated.

5.2 The DICOM file

The DICOM file (stored with suffix .dcm) contains relevant information, or metadata, about the examination. The metadata consists of tags with information. Such information are for example the size of the image, the window width and window level settings, type of anatomy being examined and convolution kernel.

For this study all images was stored as 512×512 matrices, with a reconstruction diameter of 294mm. The bitdepth was 12 (4096 gray levels) for the window settings, window width 350 and window level 35. The result from 1 of the 12 images (26mAs) is visualized in Figure 14 by using Matlab's **dicominfo** and **dicomread**.

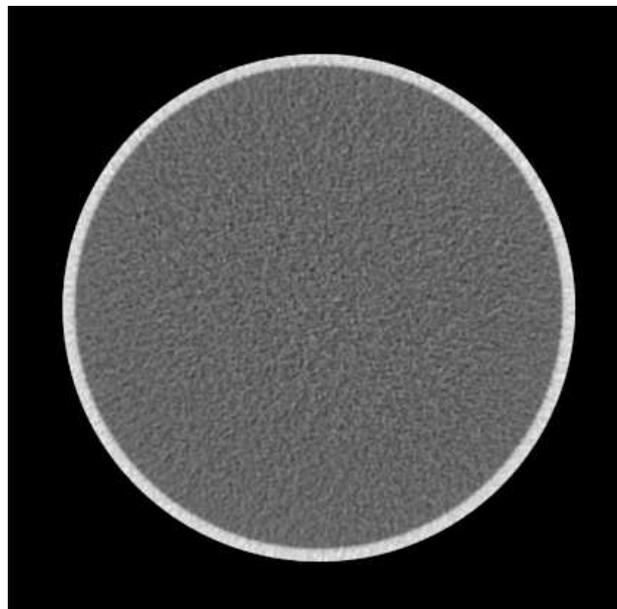


Figure 14: Captured DICOM image shows the axial 512×512 image from the water phantom at tube voltage 120kV and tube current 26mAs. The slice thickness is 3mm. The window settings of origin are window width 350 HU and window level 35 HU.

Statistics of all the images were calculated by using a large (226×226) region of interest (ROI) and are presented in Table 2.

Exposure [mAs]	Minimum pixel intensity [HU]	Maximum pixel intensity [HU]	Mean pixel intensity $\pm \sigma$ [HU]
26	-68	63	-1 ± 14
32	-54	56	-1 ± 13
40	-58	48	-1 ± 12
50	-49	42	-1 ± 11
62	-42	41	-1 ± 9
77	-38	35	-1 ± 8
96	-34	35	-1 ± 8
120	-30	33	-1 ± 7
150	-27	29	-1 ± 6
180	-24	23	-1 ± 6
216	-22	20	-1 ± 5
259	-22	21	-2 ± 4

Table 2: Image statistics for 12 different mAs-values calculated for a squared ROI-size of 226×226 . The symbol σ indicates the standard deviation.

5.3 Experiment method and implementation

The following sections will describe the methods and parameters used to design the experiment. The methods and parameters were implemented in Matlab.

5.3.1 Image selection

Images for the experiment were selected from the production of 12 different noise levels. The images are suppose to act as background images (background-known-exactly) consisted of noise of different levels. This is to receive different signal-to-noise ratio. It was desired to include a huge number of unique background images for the experiment, since repetitions of background images was undesired.

To receive the large portion of background images, 64×64 sub-images was cropped and rotated from the original images. The meaning of original images is referred to as the one in Figure 14 (and for all 12 mAs-values). For each mAs-values a set of 48 sub-images was cropped. This is illustrated in Figure 15 where 4 sub-images are cropped from the original image. Each of all sub-images was given a unique name to keep track of the location.

The start location for the production of the first 12 sub-images was at coordinate position $\Delta x=120$, $\Delta y=180$, for the original images. Then, a step of $\Delta x=20, \Delta y=0$ was taken, to create the next set of 12 sub-images. This procedure was then repeated for the new start location $\Delta x=120, \Delta y=180$.

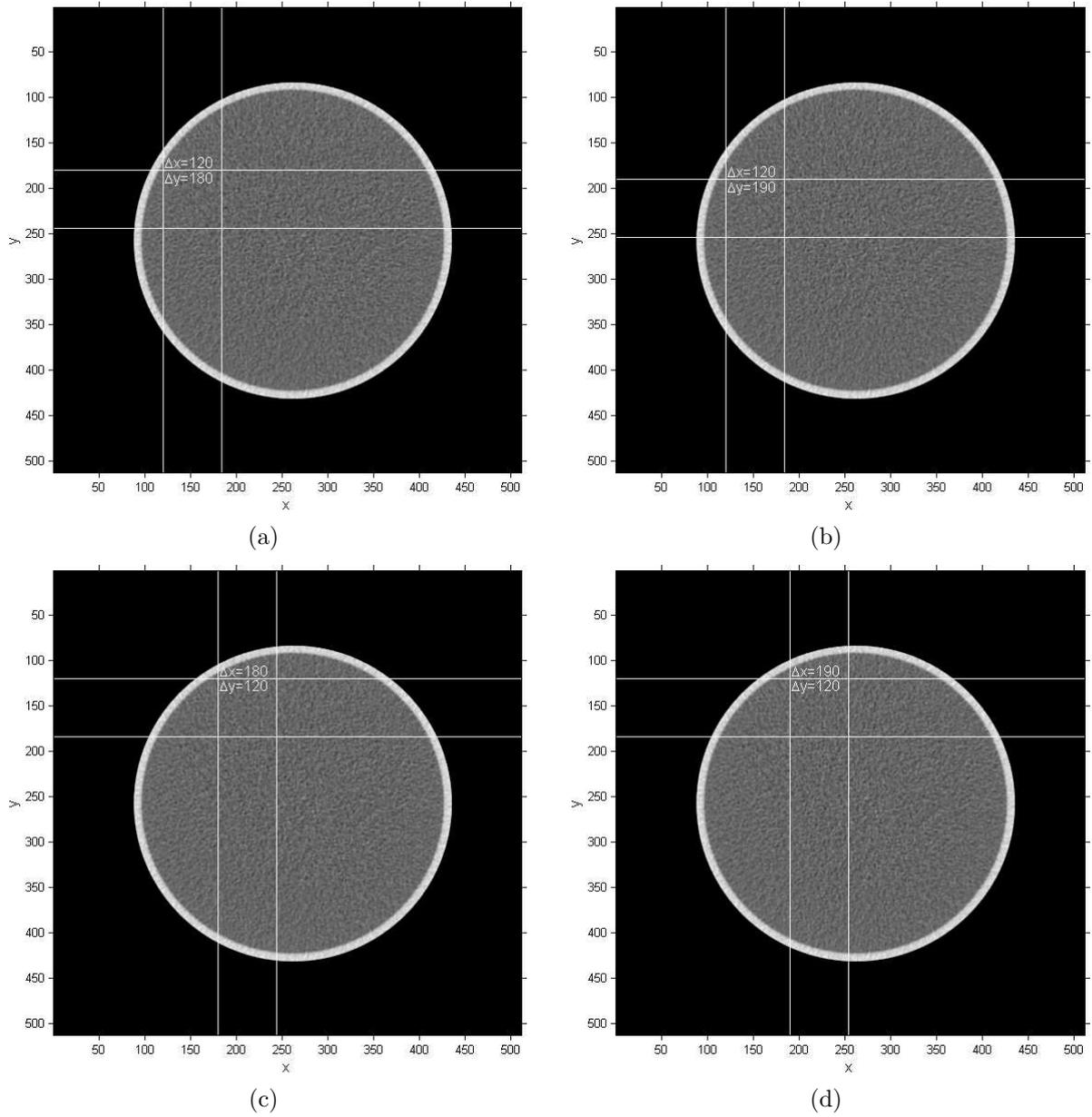


Figure 15: The four starting location for the cropping process of 48 sub-images. The size of one cropped sub-image is 64×64 .

Each of all 48 cropped sub-images, for each value of mAs, was rotated 3 times (anti-clockwise) with a step-wise rotation angle of 90° . By start and including 0° , 4 sub-images was received for each cropped sub-image. This resulted in a total amount of $48 \times 4 \times 12 = 2304$ unique sub-images. The rotation process is illustrated in Figure 16.

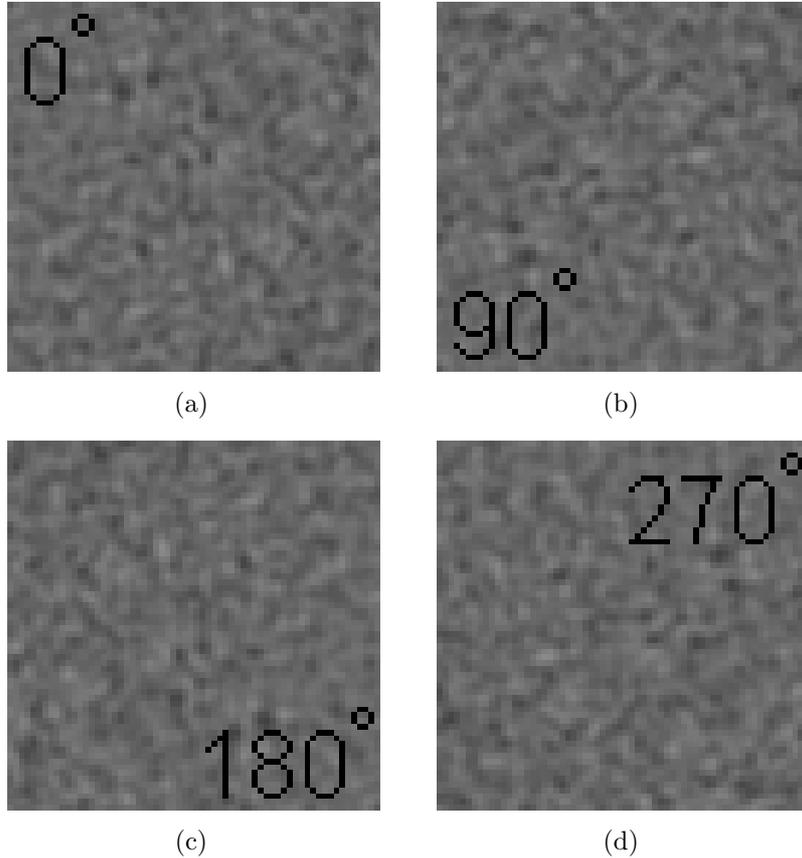


Figure 16: Anti-clockwise rotation of the cropped sub-image where (a) 0° , (b) 90° , (c) 180° and 270° .

5.3.2 Synthetic nodule simulation

Simulation of the object were done according to the model specified by [22, 23] with some modification to fit this experiment. The object represents the signal to be detected in this experiment (signal-known-exactly), and describes a nodule. The nodule model is a radially symmetric two-dimensional mask with varying contrast. The model has been used in several studies with successful outcome and supports the assumption of being mistake for a lesion in a liver [23]. Mathematically the function is given by the expression

$$c(r) = \begin{cases} C \left(1 - \left(\frac{r}{R}\right)^2\right)^n, & r \leq R \\ 0, & r > R \end{cases} \quad (12)$$

where $c(r)$ describes the amplitude as a function of the radial distance. The letters R and C represents the adjustable radius and mask peak contrast value to receive a desired

size and contrast of the nodule. The exponent n is a positive number and reflects the edge characteristics of the nodule. The selection of values for R and C was initially done through a 2AFC pilot validation to fit suggested recommendation about correct scores. The latter will be discussed further on later as the selection of exponent n .

5.3.3 Selection of mask radius

The experiment includes three different diameters, for one contrast of the simulated nodule. The diameter of the nodule was defined as full-width-at-fifth-maximum (FW5M) for the value $n=1.5$, see Figure 17(a). The selection of the value $n=1.5$, fits the lung-tumour profile [24], but has also been reported utilized for the simulation of liver lesion for discrimination between real and simulated lesions [23]. The latter support the $n=1.5$ selection, why this value was used in the experiment.

By assuming a desired start nodule diameter of 10mm, the three diameters of the nodule, calculated as FW5M, were selected to 4mm, 7mm and 10mm (after rounding to nearest integer). The corresponding radius in millimeter is $R_{mm}=2.30\text{mm}$, $R_{mm}=4.02\text{mm}$ and $R_{mm}=6.32\text{mm}$. The selection of the two diameters 4mm and 7mm has been done arbitrary to fit the experiment. The nodule profiles are plotted in Figure 17(b).

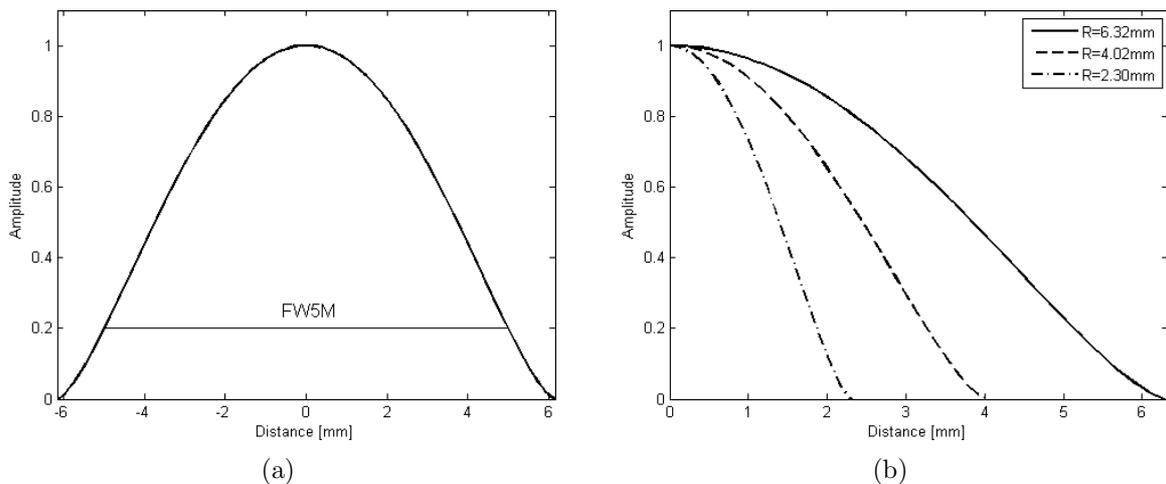


Figure 17: The nodule profiles where (a) shows how nodule diameters are calculated as the FW5M. (b) Solving eq.(12) at FW5M for $R=2.30$, $R=4.02$ and $R=6.32$ gives the desired diameters of 4mm, 7mm and 10mm.

R_{pixels}	R_{mm}	$r(R_{mm})_{FW5M,mm}$	$D(R_{mm})_{FW5M,mm}$	$D_{RTNI,mm}$
1	0.57	0.47	0.90	1
2	1.15	0.93	1.90	2
3	1.72	1.40	2.80	3
4	2.30	1.86	3.70	4
5	2.87	2.33	4.70	5
6	3.45	2.79	5.60	6
7	4.02	3.26	6.50	7
8	4.59	3.73	7.50	8
9	5.17	4.19	8.40	8
10	5.74	4.66	9.30	9
11	6.32	5.12	10.20	10

Table 3: Calculating diameters according to FW5M where R_{pixels} indicates the radius in pixels, R_{mm} radius in millimeter, $r(R_{mm})_{FW5M,mm}$ radius in millimeter at FW5M, $D(R_{mm})_{FW5M,mm}$ diameter in millimeter at FW5M and $D_{RTNI,mm}$ selected diameter in millimeter rounded to nearest integer (RTNI).

Implementation of the three simulated nodules are shown in Figure 18 where the contrast C is reduced for three different selections of peak contrast, for the two-dimensional mask.

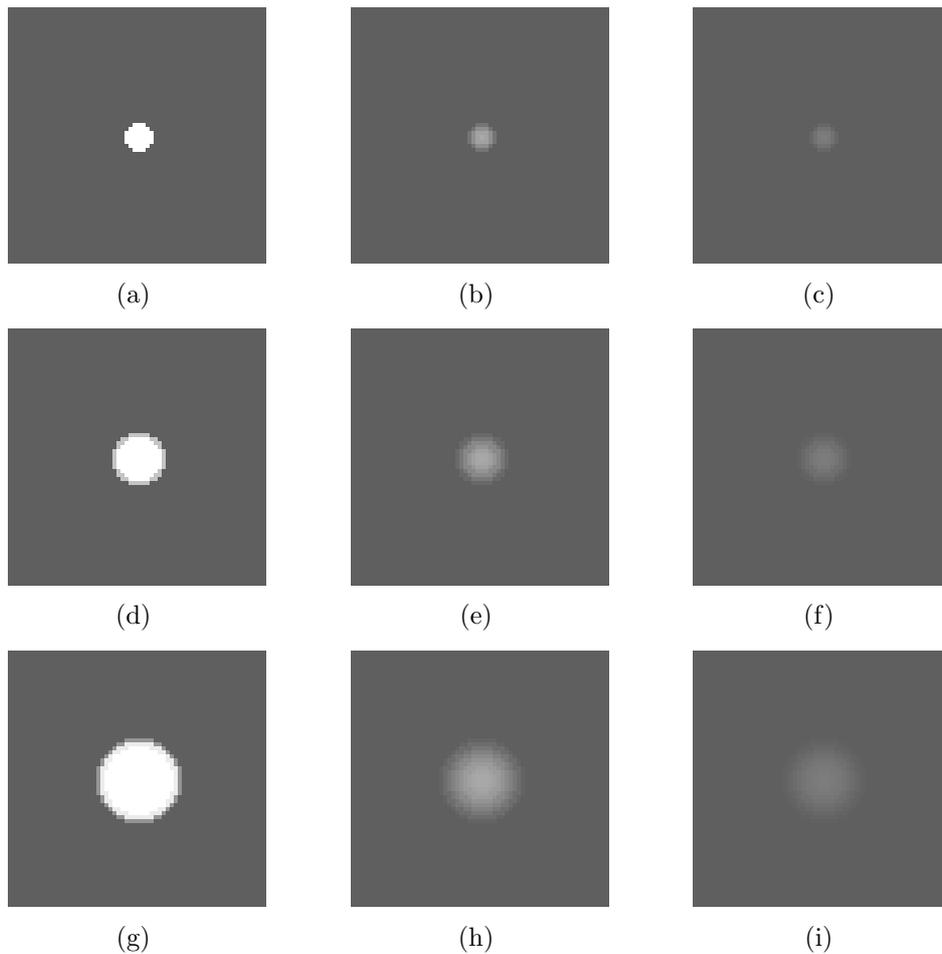


Figure 18: The simulated nodules inserted in a homogeneous background (-10 HU). The implementation is received via Matlab and illustrates contrast reduction for the steps $C=1$, $C=0.1$ and $C=0.01$ from eq.(12) where (a-c) 4mm nodule, (d-f) 7mm nodule and (g-i) 10mm nodule. The contrast $C=1$ corresponds to a mask peak contrast value of 1004 HU, $C=0.1$ to 91 HU and $C=0.04$ to 31 HU.

5.3.4 Nodule insertion

To complete images for the experiment, the synthetic nodule for the different diameters were inserted in the background images according to the expression

$$I(x, y) = c(r)I(x, y)_{noise} \quad (13)$$

where $I(x, y)$ represents a complete image consisted of a nodule and noise [24]. Each diameter of the nodule were selected and aimed at a specific range of four different mAs-values. The diameter of 10mm was inserted in 26mAs-40mAs, 7mm was inserted in 50mAs-120mAs and the diameter of 4mm in 150mAs-259mAs. By increasing the mAs-values, and thereby increasing the signal-to-noise ratio, the nodule is supposed to become more visible for detection.

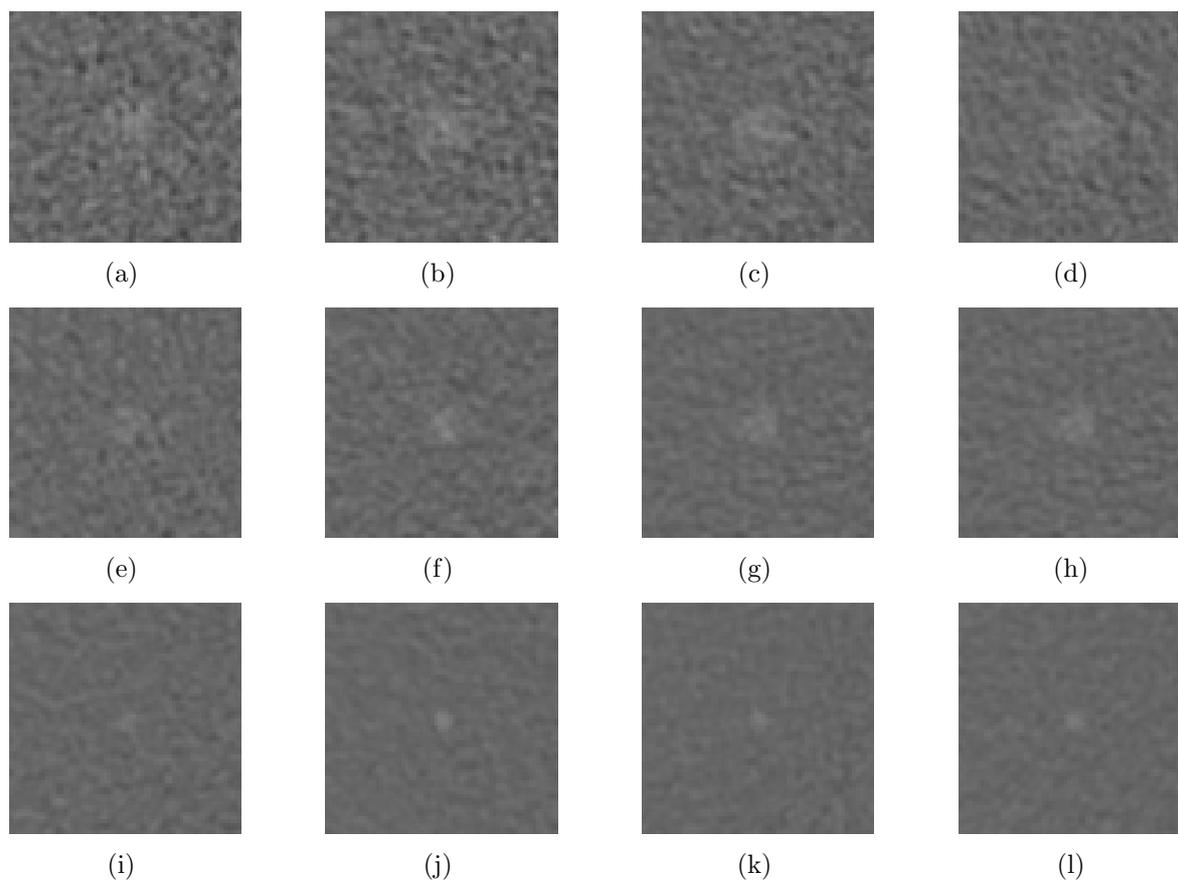


Figure 19: Nodule insertion according to eq.(13) where (a-d) nodule size 10mm are inserted in 26mAs-50mAs, (e-h) 7mm inserted in 62mAs-120mAs and (i-l) 4mm inserted in 150mAs-259mAs.

The impact of the overlapping noise will change the characteristic of the nodule. By adjustments (decrease) of the contrast for the nodule it will be less visible. The change in characteristic of the nodule profile for a contrast reduction is illustrated in Figure 20.

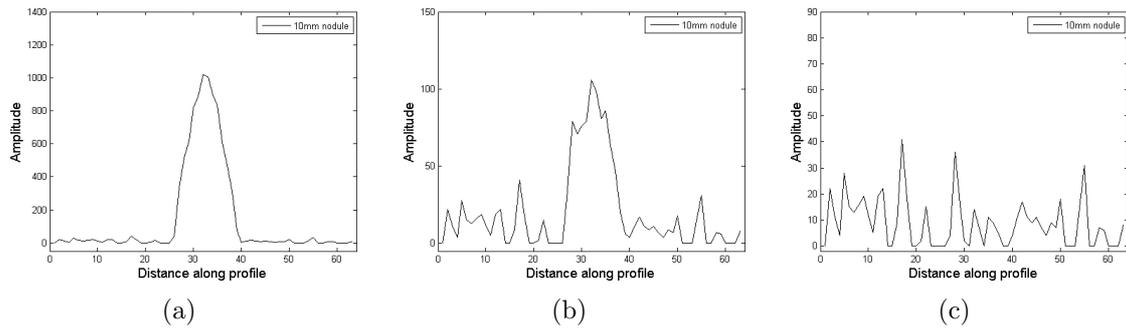


Figure 20: One-dimensional characteristics for the nodule 10mm in 26mAs with reduced contrast where (a) $C=1$, (b) $C=0.1$ and (c) $C=0.01$ where the signal is "drowning" within the noise.

5.3.5 Nodule localization

To fulfil the SKE/BKE case a complete control of the localization for the nodule insertion must be received. The nodule was randomly inserted in one of the two background images by using Matlab's **randi**. The command **randi** returns a random integer between a specified interval. To keep track on the background images consisted of a nodule it was given a unique name. In this experiment the localization of the nodule were always located in the center of the noisy images. Suggested by [16] the observer should receive information about the signal to be detected e.g. size, shape and intensity. A reference object (nodule) was therefore included, and placed outside the image in order to give the observer an opportunity to get information about the nodule. To define the localization of the nodule a (black & white) Toto circle [25] was used to provide a reference mark about the nodule in the image [16]. The size of the Toto circle was calculated according to the expression

$$d = 1.2\sqrt{w^2 + h^2} + offset \quad (14)$$

and describes the diameter of the Toto circle in millimeter. The letters w and h represents the width and height of the nodule. The offset was used to prevent too small Toto circles for the smallest diameter (4mm) of the nodule.

To prevent Toto circles with too much differences in intensity, in comparison to the noise, the intensity of the Toto circle was adapted and calculated according to the expression

$$i = \bar{I}(x, y)_{noise} \pm \sigma_{noise} \quad (15)$$

where i represent the two adapted intensities for the black & white Toto circle. The Toto is illustrated in Figure 21 for three different sizes and intensities.

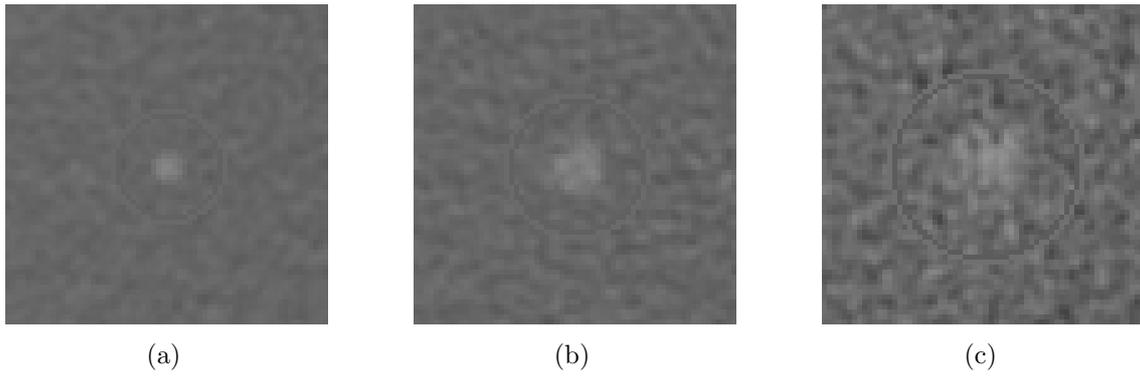


Figure 21: The black & white Toto circle are adapted to the size of the nodule and noise fluctuation according to eq.(14) and eq(15). The Toto defines the diameter to find the nodule within.

5.3.6 2AFC image compound

The compounded 2AFC image, designed according to the model of the 2AFC, that were presented to the observer is shown in Figure 22. The image, consisted of two unique background images (noise) whereas only one contains a signal (nodule) with a diameter of 4mm, 7mm or 10mm. One reference object is placed (above) outside of the noise fields to give the observer information about the size and shape. The adapted Toto circle defines the area to find the nodule within. The intensity of the background has been selected arbitrary to -10HU. The size of the -10HU background image is a 512×512 squared matrix.



Figure 22: The compounded 512×512 image consisted of a field with noise, a field with noise plus signal, two (above) reference objects and Toto circles. The signal is located in the left image of noise.

5.4 Pilot validation

To find the contrast threshold giving a recommended fraction of correct scores, a pilot validation was implemented. The pilot validation included four different human observers for their subjective response. Recommended value P , of correct scores should fit the range $0.85 \leq P \leq 0.95$ or equivalent, the detectability index d' $1.5 \leq d' \leq 2.4$ for best experimental results [16].

5.5 Image enhancement tool

A set of 24, 64×64 sub-images were produced and sent to Sapheneia as basis for the production of parameter files. One set consisted of two images, one with noise, and the other with the identical noise and an inserted nodule, for all exposure values, i.e. twelve different exposure values. The parameter files are the basis for the enhancement process of images and contains the tools needed for enhancing images. The contrast of the nodule were selected from pilot validation.

A unique parameter file were designed for each of seven mAs-values 26, 32, 40, 50, 77, 96 and 180mAs. The other five mAs-values were excluded due to problems with the implementation. All sub-images, for these seven mAs-values were processed with Clarity™ Tuning Tool which resulted in 1344 unique sub-images.

5.6 Evaluation of unenhanced and enhanced images

The evaluation were performed with a monochrome monitor EIZO RadiForce Gs220-CL (serial:181880) with the resolution 1280×1024 . The monitor was calibrated according to DICOM supplement 14 [26]. The ambient light was approximately 35-40 lux and varied between the observers.

For the image presentation and storage of the observer data a DICOM compatible program viewDEX 2.0 was used. Images were initially presented in the FIT mode where the image is scaled to fit in the canvas. The default window settings was window width 200 and window level 20. Alternative window settings were given as window width 150 and window level 20. These two window settings has been selected arbitrary after the pilot validation. The observers also assessed the zoom in/out and window/level functions during examination. The use of window/level function limits the bit depth to 10-bits. The presentation of images was unique since random mode was set as default. Each participant was given unlimited time for the examination and could at any time interrupt the examination, for further examination later. The distance from the monitor and the observer has been varied since this was eligible. A two-alternative forced choice was given, left or right. The observer were forced to choose either one of the two, to continue the examination. The analysis of data were received from a stored log-file assessed via viewDEX.

The evaluation was divided into two groups, images unenhanced and enhanced with Clarity™ tuning tool. Each participant was given an introduction followed by a training session in order to familiarize the observers with the appearance of a nodule and its size,

for the different noisy environments. The training session consisted of 30 sequenced images (10 for each diameter of the nodule). The observer assessed the correct answer during the whole training session.

5.7 Mathematical image analysis

To perform a mathematical analysis the signal-to-noise ratio (SNR) was estimated. The largest nodule (10mm) was selected [27] in order to analyze the relationship between the signal and noise, before and after they were enhanced. For the analysis, a ROI of size 11×11 pixels were selected inside the nodule and eight smaller ROIs of size 4×4 pixels outside for average calculations of the noise, see Figure 23.

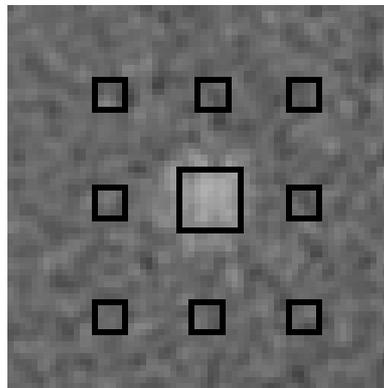


Figure 23: SNR measurement for 10mm nodule where the black squares indicates the selected ROIs.

The SNR estimation were performed for ten different images, for each seven exposure levels, before and after they were enhanced, to receive average values. This were accomplished since the noise differ from different images. The SNR were calculated according to the expression

$$SNR = \frac{1}{10} \sum_{i=1}^{10} \left(\frac{\overline{ROI}_{nodule} - \overline{ROI}_{noise}}{(\sigma_{nodule} + \sigma_{noise})/2} \right)_i \quad (16)$$

where \overline{ROI} indicates the mean value and σ_{nodule} , σ_{noise} the standard deviation for the nodule contrast and noise contrast respectively [27]. The values of σ_{nodule} , σ_{noise} are calculated as $\sigma = \sqrt{var}$ where "var" indicates the variance of contrast for each image at the selected ROI.

The efficiency of the human observer was calculated according to expression

$$\eta = \left(\frac{d'}{SNR} \right)^2 \quad (17)$$

where d' is detectability index and SNR is the measured signal-to-noise ratio described above [16].

6 Results

In this chapter, the produced data from the experiment are analyzed and presented. It includes results from the pilot validation, evaluation of unenhanced and enhanced images and the mathematical analysis.

6.1 Pilot validation

The average value of four observers gave the recommended fraction of correct scores P in the pilot with two deviant values, $P=80\%$ for the 4mm nodule at 150mAs and $P=83\%$ for 10mm nodule at 32mAs. The four observers included two engineers and two medical physicists.

The results from the pilot validation, presented in Figure 24, are the basis for the selection of contrast for the nodule, in the evaluation of unenhanced and enhanced images.

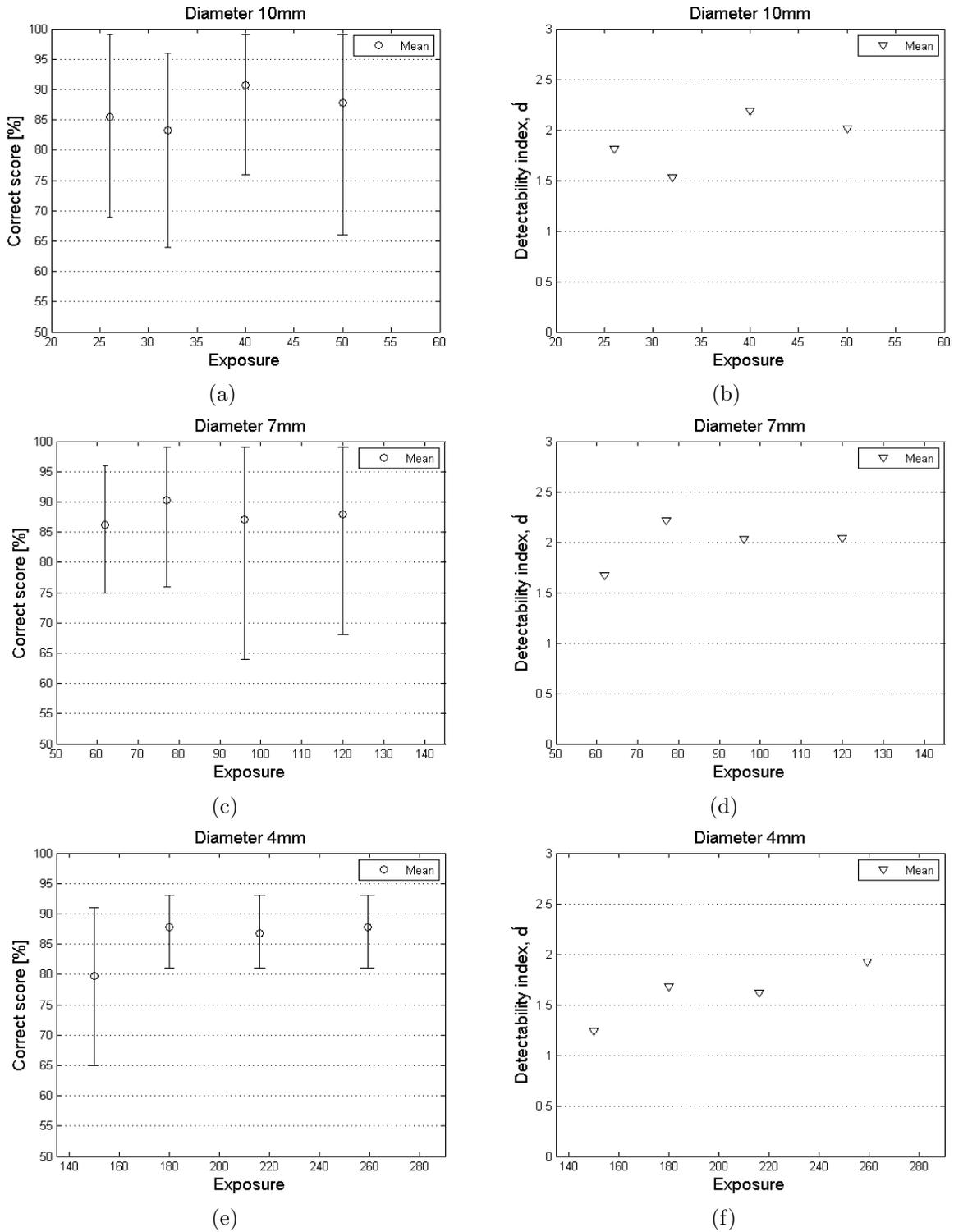


Figure 24: The results from the pilot validation in purpose to find a suitable contrast for the experiment where (a,b) mean, minimum and maximum calculations of correct scores P, and mean value calculations of d' for 10mm nodule, (c,d) 7mm nodule and (e,f) 4mm nodule.

6.2 Image analysis and measurements

The results from the pilot validation resulted in a peak contrast value of $C=0.008$ with the corresponding range between 0-8HU. This contrast were used for unenhanced and enhanced images, for all exposure levels.

When images are enhanced, the amplitude is changed with the exposure. To illustrate the effect of applying the software, noise removed plots has been used. Figure 25 show one-dimensional plots for images unenhanced and enhanced for the comparison of the changed amplitude. For this illustration, two images were used whereas one consisted of noise, and the other with the identical noise and the inserted nodule. The noise was then removed by subtraction from the image consisted of nodule and noise. This noise subtraction has been accomplished for the three different sizes, for the arbitrary selected exposure values 26mAs, 77mAs and 180mAs.

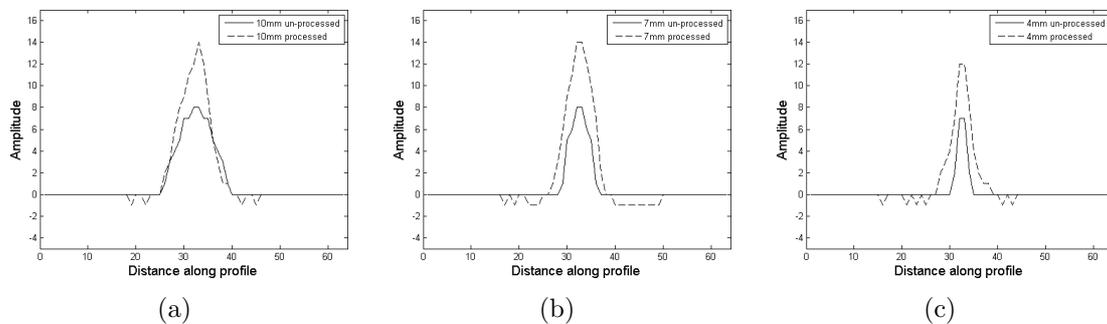


Figure 25: One-dimensional noise removed plots shows the change in amplitude before and after enhancing software is used where (a) 10mm diameter nodule for 26mAs, (b) 7mm for 77mAs and (c) 4mm for 180mAs. The new amplitude ranges from 0-12HU for 10mm, 0-14HU for 7mm and 0-14HU for 4mm.

From the 4mm and 7mm nodule it can be seen that the enhanced images have received "new" diameters. The earlier circular symmetric nodules has been "smeared" and the symmetry of the nodule does not exist. To illustrate this, two-dimensional and three-dimensional plots are presented in Figure 26 and Figure 27.

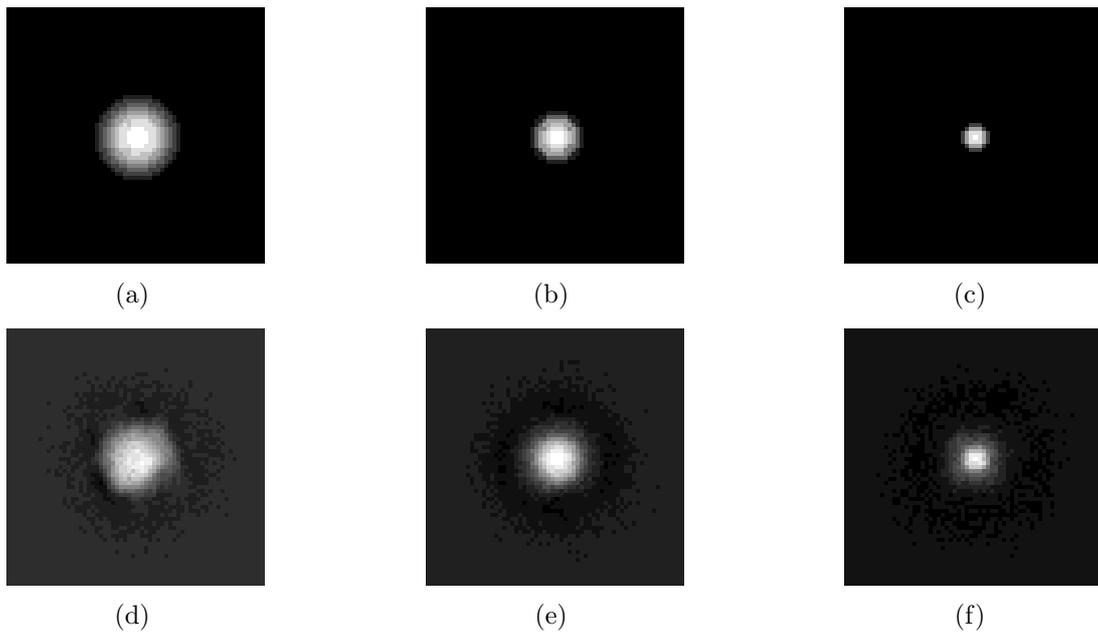


Figure 26: Two-dimensional noise removed plots for the selected contrast $C=0.008$ where (a,d) nodule size 10mm before and after image-enhancing software is applied, (b,e) 7mm and (c,f) 4mm.

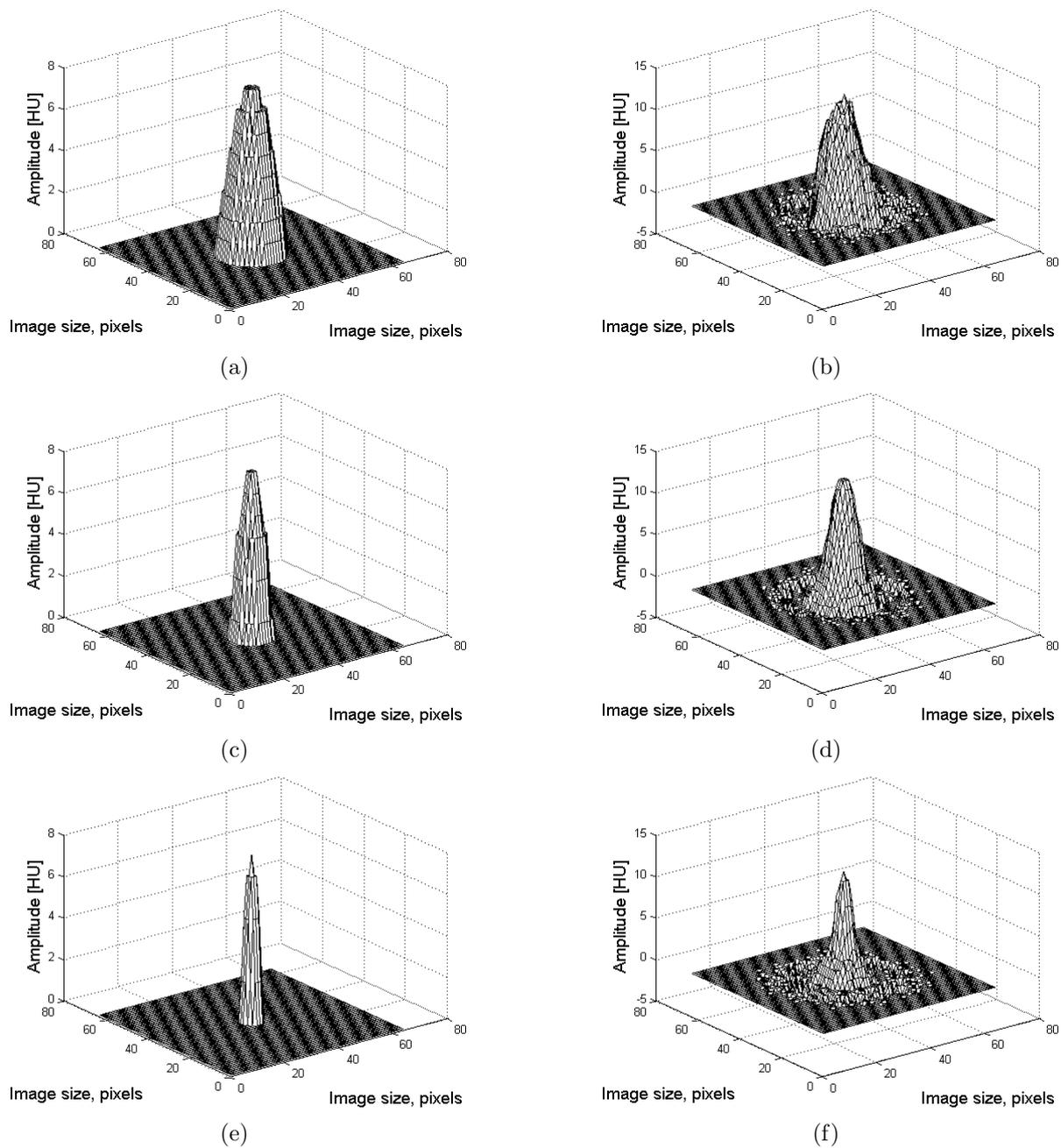


Figure 27: Three-dimensional noise removed plots to illustrate the changed amplitude and diameter for the comparison where (a,b) 10mm nodule for 26mAs, (c,d) 7mm nodule for 77mAs and (e,f) 4mm nodule for 180mAs.

6.2.1 Mathematical image analysis

The results from the estimated SNR measurements and human observer efficiency, η , is presented in Table 4 for unenhanced and enhanced images. The human observer efficiency is calculated as the average score for the participants in the evaluation. The gradual increase of exposure indicates an overall higher SNR and η for enhanced images.

Exposure	SNR unenhanced image	SNR enhanced image	$\bar{\eta}$ unenhanced image	$\bar{\eta}$ enhanced image
26	1.37	1.48	1.56	3.45
32	1.49	1.68	1.46	2.38
40	1.61	1.76	2.13	2.22
50	1.83	2.55	1.53	1.94
77	1.63	1.67	1.51	3.16
96	1.81	1.89	1.10	3.03
180	1.65	1.95	0.99	1.45

Table 4: Calculated SNR and average human observer efficiency $\bar{\eta}$ for unenhanced and enhanced images.

From Table 4 it can be seen that the SNR and average human observer efficiency depends on the size of the nodule and the exposure. For the 10mm diameter nodule, the relative SNR was increased with 8% (26mAs), 13% (32mAs), 9% (40mAs) and 39% (50mAs) for enhanced images. Why the SNR for 40mAs is increased by 9%, and thereby abrupt the ascending trend, can be related to the random noise fluctuations. From Table 2, a comparison between minimum values indicates a lower minimum value for 40mAs in comparison to 32mAs. The few existing data for the 7mm diameter nodule also indicates a relative increased SNR with 2% (77mAs) and 4% (96mAs) for enhanced images. Finally the 4mm diameter nodule shows an 18% higher SNR for enhanced images.

A comparison between the efficiency, calculated as the relative human observer efficiency, for unenhanced and enhanced images shows that for the 10mm nodule the relative human observer efficiency is increased by 2.21 (26mAs), 1.63 (32mAs), 1.04 (40mAs) and 1.27 (50mAs). For the 7mm and 4mm nodules the relative human observer efficiency is increased by 2.09 (77mAs), 2.75 (96mAs) and 1.46 (180mAs).

6.3 Evaluation of unenhanced and enhanced images

Seven observers were participating in both evaluation groups whereas two engineers, and five medical physicists. The results from the evaluation shows data from five of them, since two were excluded due to uninteresting reported results. One of the excluded observers scored 99% as average before the images was enhanced, and the other differed too much from the recommended range of correct scores with an average of 48%. The results for the 10mm nodule, for images before and after enhancing software is applied, is presented in Figure 28.

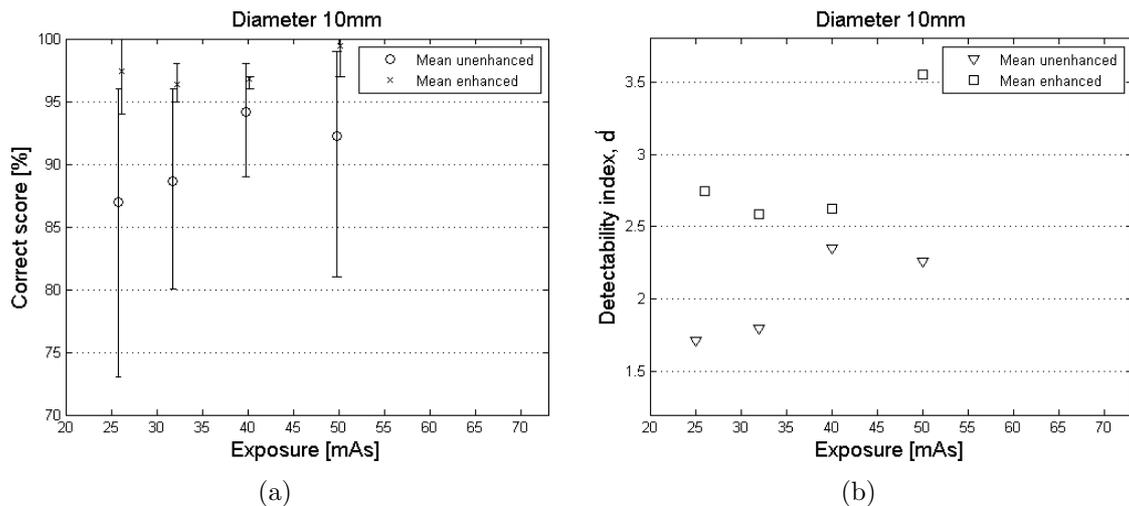


Figure 28: The results for the 10mm diameter nodule from examined images before and after enhancing software is applied where (a) minimum, maximum and mean values of correct scores for the five observers and (b) the detectability index.

For enhanced images, for all exposure levels, an overall higher scoring rate is reported. For the exposure 26mAs, the average of correct score is increased by 11%, 32mAs with 7%, 40mAs with 3% and 50mAs with 7%. However, noticeable is the mutual reduction for minimum and maximum values between all the five observers, for unenhanced and enhanced images. For the 10mm diameter nodule, a comparison of the range of minimum and maximum values, for all the exposure levels indicates a reduction. For the exposures 26mAs to 50mAs, the range between minimum and maximum values, where minimum and maximum values for all observers were included, was reduced with 74% for 26mAs, 81% for 32mAs, 89% for 40mAs and 83% for 50mAs.

The results for the 7mm nodule are presented in Figure 29. For this nodule size, the results are only compared for unenhanced and enhanced images for two exposures, 77mAs and 96mAs, due to the complication in designing parameter files for 62mAs and 120mAs.

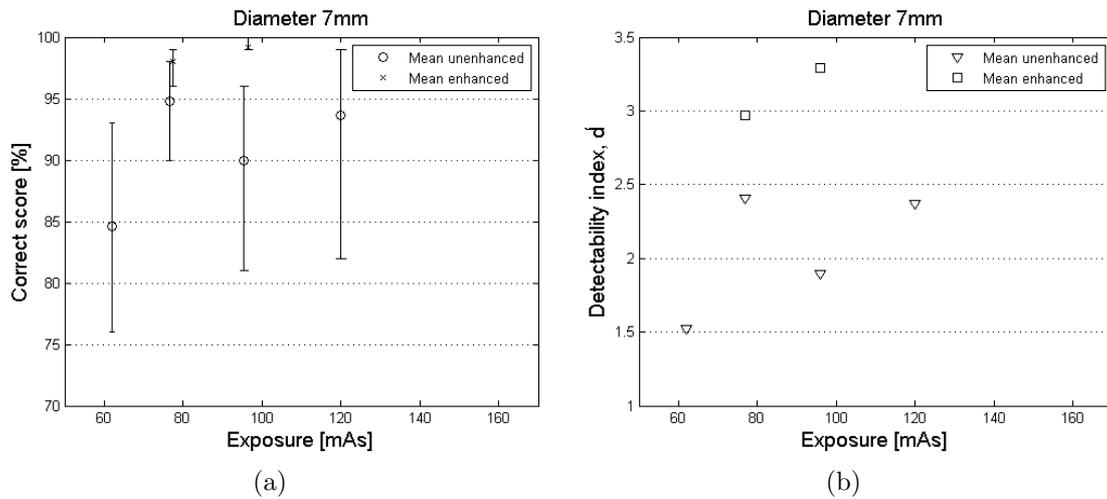


Figure 29: The results for the 7mm diameter nodule from examined images before and after enhancing software is applied where (a) minimum, maximum and mean values of correct scores for the five observers and (b) the detectability index. The comparison was only possible for 77mAs and 96mAs due to the missing parameter files which could not be designed by Sapheneia.

From the few data, it still can be seen that it follows the trend as for the larger 10mm diameter nodule. For the exposure 77mAs the average of correct scores was increased by 3% and 9% for 96mAs. As for the 10mm diameter nodule, the mutual length of minimum and maximum values are reduced with 63% for 77mAs, 73% for 96mAs for enhanced images.

For the 4mm nodule, only one data exist, i.e 180mAs due to some problems in designing parameter files for the exposure values 150mAs, 216mAs and 259mAs. It can be seen from Figure 30 that the average results, for the unenhanced images, follow an ascending trend with the exposure levels. This trend is not shown for the 10mm and 7mm nodule. However, the 4mm diameter nodule also indicate a smaller range of minimum and maximum values with a reduction of 65% and the average of correct scores are increased by 9.5%.

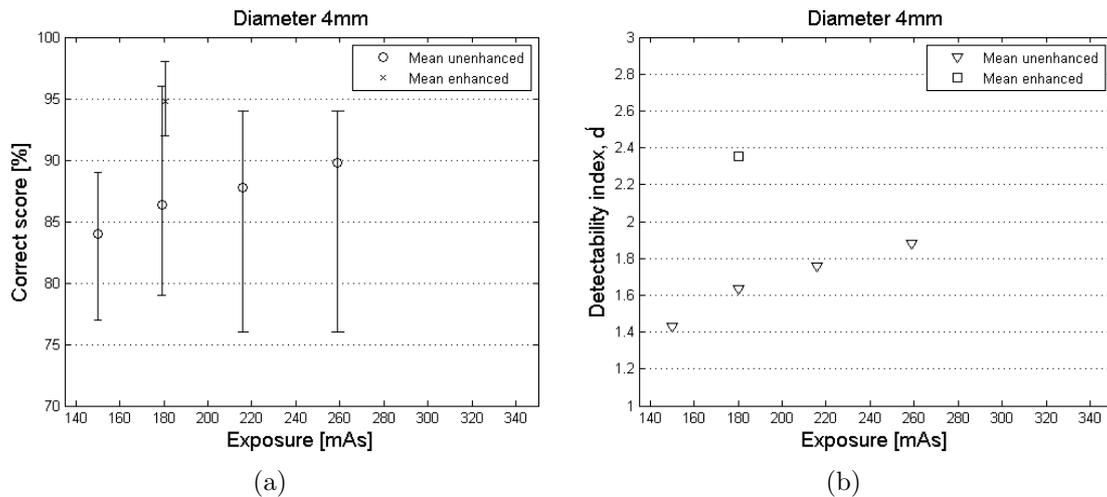


Figure 30: The results for the 4mm diameter nodule from examined images before and after enhancing software is applied where (a) minimum, maximum and mean values of correct scores for the five observers and (b) the detectability. Here, the comparison was only possible for 180mAs since images for 150mAs, 216mAs and 259mAs could not be enhanced due to the missing parameter files.

A statistical analysis was performed to assess the significance for the reported results. The analysis was performed for all the existing reported mean values, before and after the image was enhanced, i.e. 26, 32, 40, 50, 77, 96 and 180mAs. The test of significance was performed with a paired t-test for a one tail test since the distribution of differences is assumed to be a t distribution. The α -level for the paired t-test was chosen to $\alpha=0.05$. The result of this test indicates a statistical significance with $t_{observer} > t_{table}$ for all tested mean values. The reported p-values ($p < \alpha$) for the t-test confirms the separation of two tested mean values, for the cases unenhanced and enhanced images, with a probability of 95%. This means that the observers performed better when the images was enhanced in comparison to the unenhanced case.

7 Discussion

Since the implementation of the assignment for this thesis work were revised, this chapter is introduced with a short summation about the work of the thesis. Then, the genuineness of the selected parameters, results and implementation are discussed in terms of accuracy and validity, with parallels drawn to measurements from an anatomical liver.

7.1 The work of this thesis

At the beginning of this work, the task were to analyze the structures of the algorithms used in the image-enhancing software. The assignment consisted of analyzing if it was possible to increase the detectability for simple artificial detection tasks. Since the involved company considered the software as trade secret, the task were revised. Instead, they were willing to share a close-coded limited version of their software, which resulted in this evaluation work. The purpose of this thesis has never been changed, only the implementation of the assignment.

Since there were no prior experience in designing forced choice tasks at the hospital, much effort were disposed trying to adapt and design the 2AFC methodology along the path of this work, to fit this task. This were mainly done in consultation with supervisors at Södra Älvsborgs hospital and Sahlgrenska University hospital.

7.2 Validity of selected images

The basis for this thesis experiment was that the selection of images (cropping process) from the water phantom study, were assumed to occupy the same homogeneous property, for that of an anatomical liver. The liver was used since it is a relatively homogeneous area within the human body. To ensure this assumption a validity test were accomplished.

For the validity test, data were collected from one patient where the identification information was removed. No contrast agent had been used during the CT liver examination for this patient. The CT settings were 120kV at the exposure of 147mAs for kernel "bf31f".

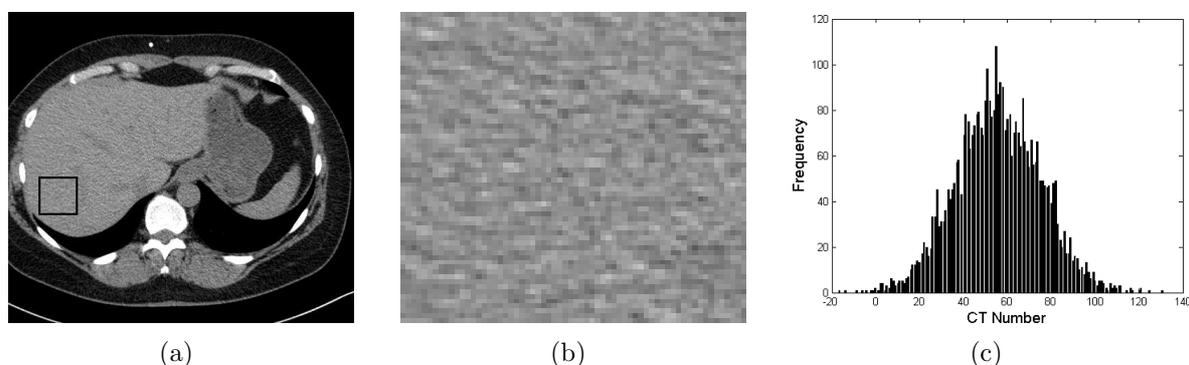


Figure 31: The process in capturing data from a liver where (a) 3mm thick liver slice which consists of a 64×64 arbitrary marked ROI, (b) the cropped 64×64 sub-image and (c) the histogram of CT numbers spread around the value of nearly 60.

One 64×64 sub-image were selected arbitrary from the liver and statistics was compared with the first four cropped sub-images from the water phantom 96mAs, 120mAs, 150mAs and 180mAs, see Table 5.

Type of examination/exposure	Min pixel intensity [HU]	Max pixel intensity [HU]	Mean pixel intensity $\pm \sigma$ [HU]
Water 96mAs	-22	28	0 ± 7
Water 120mAs	-22	21	0 ± 7
Liver 147mAs	-17	131	5 ± 19
Water 150mAs	-24	20	0 ± 6
Water 180mAs	-18	17	0 ± 5

Table 5: Statistics from water phantom images for the comparison with that of a real liver. The spread of CT numbers (min/max pixel intensity) and fluctuation (σ) are larger in comparison to the water phantom. A conclusion drawn from this statistics is that the water phantom are more homogeneous then the liver.

7.3 Nodule size and contrast

The size of the nodule was in the experiment chosen to 10mm, 7mm and 4mm in diameter. With a starting point of nodule size 10mm, the other two were selected arbitrary with an equal distance between each other. The selection of the 10mm nodule were raised from the chest radiography, were nodules smaller then 10mm are rarely reported and therefor resembles the minimum size for detection [22, 24]. However, this statement is not justified in CT coherence since smaller nodules are detectable. The study reported by [23], investigates the possibilities for the discrimination between real and simulated liver nodules/lesions for pediatric CT within the size of 2-6mm in diameter.

The contrast and the size of the 4mm and 7mm nodule were in this evaluation not selected with respect to the relationship that exists. Roughly, there should be a constant relationship between the contrast, area of object and SNR [1]. If the diameter is reduced it must be compensated either by increasing the contrast or the SNR (number of photons). To fit the constant relationship the arbitrary selected diameters 4mm and 7mm should be selected to 4.2mm and 6.5mm to maintain a constant relationship with respect to the contrast and size of the largest diameter.

Since this experiment only was aimed at detection tasks no further reflection about the contrast selection were done. According to [23] the reported 2-6mm liver lesions ranged from -20 to -41HU, in comparison to this thesis contrast selection of 0-8HU. An attempt to imitate the more darken intensity is illustrated in Figure 32.

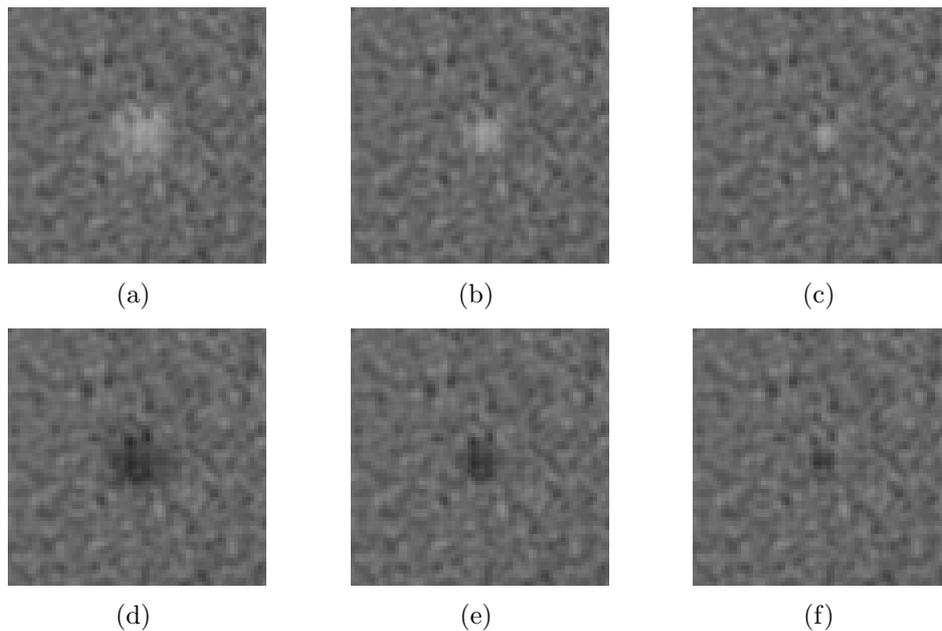


Figure 32: Simulated nodules for 26mAs where (a-c) 10-4mm nodule ranging from 0 to 58HU and (d-f) 10-4mm nodule ranging from -58 to 0HU to imitate that of a real liver lesion proposed by [23]. How the selection of the darken pixel intensity, that reflects real lesions in the liver would affect detection task scores, remains to be answered.

7.4 Implementation accuracy

The representation of a CT number are often (if not always) an integer. For the implementation of the nodule insertion in Matlab this resulted in CT numbers with decimals since the mathematical expression for the nodule took numerical values between 1-2. After the nodule insertion all values were rounded to nearest integer with a corresponding accuracy of ± 0.5 HU.

7.4.1 Insertion procedure

The nodule was inserted in one of the two noisy images by using Matlab's command **randi**. The accuracy shows that it does not produce the equal prior probability that were assumed. The outcome of 12 cases, for 96 repetitions in each case, indicates (almost) an equal prior probability with an overall average of 49 versus 47 with varying probability from case to case.

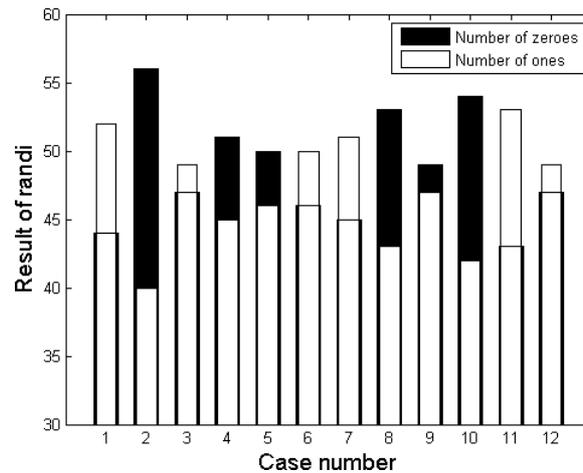


Figure 33: The result from testing the command **randi** shows an outcome of 12 cases where each case produces 96 repetitions. Case number one indicates that for all 96 images, for one exposure, the nodule was inserted in the left image 44 times and in the right image 52 times.

7.4.2 Evaluation of the 2AFC implementation

It should be mentioned that the participants of the evaluation (including the pilot validation) did not perform the examination of images under the same environmental conditions. The pilot validation was accomplished at three different locations, with three different monitors. The ambient light was not measured during the pilot and therefore lacks accuracy in comparison to the evaluation of unenhanced and enhanced images. For the evaluation of unenhanced and enhanced images, one observer did not perform the examination of images under the same environmental conditions as the other four participants. Another location and monitor were used.

Except for the two preset window and level settings, the observer was able to alter window and level settings during the examination of images. A more proper and controlled situation would have been to only use the two preset window settings for a more realistic and clinical situation, since radiologists rarely use different types of window settings [28]. The purpose of including an extensive use of window and level settings was that the image contrast would not be a limiting factor for the detection task. Instead, the detection would always be limited by the detail contrast.

8 Conclusion

The conclusions drawn from this work are summarized in this chapter. Recommendations for further possible investigation is presented.

8.1 Summation

This thesis investigates the influences of applying an image-enhancing software for computed tomography images. To develop different noisy environments, a water phantom is used and exposed with varying number of quanta. Creation and insertion of synthetic low-contrast objects serve as stimuli for detection tasks. To assess observer performance, the methodology of the two-alternative forced choice was used.

For all the diameters of the nodule, it was concluded that the overall average score was increased. Also, it was concluded that a remarkable higher and uniform discrimination amongst the observers was received. For the 10mm diameter nodule, it was concluded that a dose reduction of 11% for 26mAs, 7% for 32mAs, 3% for 40mAs and 7% for 50mAs, calculated as the average, was possible. By judging from the few existing data for the 7mm and 4mm diameter nodules, the results also indicates a higher average score with possible dose reduction of 3% for 77mAs, 9% for 96mAs and 9.5% for 180mAs. However, it should be mentioned that some of the observers scored 100%, for the evaluation of enhanced images, and further dose reductions may be possible.

8.2 Recommendations and future work

To complete the results for enhanced images, it is recommended to get in contact with Sapheneia for the production of parameter files for nodule sizes 7mm and 4mm, i.e. 62mAs, 120mAs, 150mAs, 216mAs and 259mAs.

High scoring rate, for unenhanced images indicates a possible reduction of contrast for this detection task experiment. It is therefore suggested that the test is implemented again, with the peak mask contrast adjusted to $C=0.006$ or $C=0.007$ for further investigation of possible and potential dose reductions.

Since this experiment includes synthetic nodules and unrealistic background images, it is very difficult to draw parallels with real tissues. Disregarding the nodule, the validation test indicated that the liver was more inhomogeneous than the water phantom used in this experiment. Further investigations could therefore include livers from patients with inserted synthetic nodules.

References

- [1] A.B. Wolbarst, *Physics Of Radiology 2:nd edition, 2000, Medical Physics Publishing Corporation*
- [2] R.Neubeck, The unity of radiology, *Kompendium i Röntgenteknologi 1896-2006*, Internet: <http://www.radiol.uu.se/utbildning/rtgsjuksk/Rtgteknologi/Rtg-teknologi-hela.pdf>
- [3] National Encyklopedia, National Encyklopedia internet service, Internet: <http://www.ne.se>, [Available at Internet 2009-09-26]
- [4] U. Nyman, W. Leitz, M. Kristiansson and P.Å. Pålstorp, *Stråldosreglering vid kroppsdatortomografi - bakgrund till dosregleringsprogrammet OmnimAs*, Internet: <http://www.stralsakerhetsmyndigheten.se/Ovrigt/Sok1/?q=2004:12>
- [5] C.A. Carlsson, G.A. Carlsson, Department of radiation physics, Linköping, *Basic physics of X-ray imaging*, Internet: <http://www.imv.liu.se/radiofysik/pdfs/rep8.pdf>
- [6] G. Dougherty, *Digital Image Processing for Medical Applications 1:st edition, 2009, Cambridge University Press*
- [7] A.C. Kak, M. Slaney, *Principles of computerized tomography imaging, 1999*, Internet: <http://www.slaney.org/pct/>
- [8] M. Sandborg, *Computed tomography: Physical principles and biohazards, 1995*, Internet: <http://www.imv.liu.se/radiofysik/pdfs/Rap81.pdf>
- [9] L.G. Månsson, *Methods for the evaluation of image quality: a review, 2000*, Oxford Journals
- [10] T. Godås, *Riktlinjer för utformning av strålskyddsprogram för transportörer av radioaktiva ämnen*, Internet: <http://www.stralsakerhetsmyndigheten.se/>, Strålsäkerhetsmyndigheten, Rapport 2004:10
- [11] R.C. Gonzales, R.E. Woods, *Digital Image Processing 2:nd edition, 2002, Prentice Hall*
- [12] M.Båth, *Imaging properties of digital radiographic systems: development, application and assessment of evaluation methods based on linear-systems theory, 2003*, Doctoral Thesis, Göteborg University Sweden, 2003
- [13] N.A. Macmillan, C.D. Creelman, *Detection theory: A user's guide, 2:nd edition, 2005*
- [14] L.G. Månsson, *Evaluation of radiographic procedures: Investigations related to chest imaging, 1994*, Doctoral Thesis, Göteborg University Sweden, 1994
- [15] A.E. Burgess, *The Rose model, revisited, 1998*, Optics Infobase, vol. 16, March 1999

-
- [16] A.E. Burgess, *Comparison of receiver operating characteristic and forced choice observer performance measurement methods*, 1995, Med. Phys, May 22, 1995
- [17] J.A. Hanley, B.J. McNeil, *The meaning and use of the area under a receiver operating characteristic (ROC) curve*, 1982, Radiology, vol. 143, April 1982
- [18] Sapheneia, *Clarity Tuning Tool, version 2.73*, Limited software version for enhancing images, Borrowed from Sapheneia
- [19] Matlab, *Matlab & Simulink student version R2009b*, Downloaded at internet: <http://www.etek.chalmers.se/>
- [20] M. Håkansson, M. Båth, S. Svensson, S. Zachrisson, A. Svalkvist, L.G. Månsson *viewDEX, viewDEX version 2.0*, Java-based DICOM-compatible software for observer performance, Borrowed from Södra Älvsborgs hospital
- [21] L.Karlsson, Siemens, *Mail, 13 Oct 2009*
- [22] A.E. Burgess, X. Li, C.K. Abbey, *Nodule detection in two components noise: toward patient structure*, 1997, SPIE, vol 3036
- [23] C.L. Hoe, E. Samei, D.P. Frush, D.M. Delong, *Simulation of liver lesions for pediatric CT*, 2005, Radiology, vol. 238, February 2006
- [24] M. Båth, M. Håkansson, S. Börjesson, S. Kheddache, A. Grahn, M. Ruschin, A. Tingberg, S. Mattsson L.G. Månsson, *Nodule detection in digital chest radiography: introduction to the radius chest trail*, 2005, Radiation Protection Dosimetry, vol 114, 2005
- [25] W. Huda, K.M. Ogden, E.M. Scalzetti, D.R. Dance, E.A. Bertrand, *How do lesion size and random noise affect detection performance in digital mammography*, 2006, Science Direct
- [26] Digital Imaging and Communications in Medicine (DICOM), *DICOM Part 14: Grayscale Standard Display Function*, Standard monitor calibration protocol, Internet: ftp://medical.nema.org/medical/dicom/2007/07_14pu.pdf
- [27] D.J. Tward, J.H. Siewerdsen, M.J. Daly, S. Richard, D.J. Moseley, D.A. Jaffray, N.S. Paul, *Soft-tissue detectability in cone-beam CT: Evaluation by 2AFC tests in relation to physical performance metrics*, 2007, Med. Phys, November 2007
- [28] M.Håkansson, Södra Älvsborgs sjukhus, *Personal communication, 17 Dec 2009*