Alma Mater Studiorum · University of Bologna

**School of Science**
**Department of Physics and Astronomy**
**Master Degree in Physics**

# Angiographic systems performance evaluation through a Channelized Hotelling Observer model

**Supervisor:**
**Prof. Nico Lanconelli**

**Co-supervisor:**
**Dr. Marco Bertolini**

**Submitted by:**
**Elena Cantoni**

Academic Year 2020/2021

**Abstract**

Medical image quality assessment is a challenging subject that depends on measurable and objective physical properties and long time consuming conditions related to the external perception and human subjectivity. To overcome this issue and to include the human observer in the images validation, observer models are typically used. In this thesis work the Channelized Hotelling Observer (CHO) model is used to test three human observer's ability in terms of performance in signal detection task on large set of fluoroscopic and angiographic images. Three different systems are used to acquire the images: Siemens Artis Q, Philips Azurion 7 C20 and GE Discovery IGS 740. The last two are angiographic machines available in Reggio Emilia's hospital used to acquire part of the images of the project. The remaining ones come from an EFOMP's pre-existing dataset. So, further series of acquisitions are performed on two distinct test objects: TO12 and TO16, that are phantoms appositely realized for contrast-detail (CD) analysis. From CHO application, an $\alpha$ internal noise parameter is estimated and it is used to graph CD curves that are robust tools for evaluating displayed image quality. After a statistical validation of the chosen CD curves, the optimal $\alpha$ are given: in Siemens with TO12, $\alpha = 3$ for fluoroscopy and $\alpha = 1$ for angiography are chosen, in Philips and GE with TO16, phantom limits are achieved in fluoroscopy, in angiography Philips selects $\alpha = 20$ and GE $\alpha = 10$. Comparisons between acquisitions using their optimal CHO CD curves are made and consistent considerations about their statistical significance are done. Indeed, a figure of merit, depending on the contribution of the Kerma Area Product and image quality given by the CD curve, is calculated. The most performant acquisition on Siemens is that with low dose, FOV = 22 cm , SID = 110 cm and pre-filter of 0.6 mm. The best one in angiography between Philips and GE is the low dose, FOV = 32 cm, SID = 119 cm GE acquisition.

## Sommario

La valutazione della qualità di immagini mediche è un argomento dibattuto ed intrigante che non dipende solo da proprietà fisiche misurabili ed oggettive, ma anche da condizioni esterne legate alla percezione e soggettività umana. Includere l'osservatore umano nel processo di validazione delle immagini richiede molto più dispendio di tempo e per ovviare a ciò vengono introdotti modelli ad osservatore. In questo lavoro di tesi viene utilizzato il modello ad Osservatore Hotelling Canalizzato (CHO) per testare l'abilità di tre umani nel riconoscere il segnale su campioni di immagini fluoroscopiche ed angiografiche. Tre diversi sistemi sono stati utilizzati: Siemens Artis Q, Philips Azurion 7 C20 and GE Discovery IGS 740. Le ultime due sono macchine angiografiche disponibili presso l'ospedale di Reggio Emilia da cui parte delle immagini del progetto sono state acquisite. Le restanti provengono da un preesistente campione di immagini di EFOMP. Perciò, diverse serie di acquisizioni sono state ottenute servendosi di due fantocci: TO12 e TO16, entrambi realizzati appositamente per condurre analisi sul contrasto-dettaglio (CD). Applicando il CHO, un paramentro di rumore interno $\alpha$ viene stimato ed utilizzato per graficare le curve CD che sono validi strumenti per valutare la qualità delle immagini. A seguito di una validazione statistica delle curve CD, l' $\alpha$ ottimale viene fornito: per Siemens con TO12, $\alpha = 3$ per la fluoroscopia ed $\alpha = 1$ per l'angiografia, per Philips e GE con TO16, la fluoroscopia è alterata dai limiti del fantoccio, l'angiografia seleziona per Philips un $\alpha = 20$ e per GE un $\alpha = 10$. Vengono fatti confronti tra le curve CD ottimali del CHO con l'obiettivo di avanzare considerazioni sulla loro significatività statistica. Dunque, viene calcolata una figura di merito dipendente dal prodotto del kerma per l'area (KAP) e dall'area sottostante la curva CD ottimale. L'acquisizione più performante per Siemens è quella definita da un basso livello di dose, FOV = 22 cm, SID = 110 cm ed un prefiltro di 0.6 mm. Quella migliore tra le angiografie di Philips e GE è la serie GE caratterizzata da bassa dose, FOV = 32 cm e SID = 119 cm.

# Contents

# Introduction

Medical image quality assessment is a challenging subject that must not only include the investigation of measurable physical properties of medical images such as contrast, noise and resolution. Understanding medical images and making a more comprehensive judgment of their quality also requires knowing how to answer such a question: how well can the desired information for a given task of interest be extracted from the images? The image quality process needs absolutely to consider the external contribution given by the humans when they observe the digital image. One way of assessing a human observer performance on the specific task of identifying the presence of a signal in an image is through contrast detail (CD) curves. This process is a strong method for evaluating displayed image quality. The concept behind this technique is psychophysical, i.e. includes the involvement of the human observer subjectivity in image quality chain.
Performing such a task that goes beyond objective features and includes interference due to human performance in a visual task can be very time consuming and to overcome this inconvenience observer models have been implemented. They attempt to classify digital images by trying to predict human visual detection. In this project a specific model observer is used: the Channelized Hotelling Observer model. This type of model consists of channels (in this case are 40 Gabor channels) that are tuned to spatial frequencies, orientations and phases with the aim to contain the useful selected information abut the images. The model in use is linear and it returns a response that contemplate the presence of an additive element recognized as the internal noise attributed to the human vision perception. The deep meaning of this variable is to represent the fluctuation of the human decision maker who is not able to make the same choice on the same set of images in repeated attempts, especially in low contrast area and for small details. It happens because the observer's internal psychological conditions and external environmental interferences come into play and may deviate him from making the same choice as in the previous trial.
The model was therefore applied to the available datesets and a study of the performance of the systems with regard to imaging detection was carried out. A figure of merit, capable of quantitatively describing when a system is able to perform a given task, is estimated. This will take into account the ability to detect details at the lowest possible contrast and consider the amount of dose delivered to the phantom.
This thesis is organized as follows: a theoretical introduction about the fluoroscopic imaging and concerning the meaning and the utility of a model observer, going in the detail for Channelized Hotelling Observer (CHO) model, are given. Consequently,

a brief description about test objects used and geometrical setups performed are supplied. Thus, the analysis follows a logical and chronological flow that includes at first the estimation of contrast detail curves for both human vision and model estimation for all acquisitions on each machine and for each test object. Then a physical interpretation is given on their trend studying also the limits and the conditions of applicability of the model. Then, the parameter $\alpha$ will be identified for which the curve estimated by the CHO model is as close as possible to the human CD curve. Finally, the discussion will continue with a statistical significance analysis between the optimal CD curves estimated by the model with the aim to derive consistent results and to be able to make considerations on the set parameters, showing the possible limits encountered during the process and drawing conclusions on the most performing angiographic and fluoroscopic acquisitions.

# Chapter 1

# Medical Imaging in pills

The Medical Imaging history began at the end of the 19th century, exactly in 1895 with Wilhelm Conrad Roentgen's discovery of the X-ray. This finding was expected to be revolutionary because it made looking inside the human body to possible detect and characterize diseases. Over the course of the years, the fields of application were deepened; in the last half of the twentieth century a great acceleration in advancements was recorded , motivated by the new needs of healthy care providers but, mainly, this revolution is referred to advances in the understanding of physical phenomena on which the imaging technologies are based on and to the growing availability of computational power and high-speed data communications. The development of the vast field of medical imaging allowed to perform different diagnostic techniques based on various physical characteristics. Radiography was the first medical imaging technology that used the power of an X-rays source directed to the patient and detected on the other side by a detector.
In order to generate the majority of medical images, it is fundamental that the supplied energy can penetrate the tissues interacting with them. If the energy is capable of passing through the body without a trace, the detection is void of useful information regarding the anatomy and the image construction is inefficient. Thanks to the beam attenuation, caused by different scattering or absorption within the tissues, image formation happens. The attenuation properties change according to the hit material, and it is different if it is considered a bone or a soft tissue. This distinction is the key to the beam's heterogeneous distribution that emerges from the patient [7].
The power of the electromagnetic spectrum is exploited in medical imaging. The wide range of radiations permitted a large number of innovative techniques moving from the non-ionizing radiations with the radiofrequency in magnetic resonance imaging (MRI) to ionizing radiations like X-rays, used in mammography, fluoroscopy and computed Tomography (CT), and gamma rays flaunted in Nuclear Medicine.
From a diagnostic point of view, when a medical image is acquired, it is essential to take into account a balanced compromise between the technical quality of the image and the condition of its acquisition. So, further considerations must be made: for instance, a high level of radiation dose gives, indeed, a better image, but the safety of the "environment" is compromised. Of course, patient safety and

parameters that influence the acquisition must be considered when having medical images; therefore, excessive patient dose in the pursuit of an optimal image is not acceptable.

## 1.1 Fluoroscopic Imaging

Fluoroscopic imaging differentiates itself from conventional RX imaging because it allows real-time X-rays dynamic images with high frame rates and a low dose per image. From the scintillator screens of the late of 1950s to the delivery of flat panel detectors of now, the performance of fluoroscopy enhanced exponentially. A more compact size, a larger field of view (FOV) with a high spatial resolution, a larger dynamic range and a greater stability are some of the many applied improvements. This technology turns out to be decisive for both diagnostic and therapeutic medical procedures [7].

### Fluoroscopic machine chain

The generic structure of a fluoroscopic machine (as shown in Figure 2.2) is constituted by some key components which are essential in system functionality.

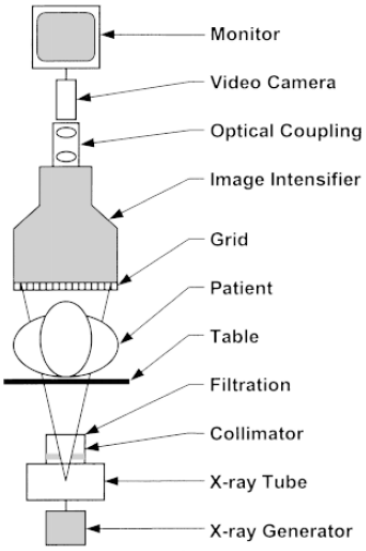



Figure 1.1: Fluoroscopic components' chain [20].

#### X-rays tube

The X-rays tube, for its name, converts the electrical energy provided by the generator into an X-rays beam. The process consists of the generation of electrons from a heated filament, accelerated toward a tungsten layer positively charged. The interaction between electrons and the anode results in the emission of X-rays. The tungsten target area, hitten by the particles, is strictly referred to the focal spot. The smaller it is, the more it is preferred because the geometric unsharpness is

minimized. In general the proper characteristics of this item are selected depending on the specific clinical usage. Interventional procedures are an instance, where an X-ray tube with a large heat capacity is necessary, so that heat dissipation is more easily improved, due to the high image recording activity.

#### Collimator

The outgoing X-rays beam must be controlled, especially in terms of directionality. The part which makes this kind of work is the collimator. In fluoroscopy it is usually motorized giving the chance to change the size in response to adjustment of the SID ( source to image distance) or FOV dimension. This component limits the beam to a dimension no larger than the field of view. In addiction, it reduces the tissue volume exposed to the radiation, decreasing, consequently, the scattering effect and improving the image contrast.

#### Filtration

Most fluoroscopy systems arrange filters to provide beam shaping and to attenuate low energy X-rays from the beam. Two different filters are described: the equalization filters are valid in glare reduction and light exposure equalization to the video camera. They are made by lead-rubber or lead-acrylic sheets and their main effect is appreciable in ABC ( Automatic Brightness Control) improvements. The other filters' family concerns with filtration materials that reduce the amount of low energy X-rays that, having such low energy, cannot pass through the body and are totally absorbed by it, aggravating the impact of the incident radiation. The most common added filtration material in fluoroscopy is copper, it is normally used for high dose procedures such as angiography and interventional applications, it allows to obtain high image contrast operating with lower kV values and dose levels [20].

The photon beam, at this point, after crossing the patient bed and the pad, interacts with the matter. The support where the patient is lying is studied to minimize the X-ray attenuation and, so, to reduce the radiation exposure. A fluoroscopic acquisition, with a duration of about 10 minutes and 30 fps (frames per second) is capable of producing thousands images using one thousandth of the X-rays dose provided in radiography. It is precisely for this reason that detectors with high sensitivity are needed to pick up low dose, real time and longer duration insights.

### 1.1.1 Image Intensifier, Based Fluoroscopy System

The X-rays Image intensifier is a very powerful and sensitive detector that covers two main functions:

- to intercept the X-rays beam and to convert it into visible light photons.
- to amplify the outgoing signal with a high gain in luminance.

The beam crossing patient's body could pass through the anti scatter grid and then interact with the intensifier. The presence of a grid reduces the scattering effect but increases a little bit the radiation exposure. In fact, when it is allowed, it is preferred the removal of the grid because patient exposure can be reduced by about 50% [20].
The Figure below shows the Image Intensifier structure highlighting three main components that are contained in a vacuum housing which keeps air out and allows unimpeded electron flow.

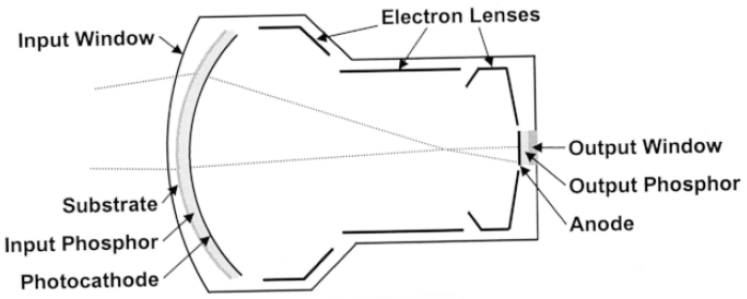



Figure 1.2: Image Intensifier scheme [20].

Firstly, the input screen is encountered and it consists of four parts. The outermost layer is the vacuum window made by a thin Aluminium sheet, following, there is a support layer slim enough to allow most X-rays to pass through it and strong enough to sustain the other two parts, the input phosphor and the photocathode. The input phosphor absorbs the X-rays and convert their energy into visible light. A thicker layer means an high X-rays absorption efficiency, but at the expense of spatial resolution, so a fair compromise should be sought (thickness about 300-450 $\mu$m) [22]. The main material used is the cesium iodine ($CsI$) which is shaped in long needles crystals that become light pipes, facilitating the light travel from the support layer to the photocathode. The fourth and last layer is usually made of antimony-cesium ($SbCs_3$), when the visible light strikes it, a certain amount of electrons are generated. Approximately, for a single 60 kV X-ray photon absorbed in the input phosphor, the photoelectrons leaving the photocathode are 200.
The photoelectrons are accelerated in the same direction by an electric field induced by the high voltage between the cathode and anode. Furthermore, in the electron optic chain they are focused thanks to the presence of an electron lenses' series. As the electrons bounce from one electrode to the next, they rapidly increase their kinetic energy due to the large voltage difference. Travelling in the Image Intensifier cavity, they reach the output phosphor, a thin (4 to 8 $\mu$m) layer of zinc cadmium sulfide doped with silver. In this phase, each electron causes its conversion in 1000 light photons out of the output phosphor.
Because of the different surface's shapes of the components (curved for the input phosphor and planar for the output one) the final image is subjected to a pincushion distortion.
The output phosphor resulting image is much smaller than the initial one because of the focus. To preserve the resolution a fine grain phosphor is used. At the same time, a reduction in size leads to a light intensity amplification.

#### Optical Coupling

The optical coupling system is the nexus that conveys the light generated in the Image Intensifier to the video camera and other image recording devices.
Its structure provides a beam-splitting silvered mirror, which directs a part of the light beam to an external device for image registration and the remainder to a video camera. According to the diameter dimension of the circular aperture of the lens, the amount of light supplied differs. Adjustment of the lens aperture on Image Intensifier-based fluoroscopy systems markedly changes the system's gain, which has an important effect on its performance [7].
When the size is small, much of the light is blocked, and the ABC system achieves a higher dose level to give a sufficient light to the camera, in this case, a low noise fluoroscopic image is created. Enlarging the hole, the obtaining image will be affected by a higher noise, with a benefit in terms of exposure.

#### Video Camera and Recording device

The television system provides a video camera that converts the light from the Image Intensifier to a voltage signal and a monitor that forms the output image.
The first component encountered in the image formation process is a photoconductor TV target whose behaviour is described by a high electrical resistance in the dark, which is reduced by varying the light intensity. Increasing the light, more electrons are produced and the video signal is higher, that is defined as a voltage versus time waveform. Some rhythmic pulses are emitted in order to synchronize the raster scanning in video camera and monitor. The electron beam, striking the monitor phosphor,generates a light flux. Finally, the raster scan on the monitor is done in synchronous with the scan of the TV target.

### 1.1.2 Flat Panel Detector, Based Digital Fluoroscopy System

The Image Intensifier (II) technology and its associated accessories made the history until the advent of FPD, Flat Panel Detector. The improvements are mainly related to its better stability, lower patient radiation doses and wider dynamic ranges. From a geometric point of view, the FPD consists of Thin Film Transistor (TFT) arrays which comprise many detector elements arranged in rows and columns packaged to form a square or rectangular area. The X-rays conversion happens, on the light of facts, according to an indirect or direct detection.
In the first case, observing Figure 1.3, the absorbed beam is converted to light in the adjacent layered phosphor and the photodiode, an electrical component of the detector element (DEL), converts, again, that light into a corresponding charge.
The direct method involves a semiconductor layer, typically selenium, which directly converts X-rays in charge. The flat panel image receptor replaces II, optics, camera systems in a much lighter and smaller package and it does not require an external TV camera, thus avoiding each possible source of noise due to the electronic components.
It has a light sensitive area and an insensitive region where the electrical devices are located.This defection could be compensated using interpolation, however the image is subjected by some artefacts. Another limitation of this system is about the

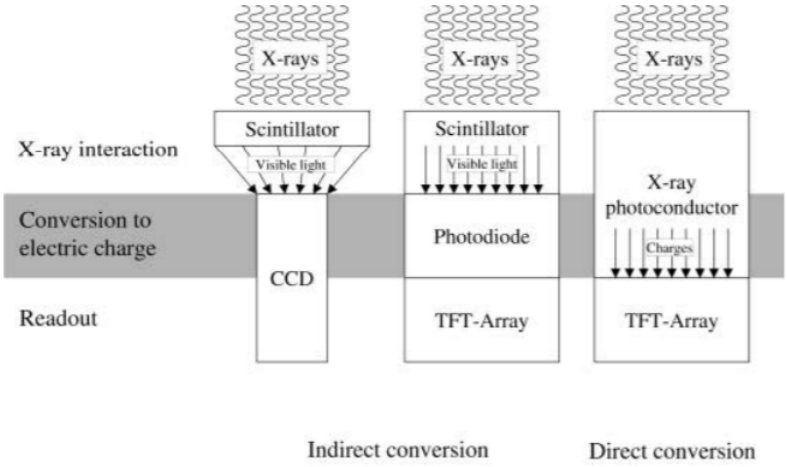



Figure 1.3: FPD scheme: indirect and direct methods [11].

spatial resolution because it is strictly related to the size of the detector elements (DELs) and to the binning effects. DELs allow an improved decrease of data rate, but cause disadvantages in terms of spatial resolution because of the larger surface area (FOV).
The FPD fluoroscopy systems do not need to increase the radiation supplied to the patient each time the FOV is reduced, and this is a great result, mainly for radioprotection purposes. Nevertheless, with smaller FOV, magnification of the surface area could make the image noise more apparent to the eyes of the observer, and, in this case, a minimal increase of radiation is used to reduce the optical perception of noise [17].
The equipment used in this project are based on systems using FPD.

## 1.2 Fluoroscopic and cineangiographic procedures

The modern angiographic equipment allows to execute complex procedures progressively optimized. Referring to the procedures used in this project, it is first of all necessary to underline how important it is to take into account the technical parameters of a specific acquisition, i.e. those terms that quantify the dose delivered to the patient and affect the quality of the digital image obtained.
The common fluorographic image is a single recorded image acquired with a flat panel as image receptor, the cineangiographic notion means a series of fluorographic images. Indeed, the fluoroscopy is a real-time imaging technique generally made with a continuous X-ray beam imparted on the region of interest. This procedure is used to observe moving biological parts for a relatively long period of time, from seconds to a handful of minutes [16].
In the course of years, several innovations were performed, the main scope is to optimize the performance of the equipment, keeping the doses as low as reasonably achievable and providing the patient with the lowest amount of radiation consistent with the image quality requirements of the clinical procedure. Initially, the fluoroscopic technique was made without the intent of preserving the images, but only to

reduce the dose irradiated while decreasing the time of fluoroscopy. A *last-image-mode* feature becomes very useful because the last acquired image or set of images are displayed on the monitor and can be examined as long as necessary using no additional radiation and it is recommended when the information captured about the anatomical part of interest is sufficiently accurate, avoiding the need to acquire fluorographic images which enlarge the amount of dose provided. In addition, modern fluoroscopes emit X-rays as a sequence of short pulses providing substantial dose saving and offering better image quality at the same average dose rates [14].
The variables of acquisition, like the number of images in each series and frame rates (from 0,5 to 30 $frames/s$), depend on the type of clinical questions.
An overall consideration about the dose level of the two previously described techniques can be done in terms of order of magnitude: the amount of dose required to acquire a fluorographic image is about 10 times grater than in a fluoroscopic acquisition modality.

## 1.3 Image quality in-depth fluoroscopic study

The concept of image quality moves around the nature of its subjectivity. What we want to reach is not a fine image observed but rather how well it expresses functional and anatomical helpful information to perform a deep clinical use. The more performant characterizing parameters that describe the image quality should measure the effectiveness of a medical image for a well intended purpose.
The right compromise to obtain a proper medical image is found considering the spatial resolution, the temporal resolution and the contrast resolution of the image without avoiding the effects caused by the imparted dose.
The spatial resolution outlines the detailed specificity which can be observed on an image. That is, it describes how small an object can be seen on an imaging system. Foremost, the pixel size plays an essential role because it is the first visual variable that could set a limit on what the image shows. An image's detail cannot be resolved if its size is smaller than the pixel size, so it is true that while an image with small pixels has the potential to offer a high local spatial resolution, it is also true that many limiting factors affect and pitch in, significantly, the overall spatial resolution.
Concerning the fluoroscopy field of application, the temporal resolution is optimal, mainly because of the high skill to provide real-time images. The eventual artefact which can be verified is lag, and it implies an overlap of the previous image data with the successive frame causing blurring. Whereas sometimes it can be beneficial, in the case of unrecording fluoroscopy process, it becomes a great issue when dynamic sequences are stored, like Digital Subtraction Angiography (DSA).
At this point, a few typical variables to keep in mind during the acquisition of an image, not neglecting its quality and within the dose saving condition, are the ones detailed as follows.

### Field of View

The Field Of View (FOV) is the dimension of the same anatomic region included in a scan. It is expressed in millimetres or centimetres and, usually, it is set accordingly to the anatomical region of interest from 20cm to 40 cm of side.

### Source-to-Image Distance

The Source-to-Image Distance (SID) is a geometric measure taken directly in the room, and it indicates the distance between the source of the x-rays (usually the focal spot of the x-ray tube) and the image detector. It affects the relative intensity of the radiation that reaches the image detector and affects the geometric properties of the image, changing its overall quality. In practice, the SID does not exceed 120 cm.

### Frame rate

The frame rate (fps) expresses the number of frames per second, and this parameter is the metre of judge per excellence of temporal resolution. A high frame rate provides a high temporal resolution, resulting in a significant dose increase for the patient and personnel. This factor can move from 30 fps to a couple of fps.

The image quality review is characterized by many other physical parameters like voltage and current, but that is not all: the contrast resolution influences the quality of an image from a more subjective perspective. It refers to the ability to discover slight variations in greyscale and discern them from the noise in the image [7].
Assuming that all imaging modalities are based on the idea that there are differences in physical properties between distinct tissue types, the contrast resolution concept refers precisely to the minimum contrast that can be used to observe the differences between detail and its surroundings. It is strictly dependent on the noise and spatial resolution properties of the image.
In fluoroscopy, it is a measure of image noise by viewing contrast detail phantoms under specific modality conditions. It is necessary to find the right compromise between the features which characterize a proper medical image.
Using a high exposure rate, indeed, the contrast resolution increases, allowing a more accurate detail detectability, but the dose level raises drastically; we can apply the same concept to spatial resolution. A smaller FOV ensures a higher magnification of the anatomical area of interest by favouring a higher spatial resolution. Generally, it is always better to minimize geometric magnification in prolonged procedures; this can be done by keeping the body close to the image receptor [14].
The fluoroscopic machines tend to support the possibility of choosing the level of dose to be delivered. This variable is also very influential on the study of image quality: a low-dose acquisition will be very noisy, while, by delivering a high dose, low-contrast details can be observed more accurately.

A series of quality metrics are developed to give a qualitative evaluation of an imaging system and describe image information, illustrate in Figure 1.4, using canonical

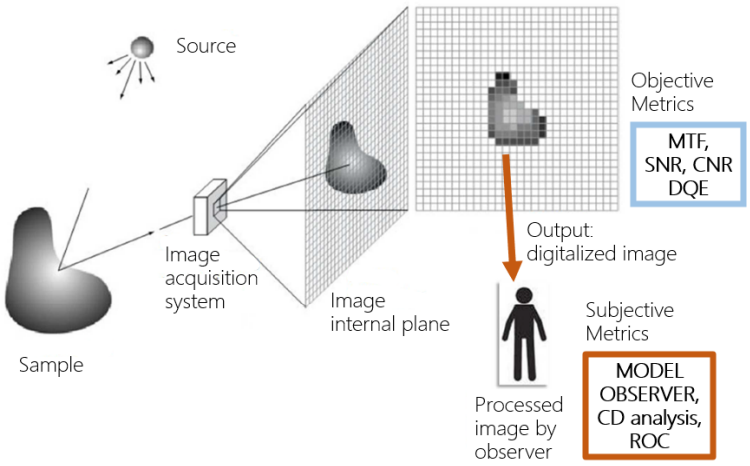



Figure 1.4: Image acquisition chain and metrics to make evaluation performance.

attributes that can be measurable. These parameters are the brightness, contrast, sharpness, blur, distortions and the noise of an image, and others that are not measurable. These latter ones fall in the field of perceptive quality, for instance, the physiology and the visual human perception of each measurable factor.At this stage, the metrics can be organized into objective and subjective.

### 1.3.1 Objective Metrics

An objective metric is used to evaluate an image through its parameters only with the aid of mathematical formulae. These metrics become helpful in the detectability of low-contrast objects.

#### Modulation Transfer Function (MTF)

The MTF is a metric indicating the fraction of signal visible in the images for each spatial frequency. When it is reduced, the blur enhances and the image quality decreases. It is considered an accepted standard for the description of spatial resolution. The signal amplitude ( in %) decreases, increasing the spatial frequency (written in *cycles/mm*).

#### Signal to Noise Ratio

The SNR is a measurement of the noise of a system and it is described by the following formula:

$$SNR = \frac{Signal}{Noise} = \frac{\sum_i (x_i - \bar{x}_{BG})}{\sigma_{BG}} \tag{1.1}$$

The equation is a ratio with a numerator the signal integrated over the entire dimensions of the object of interest equal to the difference of the mean signal $x_i$ and the mean background $\bar{x}_{BG}$ and, as denominator, the standard deviation $\sigma_{BG}$ in

the homogeneous background region of the image.
The SNR is one of the most meaningful metrics that indicates how well the observer can see an object. The *Rose Criterion* uses the SNR to say that a detail which has a SNR $\geq 5$ is consistently recognized [19] [6].

#### Contrast to Noise Ratio

The CNR is a good indicator of detectability. Its value equals the ratio between the differential intensity of average greyscale in the signal ROI and background ROI over the standard deviation in background ROI.

#### Detective Quantum Efficiency

The concept of spatial and contrast resolution are linked by other metrics, the Detective Quantum Efficiency (DQE). It is an expression of the system performance and explains the capability of a detector to extract information from a beam of radiation; it is a good metre of comparison between different systems.

$$DQE(f) = \frac{SNR^2_{OUT}}{SNR^2_{IN}} = \frac{k[MTF(f)]^2}{N\ NPS(f)} \tag{1.2}$$

The Eq.1.2 shows how the DQE is written in function of the MTF(f), a model of signal processing, and NPS(f), which describes the image system noise processing.

### 1.3.2 Subjective Metrics

Using subjective metrics means making an image evaluation that considers the the Human Visual System perception (HVS). The observer's performance in the entire scheme of assessing image quality is essential to be included.

#### Contrast-Detail analysis

The contrast-detail (CD) curve is a robust tool for evaluating displayed image quality. It is a visual and subjective technique that combines the spatial and contrast resolution concepts to assess a system performance in terms of observability.
In order to plot the contrast-detail curve, a series of factors must be taken into consideration. The psychophysical features of the observer, the viewing conditions, like the light of the environment, with the physical characteristics of the display must be contemplated [23].
In the following Figure 1.5, a CD diagram is illustrated. As can be seen, the disk diameter increases towards the left, whereas the contrast of each disk declines from top to bottom.
The three images in the figure, from left to right, are respectively noisier, even if the spatial resolution remains the same; by adding noise to the image, the contrast resolution is reduced, and the detection of small and less contrasted details becomes progressively more difficult. The orange line represents the CD curve as

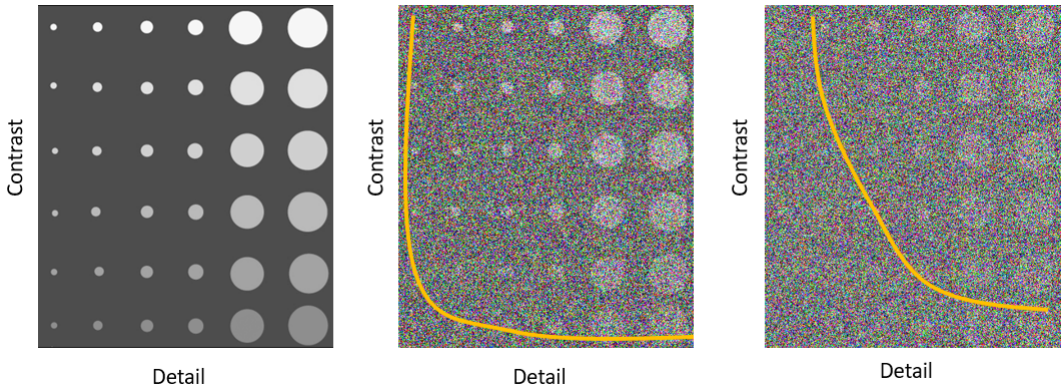



Figure 1.5: Contrast-detail diagrams and curves. On the left, it is shown the noiseless CD phantom. Moving to the right a noise's component in gradually added, losing in resolution. The orange CD curve stays, like a demarcation, to indicate the limit of visibility of disks. The higher is the noise, the larger is the number of disks not detectable.

the minimum detectable contrast in function of the diameter; it is a threshold of visual detectability that separates the visible disks from those that cannot be seen.

### ROC analysis

As far as image quality is concerned, the information about the human interpretation is a valuable aspect from a clinical and diagnostic point of view. This aspect can be tested with the Receiver Operating Characteristic (ROC) curve, whose performance includes both the quality related to image data based on a specific task and the subjective skills of the observers in assessing data. The ROC analysis does not neglect the influence given by the perceptual characteristics of the human who will interpret the image, and it considers the critical confidence level which allows distinguishing a "true positive" (TP) or "true negative" (TN) images.
In this way, some observer-dependent quantities are defined and used to describe the ROC curve. According to the confidence level, metrics like sensitivity and specificity are adapted.

$$TPF = Sensitivity = \frac{TP}{TP + FN} \tag{1.3}$$

$$TNF = Specificity = \frac{TN}{TN + FP} \tag{1.4}$$

The sensitivity, in Eq.1.3, defined as the true positive fraction (TPF), is the probability that an assumed "positive" image will be classified correctly as "positive" by the observer; the specificity (Eq.1.4), i.e. the true negative fraction (TNF), is the probability that an assumed "negative" image well be appropriately recognized. So, a trade-off between sensitivity and specificity is given by the combination between the imaging system and the observer achievement [3].
In theory, the true state of an object should be known, so who observes the images

should be able to compare them with the truth, but unfortunately, establishing diagnostic truth in clinical imaging is not easy. The decision criterion is not just one feature but an overall result that considers a lot of subjective and objective factors. In principle, ROC analysis can be led by applying a two-stage classification task; Figure 1.6(a) shows two distinct distributions based on a $2 \times 2$ decision matrix. The vertical black line indicates the internal binary decision threshold set by the observer; it can be consciously varied according to the degree of confidence the observer requires in his judgement before reporting the signal as seen or not. A high confidence level will imply that a high visual stimulus, i.e., a high contrast signal, is required before the observer reports the signal as seen, whereas if the observer employs a low degree of confidence, then signals with much lower contrast are seen [15].

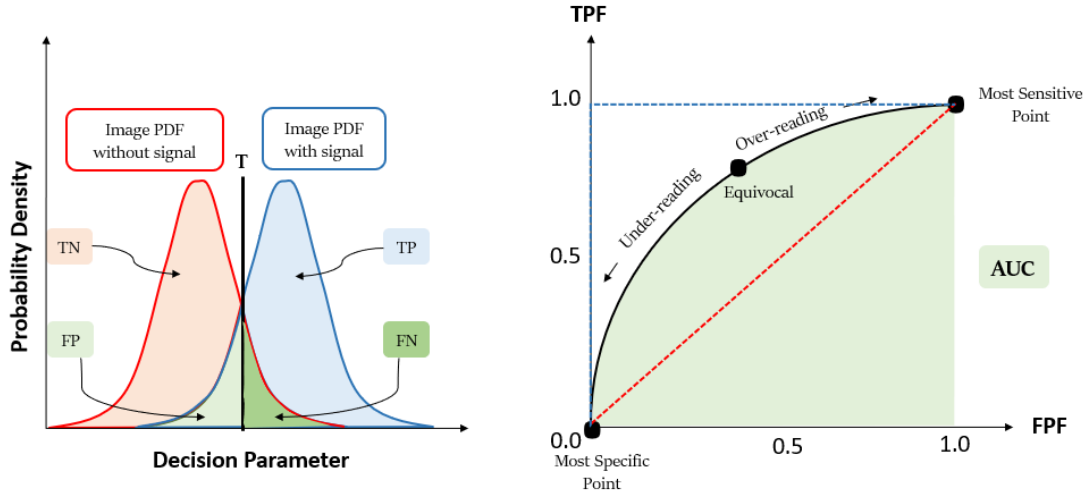



Figure 1.6: (a) PDF distributions with *true positive* (TP),*true negative* (TN), *false positive* (FP), *false negative* (FN) terms and threshold (T) decision parameter. (b) ROC curve graphical analysis and Area Under the Curve (AUC in light green.)

A curvilinear graph is plotted with TPF, the sensitivity, on the y-axis and the False Positive Fraction (FPF), equal to $[1 - TNF]$, on the x-axis to produce a ROC curve.
In Figure 1.6(b), the sensitivity increases and an over-reading is recorded if the decision parameter is moving toward the right on the curve. An inverse phenomenon happens while moving to the left: more specific details are measured, and an under-reading is verified [21].
One of the ROC curve characteristics is the capability to be used as a comparison in terms of system and observer performance. The straight ROC curve means the worst performance, that in Figure 1.6(b) coloured in dashed red, so when there is the 50% to getting true positives and just as much of false negatives. Differently, the perfect diagnostic test is given when the detection is always provided; in this case, the curve is the closest curve to the y axis (in dashed blue on the plot).
The accuracy is interpreted as a measure of diagnostic performance and depends on

several factors related to the difficulties in classification and the observer's ability. Some of the elements influencing the accuracy in ROC analysis and causing biases are the degree of confusing possibilities, the lesion size, the experience and training of the image reader.
In Figure 1.6(b), the highlighted green region means the area under the ROC curve (AUC) that is often used as a metric for describing the overall performance in ROC analysis, the range of values goes from a minimum of 0.5 in case of pure guessing (the worst situation) to a maximum of 1.0 which represents the perfect test.
A more extensive discussion of the strength of the AUC will be made in Chapter 2.

## 1.4 Radiation dose in fluoroscopy

In fluoroscopic systems, an high patient radiation dose is imparted, especially in those cases where the procedures require enough prolonged time. However, the incident radiation dose depends on many aspects, like the type of examination, the patient size, the equipment and the technique used. The main method used to evaluate the impact of radiation in fluoroscopy exploits a tissue-equivalent phantom of PMMA (polymethyl methacrylate) placed in the action field, so that a measurement of the entrance skin dose is done for every operating modes.
Dose monitoring is routinely performed, and it is essential in each procedure intending to respect the dose limits imparted for reducing the undesired effects due to higher and more prolonged exposures and by rational regulations.
Using an external probe dosimeter, the operator obtains a consistent measurement of air kerma rate. The amount of radiation in the air per minute is completed at the Interventional Reference Point (IRP), the specific point in space relative to the fluoroscopic gantry during the procedure [16] The kerma rate also called the cumulative kerma rate, underestimates the peak skin dose. The limitations are related to the fact that this magnitude does not account for backscattered radiation; the x-rays beam could not impinge upon a single area, causing an uncontrolled spread over multiple areas, and the skin could be closer or farther to the dosimeter position.
Several arrangements can be made to minimize the incident dose[14]. Some techniques include heavy beam filtration, the introduction of low frame rate pulsed fluoroscopy, the use of low doses providing higher kV and lower mA, the positioning of the X-rays tube as far from the patient as possible, thus keeping low the magnification factor, the use of less fluoroscopy time, the exploitation of last-frame-hold feature (described in Section 1.2) and a lot of other good practice measures.

# Chapter 2

# Observer Model

In the description of the imaging chain, the position occupied by the human observer is essential to get a consistent final overview about the characterization of image quality. What it is possible to extract from the human contribution is how well a desired information can be obtained from the images by the human eye. In practice, a task-based medical image quality assessment is conducted. So, the observers' figure plays an essential role in the imaging system because, through the imaging technologies, they are the final users of the equipment.

## 2.1 Human visual quality perception

When the observer's component is added to the analysis' path of image quality assessment, a series of issues, due to subjective judgment, must be taken into account. The specialist's performance has to contemplate the chance of mistake: false positive or false negative decisions, for examples, are some of the main intuitive inaccuracies that can be attributed to perceptual and cognitive factors [12].
Rating the quality of visual stimuli becomes a complicated task which necessarily needs to come to terms with different distortions and visual context, often aggravated by complex patterns. As a consequence of this, sometimes great discrepancies of quality scores between human observers are achieved. Various aspects influence the quality judgment. The external factors refer to the environment's leverage, where the quality assessment takes place. Some of these are, as instance, the distance to the monitor where the images are observed, the type of monitor chosen to make the observations that should be the same for each human observer and the light's level in the environment. The external impact has a relevant weight on quality perception, in fact, it is at least desirable that experiments of subjective quality are conducted, as far as possible, under the same conditions for each observer.
Instead, the internal factors are related, directly, to the observer's subjectivity. Going into more detail, there are influential subjective-dependent elements like detection, preference and confidence concepts that contribute differently for each subject performing the quality test.
The peculiarity of the observers can be noted in distinguishable viewing behaviours; humans who analyse series of medical images, have their own likelihood of detecting distortions depending on their appearance in relation to the content. Thus, some

observers may detect distortions that others do not detect and this aspect influences the provided quality score. In this assessment, the concept of preference is significant because of the subjectivity of visual quality perception. The distortion's interpretation differs between humans and, as a result, there is disagreement between the rating provided by different observers in relation with their preference of artefacts. Indeed, judging image quality is a difficult work because it strictly depends on stimuli and human-dependent factors. Distortion types, distortion distributions and the interaction between distortions with the underlying image content should be evaluated and interfaced with the personal background and expectations of the observer in relation with a specific image. This skill is known with the name of confidence [8].
Understanding the capabilities of the human visual system and what measures can be taken to improve the conditions and the environment in which medical images are interpreted, is becoming a deeply rooted topic of viewing medical images. If we understand what the human visual system is capable of, we can tailor the display of medical information to take advantage of these perceptual capacities [12].

## 2.2 Visual Tasks

In the medical imaging field, investigations of how well the human figure performs visual tasks are provided and continuously improved. The tasks involved can be broadly categorized in estimation and classification tasks. An estimation task concerns the quantification of one or more parameters that give a description about the object in questions, in the form of raw data. Typically, the most common parameters used in estimation performance are the size and the location of a particular region of interest, like, for example, a tumour. Instead, a classification task prompts the observer to make a decision between different classes regarding the data.
These two types of tasks are strictly correlated, in fact an estimation can be thought as a classification whose the number of classes corresponds to the number of values, the estimated parameters can assume [3].
So, the figure of a decision maker or observer is invoked to perform these tasks. Assuming that the decision in question is made without including guessing or randomness, in the simple binary case, what the observer is called upon to do, is to classify data into states of truth that can be designated as $T_1$ and $T_2$. Under binary classification tasking, the set of operations performed from the object to the decision is described in Figure 2.1.

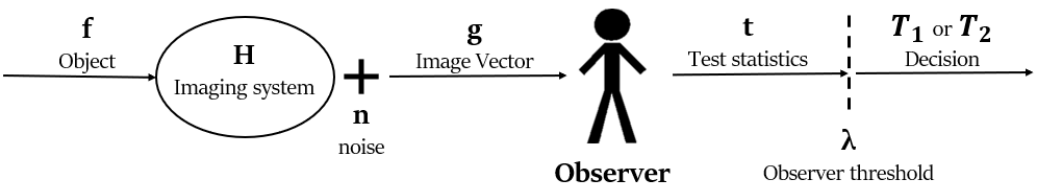



Figure 2.1: Set of operations' diagram, from object to data classification.

Thershold $\lambda$ corresponds to the decision threshold shown in Figure 1.6(a).

### 2.2.1 2-AFC and SKE/BKE binary detection task

In light of what has been said so far, it is convenient to exploit classification and decision tasks that limit the misclassification or probability of human error, due to the sensitivity and subjectivity of the observer.
One strategy for modeling decision-making in medical imaging consists of the possibility to choose between two alternative images. This task, called 2-AFC (Two Alternative Forced Choice) appears to be a simplification in many situations because the degrees of freedom are reduced. So, two different stimuli are presented in random or quasi-random order. Moreover, in this project, a SKE/BKE (*Signal Known Exactly/Background Known Exactly*) binary detection task is performed. This method assures the presence of signal or background images. Knowing this preliminary information, the observer has a good starting point for making the choice among the two alternatives.
The binary decision task of this project consists of two classes that are signal present and signal absent.
Since the human observer is required to choose one of the two proposals, the percentage of correct and incorrect answers must be taken into account. This is described by a psychometric function, which shows a non-linear relation between the physics and the observer's psychology and measures the relationship between some physical characteristics of the stimulus (e.g. contrasts) and response performance [10].

## 2.3 Model Observers

When a diagnostic decision involves a human observer, medical image quality studies should take into account the interference due to human performance in visual task. Because of the long time consuming and the high requirement of personnel that acts as observer, new faster methods are proved with the aim to get efficient results in terms of numbers and time. Indeed, computer model observers are algorithms which mimic the human performance in specific visual tasks supplying the metrics which better represents the quality of the image of interest.
In the course of the years, great developments were made, and thanks also to the rapid search for innovative instruments, these observer models have acquired more and more properties that refined the study and quality of images of a medical system. So, different observer models can be defined on the basis of used visual task and of the chosen classes of images (signal and background) and, even more so, they are classified according to their purpose of use. Some of these are, for example, image discrimination models, which try to predict human visual detection of a signal, models for detection in noise, that explore the human performance in visual noise and hardware performance models that cause the observer model to extract as much statistical data as possible from a certain task [9].

### 2.3.1 Ideal Observer

The Ideal Observer is defined as *"the observer that utilizes all statistical information available regarding the task to maximize task performance as measured by Bayes risk or some other related measures of performance"* [1].
In binary detection theory, indeed, it is considered as a figure that exploits everything available in an image to do the best classification performance.
Compared to the human (real) observer, the possibility of error excludes any erroneous source due to additional noise or to decision process, limiting errors only to the variability of the system measurement of objects and noise classes. Basically, through the ideal observer, one has the opportunity to show off theoretical tools that have a strong impact for assessing imaging system performance using raw data. It can be said that the performance of the ideal observer reflects all the information contained in the measured data because the result of the measurements is independent on the order of execution, display or processing and this guarantees desirable optimization.
The ideal observer responds to the *Likelihood ratio* statistical test:

$$\Lambda(\mathbf{g}) = \frac{pr(\mathbf{g}|T_2)}{pr(\mathbf{g}|T_1)} \tag{2.1}$$

where $pr(\mathbf{g}|T_i)$ are the hypotheses.
This method is the most optimal in terms of minimising decision errors. It requires a complete overview of the raw data information and that the ideal observer knows the threshold set at each decision trial in advance. This is why the ideal observer is considered an upper aid to improve human perception of images [3].
The ideal observer in a statistical-decision-theory approach, gives a measure of its performance that can be used as a basis of comparison in absolute terms when human performance is particularly high or low.

### 2.3.2 Linear Model Observer

Typically, with respect to image discrimination models, models for detection in noise make use of linear decision strategies between the signal and background choice. In fact, this class of models are based on the statistical decision theory. Depending on the different amount of knowledge of signal and noise given to the observer, various models have been developed.
Focussing the attention on the linear observers, they follow a common procedure to obtain performance for a signal detection task. On each trial, a fixed template is correlated with the supplied data to give rise to a decision variable. A decision is made by comparing decision variables from all possible signal locations [24].
The general form of response for linear model observer is given by a linear transformation between the vector template ($w$) and the image vector ($g$):

$$\lambda_i = \sum_{n=1}^{N^2} w_n g_n = \mathbf{w}^t \mathbf{g}_i \tag{2.2}$$

At this point, in Eq.2.2, $\lambda_i$ represents the scalar response of the model to $i$th location and symbolizes, in other words, the decision variable.

### 2.3.3 Channelized Model

A significant aspect concerns the preprocessing of the images of interest through a set of channels that are tuned to some specific spatial frequencies, orientations and phases. The usage of channels has the double function of better predicting the human observer behaviour (since our visual system performs better at certain frequencies) and reducing the dimensionality of the ideal observer computation.
With respect to the classic linear observer model, in channelized models the given response comes from the correlation among each channel, the data at the possible signal location and the scalar outputs per channel assuming the following shape

$$\eta_{ij} = \mathbf{v}_j^t \mathbf{g}_i. \tag{2.3}$$

$\eta_{ij}$ defines the response referred to the $j$th channel to the $i$th location; $\mathbf{g}$ is the known image vector and $\mathbf{v}$ is the spatial weights vector of $j$th channel. Returning to the theory of linear model, the total response is given in form of Eq. 2.2 where instead of $\mathbf{g}$ is used $\eta$ and a weighting factor $c$ is applied to the response of the $j$th channel:

$$\lambda_i = \sum_j^{N_{ch}} c_j \eta_{ij} = [\mathbf{Vc}]^t \mathbf{g}_i = \mathbf{w}^t \mathbf{g}_i \tag{2.4}$$

Eq. 2.4 uses $\mathbf{c}$ as the weights column vector for each channel and $\mathbf{V}$ as the matrix channel.
The channels can be organized in two main groups according to their utility. "Anthropomorphic" channels are those that more imitate the characteristics of the human visual system into a model observer, like for example the spatial-frequency channels which refer to the various sensitivities of the human visual system to distinguish among different ranges of spatial frequencies. There is another channels' category which moves to the concept of efficiency: "efficient" channels reflect their attribute if the resulting channelized observers approximate well the human performance and reduce the data dimension serving as few channels as possible.The set of efficient channels can be estimated by considering the physics of the image system. The amount of collected physical information should be as efficient as possible avoiding redundant informations. So, a careful selection of efficient channels is essential in the performance estimation of an ideal observer.

#### Gabor channels

One type of known channel is called Gabor channels. This filter is very common in image processing, texture analysis and feature extractions and it can approximate well the visual human perceptions. Not surprisingly, it totally embodies the "anthropomorphic" channel meaning.
A Gabor filter belongs to the family of band pass filters, in fact it has the property to select a specific range or band of frequencies and rejecting the others. Thinking in two dimensions, Gabor channels can be approximated, in a spatial domain, by

multiplying a sinusoidal wave by a Gaussian function:

$$C(x,y) = exp(\frac{-4ln2((x-x_0)^2+(y-y_0)^2)}{w_s^2})cos(2\pi f_c((x-x_0)cos\theta+(y-y_0)sen\theta)+\beta) \tag{2.5}$$

where the Gaussian function determines the passband of the channels ($w_s$) and the sinusoidal wave collect all the information related to the frequency ($f_c$), orientation ($\theta$) and phase ($\beta$) [18].
The number of channels is chosen based on the characteristics of the images set. The aim is to reduce the dimensionality of the system as much as possible, but without the loss of significant strong information from the images in question.
Below, in the figure, a 40-channel Gabor filter can be observed: namely, the 40 channels are generated from 5 frequencies, 4 orientations and 2 phases.

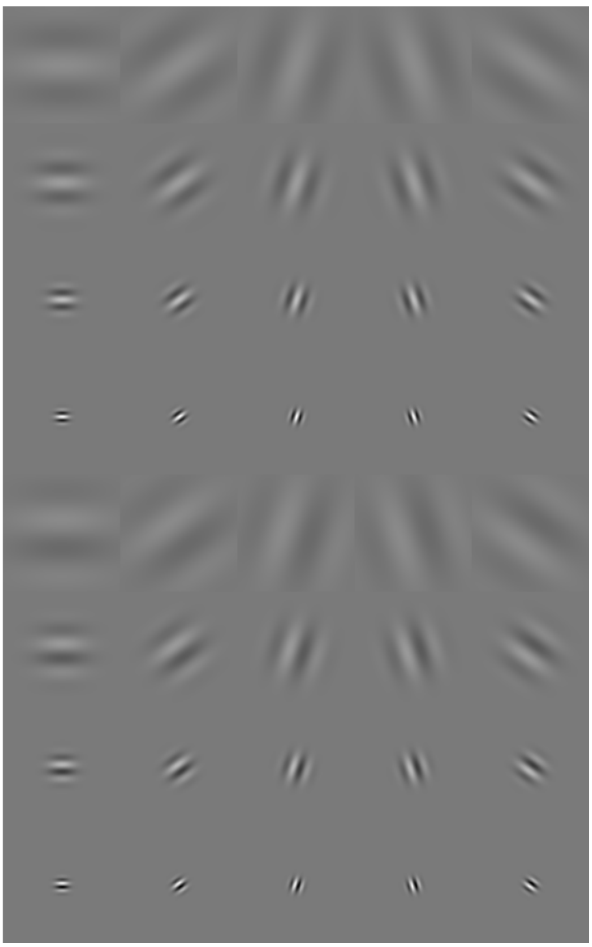


Figure 2.2: Gabor 40 channels. These channels allow to reduce the problem dimensionality and come out of the combination of 5 different frequencies, 4 orientations and 2 phases.

## 2.4 Channelized Hotelling Observer Model (CHO)

From linear discriminant analysis and from the usage of a set of channels an important model to assess the image quality of a system, is given: the Channelized Hotelling Observer Model (CHO).
This model returns a response in the form described by Eq. 2.1 plus an internal noise component $\epsilon$ referred to the intra and inter differences among human observers. The Hotelling observer template consists of the inversion of the covariance matrix of the image classes. The inversion could be very heavy due to the large dimension of the covariance matrix (i.e. for a $100 \times 100$ images the covariance matrix is comprised of $10^4$ elements), a channel mechanism is involved in order to reduce the amount of processed information. So the image is processed using a set of channels that are tuned to give spatial frequencies of interest, orientations and phases in the Gabor case.
For CHO , the main form of its template takes this shape:

$$\mathbf{w}_{CHO} = \mathbf{K}_c^{-1}(\overline{\mathbf{g}_{c1}}) - \overline{\mathbf{g}_{c2}}). \tag{2.6}$$

In the expression above, $\mathbf{g}_{ci}$ is the mean value of the channelized image vector of the class $i$ and $\mathbf{K}_c$ is the channel covariance matrix that is in turn defined as

$$\mathbf{K}_c = \frac{1}{2}(\mathbf{k}_{c1} + \mathbf{k}_{c2}) \tag{2.7}$$

where $\mathbf{k}_{ci}$ is the convariance of the channelized image vector associated to its respective class [2].

### Internal noise

Technically, the performance of a model observer, linear or not, is higher than the one of a human observer. But, because the advantage to use MOs stays in their capability to predict the human observer performance, the comparison between human and model should return the highest affinity.
To enable this, various expedients, among which the set of channels, already mentioned and discussed in subsection 2.3.3, have been introduced to degrade the system and bring it closer to human visual perception. One important degradation method is the inclusion of internal noise. Internal noise is a known component of human inefficiency in perceptual tasks and it describes, in terms of intrinsic stimulus variability, receptor sampling errors and loss of information during neuronal activity [24].This variable embodies the fluctuation proper to the human decision maker that is not able to do the same choice on the same set of images in repeated trials caused by different visual conditions and psychological interferences.
Many applications to model the noise contribution are studied: in particular, up to now, the most used configuration provides a noise addiction to the output of the individual channels. The basic property of the channel internal noise method is that the channel template is computed through covariance due to both internal and external noise.
So, taking into account the independent contribution of each channel the decision

variable on each trial is then calculated considering a weighted combination of the channel results ($g_k$) and their internal noise ($E_k$):

$$\lambda_0 = \sum_{k=1}^{N} w_{CHO_k}(g_k + E_k) \tag{2.8}$$

where

$$w_{CHO_k} = (\mathbf{K}_{ext} + \mathbf{K}_{int})^{-1}(\overline{\mathbf{g}_{c1}} - \overline{\mathbf{g}_{c2}}) \tag{2.9}$$

which is composed by a background covariance matrix $\mathbf{K}_{ext}$ and by the internal noise covariance matrix $\mathbf{K}_{int}$. The internal noise, described by a normal distributed standard deviation $\sigma = \sqrt{\mathbf{K}_{int_k,k}}$, can be estimated using two distinct approaches [24]: the uniform method provides a covariance matrix proportional to the identity matrix. This is a considerable constraint for the internal noise because it will be the same for all the channels, in fact , it is preferable a non uniform method that defines the internal covariance matrix as elements in the diagonal of external noise covariance matrix (Eq. 2.10).

$$\mathbf{K}_{int} = diag(\mathbf{K}_{ext}) \tag{2.10}$$

This way it allows to obtain different noise contributions for each channel. Furthermore, the power of the model lies in being able to make different decisions even for the same images in repeated tasks.
This human behaviour is included in the model by summing to the scalar response an internal randomized noise variable $\alpha$.
The operation is done on the basis of this mathematical formula:

$$\lambda = \lambda_0 + \chi\alpha. \tag{2.11}$$

In Eq. 2.11, $\lambda$ is the final result composed by the response $\lambda_0$ not affected by the human noise (Eq. 2.8) and the internal noise variable $\alpha$ multiplied by $\chi$, a random variable which follows a normal distribution and is given by the square root of the variance of the class with no signal.
The alpha parameter needs to be tuned in order to get a consistent result and to approximate the human visual performance [2].

### Model performance

The channelized Hotelling Observer template can be considered the best template (under Gaussian conditions) that can be derived from a linear combination of channel profiles [3]. Therefore, it is important to consider some aspects in terms of model performance estimation and about human-model comparison.
In other words, we want to study which is the optimal metrics to describe performance in signal detection task; this entity is called figure of merit (FOM) and measures, quantitatively , how well a given task is performed under the investigated condition. Taking into account to model binary classification task, a test statistics is performed and it returns , depending on the selected class, its associated probability density function, thus, according to a fixed threshold, determining the amount of

false negative and false positive fractions.
On the basis of what explained in Chapter 1 section 1.3.2, plotting TP fraction values and FP fraction values, Area Under Curve (AUC) metric is computed. Possible values can vary between 0.5 and 1, i.e. ranging complete guessing to total certainty. So the area under ROC curve can be used as FOM in performance assessment for CHO model. If the decision variable is normally distributed, it takes the shape of

$$AUC = \frac{1}{2} + \frac{1}{2} erf(\frac{SNR}{2}). \tag{2.12}$$

Observing the Eq. 2.12, the $erf$ indicates the error function and there is an important parameter which should be mentioned: the Signal to Noise Ratio (SNR).
The Hotelling Observer model is a linear observer that maximizes the SNR; indeed, under Gaussian approximation, the SNR can be completed as:

$$SNR = \frac{\overline{\lambda_1} - \overline{\lambda_2}}{\sqrt{\frac{\sigma_1^2}{2} + \frac{\sigma_2^2}{2}}} \tag{2.13}$$

The SNR is expressed in terms of means and standard deviation of classes (in our case: signal and no signal).
The second aspect regards the similarity between model and human outcomes. There is therefore a need to model a psychophysical relationship typical of the human observer regarding the detection task. In order to study how stimuli and model responses are related, a psychometric mathematical function is exploited.

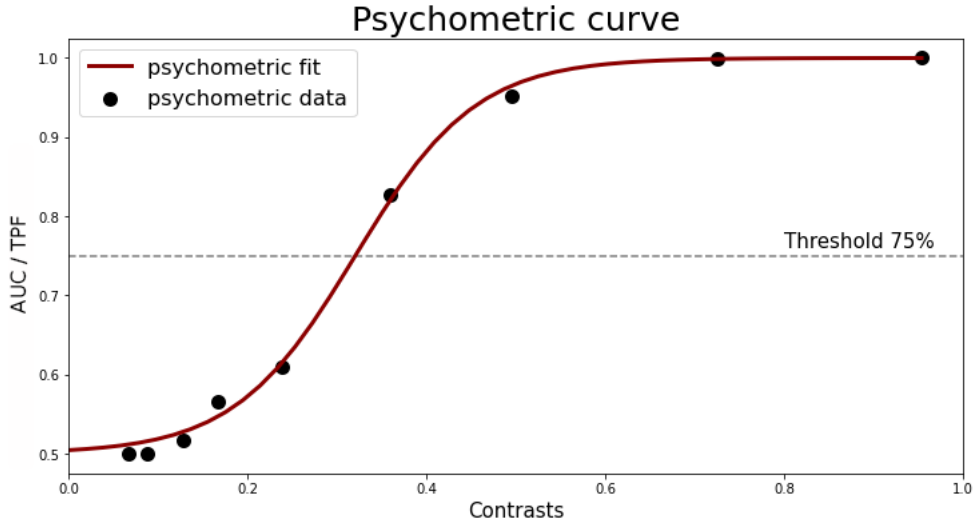



Figure 2.3: Fitted psychometric curve (red) estimated on the basis of give data (black dots). The dashed line define the visibility threshold set to 75%.

As shown in Figure 2.3, this curve appears to be a sigmoid described in this way:

$$y(x) = \frac{1}{(1 + exp(-b_1(x - b_2)))}. \tag{2.14}$$

In the equation $y$ corresponds to the AUC or TPF response in case of CHO model and human observation respectively, and $x$ to the object's contrasts. The fitting parameters are $b_1$ and $b_2$.
The physical meaning that lies behind the psychometric curve can be illustrated as a gradual degradation of performance due to the presence of different influential sources of uncertainty that the subject, firstly the human decision maker but also the mathematical observer, can encounter.
In the shown example, the confounding factor which hinders the binary classification task is the decreasing contrast in the signal present images.
Psychometric functions are able to contain a lot of information about technical aspects of the single acquisitions. From a physical point of view, it is desirable to find a single value which can summarize well the system performance level. This stimulus level is a threshold that commonly, for psychometric function which goes from 0.5 to 1, is around 0.75.

# Chapter 3

# Setup and Methods

The aim of this project is a quantitative analysis of the image quality of angiographic systems available in Reggio Emilia's hospital through a specific model known for taking into account the human contribution in assessing image quality. The project consists primarily of:

1. Verification of the chosen CHO model and its applicability to given experimental data acquired on TO12 test object;

2. Quantitative analysis of extracted contrast detail curves of TO16 test object from two different angiographic machines present in Reggio Emilia's hospital and discussion about limitations and opportunities;

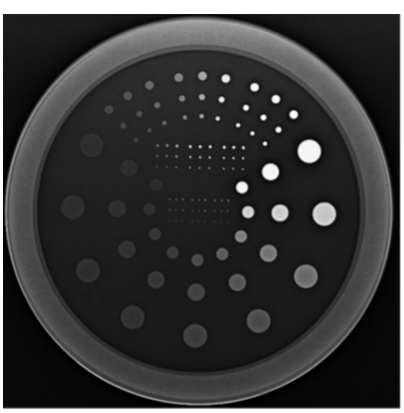
3. Statistical considerations and estimation of an appropriate figure of merit for TO12 and TO16 acquisitions.

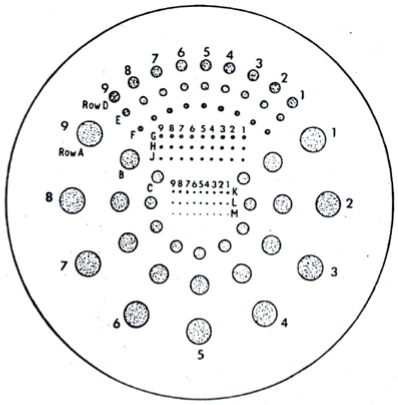
Finding out how the Channelized Hotelling Observer model works in practice is a very important step towards an understanding, from a quantitative point of view, of how the human component influences choices in the image quality assessment process. Therefore, after an accurate data analysis, deep considerations about the quality of images acquires with a specific system can be done.

## 3.1 Phantoms

First of all it is essential to introduce the test object used in our digital image acquisitions. The set of acquired images refers to a particular Threshold Contrast-Detail Detectability (TCDD) phantom suitable for fluoroscopy system. In fact it is sensitive to noise, gray-scale and unsharpeness of the specific evaluated system.
For the first objective a Leeds TO12 phantom[1] is used. In Figure 3.1 the phantom in object is illustrated like a circular plate of diameter $180mm$ and thickness $10mm$. The details are exactly 108 organized in 12 diameters and 9 contrasts. The table contained in Figure 3.1 shows the diameters in mm and records the minimum contrast and the maximum contrast associated to each details. It can be seen that the ranges of contrasts are organised in groups of three adjacent diameters.

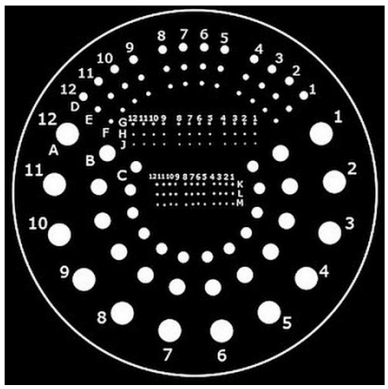
[1] https://www.leedstestobjects.com/index.php/phantom/to-12/





| REFERENCE | A | B | C | D | E | F | G | H | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Diameter | 11.1 | 8 | 5.6 | 4 | 2.8 | 2 | 1.4 | 1 | 0.7 | 0.5 | 0.35 | 0.25 |
| Contrast min | 0.0043 | 0.0043 | 0.0043 | 0.0086 | 0.0086 | 0.0086 | 0.0155 | 0.0155 | 0.0155 | 0.0674 | 0.0674 | 0.0674 |
| Contrast max | 0.0674 | 0.0674 | 0.0674 | 0.128 | 0.128 | 0.128 | 0.238 | 0.238 | 0.238 | 0.954 | 0.954 | 0.954 |

Figure 3.1: TO12 test object. The figure on the left shows its different contrasted details, the figure on the right is a detail layout with a schematic intuitive nomenclature. The table below contains all the technical specifications of the existing phantom.



| REFERENCE | A | B | C | D | E | F | G | H | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Diameter | 11.1 | 8 | 5.6 | 4 | 2.8 | 2 | 1.4 | 1 | 0.7 | 0.5 | 0.35 | 0.25 |
| Contrast min | 0.0016 | 0.0016 | 0.0016 | 0.0021 | 0.0021 | 0.0021 | 0.0032 | 0.0032 | 0.0032 | 0.013 | 0.0172 | 0.0215 |
| Contrast max | 0.0674 | 0.0674 | 0.0674 | 0.0876 | 0.0876 | 0.0876 | 0.128 | 0.128 | 0.128 | 0.573 | 0.726 | 0.954 |

Figure 3.2: TO16 test object. The figure is a detail layout with a schematic intuitive nomenclature. The table below contains all the technical specifications of the existing phantom.

For the second part of the work a similar test object is chosen, a Leeds TO16 phantom[2]. This test object, unlike the previous one, is bigger (it measures 250 mm of diameter and 10 mm in thickness) and consists of 144 details divided into 12 different diameters and 12 contrasts defined in the ranges shown in the table of

[2]https://www.leedstestobjects.com/index.php/phantom/to-16/

Figure 3.2.

## 3.2 Acquisition setup

The angiographic digital images used as input datasets for the analysis are acquired using different angiographic mechanisms.

### 3.2.1 TO12 test object on Siemens Artis Q C-arm

The first dataset used to test the model is given a priori, more specifically it is provided by the European Federation of Organisations for Medical Physics (EFOMP). The angiographic images of interest were acquired using Siemens Artis Q C-arm machine capturing TO12 test object sample.

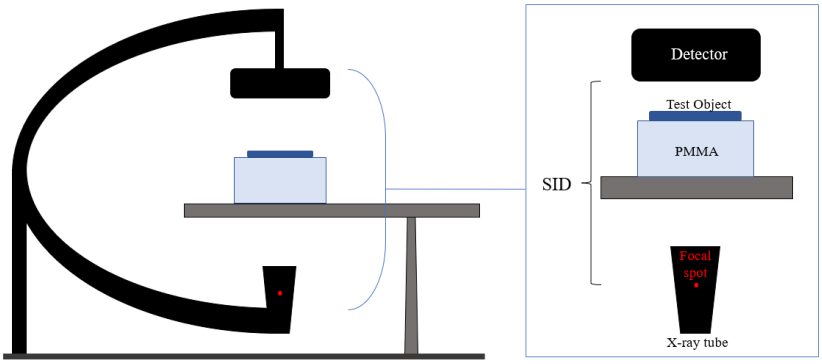



Figure 3.3: EFOMP setup acquisition.

As can be seen in Figure 3.3, the test object is posed over a thick layer of polymethyl methacrylate placed, in turn, on a couch.
From a geometric point of view, the SID, i.e. the distance from the source (or rather from its focal point) to the detector, the measurements relating to the isocentre and the dimensions of the materials used for acquisition were provided and they are reported in Table 3.1.

Table 3.1: Siemens setup specifications.

| | |
|---|---|
| **Name** | Siemens Artis Q C-arm |
| **Isocentre** | 75 cm |
| **PMMA thickness** | 20 cm |
| **SID** | 110/120 cm |
| **Source to PMMA distance** | 65/55 cm |
| **Phantom** | TO12 |

For study purposes, the collected image packages are the result of acquisitions that differ depending on the choice of parameters used, such as acquisition program,

dose level and field of view (FOV) size.
A table summarising the characteristics of each acquisition is provided (Table 3.2).

Table 3.2: Siemens acquisition's measurements

| **SIEMENS ARTIS Q TO12** | **1** | **2** | **3** | **4** | **5** | **6** | **7** |
|---|---|---|---|---|---|---|---|
| **Program name:** | Cardiac | Caridiac | Cardiac | Cardiac | Cardiac | Cardiac | Cardiac |
| **Mode:** | Low | Low | Low | Low | Medium | High | Low |
| **PMMA thickness (cm):** | 20 | 25 | 20 | 20 | 20 | 20 | 20 |
| **kV:** | 74.6 | 80.0 | 76.1 | 76.1 | 78.8 | 70.9 | 78.8 |
| **mA:** | 200 | 237 | 412 | 412 | 113 | 240 | 118 |
| **ms:** | 5,2 | 5,2 | 3,5 | 3,5 | 5,2 | 5,2 | 5,2 |
| **frame rate (fr/s):** | 15/s | 15/s | 15/s | 15/s | 15/s | 15/s | 15/s |
| **KAP**($dGy \cdot cm^2$) | 6,65 | 21,1 | 11,7 | 11,6 | 9,61 | 19,6 | 8,34 |
| **system prefilter material:** | Cu | Cu | Cu | Cu | Cu | Cu | Cu |
| **system prefilter thickness (mm):** | 0,6 | 0,3 | 0,3 | 0,3 | 0,3 | 0,2 | 0,3 |
| **FOV (cm):** | 22 | 22 | 22 | 22 | 22 | 22 | 22 |
| **SID (cm):** | 110 | 110 | 110 | 110 | 110 | 110 | 120 |
| **Acq. Type** | Fluoro | Fluoro | Angio | Angio | Fluoro | Fluoro | Fluoro |
| **#images** | 500 | 530 | 301 | 301 | 535 | 527 | 518 |

### 3.2.2 TO16 test object on Philips Azurion 7 C20 FlexArm

The acquisitions done with TO16 phantom were acquired directly at Santa Maria Nuova hospital in Reggio Emilia.
In contrast to the TO12 image set, the arrangement of the phantom in the apparatus between source and detector respects what is canonically called the "good geometry condition". In the Figure 3.4 it is possible to observe this favourable configuration and its setup specification (Table 3.3). The good geometry condition assure to

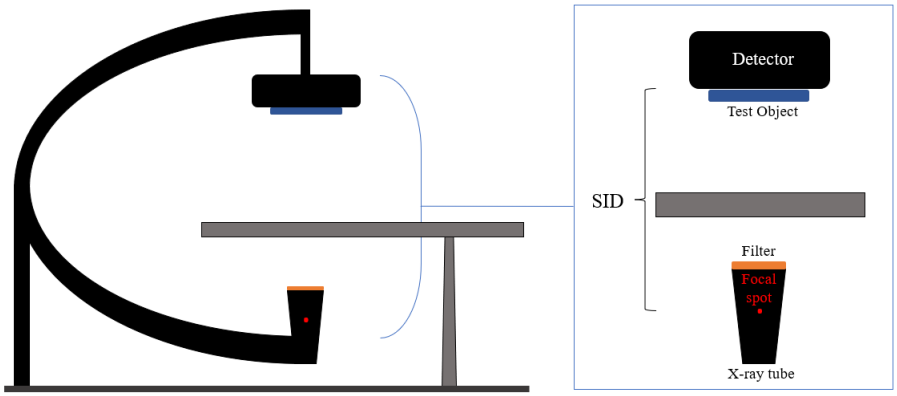



Figure 3.4: Philips setup acquisition.

perform the acquisition which minimizes the occurrence of external artefacts.
Looking at the configuration used by EFOMP (Figure 3.3), an initial consideration can be made about the use of PMMA, which, due to its considerable thickness,

Table 3.3: Philips setup specification.

| | |
|---|---|
| **Name** | Philips Azurion t C20 FlexArm |
| **Filter material** | Cu |
| **Filter thickness** | 1/0.5 mm |
| **SID** | 119/100 cm |
| **Phantom** | TO16 |

generates a phenomenon of scattering of the incoming photons which degrades the quality of the beam incident on the object and therefore reduces the final quality of the image. Furthermore, in order to consider every element that alters the final quality of the image, it is necessary to evaluate the position of the phantom between the source and the detector. The photons passing through the phantom must cross a substantial distance before reaching the detector. The incidence of air, albeit in a lesser form than the PMMA layer, generates further fluctuations of the photons causing a noisier and more degraded image.
So, a series of enhancements, to obtain a "better" image, are listed and comprise for example the reduction of photon scattering substituting the thick PMMA layer with thin filters of copper located directly in contact with the source surface. Both PMMA and Cu cause the same attenuation on the beam as there would be with a patient, but in the case of PMMA there is abundant scattered radiation in the medium which further degrades the image quality. In addition, to avoid further deviations, the phantom was placed directly in contact with the detector.
However, it is important to note that the geometry that best simulates the real situation is that of Figure 3.3. Choosing to operate in "good geometry conditions" was a reasoned choice on the basis of the TO16 test object used, which, unlike TO12, is less contrasted and would not have allowed an interesting analysis.

In Table 3.3 the main geometric parameters are shown for the acquisitions and in Table 3.4 all the characteristics and technical parameters used are illustrated (i.e. dose level, voltage, current, frames per seconds, Kerma Area Product, Field Of View, Source to Image Distance and the number of acquired images per acquisition).
The Kerma Area Product (KAP) is a method of radiation dose monitoring used in radiographic and fluoroscopic studies. It is an indicator of the radiation dose involved in the acquisition procedure.
In fact, it is calculated as the product of dose and beam area (in our case the unit of measure is $dGy \cdot cm^2$), and it is measured using an ionization chamber placed in proximity of the x-ray tube/collimator apparatus . This measure can be divided by the area of exposure (in $cm^2$) to give the air kerma, that is the kinetic energy released per unit of mass:

$$\frac{dE_k}{dm}. \tag{3.1}$$

Consequently, this can be used to calculate the accumulated dose for skin, which is important to measure in interventional and fluoroscopic procedures for radioprotection purposes.
During the acquisition phase it was necessary to perform some steps to position the

Table 3.4: Philips acquisition's measurements

| | Acq. Type | Program name | Dose level: | kV: | mA (mAs for angio) | frames/s | KAP ($dGy \cdot cm^2$) | Filter Thickness (mm) | FOV (cm) | SID (cm) | # images |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | Angio | Abdomen | medium | 80 | 2.0 | 6 | 7.04 | 1 | 31 | 119 | 422 |
| **2** | Fluoro | Abdomen | high | 65 | 3.5 | 15 | 3.43 | 1 | 22 | 119 | 450 |
| **3** | Fluoro | Abdomen | medium | 68 | 4.7 | 15 | 1.67 | 1 | 22 | 119 | 450 |
| **4** | Fluoro | Abdomen | low | 69 | 2.9 | 15 | 1.37 | 1 | 22 | 119 | 450 |
| **5** | Fluoro | Abdomen | high | 64 | 2.8 | 15 | 4.86 | 1 | 31 | 119 | 450 |
| **6** | Fluoro | Abdomen | medium | 66 | 3.8 | 15 | 2.35 | 1 | 31 | 119 | 450 |
| **7** | Fluoro | Abdomen | low | 67 | 2.4 | 15 | 1.28 | 1 | 31 | 119 | 404 |
| **8** | Fluoro | Abdomen | medium | 66 | 3.9 | 15 | 1.92 | 1 | 22 | 100 | 450 |
| **9** | Fluoro | Abdomen | medium | 65 | 3.1 | 15 | 2.89 | 1 | 31 | 100 | 450 |
| **10** | Fluoro | Abdomen | medium | 63 | 2.2 | 15 | 1.20 | 0.5 | 31 | 119 | 450 |
| **11** | Fluoro | Abdomen | medium | 64 | 2.8 | 15 | 0.77 | 0.5 | 22 | 119 | 450 |
| **12** | Fluoro | Abdomen | medium | 62 | 1.7 | 15 | 1.46 | 0.5 | 31 | 100 | 450 |
| **13** | Fluoro | Abdomen | medium | 63 | 2.2 | 15 | 0.69 | 0.5 | 22 | 100 | 401 |

couch at the isocentre and to centre the phantom to the image acquisition area.
The geometry described above is facilitating from this point of view, but it is necessary to consider different behaviours at the moment of acquisition due to the dimension of the FOV, the SID and , from a temporal point of view, due to the number of desired images acquired per seconds ()$15\frac{fr}{s}$).

### 3.2.3 TO16 test object on GE Discovery IGS 740

Another angiographic system used to perform the study in place is GE Discovery IGS 740. Because of the acquisitions were done using TO16 Leeds, the setup configuration is the same of that used for Philips machine. For this reason, the geometry shown in Figure 3.4 in the previous section is also valid for this acquisition set. The spatial and technical characteristics of these acquisition are shown in Table 3.5.

Table 3.5: GE setup specification.

| | |
|---|---|
| **Name** | GE Discovery IGS 740 |
| **Filter material** | Cu |
| **Filter thickness** | 1/0.5 mm |
| **SID** | 119/100 cm |
| **Phantom** | TO16 |

In Table 3.6 are illustrated the fixed parameters used in each GE acquisition. A pre-filtration is automatically imposed and it corresponds to $0,3mm$ of Cu.
The associated flat panel is a $45cm \times 45cm$ Csl(Tl) detector coupled with an a-Si

Table 3.6: GE acquisition's measurements.

| | Acq. Type | Program name | Dose level: | kV: | mA (mAs for angio) | frames/s | KAP ($dGy \cdot cm^2$) | Filter Thickness (mm) | FOV (cm) | SID (cm) | # images |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | Angio | Abdomen | low | 67 | 52.0 | 6 | 2.66 | 1 | 32 | 119 | 451 |
| **2** | Angio | Abdomen | normal | 67 | 93.0 | 6 | 4.75 | 1 | 32 | 119 | 451 |
| **3** | Fluoro | Abdomen | low | 73 | 0.3 | 15 | 0.63 | 1 | 32 | 119 | 450 |
| **4** | Fluoro | Abdomen | normal | 70 | 1.1 | 15 | 1.83 | 1 | 32 | 119 | 450 |
| **5** | Fluoro | Abdomen | low | 73 | 0.6 | 15 | 0.38 | 1 | 20 | 119 | 451 |
| **6** | Fluoro | Abdomen | normal | 68 | 2.2 | 15 | 1.05 | 1 | 20 | 119 | 450 |
| **7** | Fluoro | Abdomen | normal | 70 | 0.8 | 15 | 1.73 | 1 | 32 | 100 | 450 |
| **8** | Fluoro | Abdomen | normal | 68 | 1.5 | 15 | 1.05 | 1 | 20 | 100 | 450 |
| **9** | Fluoro | Abdomen | low | 70 | 0.3 | 15 | 0.13 | 0.5 | 20 | 119 | 330 |
| **10** | Fluoro | Abdomen | normal | 66 | 1.1 | 15 | 0.48 | 0.5 | 20 | 119 | 366 |
| **11** | Fluoro | Abdomen | low | 73 | 0.2 | 15 | 0.29 | 0.5 | 32 | 119 | 450 |
| **12** | Fluoro | Abdomen | normal | 67 | 0.4 | 15 | 0.76 | 0.5 | 32 | 100 | 450 |
| **13** | Fluoro | Abdomen | normal | 66 | 0.8 | 15 | 0.45 | 0.5 | 20 | 100 | 451 |

matrix characterized by a nominal pixels pitch of $0.200mm$. As can be observed from the Table 3.6, two possible levels of dose (normal and low) and two of the four available FOVs (20 and 32) are used.

## 3.3 CHO model implementation

A pre-existing home made Matlab script was used to perform our analysis. Some modifications and enhancements were added to the original code.
In order, the schematic mechanism of the software consists of cropping signal and background ROIs from DICOM image sequences, these ROIs, saved in .tiff format, are cropped with dimensions conforming to the diameters of the signal disks. In our projects 144 or 108 sets were obtained, that is the number equivalent to the details of the TO16 and TO12 phantoms, respectively.
Then we performed through an affine transformation process that aims to match the centre of disks of a reference fluoroscopic image, from those acquired, with a mammographic acquisition of the same object which is better resoluted. In this way, it was possible to spatially map the acquired images and make the cropping process more precise.
The main issue encountered in this phase concerns the capability to cut, effectively, background ROIs; this is because they are extracted as close as possible to the sub-image of the associated signal despite the fact that the background spaces are narrow. This is, therefore, why great accuracy is sought in the definition of ROIs.
Once the sub-images are obtained, a human test can be performed in order to tune the model.

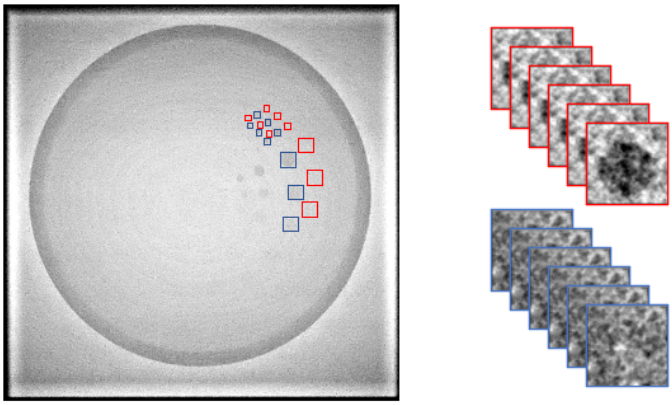


Figure 3.5: TO16 fluoroscopic image and signal (red) and background (blue) extracted ROIs.

For this task, a GUI interface shown in Figure 3.6 is returned. It represents the 2-AFC test (subsection 2.2.1): in practice a reference sub-image is given for each diameter and two alternative choice are supplied. The task is to select which image is definitely a signal between the two independent alternatives ( one with signal and one without signal). The pictures in GUI interface are magnified to $128 \times 128$ pixels using a bilinear interpolation [2].

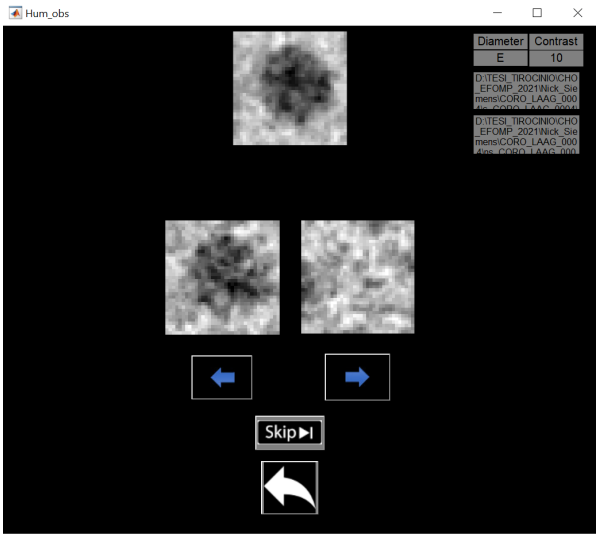



Figure 3.6: GUI 2-AFC test interface.

### 3.3.1 Observer's properties and collocation

To execute the analysis three human observers (2 juniors and 1 senior) are chosen to participate. They were asked to identify the image containing the signal by putting them all under the same observation conditions.

1. The monitor used to observe the images was the same for each observer.

2. The light in the environment was checked with a luxometer in range $[11; 15]lux$.

3. The distance between eyes and monitor was fixed to $50cm$.

4. The pictures in GUI interface are located at eye level.

The test requires the human observer to see, at least, six ROIs for each contrast of each diameter, starting with the highest contrast and ending with the lowest one.
In the case of human estimation, the analysis continues by estimating the fractions of true positives from which the psychometric curve can be obtained. In case of CHO CD curve estimation, the value in Eq. 2.12 is used.
The psychometric curves are of great importance in this phase of fitting process because they are used to extrapolate those contrasts at 75% visibility which are then used to plot CD curves.
A schematic diagram illustrates the main steps of the followed procedure (Figure 3.7).

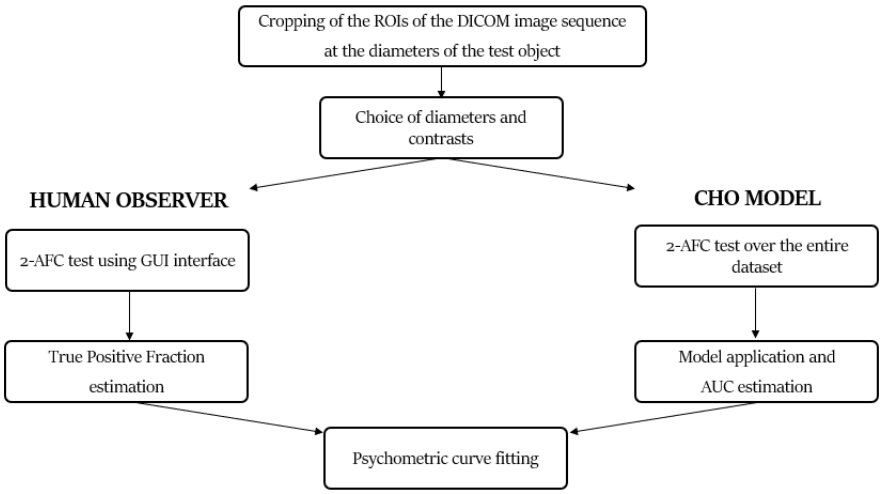



Figure 3.7: Schematic diagram about the steps followed in human observation and CHO model procedures.

### 3.3.2 Contrast points fitting

This part of the project can be considered the main junction of many future considerations. A series of improvements have been added to the original code with the idea of making it more adaptable for any type of data, going as far as possible, to consider hypothetical data sets from a physical point of view and finding the best adaptations.
The shown Matlab code highlights the main steps that have been performed.

```
fo = fitoptions('Method','NonlinearLeastSquares',...
        'Lower',[0, -Inf],...
        'Upper',[C_nom(1), Inf],...
        'StartPoint',[0.5256,0.2283]);

ft = fittype('(0.5./(1+exp(-(x-b2).*b3)))+0.5','options',fo);
        [curve2,gof] = fit(C_nom,AUC_defG,ft);
R_square(diam-handles.InitialDiam+1,1) = gof.rsquare;

% y estimation with fit parameters
x = linspace(C_nom(1,1),C_nom(handles.Contrmax-handles.
    Contrmin+1,1),50);
y = (0.5./(1+exp(-(x-curve2.b2).*curve2.b3)))+0.5;

% R square study:
if gof.rsquare>0.75
y_diff75 = abs(y-0.75);
int_y_diff75 = round(y_diff75,4).*100000;
ymin = min(y_diff75);
int_ymin = round(ymin,4).*100000;
ind_y75 = find(int_y_diff75 == int_ymin);
x75 = double.empty(0,length(ind_y75));
for el = ind_y75
        ind_el = find(ind_y75 == el);
        x75(ind_el) = x(el);
end
contr75 = mean(x75);

% if R square less than 0.75, study the mean of values and
    attribute the preferred contrast

else
        if median(y)>0.8
                contr75 = C_nom(handles.Contrmax-handles.
                    Contrmin+1);
        elseif ((median(y) > 0.7) && (median(y)<0.8))
                contr75 = x(25);
        elseif median(y)<0.7
                contr75 = C_nom(1,1);
        end
end

C_visCHO(diam-handles.InitialDiam+1,1) = contr75;
```

The first lines of the code show a fitting process; it is properly implemented with a non linear least square method setting optionals parameters like starting points and upper and lower extremes of the fitting estimated parameters, high attention is

posed on the first parameter $b2$ that is necessarily contained in [0,C_nom(1)] where C_nom(1) is the maximum nominal contrast of the phantom.
The fitting was made using a sigmoid function because it optimizes at best the psychometric function (Figure 2.3), that, as already discussed, reflects the performance's degradation in terms of AUC/TPF and associated contrasts.
As mentioned above, in the case of the CHO model, the psychometric function from which the contrast of interest is derived, is described as a function of a range of contrasts defined between the extremes contrasts of the phantom used as test object and the AUC. The latter is estimated on the basis of Eq 2.12 which, to be calculated, requires as a known parameter (and in turn calculable) the SNR (Eq 2.13).
A slight difference can be read in the realisation of the psychometric curve when it refers to human observations. Compared to the $y$ = AUC used by the model to draw the psychometric curve, for the human the $y$ becomes equal to the fraction of true positives. This is because, if with the model the AUC is estimated on the basis of all analysed sub-images, this is not the case for human observations for which only a part of the signal and background ROIs are observed to give a result. Moreover, another consideration that should not be overlooked is the visibility threshold, which in the case of the human observer is difficult to estimate and cannot be varied because it depends on further internal and external variables. Therefore, the psychometric curves for each diameter are made using the TPF defined in a range between 0.5 and 1 as the AUC.
The used fitting function is:

$$y(x) = \frac{0.5}{(1 + exp(-b_1(x - b_2)))} + 0.5. \tag{3.2}$$

The decision to employ a reference function with the aspect of Eq. 3.2 rather than the classical form in Eq. 2.14 moves around the goal to find the researched values $y(x) \simeq 0.75$.
This operation was executed passing each contrast for each single diameter. In fact the fit takes as inputs the nominal contrasts of the test object and their AUC/TPF values. In this way, the psychometric fitted curves match to the number of analysed diameters.
Returning to the code, a vector $x$ was defined. It contains 50 linearly distributed values between the minimum and maximum values of the nominal contrasts associated to the specific diameter. It was used, also taking advantage of both parameters estimated by the fit, to calculate $y(x)$. Therefore, in order to extract the contrast at 75% of visibility, the behaviour of $y(x)$ was studied by first exploiting the correlation parameter $R^2$ as a binding condition.
If the fist condition is true, so the $R^2 > 0.75$ , it is assumed that the data and its associated fit were well executed. The main idea is to find which element in $x$ best satisfies the condition $y(x) \simeq 0.75$ in order to extract the corresponding contrast. Therefore, the process continues with the estimation of a vector containing the absolute value of $(y(x) - 0.75)$, from this vector its minimum was extracted and the final desired contrast was found as a single value or as a mean of all the contrasts with the same absolute minimum difference.

Otherwise, in the case of $R^2 < 0.75$ a different approach was used. This situation happens when the diameter is seen for every contrast or never detected for any contrast .
At this point, the nature of the data was investigated, speculating on its conformity. In particular, if the data is not accurately fitted, from observational experience, it may be due to the fact that the individual sets of values being considered are contained within narrow ranges. This happened when the diameter was seen for every contrast or never detected (hence the values were all close to 1 or to 0.5). This certainly does not adequately characterize the sigmoid defined by the equation in Eq 3.2 that needs data over a wide range of values.
This phenomenon can in turn be organised in three different cases. If the data are typically "high", i.e. there are AUC values in $[0.8, 1.0]$, it could mean that all details are always visible on average. If the data are "medium", i.e. with AUC in $[0.7, 0.8]$, then they fall on average in the range of 75% visibility. Finally, if the data tends to be "low", i.e. with AUC in $[0.5, 0.7]$ this means that the probability of appreciating more or less contrasted diameters is low.
Therefore, it was decided to use the measure of the median of $y(x)$ to define which class to attribute the data set. If the median is greater than 0.8 then the minimum nominal contrast is attributed, if it is between 0.7 and 0.8, then the central contrast of the $x$ vector is used, which is theoretically assumed to be the contrast associated with an AUC of 0.75, otherwise, if the median is less than 0.7, the maximum nominal contrast of the test object is attributed.
A schematic annotation about the followed fitting steps is given in Figure 3.8.

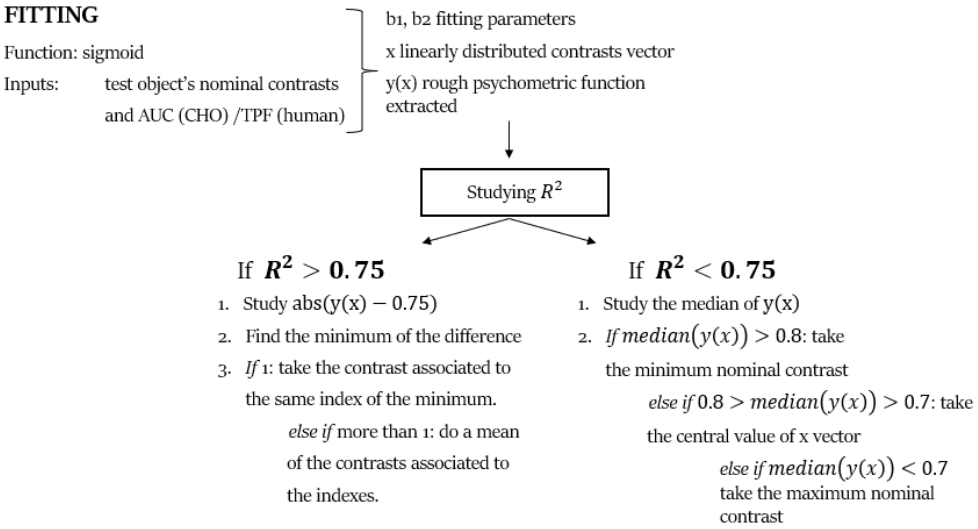



Figure 3.8: Schemae about the steps followed in human observation and CHO model procedures.

Following these steps, an accurate point selection can be logically performed. Once the contrasts of interest have been obtained for each diameter, the contrast detail curve can be estimated.
The study behind CD curves was realized with the aim to verify the applicability of

CHO model and to investigate about the relationship between human, model and especially the internal noise $\alpha$ component.

### 3.3.3 CHO $\alpha$ optimization

The choice of the internal noise component is a controversial issue that takes into account not only physical factors, but also and above all psychophysical factors. This noise component, as can be seen from Eq. 2.11 is additive to the model response and depending on the system used, the environmental conditions and the disposition of the human observer, has a wide variability.
Running the main script, from the graphical interface it is possible to set this parameter and estimate, applying the model, the CHO curve associated with the chosen $\alpha$.
The obtained curves were analysed with the aim to detect the preferential $\alpha$ which minimizes the distance between CHO curve and human curve.

```
def meanobs_minimization(df_cd, p_mean_std, txt_files, n_alpha, list_alphas, w):
    """
    Finds CHO curve which minimizes the mean observer curve

    Parameters
    ----------
    df_cd :         contrast-detil curve dataframe
    p_mean_std :    dataframe filled with mean and std of the same diameter
                    points seen by different observers
    txt_files :     txt dataframe
    n_alpha :       int value representing the number of rows where the word
                    'alpha' appears
    list_alphas :   alpha names series in dataframe
    w :             weighting factor

    Returns
    -------
    df_points_curvemin :    dataframe including the minimizing CD dataset referred
                            to averaged human observer curve
    df_table_points_curvemin :  dataframe of minimum distances curves and the
                                referred weighted distances

    """
    m_points_diff = np.ndarray((int(n_alpha), len(df_cd)))
    points_diff = minimization.tot_distances(
        m_points_diff,txt_files,n_alpha,df_cd,p_mean_std)

    m_points_sum_dist = np.ndarray((int(n_alpha), 1))
    df_points_sum_dist = minimization.tot_weighted_sum(
        m_points_sum_dist,points_diff,txt_files, n_alpha,w)

    df_points_curvemin, df_table_points_curvemin = minimization.minimum(
        df_cd, df_points_sum_dist, range(
        0, 1), list_alphas)

    return df_points_curvemin, df_table_points_curvemin
```

Once some different $\alpha$ curves are collected, the procedure consists in the estimation

of distances between contrast points associated to the same diameters of human and CHO curve. Then, after distances' sum calculation, it is defined which model curve best mimics the single human observer curves.
Because three different human observers are involved in this research, it was necessary to do a mean between points in order to obtain the human observed averaged curve. The same procedure of the single human curves was followed: the distances were estimated (*minimization.tot_distances*), the sum was performed (*minimization.tot_weighted_sum*) and the identification of the minimum (*minimization.minimum*) for studying the $\alpha$ for which the CHO curve is closest to the human was completed.
The shown extract of the code illustrates the function created to find the $\alpha$ that minimizes the CHO values distance from the values of averaged observer CD curve. To learn more about this function and its dependencies and for details of the followed procedure for extracting the best $\alpha$, please refer to the home made repository on GitHub[3].

### 3.3.4 Statistical inference

Once the best $\alpha$ is experimentally calculated it is important to prove this assumption from a statistical point of view.
What we expect is that the averaged human curve and the CHO curve with the $\alpha$ optimized can be compared in terms of statistical distribution.
One of the most known non-parametric hypothesis tests was chosen to study the data of interest, the Wilcoxon signed-rank test. This decision mainly depended on the characteristics of the samples in use.
We are dealing with a data set that is affected by subjectivity and it is made by a low number of elements (at maximum 12 elements referred to the chosen contrast referrred to each diameter). So, it is difficult observing a gaussian behaviour of the samples of interest, thus, just in case we preferred to use the Wilcoxon test.
The choice of the Wilcoxon test is therefore suitable for this nature of data. It requires paired samples and we expect to observe non statistical difference.
The null hypothesis $H_0$ means the two samples have the same distribution, the $H_1$ hypothesis statistically supports the opposit, that the samples have significant discrepancy about their distribution.
This behaviour is tested using the *p value*. Basically, it expresses how plausible it is that the observed data are obtained if the null hypothesis is true.
The significance of the test is fixed to 95%.
In the case of non-significant differences, it is useful and interesting to study how strong the predictive power of the applied model actually is. In particular, how much is the difference between the observed values (human) and those estimated by the model (CHO).
The determination's coefficient or $R^2$ is used in this project as a goodness' index and it measures the strength of linear relationship between the variables of the regression model. In order to obtain and contextualize this entity, the $r$ correlation coefficient and the p value associated to $r$ are calculated and given.

[3]https://github.com/Elena-Cantoni/CHOanalysis

If the determination coefficient is a quantity able to describe more globally the relationship between observations and model, the correlation coefficient is at the origin of understanding of this link because it provides more precise information, strictly speaking, about the form and essence of the data.

$$r = \frac{\sum[(x_i - \overline{x})(y_i - \overline{y})]}{\sqrt{\sum(x_i - \overline{x})^2 * \sum(y_i - \overline{y})^2}} \tag{3.3}$$

The closer $r$ is to 0, the weaker the linear relationship is. A positive $r$ value describes a positive correlation, so the variables grow in parallel, inversely in case of negative $r$. Values close to 1 or -1 indicate a strong correlation between data. In this project the statistics was performed in Python and it was used the *linregress* package to estimate the correlation coefficient, the p value and other parameters associated to the regression line. It was defined an appropriate 'correlation' function taking as inputs the CHO curves changing $\alpha$ and the the observer curves, that, depending on the purpose will be the curves. In our case the focus is on the correlation between the optimal $\alpha$ curve and the average curve of the humans.

```
def correlation(h_dataset, a_dataset, humans, alphas):
    """
    Estimates the correlation parameters obtained between human and CHO model response.
    A least-squares regression method is used.
    The results are collected in a dataframe.

    Parameters
    ----------
    h_dataset : CD human curve dataframe
    a_dataset : CD CHO curve dataframe
    humans : series of the title names of the observer curves
    alphas : series of the title names of the different CHO curves
             with different alpha

    Returns
    -------
    df_corr :   3-D dataframe containing correlation parameters
                (n° humans, n° parameters,°n alphas)

    """
    corr = np.ndarray(
        (((len(humans))), 6, len(alphas)))
    hum = -1
    for h in humans:
        hum += 1
        al = -1
        for a in alphas:
            al += 1
            slope, intercept, r_value, p_value, std_err = linregress(
                a_dataset[a], h_dataset[h])
            parameters = (slope, intercept, r_value, p_value, std_err)
            corr[hum, :5, al] = parameters
            r_sq = r_value**2
            corr[hum, -1, al] = r_sq
    df_corr = pd.DataFrame(corr[0, :, :], index=[
                           'slope', 'intercept', 'r value', 'p value', 'std', 'r2'])

    return df_corr
```

### 3.3.5 CD curve and dose radiation: a specific Figure of Merit

When a Quality Control of a medical system is performed, so if CHO CD curves are to be included in routine Quality Controls (QCs), further factors should be considered.
The purpose of these controls is to ensure that the performance of the machine is maintained over time, particularly from the point of view of patient radioprotection. The dose must be kept at a reasonably low level so that it is compatible with obtaining the required diagnostic information.
The pursuit of this objective ensures that the X-ray equipment subjected to such controls meets the minimum criteria of acceptability.
Foremost, using the results of the CHO model, first remarks can be done about eventual optimization of the angiographic machine and the establishment of its stability in the course of the time.
The optimization of the fluoroscopic machine is studied with the estimation of the $\alpha$ parameter that defines the index degradation specific of the system. This kind of operation allows to improve the nexus between machine and human by providing the appropriate tools to demonstrate how the feedback of the machine can be brought back to the human interpretation who is the final user of the equipement.
The study of stability is part of the branch of periodic quality control which proves the performance of the system of interest over time. This aspect is very important to analyse because the efficiency of a machine is also measured by its performance over time and, therefore, being able to monitor its performance is essential to know and responsibly control the advantages of the system used.
A very important component that determines the quality of a system is the entrance skin dose delivered to the patient, so it is necessary to study the amount of radiation for each acquisition.

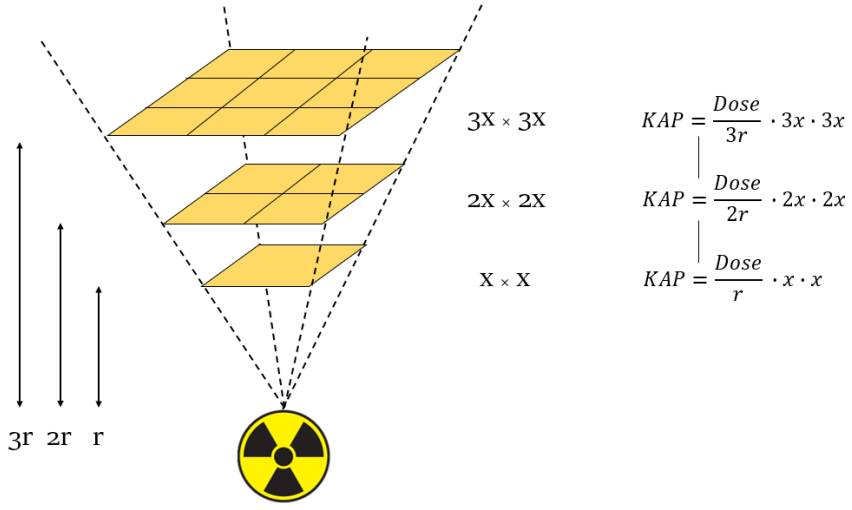



Figure 3.9: KAP physics.

A FOM was therefore designed to take into account both the dose delivered and the response returned by the contrast detail curves obtained from the CHO model.

$$FOM = \frac{1}{AUC_{CDC} \cdot KAP} \tag{3.4}$$

where $AUC_{CDC}$ stays for the area under the chosen contrast-detail curve and $KAP$ for Kerma Area Product. In this way, what we expect is that as the value of the $KAP$ and the area underneath the contrast-detail curve decrease, a higher FOM result is obtained. Thereby, it means a better performance of the machine at a fixed set up.
As it can be observed in Figure 3.9, the choice to use the $KAP$ rather than the Air Kerma stays in its property to be independent of the distance from the X-ray source, however keeping a certain dependency from the dimension's area. For this reason, with almost the same setup parameters, the KAP size was preferred as it is more stable in space and it is directly supplied by the machine.
If a figure of merit (of the shape in Eq. 3.5), dependent only on either KAP or AUC, were to be defined, it would be insufficient to provide a complete judgement on which system performs best on the required tasks.

$$FOM_{KAP} = \frac{1}{KAP} \qquad FOM_{AUC} = \frac{1}{AUC_{CDC}} \tag{3.5}$$

This is because it is not certain that an acquisition, returning a very low AUC, is acceptable in terms of the dose delivered, which, on the other hand, could be excessively high. The same can be said in the case of a very low KAP, which could be insufficient in minimum contrasts visibility (high AUC).

# Chapter 4

# Discussion and Results

The possibility of testing the Channelized Hotelling Observer model over a large series of fluoroscopic and angiographic images was a great advantage in conducting a robust and reliable study. Indeed, several considerations emerged with the aim of highlighting the relationship between the human observer's subjectivity and the system's objectivity. In the following sections, the CHO model's behaviour and its applicability to the dataset in use, or better to the dataset obtained acquiring images on a specific test object, will be treated and discussed. We will see that the choice of the phantom will have a substantial weight.

## 4.1 TO12: CHO applications

Starting from images acquired using the protocols described in Table 3.2 with TO12 test object, useful considerations for both fluoroscopy and angiography can be done.

### 4.1.1 CHO and human CD curves

Different $\alpha$ curves are extracted applying the CHO model for each protocol and three humans observed the details of interest. A visual and quantitative overview of the behaviour of these curves is shown in Figure 4.1.

First, it is crucial to clarify the relationship between the curves estimated by the model by varying the noise parameter $\alpha$. By increasing $\alpha$, the contrast detail curve gradually rises. Based on the observer model theory, a physical interpretation of this is quite intuitive: a noisy component is being added to the model template, which, depending on its magnitude, generates fluctuations that degrade its output. Consequently, it gradually decreases the model performance, and the contrast, which is 75% visibility, will be progressively higher. Therefore, it is expected that increasing $\alpha$ will raise the respective curves for the same protocol.
Looking at the contrast detail curves resulting from human observations, the variability dictated by the subjective component immediately jumps out. This phenomenon can be observed mainly for the smaller diameters, which can be more challenging to recognise when observing due to their dimensions, despite being the most contrasted diameters as annotated in Figure 3.1.

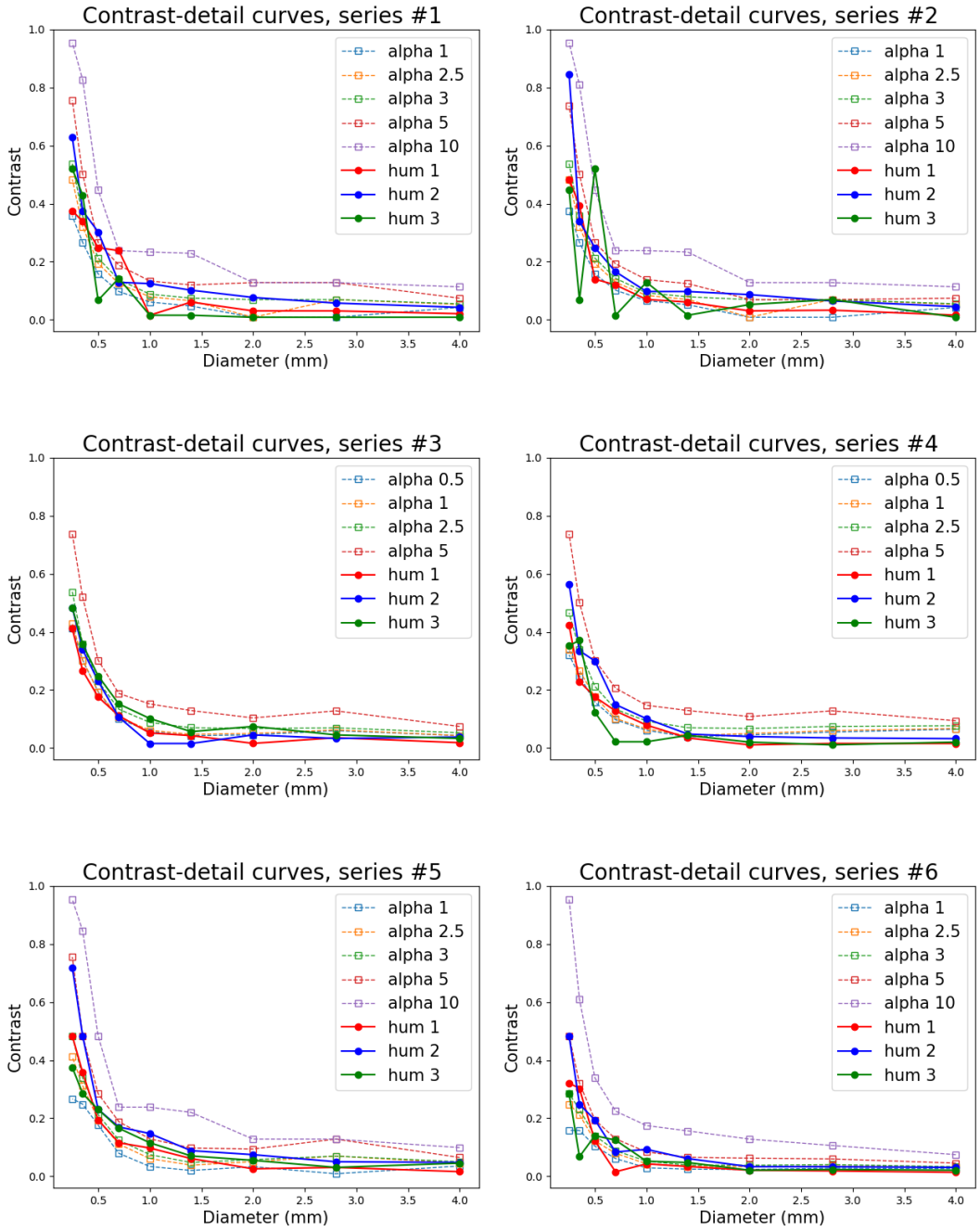



Figure 4.1: TO12 contrast-detail curves for each studied acquisition, each plot refers to a specific protocol named with the same number of series used in Table 3.2. Dashed lines represent CHO CD curves, solid lines indicate the human observer CD curves. The colours distinguish curves by varying $\alpha$ and humans.

So, it happens because as the diameter size decreases, it gets closer and closer to the pixel size. The observer may be deceived into recognising as a signal a detail that is actually a background and vice versa is likely to identify as noise what is really a signal.

The contrasts limit range is another crucial aspect that should be taken into account. To appreciate the trend of the curves just described, they must be contained within the range of contrasts defined by the model. In this case, the nominal maximum and minimum contrasts associated with the test object were chosen as upper and lower limits.
The contrast detail curves obtained from the TO12 phantom disks are within the expected range (Figure 4.2). For both the human and the model, differences in the curves can be appreciated. Therefore, it is possible to conduct more in-depth analysis from a statistical point of view.

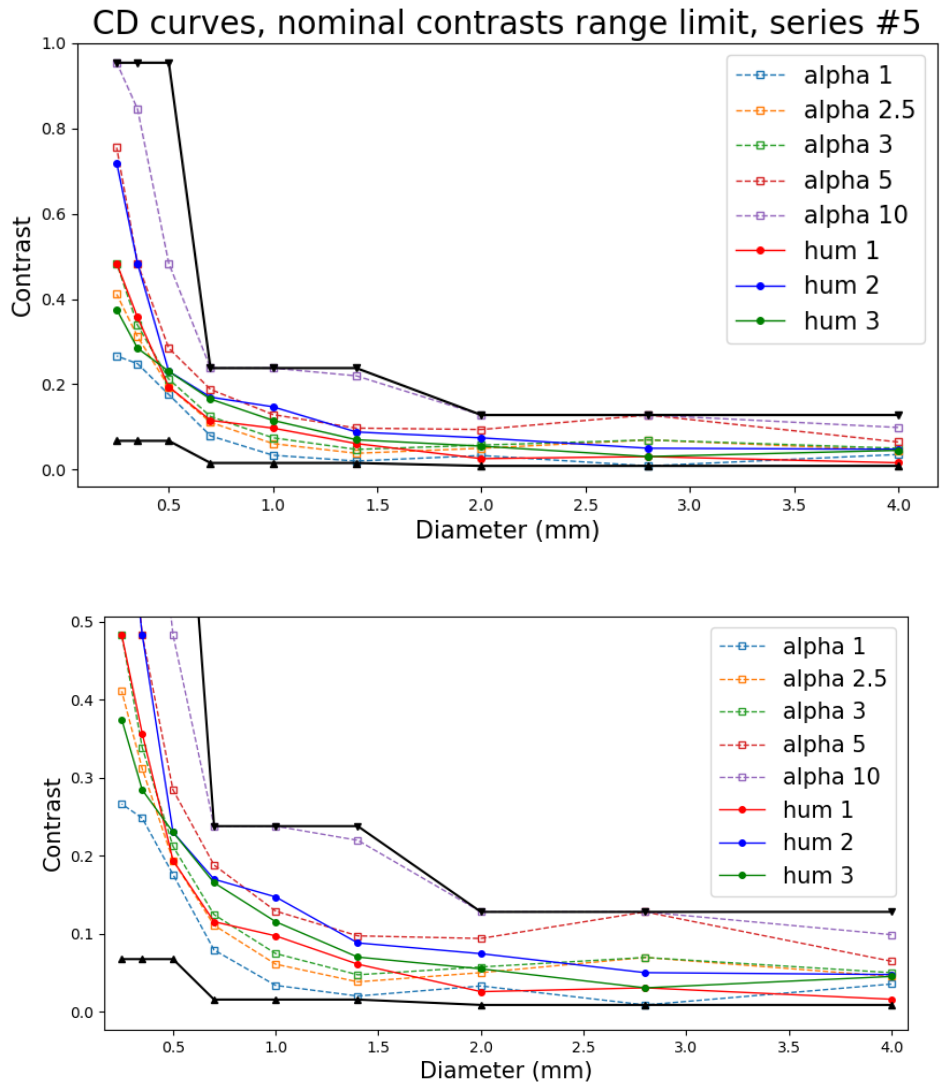



Figure 4.2: Example of contrast detail curves with contrasts limit range. The solid black lines indicate from the top to the bottom respectively the maximum nominal contrasts and the minimum nominal contrasts. The second graph shows an enlargement on the most critical area where the curves are concentrated within the range of contrasts.

### 4.1.2 Average of humans and choice of optimal $\alpha$

The psychophysical interpretation is strictly related to the human observer's choice, its decision's time and the external conditions that may influence and alter his decision. From the graphs previously illustrated, it was possible to see how the curves of humans show almost random peaks and troughs, which are the variability due to the subjective decision of each observer.
An averaged human curve is calculated to reduce the fluctuations associated with the humans. In this project, the final human CD curve is the mean from three different observers. Increasing the number of observing subjects, we expect the fluctuations are reduced.

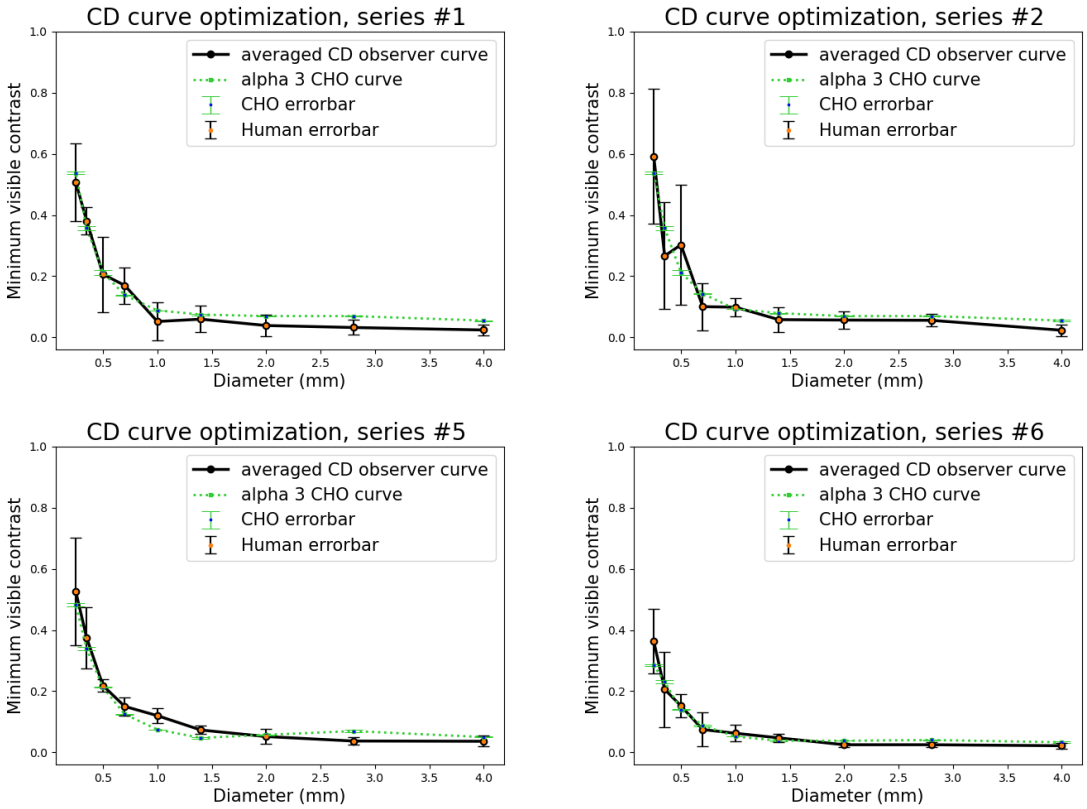



Figure 4.3: TO12 fluoroscopic acquisitions. The preferred $\alpha$ is selected for each protocol. The solid black line is the CD averaged curve with the associated error, the dotted line in green represents the CHO curve estimated with the optimal $\alpha$ with its uncertainty.

The Figures 4.3 and 4.4 show the averaged human and the optimal CHO CD curves described by a specific $\alpha$ in case of fluoroscopic and angiographic acquisitions, respectively. Quantitatively, $\alpha$ could be considered a calibration factor for the system to be tuned on the human. From model theory, a random variable is magnified by this 'tuning factor': if it increases, the random component has a greater weight over the model's response and vice versa.
The optimal $\alpha$ is chosen for each protocol. In the case of fluoroscopy (Figure 4.3), varying the parameters like the level of dose or the pre-filtering, the same resulting

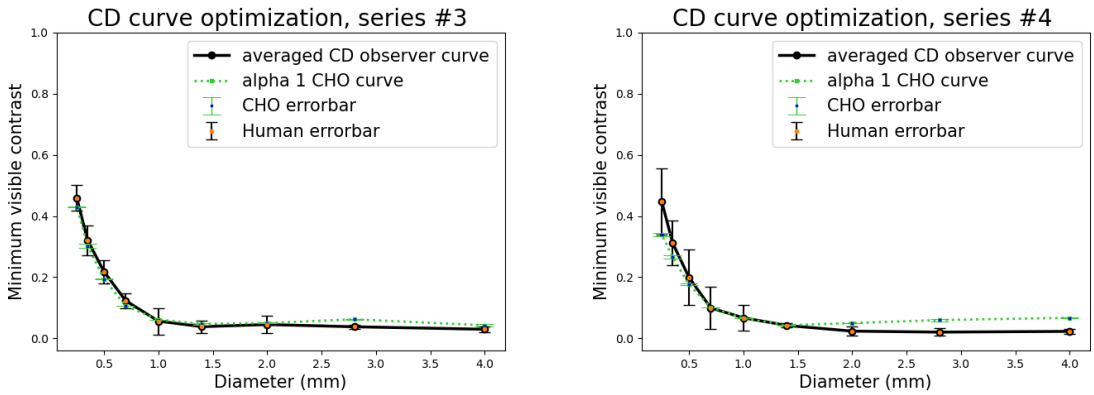



Figure 4.4: TO12 angiographic acquisitions. The preferred $\alpha$ is selected for each protocol. The solid black line is the CD averaged curve with the associated error, the dotted line in green represents the CHO curve estimated with the optimal $\alpha$ with its uncertainty.

$\alpha$ is, on average, estimated.
For angiographic acquisitions, the alpha estimated for the various protocols is generally the same value.
It is interesting to understand why this happens: further insights about the behaviour of $\alpha$ noise parameter led to consider this 'tuning' factor as a degrading correction to add to the overall medical system. It is independent of the single parameters chosen for each protocol. But, changing the mode acquisition, so passing from fluoroscopy to angiography, the system requires different methods of application, and in terms of internal noise, the machine reacts differently. The $\alpha$ factor changes varying the acquisition type, as shown in Table 4.1.

Table 4.1: Optimal $\alpha$ values.

| **Acquisition Type** | **optimal** $\alpha$ |
|---|---|
| Fluoroscopy | 3 |
| Angiography | 1 |

Both the CD curves of the average human and the model have an error bar. The error is calculated on the contrast at 75% visibility associated with each specific diameter. The uncertainty referred to the CD averaged human is the standard deviation:

$$\sqrt{\frac{\sum_i^n (x_i - \overline{x})^2}{n-1}} \tag{4.1}$$

where $x_i$ corresponds to the sample of the single observer, in this case $n = 3$.
About the CD CHO curve, its uncertainties comes from the least square fit used to estimate the 75% of visibility contrast points. So, the mean square error is used.

### 4.1.3 Validation with statistics

The visual condition of making assumptions about the similarity of distribution of the two compared curves finding the optimal noise is validated with statistic.
The Wilcoxon test is used to prove that the two samples (human and model) have the same distribution at the 95% of confidence level.
If $p < 0.05$ the differences in the observed data are statistically significant, if $p > 0.05$ the $H_0$ hypothesis is sustained and cannot be rejected. In Table 4.2 the obtained p values are reported.

Table 4.2: Wilcoxon test, significance.

| **WILCOXON** | **#1** | **#2** | **#3** | **#4** | **#5** | **#6** |
|---|---|---|---|---|---|---|
| **p value** | 0.16 | 0.57 | 0.20 | 0.98 | 0.16 | 0.50 |

As it can be seen, the probabilities are all larger than 0.05. Then, all the differences between the human and the model CD curves are not statistically significant in the 95% confidence interval. From this consideration, the contrast detail curves estimated by the model and obtained from human observation are not statically different. Therefore, with a 5% possibility of error, they can be attributed to the same distribution.
Once the absence of significance between the datasets is tested, it is interesting to understand how the contrasts, estimated by the model and those that come out from the human observation, are related.
Table 4.3 shows the correlation coefficient, the p value and the determination coefficient ($R^2$) attributed to each protocol.

Table 4.3: Linear correlation factors between 75% visibility contrasts for human and model curves extracted by each protocol.

| # | ***r*** | ***p value*** | $\boldsymbol{R^2}$ |
|---|---|---|---|
| **1** | 0.99 | $0.42{\cdot}10^{-7}$ | 0.98 |
| **2** | 0.96 | $0.39{\cdot}10^{-4}$ | 0.92 |
| **3** | 0.99 | $0.89{\cdot}10^{-9}$ | 0.99 |
| **4** | 0.99 | $0.12{\cdot}10^{-7}$ | 0.99 |
| **5** | 0.99 | $0.26{\cdot}10^{-6}$ | 0.98 |
| **6** | 0.98 | $0.80{\cdot}10^{-5}$ | 0.95 |

In this case, every p values confirm that the sought linear relationship between human and model is significant, the probability is for each protocol greater than the established threshold ($p > 0.05$). The correlation between the two variables is strong and near 1, so the visual hypothesis that the two variables increase or decrease proportionally is confirmed. The $R^2$ that is always bigger than 0.92 overall validates the high relationship between the variables.
As observed on the graphs in Figure 4.3 and 4.4, and confirmed by the experimental

data, the paired samples are linked; the example reported in Figure 4.5 shows this linear relationship.

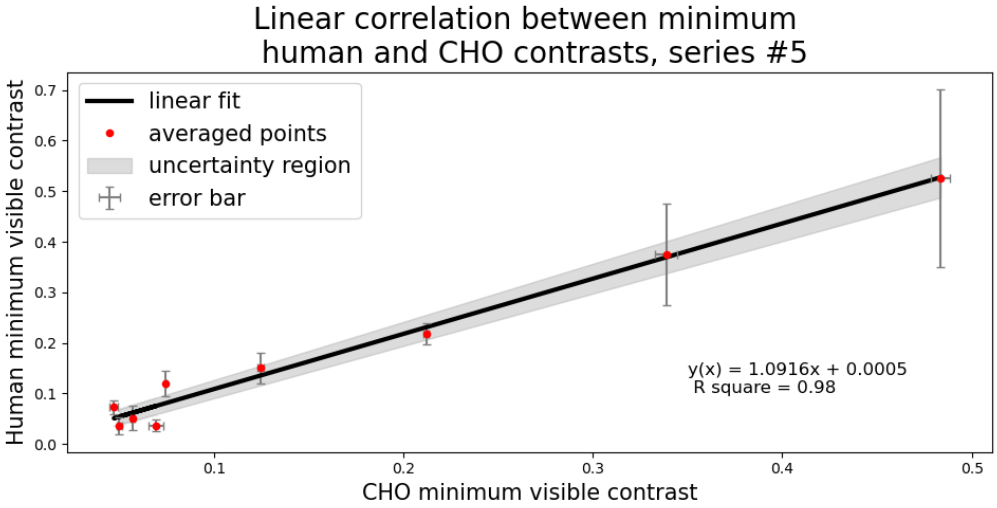



Figure 4.5: Linear relationship between human and model contrasts. The single points are the contrasts referred to a specific diameter with dimension, from left to right,space of $4mm$ to $0.25mm$.

### 4.1.4 From human to CHO model

The statistical evidence of correspondence between the CD curves estimated by the human observers and the model's application allows increased confidence in the model, which is statistically compatible with human results. Assuming this, it is, therefore, possible to carry out an optimization analysis aimed at automatically identifying the noise factor needed by a specific fluoroscopy system to faithfully reproduce (within statistical limits) the subjectivity of human vision.
Of course, the validity of the model depends on the chosen tasks and operability conditions, i.e., in the present case, for uniform and static object and background.
Taking as reference the $\alpha$ parameters listed in Table 4.1 for the two acquisition's types, a more in-depth study of the relationship between different protocols was carried out. As mentioned above, the trend of the contrast detail curves is quite explanatory regarding the quality of the images in terms of signal degradation. It is around this topic that further considerations will be made.
First of all, the Wilcoxon test was applied; this time not between the human averaged curve and the optimal CHO curve but between the CHO CD curves of different protocols calculated with the same internal noise component $\alpha$. The p value refers to the probability of belonging to the same distribution.
Following what the p values tell us, we can make considerations about the different protocols used by comparing the performance of their CHO curves.

Table 4.4: p value measurements given by Wilcoxon test between couples of $\alpha = 3$ (fluoroscopy) and $\alpha = 1$ (angiography) curves estimated for different protocols. The significant measures are highlighted in bold.

| # | **1** | **2** | **5** | **6** | **3** | **4** |
|---|---|---|---|---|---|---|
| **1** | 1 | 0.102 | **0.018** | **0.004** | **0.004** | **0.012** |
| **2** | 0.102 | 1 | **0.018** | **0.004** | **0.004** | **0.012** |
| **5** | **0.018** | **0.018** | 1 | **0.004** | **0.012** | **0.039** |
| **6** | **0.004** | **0.004** | **0.004** | 1 | **0.004** | **0.004** |
| **3** | **0.004** | **0.004** | **0.012** | **0.004** | 1 | 0.263 |
| **4** | **0.012** | **0.012** | **0.039** | **0.004** | 0.263 | 1 |

The first cross-comparison was made with fluoroscopic protocols #1 and #2, which returned a non-significant result, i.e. it was possible to see that their associated CD curves were statistically distributed in the same way.

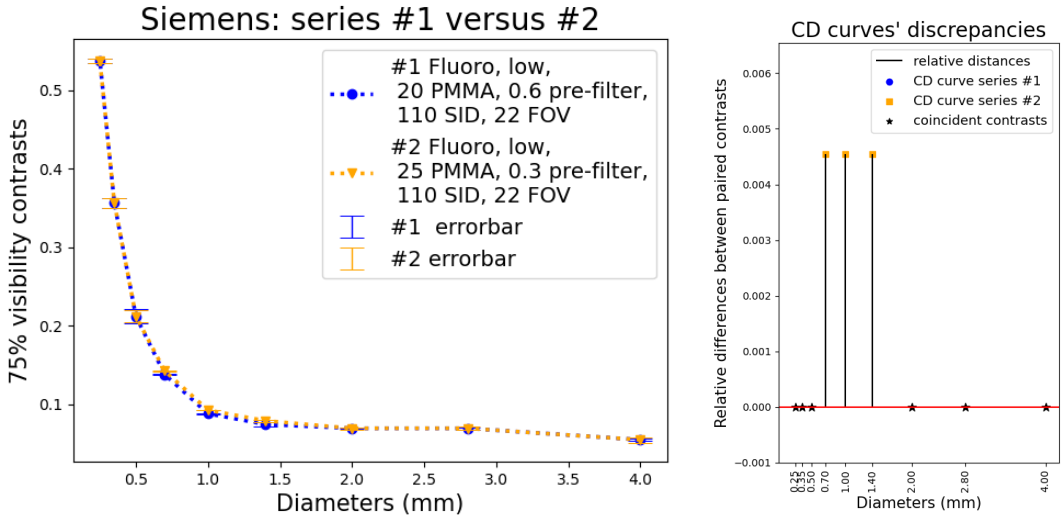



Figure 4.6: On the left: CD curves of series #1 and series #2. Each curve has its error. On the right: discrepancy's graph where the black lines indicate the distances between the paired contrasts of the curves and the coloured dots and squares indicate the upper curve for that specific point contrast. The black stars, if presents, represent the coincident values.

The graph in Figure 4.6 confirms this. It immediately shows an affinity between the two contrast detail curves that present slight discrepancies of the order of a thousandth in correspondence to the details with a diameter of $0.7, 1.0, 1.4$ mm.

Imposing the acquisition time to $30s$, the $KAPs$ of series #1 is three times smaller than that of series #2. The dose (or KAP) increase is closely linked to the pre-filtration of the fluoroscopy system that, in the studied case, is double for acquisition #1 rather than acquisition #2. If the copper plate is thinner, the radiation beam will be less shielded and altered. The 5 cm difference in PMMA thickness between the two setups is also not negligible. At the light of these results, our analysis confirms the fluoroscopy system's capability to compensate for different "patient" attenuations. For fluoroscopy, all the other results show statistical significance.

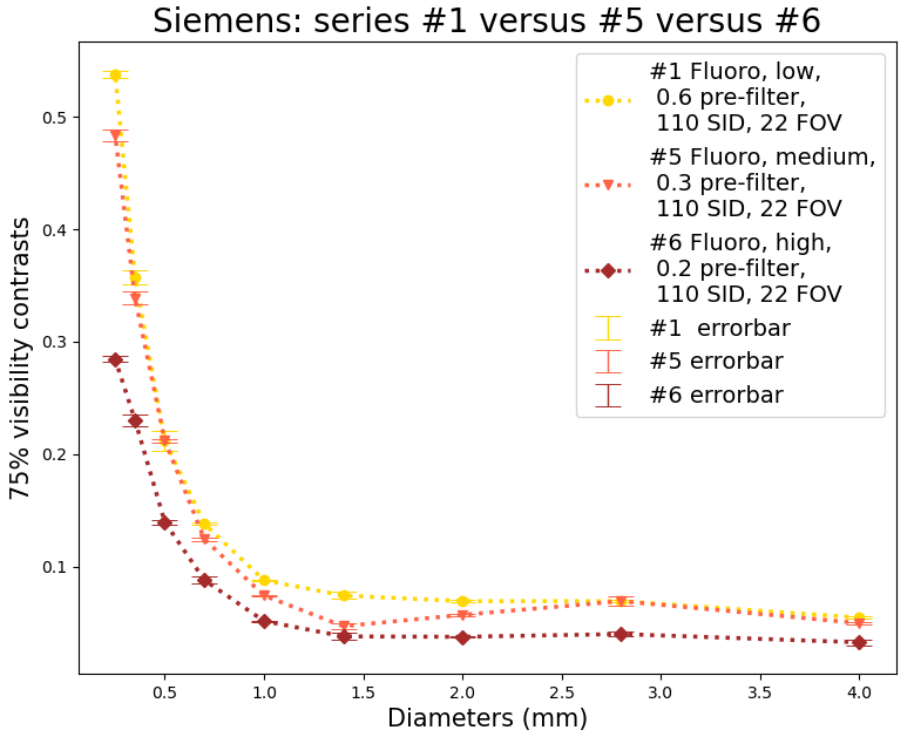



Figure 4.7: CD curves of acquisitions #1, #5 and #6. Each curve has its error.

From Figure 4.7, it is evident that the curves differ significantly and that it is easier to understand which acquisition is preferable based on the parameters set.
The various series illustrated differ in thickness of the pre-filtering and in dose level setting.
The first factor is partly influential on the degradation of the contrast detail curves, as described in the previous comparison between series #1 and series #2 in Figure 4.6.
In addition, this time, another essential factor affects the performance of protocols; this component is the pre-selected dose level that appears to be much relevant in CD curves discrimination.
When this parameter varies, there are sufficiently significant discrepancies that allow us to see that the CD curve with the best performance is the one with the highest dose, but at the expense of the amount of $KAP$, which, despite this, is very high compared to the others.

The angiographic acquisitions require a lower $\alpha$ for the optimization. Series #3 and #4 did not show a statistical difference.
In Figure 4.8 the differences achieved between these curves oscillate around zero; there is not an evident higher curve.

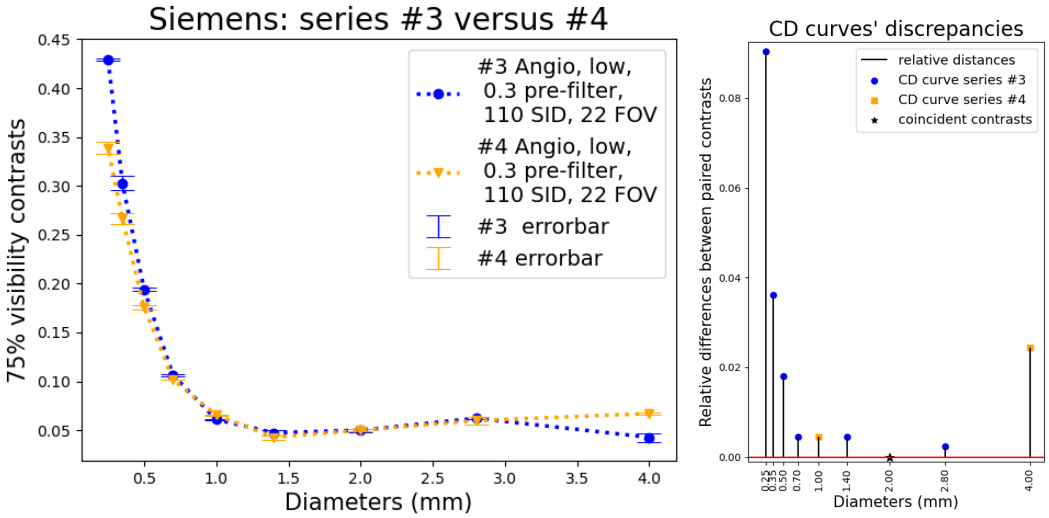



Figure 4.8: On the left: CD curves of series #3 and series #4. Each curve has its error. On the right: discrepancy's graph where the black lines indicate the distances between the paired contrasts of the curves and the coloured dots and squares indicate the upper curve for that specific point contrast. The black stars, if presents, represent the coincident values.

Only discrepancies observed between the coupled contrasts are more substantial in the proximity of extremities (i.e. concerning the smallest and largest diameters studied). In the case of the largest diameter, series #3 is slightly better than #4; conversely, for the smallest diameter, the performance is reversed. The #3 and #4 protocols are characterized by the same parameters fixed by the user like the FOV, the SID and the pre-filtration, so this comparison wants to be a countercheck of the Siemens system stability under the same acquisition parameters.
It is interesting to see how is the relationship between angiographic and fluoroscopic protocols. The comparison is made using CHO curves estimated with the respective optimal $\alpha$ parameters as shown in Figure 4.9 the angiographic CD curve has $\alpha = 1$ and the other CD curves have $\alpha = 3$.
The fluoroscopy and angiography curves' differences are statistically significant.
The contrast detail curve of series #4 falls between the high and medium dose fluoroscopic curves. Visually, it is evident that the curve obtained from series #6 is the best performing of the Siemens system in terms of signal detection.
It is essential to note that the KAP values are considerably higher than for the other acquisitions in both series #6 and angiographic series #4. This topic will be dealt with in more detail in the next section.

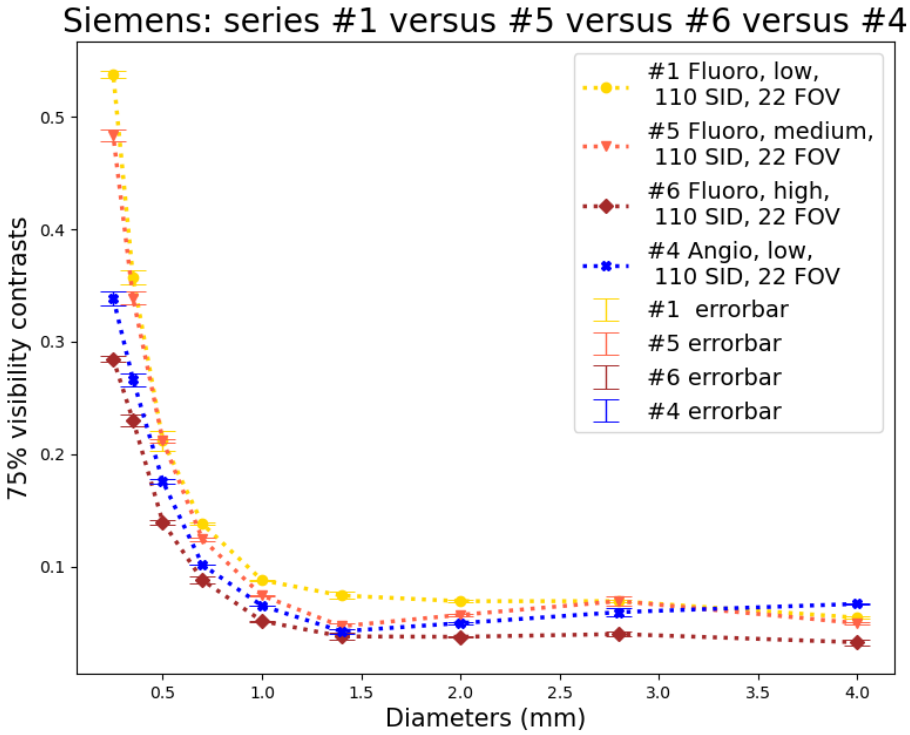



Figure 4.9: CD curves of fluoroscopic acquisitions #1, #5, #6 and of angiographic protocol #4. Each curve has its error.

### 4.1.5 FOM: the metrics of TO12 tasks performance

As the previous sections have shown, not only the ability to identify details of interest at the lowest possible contrast makes the fluoroscopy system the best. In the clinic, many limiting aspects need to be considered, such as the operating environment and the ionising radiation used to acquire the digital images, which can interfere with the patient's body and the active personnel. These are some of the factors that, if neglected, can lead to a non-optimal exam in terms of dose.
Therefore, returning to the performance of our angiography system, it is not strictly necessary for the machine to seek optimal conditions at the expense of the amount of dose delivered.
In this thesis work, multiple acquisitions were performed, highlighting how it is crucial to know the incoming radiation dose and be clear about how this affects the quality of the digital image. It was necessary to resort to metrics that would express the system's performance under examination (Eq. 3.4), taking into account the ionising radiation and the quality of the images.
In Table 4.5 the KAP and the area under the CD curve have been reported. The first one is weighed on a time fixed at 30 seconds, knowing the frame rate and the number of images for each acquisition; the second one is calculated experimentally starting from the CD curves obtained at the optimal $\alpha$. The FOM that is extracted acts as an indicator of the quality and check of the specific acquisition.

Table 4.5: Figure of Merit estimation for Siemens acquisitions

| # | AUC | KAP ($dGy \cdot cm^2$) | #images | fr/s | #images per 30s | KAP 30s | $FOM_{KAP}$ | $FOM_{AUC}$ | FOM |
|---|---|---|---|---|---|---|---|---|---|
| **1** | 0.36 | 6.65 | 500 | 15 | 450 | 5.98 | 0.17 | 2.8 | 0.46 |
| **2** | 0.37 | 21.1 | 530 | 15 | 450 | 17.9 | 0.056 | 2.7 | 0.15 |
| **3** | 0.29 | 11.7 | 301 | 15 | 450 | 17.5 | 0.057 | 3.5 | 0.20 |
| **4** | 0.29 | 11.6 | 301 | 15 | 450 | 17.3 | 0.058 | 3.5 | 0.20 |
| **5** | 0.32 | 9.61 | 535 | 15 | 450 | 8.08 | 0.12 | 3.1 | 0.38 |
| **6** | 0.21 | 19.6 | 527 | 15 | 450 | 16.7 | 0.060 | 4.7 | 0.28 |

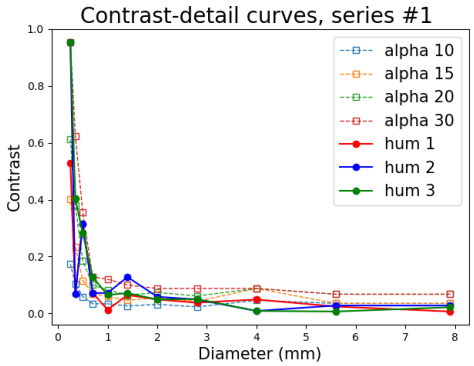
Looking at the data collected, if the FOMs were only dependent on KAP or only reliant on AUC (Eq. 3.5), we would have that the best performing protocols differ; this is because both metrics lack completeness.
If the FOM depends only on the KAP, the best performant acquisition results to be the first; it does not not go in the same way if the FOM relies on the AUC: in this case, the most performant acquisition seems to be the sixth followed by the fourth, third and fifth.
Now, also considering the FOM to judge which acquisition is better, we can declare that between series 1 and 2 of Figure 4.6, it is chosen the first series. Among acquisitions #1, #5 and #6 of Figure 4.7 the highest FOM is that of series #1; between the two angiographic acquisitions #3 and #4, we can say the final performance is similar (198 for series 3 and 0.2024 for series 4) as confirmation of system's stability. The chosen FOM, therefore, takes both factors into account. As can be seen, acquisition #1, described by a low dose level, 0.6 mm of pre-filter, FOV 22, SID 110, seems to be the most performing concerning signal detection and input dose for the Siemens angiographic system.

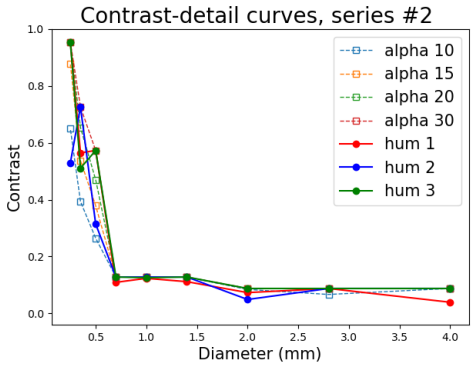
## 4.2 TO16: CHO applications

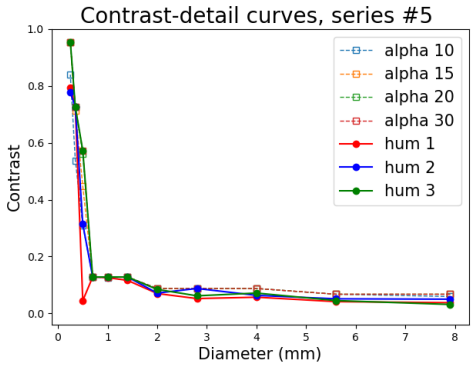
A significant part of the project was to acquire fluoroscopic and angiographic images of the TO16 test object at the Reggio Emilia hospital and process them to conduct an applicability study of the CHO model on these datasets as well.
In chapter 3 the systems used to make the acquisitions were described and, in the next sections, a discussion will be made regarding the contrast detail curves of the humans and of the CHO model, distinguishing between fluoroscopic and anangiographic acquisitions. Therefore several considerations and results on the test object used on the two systems Philips and GE will be illustrated.

### 4.2.1 CHO and human CD curves

The same steps followed for TO12 acquisitions on Siemens angiographic system were used for Philips and GE equipments, as well.
Three human observers were invoked, and some contrast detail curves, varying $\alpha$ parameter, were collected. The graphs in Figures 4.10 and 4.11 show how the human and CHO model-estimated curves are still related from a visual point of view. It is worth noting the first impression made when observing these trends. For both GE and Philips, the CD curves of the model that optimize the humans tend to be higher or, instead, require a higher internal noise component than the contrast detail curves estimated on the Siemens machine with a different test object.
The plots of Figures 4.10 and 4.11 are given as an example to show in general terms the trend of CD curves changing some parameters. The individual paired graphs on the same line refer to the most similar acquisitions for Philips on the left and GE on the right. Angiographic and fluoroscopic acquisitions were shown paired from top to bottom as follows: low and medium dose and higher FOV with 1 mmCu (series #1 for Philips and GE), high dose and lower FOV with 1 mmCu (series #2 for Philips and series #6 for GE), high dose and higher FOV with 1mmCu (series #5 for Philips and series #4 for GE) and finally normal dose and lower FOV with 0.5 mmCu (series #11 for Philips and series #10 for GE).
Indeed, another characteristic that emerges from the curve is its shape that,with respect to the CD curves in Figure 4.1, are more sharp-edged, assuming an L-shape. This behaviour aroused suspicion from the beginning, and further investigations were carried out to gain a better understanding.
At this point, to understand what governs the quality trends observed in the CD curves, it is better to proceed step by step, distinguishing between fluoroscopy and angiography analysis.
The quality of analysed images could be a relevant factor to observe a significant discrepancy between the two methods, especially in a situation like this one that returns anomalous results at first sight.



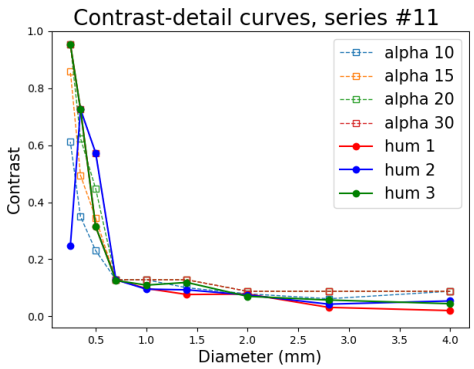
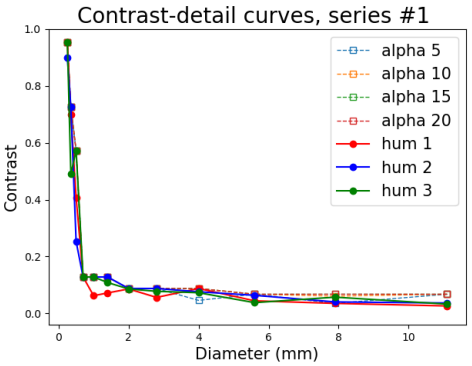
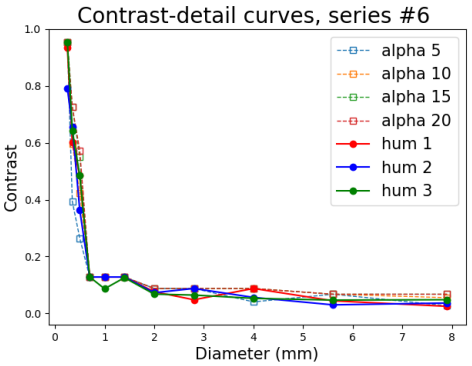
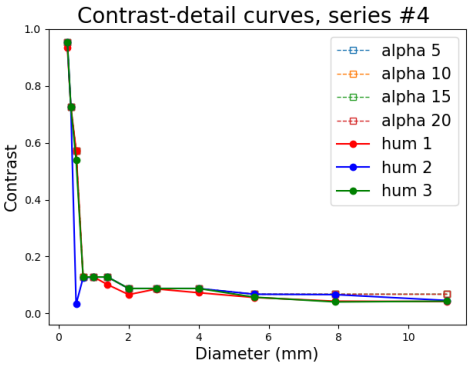




Figure 4.10: Philips TO16 CD curves for each studied acquisition, each plot refers to a specific protocol named with the same number of series used in Table 3.4. Dashed lines represent CHO CD curves, solid lines indicate the human observer CD curves. The colours distinguish curves by varying $\alpha$ and humans.



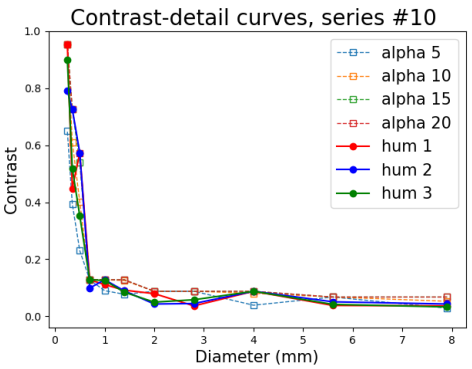




Figure 4.11: GE TO16 CD curves for each studied acquisition, each plot refers to a specific protocol named with the same number of series used in Table 3.6. Dashed lines represent CHO CD curves, solid lines indicate the human observer CD curves. The colours distinguish curves by varying $\alpha$ and humans.

### TO16 fluoroscopy acquisitions and CHO model

As a first approach, we tried to understand how the estimated fluoroscopic curves are distributed within the range of nominal contrasts fixed a priori and listed in Figures 3.2.
The figure 4.12 shows the contrast detail curves of two acquisitions, one by Philips and one by GE, within the range of nominal contrasts of the TO16 test object.
The first consideration that can be made is that the fluoroscopic CD curves tend

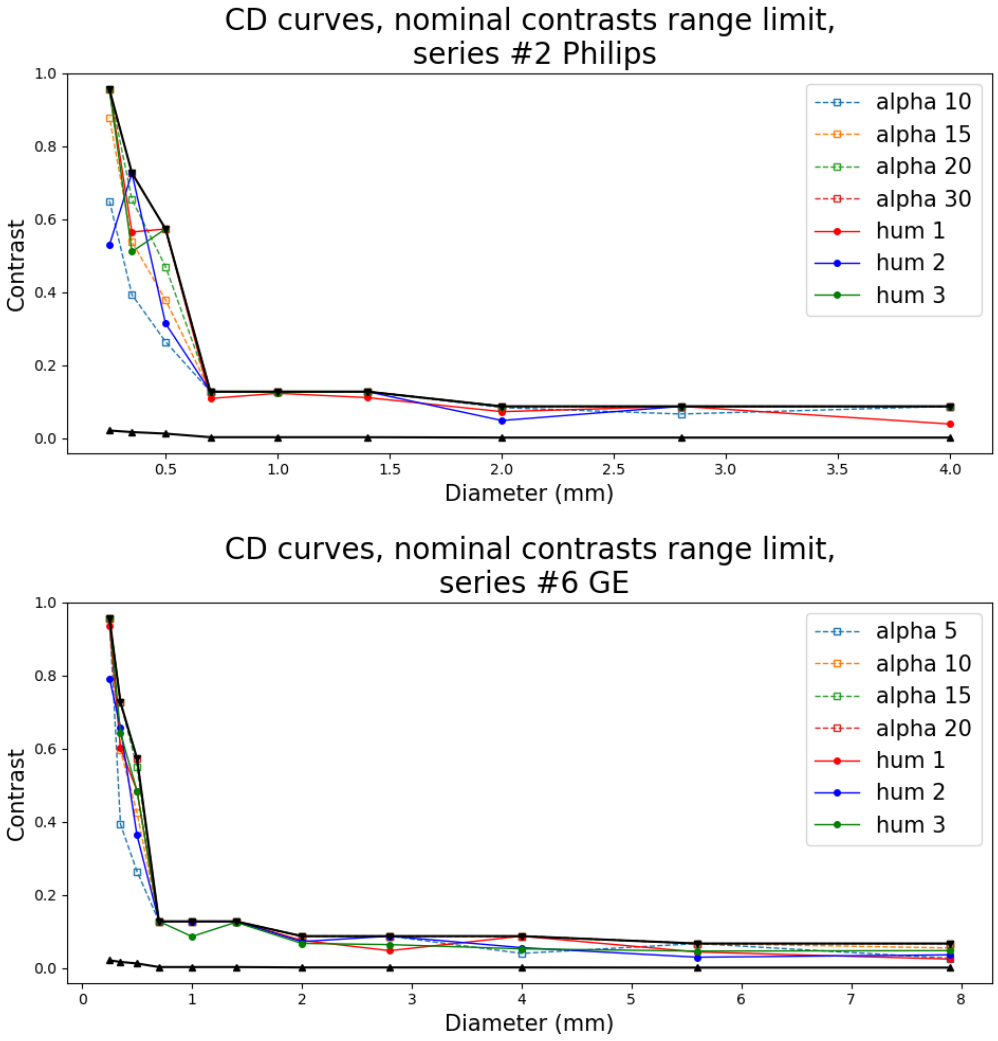



Figure 4.12: Fluoroscopic protocols (Philips series #2 on the top and GE series #6 on the bottom) are shown in the range of TO16 nominal contrasts.

to approach the upper edge of the range, thus leading to an evident squashing at the upper limit of the curves with higher alpha. Why this happens was the question investigated to give a physical explanation to this phenomenon.
It is important to remember that the model has internal limits imposed so that the curves obtained cannot go outside the range of nominal contrasts, as mentioned in section 3.3.2. These conditions have been set based on the test object used. Test

object's details depend on the diameter, and they are visible in a range of contrasts well defined by the manufacturers in fixed conditions (e.g. 70 kV, 1mmCu). If CD curves were to exceed the limit, above or below, the considerations that can be made about the setup, parameters used, and the limits of instruments used would be misleading and inconsistent since contrasts outside that range are not present in the phantom.

### 4.2.2 Test object's choice and usage limits

The reason why the curves flatten out at the upper limit and therefore the minimum nominal contrasts are much higher, both for humans and for the CHO model, lies outside the physics of the model.
The focus was, thus, on the external conditions of data acquisition and the instruments used. The choice of the type of acquisition setup was the first topic of discussion to understand if there were any hindering conditions during the acquisition phase. Operating under "good geometry" conditions (i.e. positioning the phantom close to the detector and placing the copper filtering layer completely adjacent to the RX source) helps achieve the best possible spatial management to obtain the highest performance achievable in practice.
In addition, it is fair to note that in the case of the TO12 acquisitions, the setup geometry was worse than ours. Given the evidence of qualitatively better signal detection images on which the model is effectively applied despite worse geometric conditions, one feels inclined to reject the hypothesis that geometry is the triggering cause for these acquisitions with altered performance.
Before challenging the systems' performance, for which the study with the CHO model aims to prove its stability and its optimisation, we tried to understand whether other instruments have a relevant weight on the performance of the angiographic system.
This thesis focussed the attention to the choice of the test object.
TO16, like the TO12 Leeds phantom, are typically used to perform contrast detail analyses for many systems. They are pretty similar in shape but differ one from the other for physical characteristics. In Tables 4.6 and 4.7 the diameters dimensions and their contrast measures are shown moving from the highest contrasted disk (top) to the lowest one (bottom) for each diameter.

Table 4.6: TO16 nominal contrasts declared by the manufacturer. Maximum and minimum nominal contrasts are written in italic and bold.

| **TO16** | **Diam 1** | **Diam 2** | **Diam 3** | **Diam 4** | **Diam 5** | **Diam 6** | **Diam 7** | **Diam 8** | **Diam 9** | **Diam 10** | **Diam 11** | **Diam 12** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **mm** | 11.1 | 7.9 | 5.6 | 4.0 | 2.8 | 2 | 1.4 | 1.0 | 0.7 | 0.5 | 0.35 | 0.25 |
| **Contr 1** | ***0.0674*** | ***0.0674*** | ***0.0674*** | ***0.0876*** | ***0.0876*** | ***0.0876*** | ***0.128*** | ***0.128*** | ***0.128*** | ***0.573*** | ***0.726*** | ***0.954*** |
| **Contr 2** | 0.045 | 0.045 | 0.045 | 0.0674 | 0.0674 | 0.0674 | 0.0876 | 0.0876 | 0.0876 | 0.36 | 0.573 | 0.726 |
| **Contr 3** | 0.0322 | 0.0322 | 0.0322 | 0.045 | 0.045 | 0.045 | 0.0674 | 0.0674 | 0.0674 | 0.238 | 0.36 | 0.573 |
| **Contr 4** | 0.0215 | 0.0215 | 0.0215 | 0.0322 | 0.0322 | 0.0322 | 0.045 | 0.045 | 0.045 | 0.167 | 0.238 | 0.36 |
| **Contr 5** | 0.0172 | 0.0172 | 0.0172 | 0.0215 | 0.0215 | 0.0215 | 0.0322 | 0.0322 | 0.0322 | 0.128 | 0.167 | 0.238 |
| **Contr 6** | 0.013 | 0.013 | 0.013 | 0.0172 | 0.0172 | 0.0172 | 0.0215 | 0.0215 | 0.0215 | 0.0876 | 0.128 | 0.167 |
| **Contr 7** | 0.0086 | 0.0086 | 0.0086 | 0.013 | 0.013 | 0.013 | 0.0172 | 0.0172 | 0.0172 | 0.0674 | 0.0876 | 0.128 |
| **Contr 8** | 0.0066 | 0.0066 | 0.0066 | 0.0086 | 0.0086 | 0.0086 | 0.013 | 0.013 | 0.013 | 0.045 | 0.0674 | 0.0876 |
| **Contr 9** | 0.0043 | 0.0043 | 0.0043 | 0.0066 | 0.0066 | 0.0066 | 0.0086 | 0.0086 | 0.0086 | 0.0322 | 0.045 | 0.0674 |
| **Contr 10** | 0.0032 | 0.0032 | 0.0032 | 0.0043 | 0.0043 | 0.0043 | 0.0066 | 0.0066 | 0.0066 | 0.0215 | 0.0322 | 0.045 |
| **Contr 11** | 0.0021 | 0.0021 | 0.0021 | 0.0032 | 0.0032 | 0.0032 | 0.0043 | 0.0043 | 0.0043 | 0.0172 | 0.0215 | 0.0322 |
| **Contr 12** | ***0.0016*** | ***0.0016*** | ***0.0016*** | ***0.0021*** | ***0.0021*** | ***0.0021*** | ***0.0032*** | ***0.0032*** | ***0.0032*** | ***0.013*** | ***0.0172*** | ***0.0215*** |

Table 4.7: TO12 nominal contrasts declared by the manufacturer. Maximum and minimum nominal contrasts are written in italic and bold.

| **TO12** | **Diam 1** | **Diam 2** | **Diam 3** | **Diam 4** | **Diam 5** | **Diam 6** | **Diam 7** | **Diam 8** | **Diam 9** | **Diam 10** | **Diam 11** | **Diam 12** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **mm** | 11.1 | 7.9 | 5.6 | 4.0 | 2.8 | 2 | 1.4 | 1.0 | 0.7 | 0.5 | 0.35 | 0.25 |
| **Contr 1** | ***0.0674*** | ***0.0674*** | ***0.0674*** | ***0.128*** | ***0.128*** | ***0.128*** | ***0.238*** | ***0.238*** | ***0.238*** | ***0.954*** | ***0.954*** | ***0.954*** |
| **Contr 2** | 0.045 | 0.045 | 0.045 | 0.0876 | 0.0876 | 0.0876 | 0.167 | 0.167 | 0.167 | 0.726 | 0.726 | 0.726 |
| **Contr 3** | 0.0322 | 0.0322 | 0.0322 | 0.0674 | 0.0674 | 0.0674 | 0.128 | 0.128 | 0.128 | 0.496 | 0.496 | 0.496 |
| **Contr 4** | 0.0215 | 0.0215 | 0.0215 | 0.045 | 0.045 | 0.045 | 0.0876 | 0.0876 | 0.0876 | 0.36 | 0.36 | 0.36 |
| **Contr 5** | 0.0155 | 0.0155 | 0.0155 | 0.0322 | 0.0322 | 0.0322 | 0.0674 | 0.0674 | 0.0674 | 0.238 | 0.238 | 0.238 |
| **Contr 6** | 0.0108 | 0.0108 | 0.0108 | 0.0215 | 0.0215 | 0.0215 | 0.045 | 0.045 | 0.045 | 0.167 | 0.167 | 0.167 |
| **Contr 7** | 0.0086 | 0.0086 | 0.0086 | 0.0155 | 0.0155 | 0.0155 | 0.0322 | 0.0322 | 0.0322 | 0.128 | 0.128 | 0.128 |
| **Contr 8** | 0.0065 | 0.0065 | 0.0065 | 0.0108 | 0.0108 | 0.0108 | 0.0215 | 0.0215 | 0.0215 | 0.0876 | 0.0876 | 0.0876 |
| **Contr 9** | ***0.0043*** | ***0.0043*** | ***0.0043*** | ***0.0086*** | ***0.0086*** | ***0.0086*** | ***0.0155*** | ***0.0155*** | ***0.0155*** | ***0.0674*** | ***0.0674*** | ***0.0674*** |

The first different property noted from the tables concerns the coupling of contrasts of adjacent diameters. In the case of TO12, this maintains the same range of contrasts for triplets of diameters; the same is not valid for the three smaller diameters of TO16 which differ from each other going, gradually increasing in radius, to reduce their maximum nominal contrast that touches for the third smallest diameter a 0.573 of TO16 to the detriment of 0.954 of TO12.
Except for the three smallest diameters which are characterised by different nominal contrasts, TO16 also starts to behave like triplets. The substantial difference, which probably lies at the root of the difficulty in recognising where there is and where there is no signal, is in the value assumed by the maximum nominal contrasts for the medium diameters, in particular the diameters of 4, 2.8, 2 millimetres and then 1.4, 1, 0.7 mm. This aspect highlights how the diameters in TO16 ( 0.128 for diameter 9) are much less contrasted than those in TO12 ( 0.238 for diameter 9).

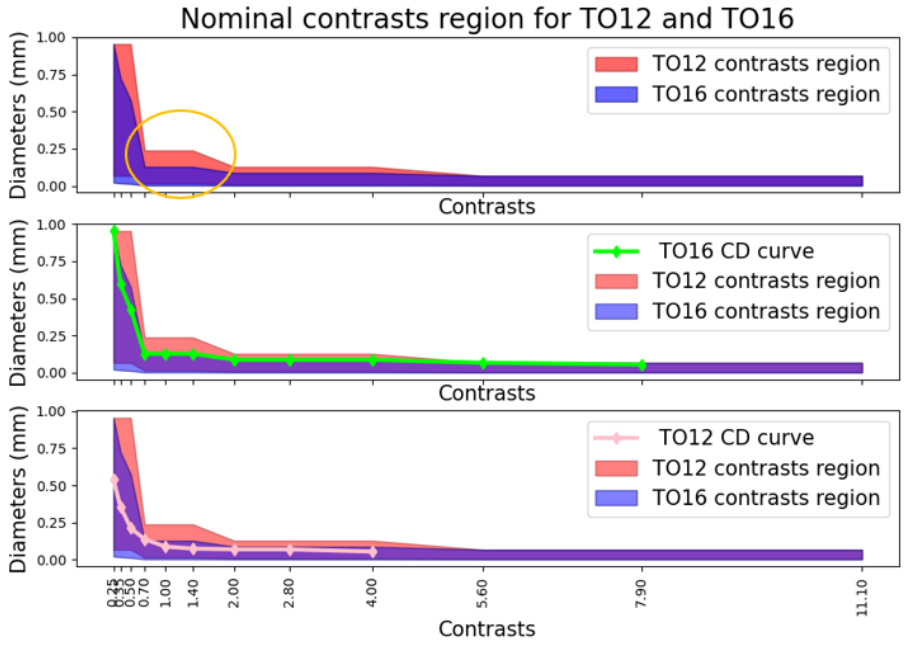



Figure 4.13: Nominal contrasts regions for TO16 and TO12 and respective contrast detail curve trends.

This behaviour can be observed better in Figure 4.13. The orange circle indicates the plateau that causes the edge step that can be observed from the experimental data. The second and third plots show what was said theoretically: because of the drastic reduction of maximum contrasts in the proximity of the fourth smallest diameter, the contrast detail curve in green of the second sub-figure is forced to adapt its trend to the maximal values available.
This fact happens because, in the phase of observation or CHO application, the human/model cannot correctly detect a signal for most of the cases. Instead, in the case of TO12, this behaviour is not observed, first of all, because the range of contrasts are higher and then because the proper positive fractions are usually comprised between the minimum and maximum threshold.
Those limitations due to the internal characteristics of the Leeds TO16 test object shrink the field of application of our CHO model. An attempt was made to estimate the optimal $\alpha$.
The procedure followed was the same as in TO12, but the considerations were more critical. The contrast detail curves at optimal $\alpha$ that have been chosen are respectively for Philips $\alpha = 20$, and for GE $\alpha = 10$. Still, this estimation loses reliability and trustworthiness since many of the points that constitute the contrast detail curves are exactly the values of the nominal maximum contrasts. Figures 4.12 confirm this.
This trend, both of the human curves and of the curves estimated by the model, prevents us from continuing in the analysis, especially the statistical one, because the curves, approaching the maximum, that is one of the nominal contrasts, get closer and closer and could erroneously resemble each other more than they are.

### TO16 angiography acquisitions and CHO model

As far as the angiography acquisitions are concerned, we follow the same procedure performed in fluoroscopy.
A prior check was carried out on the curves' shape to ensure that they were distinguishable within limits imposed by the model. In support of this, the graphs in Figure 4.14 are shown.
Although the curves estimated by the GE machine seem to be close to the upper limit, we made sure, by directly observing the experimental data, that the minimum visibility contrasts were not the same as the nominal maximum ones. Since, therefore, the points constituting the curves were distinguishable, even though they were close to the maximum, it was possible to proceed with the analysis.
In fact, thanks to the next considerations, we will see that the model applied to angiographic acquisitions can distinguish the curves better than in fluoroscopy.
The choice of optimal alpha will be consistent with the data.

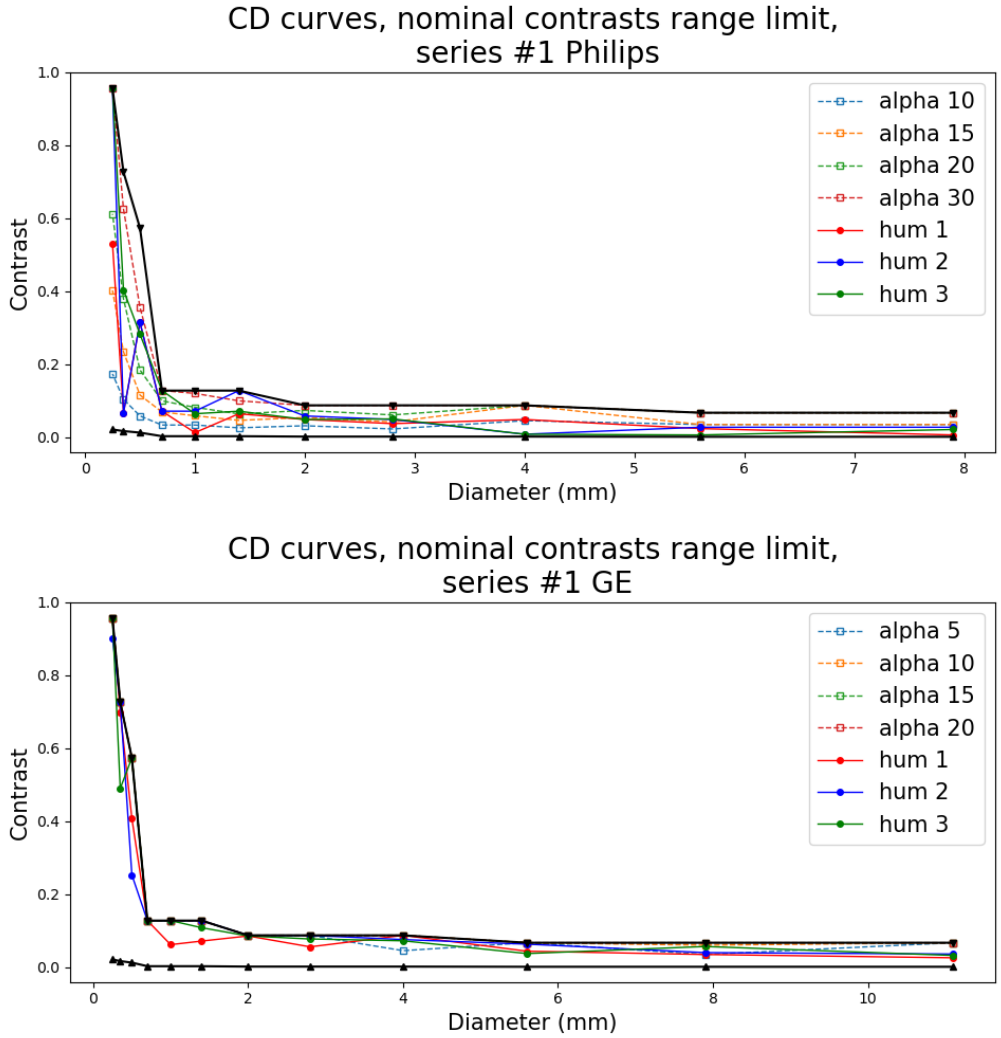



Figure 4.14: Angiographic protocols (Philips series #1 on the top and GE series #1 on the bottom) are shown in the range of TO16 nominal contrasts.

### 4.2.3 $\alpha$ optimization and statistics on angiographic protocols

An $\alpha$ optimization analysis is conducted with the aim to find the proper internal noise parameter which best performs the human averaged curve.
In Figure 4.15 the optimal selected CD curve associated to the averaged human curve are shown. Each curve has its proper uncertainty estimated as the standard deviation for human averaged CD curve and as the mean square error for CHO CD curve. On the basis of this, it is possible to say that the internal noise parameters that minimize the distances between the curves are those in Table 4.8.

Table 4.8: Optimal $\alpha$ values for Angiography using TO16 test object

| **Angiographic system** | **optimal $\alpha$** |
|---|---|
| Philips | 20 |
| GE | 10 |

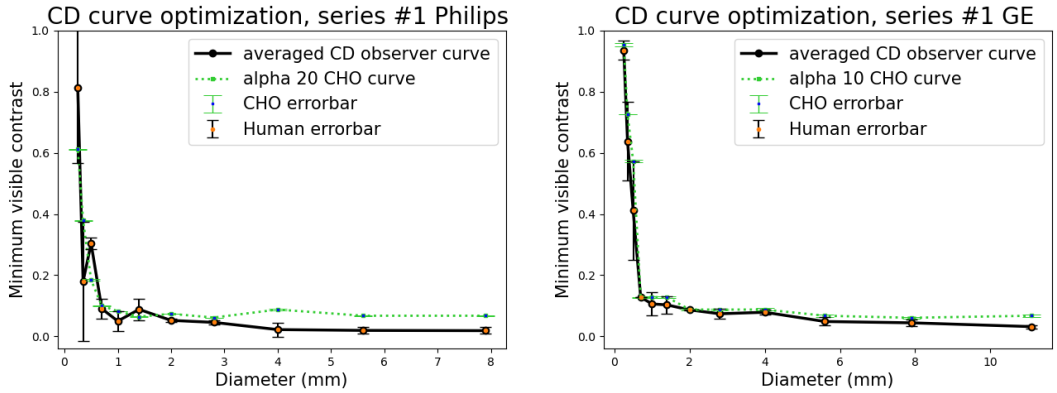



Figure 4.15: TO16 angiographic acquisitions. The preferred alpha is selected for each protocol. In black there is the CD averaged curve with the associated error, the dotted line in green represents the CHO curve estimated with the optimal $\alpha$.

It is not strange that the $\alpha$ differs changing the angiographic machine. Each system is realized to perform the tasks appropriately, but this condition does not impose that all the mechanisms must be obtained with the same fixed parameters. For example, each machine can have its proper fields of views (e.g. FOV like $22 \times 22$ cm and $31 \times 31$ cm for Philips and $20 \times 20$ cm and $32 \times 32$ cm for GE), its proper pixel dimension or its proper number of bits per pixel (8 or 16 bits).
It is more than expected that different noise parameters are estimated between various systems. It is also usual that, even with the same machine, different $\alpha$ values can be obtained depending on the type of acquisition (e.g. fluoroscopy and angiography). So, the alpha becomes an indicator of the percentage noise perceived by the human eye in reaction to the results given by the machine.
All that remains is to look for additional support from statistics. The Wilcoxon test, which estimates the p-value intending to study whether or not a result is significant, is always used. Still, in the 95% confidence boundary, the estimated p-values, calculated relating human and CHO optimal CD curves of the same acquisition, do not show statistically significant differences (Table 4.9).

Table 4.9: Wilcoxon test, significance in angiographic acquisition.

| **WILCOXON** | **Philips** | **GE** |
|---|---|---|
| **p value** | 0.47 | 0.07 |

Now it is interesting to see how the curves obtained in angiographic acquisitions are statistically related. In this phase, a new CD curve referred to series #2 in GE is introduced; about this one, any observer has visualized the images, so the taken $\alpha$ is that used for GE acquisition #1 because the angiographic machine and the type of acquisition are the same. Thus, the probability that the curves belong statistically to the same distribution is expressed in terms of p values gotten by the Wilcoxon test.

Table 4.10: p value measurements given by Wilcoxon test between couples of $\alpha = 20$ (Philips) and $\alpha = 10$ (GE) curves are estimated for different protocols.

| | **Philips** | **GE series #1** | **GE series #2** |
|---|---|---|---|
| **Philips** | 1 | 0.114 | 0.139 |
| **GE series #1** | 0.114 | 1 | 0.317 |
| **GE series #2** | 0.139 | 0.317 | 1 |

The non-significance allows us to study how the curves are disposed and related to each other: which protocols are preferable in terms of lower contrast detail visibility and how set parameters impact on each acquisition. The Figure 4.16 is very explicative, and a series of considerations can be made. Refreshing the given

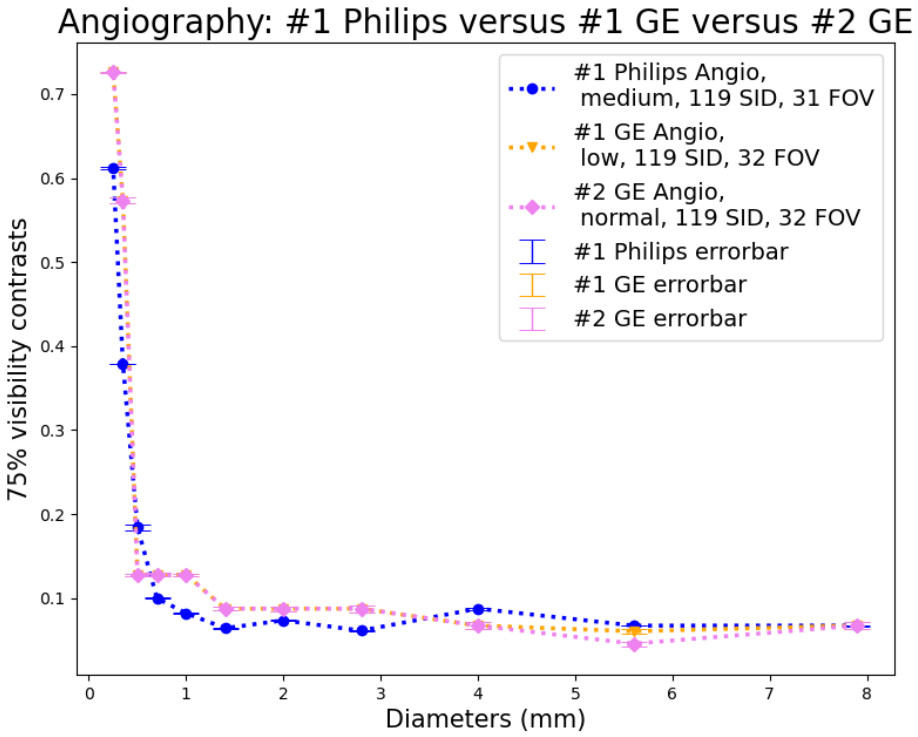



Figure 4.16: CD curves of protocols 1 acquired with Philips and protocols 1 and 2 of GE. Each curve has its error.

parameters of each acquisition (also listed in Tables 3.6 and 3.4), series #1 and #2 of GE system are remarkably similar: those protocols have the same $FOV = 32 \times 32$ cm, the same $SID = 119$ cm and the same copper filtration (1 mmCu), they differ only for the dose level that, in case of series #1 is low and for series #2 is normal. This difference is slightly visible in the figure, near the second largest diameter, where we can see that GE's series #1 needs a somewhat higher contrast to detect the detail than GE's series #2. In principle, the two curves are almost identical.

The CD angiographic curve of the Philips machine, on the other hand, is good for small diameters, which means that it can discriminate the signal better than GE. From a visual point of view, the profile of the Philips curve tends to be lower than the GE ones.
A spontaneous consideration arises on the $\alpha$ parameter: a higher $\alpha$ does not mean a worse performance; this factor is independent on the angiographic equipment performance; it represents a subjective noise index that we need to add to the model to approach the human visual perception.

### 4.2.4 TO16 angiographic FOM

To conclude, the figure of merit performed in the TO12 analysis is also calculated for TO16. In light of this fact, we can observe how angiographic acquisitions are more subjected to a higher KAP than fluoroscopic protocols. For this fact, however, we expect the FOM (dependent only on the KAP) is generally lower.
But, because a complete interpretation of the acquisition performance is a trade-off between the entrance dose and the values of the AUC of each CD curve, it is desirable to estimate a good FOM (Table 4.11) which describes how the series of interests perform the required tasks.

Table 4.11: Figure of Merit estimation for Angiographic acquisitions in GE and Philips.

| | **AUC** | **KAP** $(dGy \cdot cm^2)$ | **#images** | **fr/s** | **#images per 30s** | **KAP 30s** | **FOM**$_{KAP}$ | **FOM**$_{AUC}$ | **FOM** |
|---|---|---|---|---|---|---|---|---|---|
| **PHILIPS** | 0.64 | 7.04 | 422 | 6 | 180 | 30.0 | 0.33 | 1.6 | 0.52 |
| **GE 1** | 1.1 | 2.66 | 451 | 6 | 180 | 1.06 | 0.94 | 0.95 | 0.89 |
| **GE2** | 1.0 | 4.75 | 451 | 6 | 180 | 1.89 | 0.53 | 0.98 | 0.52 |

This comparison between angiographic acquisitions performed with different machines but with almost similar protocols provides a good starting point for reasoning in terms of acquisition stability. Philips and GE equipment, using similar parameters, show performances not significantly different. As a result, presumably, the two manufacturers have set almost similar low contrast performance for this protocol.
Based on the figures of merit, the Philips system, such as high-dose GE are less performing than low-dose GE. In terms of overall performance, the dose delivered is discriminating.

## 4.3 Limitation and future works

In the light of the facts, it must not be forgotten that the study was conducted under conditions far from those of the clinic.
The use of a phantom is undoubtedly facilitating and less onerous, especially for a research study in its beginnings. It must also be acknowledged that there are limiting factors that cannot be overlooked when comparing it with clinical application. Firstly, with a test object we operate in static conditions and not in dynamic conditions as if there were a patient. In addition, the phantom is not iodinated, as is the case in the clinic, where an iodinated contrast agent is given to the patient; it leads to a different attenuation of the radiation. Another discriminating factor concerns the observed details: in the case of the test object we know the shape and size, so we are operating in an easier condition than in the clinic.
It is important to stress, however, that to test a model it is useful to set degrees of freedom in order to have more control over its physical behaviour. Therefore, testing the optimization of a protocol and its stability on a test object does not necessarily give the same results as in clinic, but it is certainly very useful in the control phase to give a concrete and consistent answer on the choice of acquisition parameters.
However, the field remains open with the future aim of bringing this study into the clinic.

# Conclusions

In medical physics, image quality is therefore a much-discussed topic, which depends not only on objective physical properties, but also on subjective judgements of the observer. Understanding how to interpret visual perceptions is a very tempting challenge that moves to the idea of studying observer models that predict human performance in visual tasks.
In this thesis project a wide range of acquisitions' protocols and setups were involved allowing to conduct different kind of studies based on the application of Channelized Hotelling Observer model, depending on the chosen angiographic system and the parameters set.
The possibility of having many sets of data acquired on different test objects made it possible to do specific considerations about the quality control assessment of each system-phantom combination. Moreover, with the same test object, it was possible to compare some similar acquisitions (for choice of parameters and protocol) of different angiographic machines.
The pairing of the TO12 test object and the Siemens angiographic system has provided good grounds for drawing significant considerations about its signal detection performance. For both fluoroscopy and angiography, the application of the CHO model gave consistent results. Contrast detail curves optimising the mean CD curve of the observers were estimated and approved by the statistics with $\alpha = 3$ for fluoroscopic acquisitions, and with $\alpha = 1$ for angiographic acquisitions. Therefore, studies on the optimisation of the used protocol and on the stability of the machine for various acquisitions were carried out. Siemens' series #1 (fluoroscopic acquisition, PMMA thickness = 20 cm, low dose, SID = 110 cm, FOV = 22×22 cm and 0.6 mm of pre-filtering) and #2 (fluoroscopic acquisition, PMMA thickness = 25 cm, low dose, SID = 110 cm, FOV = 22×22 cm and 0.3 mm of pre-filtering) do not show statistically significant differences in their CD curves, it means that the system is performing well and, with the same cardiac protocol, the results are compatible. In fact, as expected, the dose rises as the PMMA thickness layer increases. As far as different dose level series are concerned, the results appear to be significantly different despite the fact that the same SID and FOV are set for all the considered acquisitions. This aspect makes us say that the dose level is highly discriminating on the performance of the system in terms of CD area. The signal detection is easier in high dose series #6 (with high dose, FOV = 22×22 cm, SID = 119 cm and 0.2 mm of pre-filtering) than for series #5 (that, despite series #6, differs for medium dose's level and 0.3 of pre-filtering) and #1 (that, despite series #6, differs for low dose's level and 0.6 of pre-filtering) but at the expense of KAP which is abundant and it degrades the final figure of merit. To test the stability

of the machine, the angiographic acquisitions #3 and #4, with the same parameters, proved to be non-significantly different and this is a good result to verify that the machine does not behave abnormally when the measurement is repeated under the same conditions (only the internal parameters such as voltage, current and pre-filtering, vary because they are set automatically by the system). At the light of facts, for Siemens system with TO16 phantom ,the best FOM is that of #1 fluoroscopic acquisition.
The analysis performed on TO16 test object images, acquired from the angiographic Philips and GE systems, responded differently to the CHO model application. The causes were reached in the internal characteristics of the phantom in used. The TO16 details appeared to be less detectable because their nominal contrasts are set to be potentially lower that the contrasts defined for the TO12 test object. This configuration therefore is cause of major limitation which is much evident in fluoroscopy as it has less contrasted and noisier images than angiography, primarily because the modalities delivers a lower dose to the patient. In fluoroscopy, the human found it more difficult to identify phantom details and the model could not compute a suitable alpha noise parameter from which to estimate the optimal contrast detail curve. The model, at this level of contrast, is unable to discriminate between theoretical contrast detail curves and this prevents us from making meaningful considerations.
On the other hand, interesting trends can be discerned from angiography, which tends to record higher doses than fluoroscopy. The CHO model works and optimal $\alpha$ can be estimated: alpha = 10 for the GE system and alpha = 20 for the Philips system. From CD curves, it seems that all the angiographic acquisitions are consistently performant. Thus, imposing the same parameters, GE and Philips return a certain stability in acquisition when setting the parameters. In terms of figure of merit, the most performant angiographic acquisition is the series #1 of GE because it supplies a lower amount of dose with a similar CD curve behaviour.
However, this investigation, conducted using test objects, has its limitations mainly due to the fact that it differs from clinical acquisition in factors such as the static nature of the acquisitions with respect to the dynamic nature in the case of a patient. Therefore, it is not guaranteed that it is the same thing to optimize a protocol and test its stability in the clinic, but it is certainly of great advantage to use the phantom in a quality control phase, whether periodic or not, to test the stability and optimisation of the machine and to provide consistent results and make substantial considerations on the choice of parameters that can be changed in the acquisition phases.

# Bibliography


[1] Barrett HH, Myers KJ. *Foundations of Image Science.* New York: Wiley & Sons, 2004.

[2] Bertolini, Marco, et al. *Characterization of GE discovery IGS 740 angiography system by means of channelized Hotelling observer (CHO).* Physics in Medicine & Biology 64.9 (2019): 095002.

[3] Beutel, Jacob, Harold L. Kundel, and Richard L. Van Metter. *Handbook of medical imaging.* Vol. 1. Spie Press, 2000.

[4] Bogacz, Rafal, et al. *The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks.* Psychological review 113.4 (2006): 700.

[5] Bradley, William G. *History of medical imaging.* Proceedings of the American Philosophical Society 152.3 (2008): 349-361.

[6] Burgess, Arthur E. *The Rose model, revisited.* JOSA A 16.3 (1999): 633-646.

[7] Bushberg, Jerrold T., and John M.Boone. *The essential physics of medical imaging.* Lippincott Williams & Wilkins, 2011.

[8] Engelke, Ulrich, Anthony Maeder, and Hans-Jürgen Zepernick. *Human observer confidence in image quality assessment.* Signal Processing: Image Communication 27.9 (2012): 935-947.

[9] He, Xin, and Subok Park. *Model observers in medical imaging research.* Theranostics 3.10 (2013): 774.

[10] Kingdom, F. A., Prins N. (2010) *Psychophysics: A practical introduction*, Elsevier: London.

[11] Kotter, E., and M. Langer. *Digital radiography with large-area flat-panel detectors.* European radiology 12.10 (2002): 2562-2570.

[12] Krupinski, Elizabeth A. *The importance of perception research in medical imaging.* Radiation medicine 18.6 (2000): 329-334.

[13] Krupinski, Elizabeth A., and Yulei Jiang. *Anniversary paper: evaluation of medical imaging systems.* Medical physics 35.2 (2008): 645-659.

[14] Mahesh, Mahadevappa. *Fluoroscopy: patient radiation exposure issues.* Radiographics 21.4 (2001): 1033-1045.

[15] Martin, C. J., P. F. Sharp, and D. G. Sutton. *Measurement of image quality in diagnostic radiology.* Applied radiation and isotopes 50.1 (1999): 21-38.

[16] Miller, Donald L., et al. *Quality improvement guidelines for recording patient radiation dose in the medical record for fluoroscopically guided procedures.* Journal of Vascular and Interventional Radiology 23.1 (2012): 11-18.

[17] Nickoloff, Edward Lee. *AAPM/RSNA physics tutorial for residents: physics of flat-panel fluoroscopy systems: survey of modern fluoroscopy imaging: flat-panel detectors versus image intensifiers and more.* Radiographics 31.2 (2011): 591-602.

[18] Petrov, Dimitar, et al. *Systematic approach to a channelized Hotelling model observer implementation for a physical phantom containing mass-like lesions: Application to digital breast tomosynthesis.* Physica Medica 58 (2019): 8-20.

[19] Rose, Albert. *The sensitivity performance of the human eye on an absolute scale.* JOSA 38.2 (1948): 196-208.

[20] Schueler, Beth A. *The AAPM/RSNA physics tutorial for residents general overview of fluoroscopic imaging.* Radiographics 20.4 (2000): 1115-1126.

[21] Vining, DavidJ, and Gregory W. Gladish. *Receiver operating characteristic curves: a basic understanding.* Radiographics 12.6 (1992): 1147-1154.

[22] Wang, Jihong, and Timothy J. Blackburn. *The AAPM/RSNA physics tutorial for residents: X-ray image intensifiers for fluoroscopy.* Radiographics 20.5 (2000): 1471-1477.

[23] Wang, Jihong, et al. *Contrast-detail characteristic evaluations of several display devices.* Journal of digital imaging 13.1 (2000): 162-167.

[24] Zhang, Yani, Binh T. Pham, and Miguel P. Eckstein. *Evaluation of internal noise methods for Hotelling observer models.* Medical physics 34.8 (2007): 3312-3322.