UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

Ph.D
itee
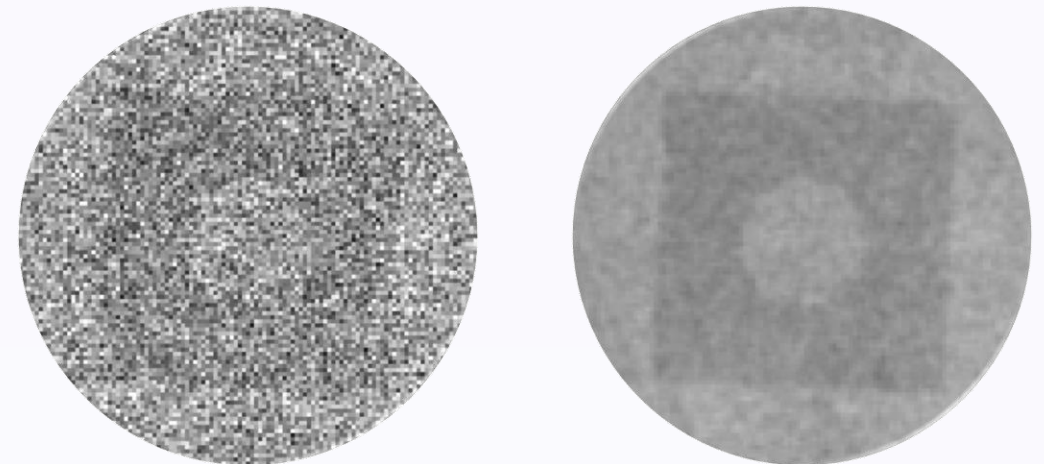INFORMATION TECHNOLOGY
ELECTRICAL ENGINEERING

Fluoroscopy provides real-time X-ray screening of patient's organs and of various radiopaque objects, which make it an invaluable tool for many interventional procedures. For this reason, the number of fluoroscopy screenings has experienced a consistent growth in the last decades. However, this trend has raised many concerns about the increase in X-ray exposure, as even low-dose procedures turned out to be not as safe as they were considered, thus demanding a rigorous monitoring of the X-ray dose delivered to the patients and to the exposed medical staff. In this context, the use of very low-dose protocols would be extremely beneficial. Nonetheless, this would result in very noisy images, which need to be suitably denoised in real-time to support interventional procedures. Simple smoothing filters tend to produce blurring effects that undermines the visibility of object boundaries, which is essential for the human eye to understand the imaged scene. Therefore, some denoising strategies embed noise statistics-based criteria to improve their denoising performances. This dissertation focuses on the Noise Variance Conditioned Average (NVCA) algorithm, which takes advantage of the a priori knowledge of quantum noise statistics to perform noise reduction while preserving the edges and has already outperformed many state-of-the-art methods in the denoising of images corrupted by quantum noise, while also being suitable for real-time hardware implementation. Different issues are addressed that currently limit the actual use of very low-dose protocols in clinical practice, e.g. the evaluation of actual performances of denoising algorithms in very low-dose conditions, the optimization of tuning parameters to obtain the best denoising performances, the design of an index to properly measure the quality of X-ray images, and the assessment of an a priori noise characterization approach to account for time-varying noise statistics due to changes of X-ray tube settings. An improved NVCA algorithm is also presented, along with its real-time hardware implementation on a Field Programmable Gate Array (FPGA). The novel algorithm provides more efficient noise reduction performances also for low-contrast moving objects, thus relaxing the trade-off between noise reduction and edge preservation, while providing a further reduction of hardware complexity, which allows for low usage of logic resources also on small FPGA platforms. The results presented in this dissertation provide the means for future studies aimed at embedding the NVCA algorithm in commercial fluoroscopic devices to accomplish real-time denoising of very low-dose X-ray images, which would foster their actual use in clinical practice.

REAL-TIME QUANTUM NOISE SUPPRESSION IN VERY LOW-DOSE FLUOROSCOPY

EMILIO ANDREOZZI

PH.D. IN INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Università degli Studi di Napoli Federico II

Ph.D. Thesis
IN
Information Technology and Electrical Engineering

Real-Time Quantum Noise Suppression in Very Low-Dose Fluoroscopy

Emilio Andreozzi

Tutors: Prof. Mario Cesarelli

Prof. Paolo Bifulco
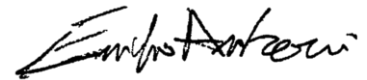
Coordinator: Prof. Daniele Riccio

XXXIII Ciclo

I hereby declare that the contents and organisation of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

……………………………….

Emilio Andreozzi

Naples, 2nd January 2021

Acknowledgements

Contents

List of Figures

x

List of Tables

Chapter 1

Introduction

1.1   The need for dose reduction in fluoroscopic procedures

Fluoroscopy is a medical imaging modality that provides continuous, real-time X-ray screening of patient's organs (also with contrast agents) and of various radiopaque objects involved in surgery (e.g. surgical instruments, catheters, wire-guides, prosthetic implants, implanted devices). This makes it appealing for image-guided surgical procedures [1 – 2], such as orthopaedic surgery, angioplasty, pacemaker and defibrillator implantation. Fluoroscopy-guided procedures are also widespread in diagnostics, e.g. angiography (i.e. the imaging of blood vessels using contrast agent), investigations of gastrointestinal tract, assessment of joints motion, spine kinematics and implanted prosthesis [3 – 5], and also in therapeutics, e.g. in image-guided radiotherapy, cancer ablation, lithotripsy [6 – 11]. Hence, fluoroscopy is undoubtedly an invaluable tool for several medical procedures and, indeed, its use in clinical practice has experienced a consistent growth in the last decades [12 – 22]. However, this trend has raised many concerns about the safety risks related to the increase in exposure to X-rays, as they may cause serious damages to human tissues and organs [23 – 26], also being important risk factors for the development of tumours. Even the low-dose procedures have recently been found to be not as safe as they were considered [24]. While there is no general consensus on the models to be adopted for the prediction of dangerous X-rays exposure effects and their related safety limits, many medical associations suggest that the use of fluoroscopy, and, generally, of all medical procedures involving the exposure to X-rays, should always be carefully evaluated and monitored in clinical practice. This is why the rigorous monitoring of the X-ray dose delivered to the patients and to the exposed medical staff has gained progressively more attention in the scientific literature, also being subject to formal regulations from national and international health organizations [27 – 29]. Considering the constantly increasing number of fluoroscopy-based procedures in clinical practice and the need to keep the overall dose delivered to the patients as low as reasonably achievable (A.L.A.R.A principle [30 – 37]), it is clear that very low-dose fluoroscopic procedures would be extremely beneficial.

1.2   Quantum-limited X-ray images

The A.L.A.R.A principle clarifies the current limitation to access such very low-dose procedures: it is not possible to arbitrarily reduce the X-ray dose beyond certain levels, as it would cause severe degradation of image quality, which could seriously hinder the physician's ability to understand the scenario and operate correctly. Indeed, the X-ray dose depends on a number of parameters and conditions, such as the X-ray tube settings (i.e. tube current and voltage), the exposure time, the distance between the X-ray source and the irradiated tissue, the anti-scatter grids and the additional filtration [38]. Generally, most of these parameters are selected to optimize determined

features of the imaged scene, thus, only the tube current and, sometimes, the exposure time can be modified to reduce the overall dose delivered to the patient. As an example, a common practice to limit the overall exposure time in fluoroscopy during surgical procedures is to turn off the X-ray source periodically and/or to use pulsed protocols, which place a limitation on frame rate though [38]. However, the exposure times are still very long and unpredictable in interventional fluoroscopy [26, 39], as they depend on the particular needs of the surgeon in each situation.

In practice, the dose is mainly limited by reducing the tube current, which implies a reduction of the X-ray radiation intensity, i.e. the number of X-ray photons that reach the detector. This low photons availability gives rise to a signal-dependent, Poisson-distributed noise, usually referred to as "quantum noise" or "Poisson noise" [40]. Quantum noise is inherent to the image formation process, thus it cannot be avoided or even limited by improving detectors technology and requires the application of proper denoising strategies in the digital domain [40]. Moreover, it can be proved that the Signal-to-Noise Ratio (SNR) of Poisson noise is equal to the square root of the expected photon count, which is directly related to the dose, so the lower the dose, the lower the SNR. X-ray images are affected also by other sources of noise, such as the electronic noise of the detectors and the readout circuitry, as well as the quantization noise involved in the analogue to digital conversion. This sources are usually modelled as additive white gaussian noise (AWGN), therefore they are considered as signal-independent. X-ray images that are mainly limited by quantum noise are referred to as quantum-limited, and this is the case of low-dose fluoroscopic images, in which quantum noise represents by far the dominant noise source [40]. Models and properties of quantum noise in raw images, as well as in images processed with luminance transformations commonly implemented in fluoroscopic devices (e.g. log-mapping and gamma correction) are detailed in Chapter 1.

1.3   Poisson noise reduction

Commercial fluoroscopic devices usually implement only simple spatial or temporal averaging. Spatial correlation, due to the Point Spread Function (PSF) of the specific fluoroscopic device, lowers the noise reduction ratio of spatial averaging, whereas temporal uncorrelation, ensured by the reduced lag times of modern flat panel detectors (less than 1 ms as compared to the 33 ms sampling time, corresponding to the maximum frame rate of 30 frames per second) [41], is usually exploited via temporal filtering in commercial devices. However, simple smoothing filters usually don't achieve acceptable results, as they introduce significant blurring effects, thus accomplishing noise reduction to the detriment of fine image details (e.g. edges, textures, etc.). Indeed, the edge blurring caused by spatial smoothing, as well as the motion blurring caused by temporal smoothing, both reduce the edge sharpness and, therefore, the perceptibility and localization of boundaries between different objects and body parts, thus impairing the physician's ability to understand the imaged scene.

The random fluctuations in luminance signals are all related to the stochastic nature of photons arrival on a detector, therefore the Poisson noise model applies to every kind of images acquired in low-photons availability conditions, such as confocal and fluorescence microscopy [42 – 43], astronomical imaging [44] and night photography [45]. The last application field, which undoubtedly draws the largest interest due to its tight relationship with consumer electronics, has encouraged the development of a number of different approaches for single images or image sequences affected by Poisson noise [46 – 54]. However, it is worth underlining that photographs belong to the class of

natural images, whose main features and scopes are profoundly different as compared to medical images [55]. Hence, predicting the efficiency of methods tailored to natural images in the denoising of medical X-ray images is not straightforward and rather requires dedicated evaluations on both images with real noise and with clinically relevant information. Indeed, real fluoroscopic noise has far more complex components (e.g. fixed pattern noise, Compton scatter noise) and it is also modified by different non-idealities of the imaging system, such as the PSF and the Detective Quantum Efficiency (DQE) of the X-ray detectors, as well as the polyenergetic spectrum of the X-ray radiation, thus showing up as not an ideal Poisson random process [56 – 59]; clinical structures are known to alter the structural information of medical images with respect to natural ones, hence, all those methods that make assumptions on representative features of the latter may fail when dealing with the former, as those assumptions may not hold for them [55]. In addition to the approaches developed within the context of night photography, some denoising strategies explicitly devised for medical X-ray images and sequences have also been proposed [40, 60 – 72, 75 – 77]. Further details on the performances of denoising strategies proposed for Poisson noise removal in images and videos are described in Chapter 5.

1.4 The low-dose and real-time constraints

One of the most appealing features of fluoroscopy is undoubtedly the real-time operation, which is essential in interventional procedures to allow the physician monitor the internal organs of the patient and promptly intervene in response to every change in the scenario. Hence, it is clear that denoising approaches devised to improve fluoroscopic image quality in these applications must be suitable for real-time implementation, which means that they should not exhibit an impractical computational burden. Although the technological development of the last decade has assuredly increased the computational power of digital systems, many denoising approaches proposed in literature are still far from being suitable for real-time implementation, mainly because they rely on optimization processes that do not ensure practical and stable computational times and, thus, the observance of the real-time constraint [60].

Considering the paramount importance of dose reduction in X-ray imaging, the ability to provide substantial image quality restoration in heavily degraded images and image sequences is compulsory to foster very low-dose procedures in clinical practice. While there is a natural inclination to project the gain in image quality obtained at lower noise levels also to higher noise levels (quantified by means of the many Image Quality Assessment (IQA) metrics and indices proposed in literature [55, 78 – 80]) to predict the effectiveness of determined approaches in very noisy conditions, this is not observed in practice, as the validity of their basic assumptions is progressively hindered by the degraded quality of the images to be restored. Therefore, the lower the original image quality, the lower the gain in image quality, independently of the denoising power of the specific algorithm. However, the decrease in image quality improvement with the degraded original image quality is dependent on the specific algorithm, in particular on the robustness of its basic assumptions to the noise, thus, as stated in the previous paragraph, it is mandatory for denoising algorithms to be tested on very noisy X-ray images, in order to assess if they could effectively enable very low-dose procedures.

1.5   Current issues in fluoroscopic image denoising

In light of the requirements and constraints described so far, a conspicuous part of the approaches proposed in literature have not demonstrated their effectiveness in providing a suitable image quality restoration for very low-dose fluoroscopic sequences. Many algorithms cannot operate in real-time [69 – 70], or are claimed to be suitable for real-time but have not been experimentally demonstrated [65– 66], while the great part of the few algorithms that achieve real-time filtering of Poisson noise have never been tested on real fluoroscopic images acquired at very low doses [53, 81]. Moreover, their performances have very often been only assessed on static scenes corrupted by simulated Poisson noise: such scenes clearly represent oversimplistic test benches for processing of image sequences, which are rather meant to capture scenes with moving objects. Indeed, a trade-off between noise reduction and edge preservation exists also for temporal filtering, hence, an effective edge-aware denoising algorithm would probably not take full advantage of the temporal dimension in order to prevent motion blur, whereas another algorithm might achieve deeper noise reduction at the cost of increasing motion blur, and it is clear that assessing their performances on static scenes, where no motion blur may occur, would not highlight the edge-awareness of the former, but rather the mere improvement in noise reduction of the latter. An example of this observation is given in Chapter 5. Moreover, the gain in image quality provided by the proposed approaches is usually measured via IQA metrics and indices that, at the best, are well-suited for natural images, but more commonly are still mere global measures of pixel-wise dissimilarities (e.g. Mean Squared Error, Peak Signal-to-Noise-Ratio), already recognized as very inaccurate predictors of the perceived image quality. The inability to provide reliable measures of perceived image and video quality in medical X-ray imaging certainly hinders the assessment and comparison of actual performances of algorithms for image quality restoration and, thus, the research of more performing approaches. The issues just described highlight that efficient and reliable solutions have not been found yet, which enable the real-time quality improvement of X-ray images in very low-dose fluoroscopy procedures so as to foster their use in clinical practice.

1.6   Dissertation outline

This dissertation presents an analysis of image quality assessment and restoration in very low-dose fluoroscopic sequences, by analysing and addressing issues of different nature that currently place limitations on the use of very low-dose fluoroscopy imaging in clinical practice. The dissertation is organised as follows:

- **Chapter 2** describes mathematical models for quantum noise statistical properties in raw and transformed X-ray images;
- **Chapter 3** details the quantum noise filtering approach the dissertation is focused on (i.e. the Noise Variance Conditioned Average algorithm), also comparing it with several state-of-the-art denoising algorithms;
- **Chapter 4** presents an interesting application of the NVCA algorithm to fluoroscopy-based, automated spine kinematics analysis;
- **Chapter 5** presents the comparison between NVCA and the current state-of-the-art algorithm for Gaussian video denoising VBM4D, adapted to Poisson-Gaussian noise reduction via a variance stabilizing transformation, also with a focus on the trade-off between noise reduction

and edge-preservation in the case study of fluoroscopic sequences for spine kinematics analysis;

- **Chapter 6** addresses the current issues on unsuitability of common IQA metrics and indices in predicting the quality of low-dose X-ray images as perceived by a human subject and presents a feasibility study on a novel edge-aware IQA index devised to guide more accurately the identification of the best trade-off between noise reduction and edge-awareness;

- **Chapter 7** presents an improved version of the NVCA algorithm, based on the separation of filtering operations in the temporal and spatial domains and the use of more performing space-variant thresholds, as well as its real-time implementation on a Field-Programmable Gate Array (FPGA) platform;

- **Chapter 8** presents an innovative approach to enable the real-time implementation of the NVCA algorithm in fluoroscopy procedures with time-varying noise statistics due to changes in X-ray tube settings, based on an *a priori* noise characterisation for a commercial fluoroscopic device, which would obviate the need for inferring noise statics after any change of tube settings.

Publications and patents

National and international Patents

- **Andreozzi, E.**, Esposito, D., Bifulco, P.; Gargiulo, G.D. 15/10/2020. Physiological parameter sensing systems and methods. PCT/AU2020/051107. Australian Patent Office.

- Sarno, A.; **Andreozzi, E.** 05/06/2020. Sistema per il tracciamento delle interazioni tra soggetti. IT Patent Application N. 102020000013369. Italian Patent and Trademark Office.

Journal papers

- **Andreozzi, E.**, Fratini, A., Esposito, E., Cesarelli, M., Bifulco, P. Toward A Priori Noise Characterization for Real-Time Edge-Aware Denoising in Fluoroscopic Devices. *BioMed Eng OnLine*. (under revision for publication)

- **Andreozzi, E.**, Fratini, A., Esposito, D., Naik, G., Polley, C., Gargiulo, G.D., Bifulco, P. (2020). Forcecardiography: A Novel Technique to Measure Heart Mechanical Vibrations onto the Chest Wall. *Sensors*, *20*, 3885. https://doi.org/10.3390/s20143885

- **Andreozzi, E.**, Carannante, I., D'Addio, G., Cesarelli, M., Balbi, P. (2019). Phenomenological models of NaV1.5. A side by side, procedural, hands-on comparison between Hodgkin-Huxley and kinetic formalisms. *Scientific Reports,* 9, 17493. https://doi.org/10.1038/s41598-019-53662-9

- **Andreozzi, E.**, Gargiulo, G.D., Fratini, A., Esposito, D., Bifulco, P. (2018). A Contactless Sensor for Pacemaker Pulse Detection: Design Hints and Performance Assessment. *Sensors*, 18, 2715. https://doi.org/10.3390/s18082715

- Sarno, A., **Andreozzi, E.**, De Caro, D., Di Meo, G., Strollo, A.G.M., Cesarelli, M.. Bifulco, P. (2019). Real-time algorithm for Poissonian noise reduction in low-dose fluoroscopy: performance evaluation. *BioMed Eng OnLine,* 18, 94. https://doi.org/10.1186/s12938-019-0713-7

- Castellano, G., De Caro, D., Esposito, D., Bifulco, P., Napoli, E., Petra, N., **Andreozzi, E.**, Cesarelli, M., Strollo, A.G.M. (2019). An FPGA-oriented Algorithm for Real-Time Filtering of Poisson Noise in Video Streams, with Application to X-Ray Fluoroscopy. *Circuits Syst Signal Process*, 38, 3269–3294. https://doi.org/10.1007/s00034-018-01020-x

- Esposito, D., **Andreozzi, E.**, Gargiulo, G. D., Fratini, A., D'Addio, G., Naik, G. R., Bifulco, P. (2020). A Piezoresistive Array Armband With Reduced Number of Sensors for Hand Gesture Recognition. *Frontiers in Neurorobotics*, 13. https://doi.org/10.3389/fnbot.2019.00114

- Esposito, D., **Andreozzi, E.**, Fratini, A., Gargiulo, G.D., Savino, S., Niola, V., Bifulco, P. (2018). A Piezoresistive Sensor to Measure Muscle Contraction and Mechanomyography. *Sensors*, 18, 2553. https://doi.org/10.3390/s18082553

- Centracchio, J., Sarno, A., Esposito, D., **Andreozzi, E.**, Pavone, L., Di Gennaro, G., Bartolo, M., Esposito, V., Morace, R., Casciato, S., Bifulco, P. Efficient Automated Localization of ECoG Electrodes in CT Images Via Shape Analysis. *IJCARS*. (under revision for publication)

- Esposito, D., Savino, S., Cosenza, C., **Andreozzi, E.**, Gargiulo, G.D., Polley, C., Cesarelli, G., D'Addio, G., Bifulco, P. (2021). Evaluation of Grip Force and Energy Efficiency of the "Federica" Hand. *Machines*, 9, 25. https://doi.org/10.3390/machines9020025

- Esposito, D., Savino, S., **Andreozzi, E.**, Cosenza, C., Niola, V., Bifulco,P. The Federica hand. *Res. Biomed. Eng.* (under revision for publication).

- Iuppariello, L., D'addio, G., Lanzillo, B., Balbi, P., **Andreozzi, E.**, Improta, G., Faiella, G., Cesarelli, M. (2019). A novel approach to estimate the upper limb reaching movement in three-dimensional space. *Informatics in Medicine Unlocked*, 15. https://doi.org/10.1016/j.imu.2019.01.005

Conference papers

- **Andreozzi, E.**, Pirozzi, M. A., Fratini, A., Cesarelli, G., Cesarelli, M.; Bifulco, P. (2020). A Novel Image Quality Assessment Index for Edge-Aware Noise Reduction in Low-Dose Fluoroscopy: Preliminary Results. *2020 International Conference on e-Health and Bioengineering (EHB), Iasi, Romania*. https://doi.org/10.1109/EHB50910.2020.9280107

- **Andreozzi, E.**, Pirozzi, M. A., Fratini, A., Cesarelli, G., Bifulco, P. (2020). Quantitative performance comparison of derivative operators for intervertebral kinematics analysis. *2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Bari, Italy*. https://doi.org/10.1109/MeMeA49120.2020.9137322

- **Andreozzi E.**, Pirozzi M.A., Sarno A., Esposito D., Cesarelli M., Bifulco P. (2020).A Comparison of Denoising Algorithms for Effective Edge Detection in X-Ray Fluoroscopy. In: Henriques J., Neves N., de Carvalho P. (eds) *XV Mediterranean Conference on Medical and Biological Engineering and Computing – MEDICON 2019. MEDICON 2019. IFMBE Proceedings*: *Vol 76* (pp 405-413). Springer, Cham. https://doi.org/10.1007/978-3-030-31635-8_49

- Polley, C., **Andreozzi, E.**, Bifulco, P., Esposito, D., Naik, G., Gunawardana, U., Gargiulo, G. (2019). Low Cost Analogue Front End for Electronic Stethoscopes Application with Silicone Enclosure. *2019 International Conference on Electrical Engineering Research & Practice (ICEERP), SYDNEY, Australia*. https://doi.org/10.1109/ICEERP49088.2019.8956972

- Pirozzi, M.A., **Andreozzi, E.**, Magliulo, M., Gargiulo, P., Cesarelli, M., Alfano, B. (2020). Automated Design of Efficient Supports in FDM 3D Printing of Anatomical Phantoms. In: Henriques J., Neves N., de Carvalho P. (eds) *XV Mediterranean Conference on Medical and Biological Engineering and Computing – MEDICON 2019. MEDICON 2019. IFMBE Proceedings*: *Vol 76* (pp 292-300). Springer, Cham. https://doi.org/10.1007/978-3-030-31635-8_35

- Pirozzi, M.A.; **Andreozzi, E.**; Magliulo, M.; Gargiulo, P.; Cesarelli, M.; Alfano, B. (2020). 3D-printed anatomical phantoms for medical imaging applications: a methodological mini-review. *Nordic Baltic Conference on Biomedical Engineering and Medical Physics (NBC 2020), Reykjavik, Iceland*, (Article in press).

- D'Addio, G., Evangelista, S., Donisi, L., Biancardi, A., **Andreozzi, E.**, Pagano, G., Arpaia, P., Cesarelli, M. (2019). Development of a Prototype E-Textile Sock. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany* (pp. 17498-1752). https://doi.org/10.1109/EMBC.2019.8856739

- Esposito, D., Gargiulo, G. D., Parajuli, N., Cesarelli, G., **Andreozzi, E.**, Bifulco, P. (2020). Measurement of muscle contraction timing for prosthesis control: a comparison between electromyography and force-myography. *2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Bari, Italy*. https://doi.org/10.1109/MeMeA49120.2020.9137313

- Esposito, D., Cosenza, C, Gargiulo, G. D., **Andreozzi, E.**, Niola, V, Fratini, A., D'Addio, G., Bifulco, P. (2020). Experimental Study to Improve "Federica" Prosthetic Hand and Its Control System. In: Henriques J., Neves N., de Carvalho P. (eds) *XV Mediterranean Conference on Medical and Biological Engineering and Computing – MEDICON 2019. MEDICON 2019. IFMBE Proceedings*: *Vol 76* (pp 405-413). Springer, Cham. https://doi.org/10.1007/978-3-030-31635-8_49

Chapter 2

Quantum noise model and statistics

2.1 Poisson – Gaussian mixture model

In an X-ray system, the number of photons that emerge from a patient and reach a single pixel of the detector plane can be modelled by a temporally stochastic Poisson process [40 – 41, 62, 82 – 83], whose probability density function (pdf) is described in (1):

$$p(n) = \frac{\lambda^n}{n!} e^{-\lambda} \tag{1}$$

where λ is the expected photon count. Simple calculations allow deriving a very important feature of Poisson distribution, namely the Expected Value – Variance Relationship (EVaR), which is reported in (2):

$$\sigma_p^2(\mu_p) = \mu_p \tag{2}$$

Therefore, the variance of the number of photons that reach a single pixel is equal to the expected photon count. This property confirms that, as already stated in the Introduction, the SNR of Poisson noise, which is defined as the ratio of the expected value and the standard deviation, is equal to the square root of the average photon count $\sqrt{\lambda}$. However, in practice, the information carried by this random process occurring at a single detector pixel is usually coded in a digital image, and particularly in the grey level of the corresponding image pixel, which is proportional to the actual photon count, thus being characterized by a modified EVaR, as reported in (3):

$$g(\lambda) = a \cdot p(\lambda) \quad \rightarrow \quad \sigma_g^2(\mu_g) = a \cdot \mu_g \tag{3}$$

where g is the grey level of the digital image pixel corresponding to the detector pixel that is reached by a number of photons described by p, and a is the coefficient of proportionality between g and p, also known as "detector gain". The EVaR clarify the signal-dependent nature of quantum noise (heteroscedasticity), which, unlike the well-known AWGN, cannot be characterized by a single, global noise variance estimate (homoscedasticity), but rather requires the estimation of the detector gain, in order to be able to estimate the local, signal-dependent noise variance from the local mean luminance.

As previously stated, X-ray images are also affected by other sources of noise that are usually modelled as AWGN. Considering that quantum noise and AWGN sources are mutually independent, the latter introduce a constant contribution to the noise variance (i.e. a noise floor), which should be included in the noise model, as shown in (4):

$$\sigma_g^2(\mu_g) = a \cdot \mu_g + b \tag{4}$$

where *b* corresponds to the overall variance of the AWGN component. This model is known as Poisson – Gaussian mixture and has been used in various denoising approaches devised for low-light images [40, 47 – 50, 60 – 61, 64, 71 – 72] which thus require the estimation of EVaR parameters (*a* and *b*), commonly referred to as noise parameters.

2.2 Effect of luminance transformations on noise model

Most fluoroscopic devices implement luminance transformations, such as the logarithmic mapping and the gamma – correction, usually to perform a white compression. This operations obviously modify the statistics of the noise and, in particular, the EVaR. Therefore, it is important to take it into account when performing the estimation of noise parameters. To analytically derive the EVaR obtained after the application of luminance transformations, it is important to express the Poisson – distributed signal at the generic pixel *g(x,y)* as an additive noise model. Although Poisson noise is inherently signal – dependent, it can be well approximated by a signal – dependent Gaussian noise, with equal expected value and variance, for not too low expected values [85 – 86]. Considering also the effect of detector gain *a*, the pixel luminance *g(x,y)* is expressed as a Gaussian distribution with variance equal to the expected value multiplied by the detector gain, which can be decomposed in a deterministic term *s(x,y)* (the signal) corresponding to the expected value, and a stochastic term *n(x,y)* (pure noise) corresponding to a zero-mean Gaussian random variable with the same variance, as reported in (5).

$$\begin{cases} g(x,y) = a \cdot p(x,y); & p(x,y) \sim Poisson\big(\lambda(x,y)\big) \\\\ g(x,y) \approx g'(x,y); & g'(x,y) \sim \mathrm{N}\big(a \cdot \mu_g, a^2 \cdot \mu_g\big) \\\\ g(x,y) \approx s(x,y) + n(x,y); & n(x,y) \sim N\big(0, a^2 \cdot \mu_g\big) \end{cases} \tag{5}$$

Indeed, deriving the EVaR for this approximated description yields the same expression found for the original Poisson – distributed pixel luminance:

$$\begin{cases} E[g(x,y)] = \mathrm{E}[s(x,y)] = a \cdot \mu_g \\\\ Var[g(x,y)] = Var[n(x,y)] = a^2 \cdot \mu_g = a \cdot E[g(x,y)] \end{cases} \tag{6}$$

Considering the logarithmic mapping operation reported in (7), it is possible to derive the related EVaR, which is outlined in (8) (to simplify the notation, the dependence on pixel position (x,y) was omitted).

$$g_{ln} = c_{ln} \cdot ln(1+g) \approx c_{ln} \cdot ln\left[s\left(1+\frac{n}{s}\right)\right] = c_{ln} \cdot ln(s) + c_{ln} \cdot ln\left(1+\frac{n}{s}\right) = s_{ln} + n_{ln} \tag{7}$$

$$Var[g_{ln}] = a \cdot c_{ln}^2 \cdot e^{-\frac{\mathrm{E}[g_{ln}]}{c_{ln}}} \tag{8}$$

Carrying out the same procedure for the gamma – correction operation (9), it is possible to obtain an approximated analytic expression of the related EVaR (10):

$$\begin{cases} g_\gamma = c_\gamma \cdot g^\gamma = c_\gamma \cdot s^\gamma \cdot \left(1 + \dfrac{n}{s}\right)^\gamma \\[2mm] E\left[\dfrac{n}{s}\right] = 0 \ll 1 \;\; \rightarrow \;\; g_\gamma \approx c_\gamma \cdot s^\gamma \cdot \left(1 + \gamma\dfrac{n}{s}\right) = c_\gamma \cdot s^\gamma + c_\gamma \cdot s^\gamma \cdot \gamma\dfrac{n}{s} \end{cases} \tag{9}$$

$$\begin{cases} E[g_\gamma] = c_\gamma \cdot s^\gamma \\[2mm] Var[g_\gamma] = c_\gamma^2\, s^{2\gamma}\, \gamma^2 \cdot \dfrac{a\,s}{s^2} = a\,\gamma^2\, c_\gamma^{\frac{1}{\gamma}} \left(c_\gamma \cdot s^\gamma\right)^{\left(2 - \frac{1}{\gamma}\right)} = a\gamma^2 c_\gamma^{\frac{1}{\gamma}} \cdot E[g_\gamma]^{\left(2 - \frac{1}{\gamma}\right)} \end{cases} \tag{10}$$

2.3 Limitations of the noise models

The presented models are based on the hypothesis that the pixel luminance can be described by a single Poisson random variable multiplied by the detector gain a and potentially superimposed to statistically independent, space – invariant Gaussian component. However, as already stated in the Introduction, this is a simple yet useful approximation of a far more complex behaviour, which is based on the assumption of monoenergetic X-ray radiation and quasi – ideal detector response. Indeed, the X-ray radiation generated by commonly used X-ray tubes has a polyenergetic spectrum, which means that the photons produced by the tube have different energies, and this has an impact on the actual statistics of the detected signals. The formation of transmission X-ray images is based on the attenuation of X-rays due to their interaction with the object to be imaged, which mainly happens via three mechanisms, namely the photoelectric absorption, the Compton scattering and the Rayleigh scattering. However, at the energies involved in common X-ray screenings, photoelectric absorption and Compton scattering dominate the overall attenuation phenomenon, which is taken into account by means of a synthetic parameter, referred to as the linear attenuation coefficient μ_l. Each material is characterized by its own linear attenuation coefficient, which is also a function of the photon energy ($\mu_l \propto \dfrac{1}{E^3}$). The exponential law in (11) defines the relationship between the average number of X-ray photons at energy E on a primary incident ray (N_{in}) and the average number of photons at the same energy on the corresponding primary ray (N_{out}) that emerges from an homogeneous object with linear attenuation coefficient μ_l:

$$N_{out}(E) = N_{in}(E) \cdot e^{-\mu_l(E) \cdot x} \tag{11}$$

where x is the optical path length of the primary ray within the object. It is clear from (11) that photons at different energies experience different attenuations and, thus, are described by Poisson random variables with expected values equal to the average numbers of emerging photons $N_{out}(E)$. In common flat panel detectors, the actual signal from the detector is obtained through a conversion of the overall

energy released by the photons that reach the detector in a voltage signal, which can be expressed as in (12):

$$V = C \cdot \int_{E_{min}}^{E_{max}} N_{out}(E) \cdot E \, dE \qquad (12)$$

where C is an energy-to-voltage conversion factor. Equation (12) highlights that by only removing the assumption of monoenergetic radiation, the expression of the signal provided by a sensing element of an X-ray detector is far from being a simple Poisson variable, and it is rather a weighted sum of the Poisson variables describing the arrivals of X-ray photons of different energies, with weights corresponding to the related photons energies. It is also important to note that single sensing elements of detectors may exhibit reduced sensitivities at certain energy levels, depending on their material and thickness, which may also experience stochastic fluctuations, commonly referred to as Swank noise [56, 88]. In addition, different sensing elements and related readout circuitry may exhibit non-uniform gain and offset terms, giving rise to the so-called fixed-pattern noise, which adds a contribution to the spatial noise variance that is proportional to the square of the local exposure [56 – 59]. Moreover, Compton scattered rays that reach a sensing element of the detector introduce an additional contribution, which depends on the interaction of the X-ray photons with other elements of the imaged object, thus usually causing a reduction of contrast and an increase of noise power at the same time [87]. Finally, the spreading of secondary quanta over neighbour sensing elements in the detector material, determines a non-negligible correlation between the luminance values of neighbour pixels, quantified by the PSF of the detector, which introduce a further alteration of noise statistics and appearance in the final image

The limitations just described give an idea of the several parameters that influence the appearances and features of real fluoroscopic noise, which certainly cannot be fully captured by the simple Poisson – Gaussian model. However, it is worth remembering that models are always partial representation of real phenomena and, above all, they should have the proper level of accuracy and complexity for their intended applications. It is clear that a model accounting for all the above-mentioned phenomena would be almost impractical to be implemented in denoising strategies for fluoroscopic noise and, therefore, simplified models with reasonable accuracy should be considered for denoising purposes. On the other hand, a denoising strategy undoubtedly has to deal with reality. For interventional fluoroscopy applications this means that, besides being tested on synthetic sequences that mimic determined features of fluoroscopic sequences, denoising algorithms must ultimately prove their effectiveness in real-time processing of actual, very low-dose fluoroscopic sequences.

Chapter 3

The Noise Variance Conditioned Average algorithm

3.1 Basic principles of NVCA denoising

The main idea behind the NVCA [40] denoising strategy comes from the observation that the undesirable blurring effects caused by the simple moving average algorithm arise when the grey values of noisy pixels belonging to different objects are averaged together. Indeed, besides certainly reducing noise, this operation literally spreads the luminance of brighter pixels over their neighbourhood, creating intermediate grey values that were not present in the original noiseless image and resulting in smoothed edge profiles. This behaviour can be easily interpreted in the frequency domain: the edges represent abrupt spatial changes in luminance, so their spectral power is mainly concentrated at relatively high spatial frequencies; the moving average filter is a low-pass (smoothing) filter, therefore it reduces the amplitude of the higher frequency components, while leaving the amplitude of lower frequency components almost unaltered, which results in the observed slower spatial variations in luminance. However, smoothed edge profiles substantially hinder the ability of localizing object boundaries, because profiles are characterised by many pixels with small, comparable luminance variations, which make it difficult for the human vision system (HVS) to determine which of them correspond to the true edge location. The same issues apply to the motion blur in an image sequence. Indeed, when an object moves in the scene, it is represented by pixels in different spatial positions within subsequent frames and, from a single pixel point of view, the same pixel represents different objects within subsequent frames, hence the instantaneous luminance values it assumes at any given time during the sequence actually belong to different distributions. Motion blur results from the time averaging of pixel luminance during objects motion, when they exit from or go in determined pixels areas. It follows that the key to prevent blurring effects when performing a moving average is to discriminate between noisy pixels belonging to the same object, which can be effectively averaged to reduce noise power, and noisy pixels of different objects, which should not be included in the same average operation. Certainly, a method to perform such a discrimination is required, which should be based on an assumption that is sufficiently robust to noise.

The NVCA algorithm is based on the hypothesis that pixels whose difference in luminance is greater than a certain multiple of the local noise standard deviation (SD) are more likely to belong to different objects, i.e. their difference in luminance is more likely to be due to the actual local contrast (the true signal) than to the spurious variations caused by the noise. Hence, NVCA performs a conditioned average in a common sliding window fashion and uses the outlined criterion to identify the neighbours pixels that belong to a different object, in order to exclude them from the average operation that produces the filtered value of the current pixel. In particular, the NVCA algorithm performs a conditioned spatio-temporal moving average, in order to take advantage of the temporal

dimension as well. Indeed, a sequence of 2D images represents a 3D matrix, thus 3D spatio-temporal masks of pixels can be defined, which are generically described as S×S×T matrices, where S and T are the mask sizes in the spatial and temporal domain, respectively. Let indicate the position of the current pixel as *(m,n,p)*, with *(m,n)* being the position indices within the frame and *p* the frame index. The NVCA algorithm considers the spatio-temporal neighbourhood of size S×S×T centred at the pixel *g(m,n,p)*, i.e. all the pixels *g(m − i, n − j, p − k)* with $i, j \in \left[-\left\lfloor \frac{S}{2} \right\rfloor ; \left\lfloor \frac{S}{2} \right\rfloor \right]$ and $k \in [\, 0;\, T - 1\,]$, and computes the differences between the grey level of each pixel in the neighbourhood and the grey level of the current pixel. Then, the algorithm estimates the local noise SD by assuming the grey level of the current pixel as the expected value of a mixed Poisson – Gaussian distribution, whose SD given by (13):

$$\sigma(m, n, p) = \sqrt{a \cdot g(m, n, p) + b} \tag{13}$$

with *a, b* being the Poisson – Gaussian noise parameters introduced in Chapter 2.

Afterward, the algorithm determines a threshold G_{th} for the grey levels differences, in order to create a binary mask *M* of Boolean values that flag the pixels to be included in the average computation. The threshold is assumed as a multiple of the local noise SD with a coefficient N_σ as shown in (14):

$$G_{th}(m, n, p) = N_\sigma \cdot \sigma(m, n, p) = N_\sigma \cdot \sqrt{a \cdot g(m, n, p) + b} \tag{14}$$

and the generic element of the binary mask M can be defined as:

$$\begin{cases} M(i, j, k) = 0 & if \ |g(m - i, y - j, p - k) - g(m, n, p)| > G_{th}(m, n, p) \\ \\ M(i, j, k) = 1 & if \ |g(m - i, y - j, p - k) - g(m, n, p)| \le G_{th}(m, n, p) \end{cases} \tag{15}$$

The filtered value of the current pixel $g_f(m,n,p)$ is then obtained by averaging the grey levels of the pixels in the spatio – temporal neighbourhood mask whose corresponding elements in the binary mask *M* are non-null, as described in (16). Figure 1 shows a schematic description of NVCA.

$$g_f(m, n, p) = \frac{1}{\sum_{i,j,k} M(i, j, k)} \cdot \sum_{k=0}^{T-1} \sum_{i=-\left\lfloor \frac{S}{2} \right\rfloor}^{\left\lfloor \frac{S}{2} \right\rfloor} \sum_{j=-\left\lfloor \frac{S}{2} \right\rfloor}^{\left\lfloor \frac{S}{2} \right\rfloor} M(i, j, k) \cdot g(m - i, n - j, p - k) \tag{16}$$

Despite its simple formulation, NVCA is able to provide significant noise reduction in real low-dose fluoroscopic sequences, while substantially preserving the sharpness and detectability of object boundaries [3 – 4, 60]. It is precisely the simplicity that made its real – time implementation feasible on programmable hardware like the Field Programmable Gate Arrays (FPGA) [61]. Indeed, NVCA does not depend upon iterative schemes, thus ensuring deterministic computational times, and exhibits a low computational complexity, as it requires only one division per pixel (the local noise SD threshold can be obtained from a lookup table), which yields an order of multiplications (big-O notation [89 – 90]) equal to $O(N)$ for an image sequence of $\sqrt[3]{N} \times \sqrt[3]{N} \times \sqrt[3]{N}$ pixels.

Despite its simple formulation, NVCA is able to provide significant noise reduction in real low-dose fluoroscopic sequences, while substantially preserving the sharpness and detectability of object boundaries [3 – 4, 60]. It is precisely the simplicity that made its real – time implementation feasible on programmable hardware like the Field Programmable Gate Arrays (FPGA) [61]. Indeed, NVCA does not depend upon iterative schemes, thus ensuring deterministic computational times, and

exhibits a low computational complexity, as it requires only one division per pixel (the local noise SD threshold can be obtained from a lookup table), which yields an order of multiplications (big-O notation [89 – 90]) equal to $O(N)$ for an image sequence of $\sqrt[3]{N} \times \sqrt[3]{N} \times \sqrt[3]{N}$ pixels.

Figure 1. Schematic representation of the NVCA algorithm

3.2   Effects of parameters choice

The NVCA filter has three parameters, namely the spatial and temporal mask sizes S, T, and the noise SD threshold multiplier N_σ, that can be selected in order to achieve the optimal results; however it is not as straightforward as it might seem. Indeed, while for the moving average algorithm the effect of increasing or decreasing the mask size, in space and/or in time, is well-known from theory of linear space- and time-invariant filters, it is not the case for NVCA, as it is rather a space- and time-variant (i.e. non – linear) filter. Certainly, it could be observed that, provided the algorithm is efficient in selecting the pixels belonging to the local noise statistics, it would better to increase the mask size, both in space and time, so as to achieve a deeper noise reduction, because more pixels would potentially be averaged together. Moreover, by using a large spatio – temporal mask, the NVCA algorithm could result in an equivalent tracking of a relatively small object that moves within the region defined by the spatial extent of the mask during the time interval defined by the temporal extent. Nonetheless, considering that the algorithm would not provide ideal performances in practice, it follows that increasing the mask size would result in an increased number of misclassified pixels, as indeed moving away from the current pixel increases the chances to run into another object, which could be confused with the current one, thus resulting in the generation of blurring effects.

In addition, the choice of the noise SD threshold multiplier N_σ should be analysed as well, as it plays a key role in the statistical assumption at the core of NVCA operating principle. Indeed, N_σ regulates the number of pixels that are classified as belonging to the same object of the current pixel and, therefore, it modulates the amount of noise reduction (the larger the number of averaged pixels, the deeper the noise reduction) as well as the amount of edge blurring (the larger the number of

averaged pixels, the higher the percentage of pixels misclassified as belonging to the local noise distribution). This can be proved in terms of probability. Considering a spatio – temporal neighbourhood that comprises a dark object over a bright background, let λ_1 and λ_2 be the expected values of the related Poisson distributions, which can be approximated by Gaussian distributions with equal mean and variance $X_1 \sim N(\lambda_1, \lambda_1)$, $X_2 \sim N(\lambda_2, \lambda_2)$. Let assume that the current pixel belongs to the dark object and define a confidence interval as a function of the parameter N_σ :

$$\left[\lambda_1 - N_\sigma \sqrt{\lambda_1}; \quad \lambda_1 + N_\sigma \sqrt{\lambda_1} \right] \tag{17}$$

The probability that a value from the dark pixels distribution is correctly classified as belonging to the same distribution, i.e. the True Positive rate (TP) is given by (18):

$$TP = P\left(\lambda_1 - N_\sigma \sqrt{\lambda_1} \leq X_1 \leq \lambda_1 + N_\sigma \sqrt{\lambda_1}\right) = F_1\left(\lambda_1 + N_\sigma \sqrt{\lambda_1}\right) - F_1\left(\lambda_1 - N_\sigma \sqrt{\lambda_1}\right) \tag{18}$$

where $F_1(x)$ is the Cumulative Distribution Function (CDF) of the dark pixels distribution, defined as:

$$F_1(x) = \frac{1}{2}\left(1 + erf\left(\frac{x - \mu_1}{\sqrt{2}\sigma_1}\right)\right) = \frac{1}{2}\left(1 + erf\left(\frac{x - \lambda_1}{\sqrt{2\lambda_1}}\right)\right) \tag{19}$$

Therefore the TP can be described as in (20):

$$TP = \frac{1}{2}\left(erf\left(\frac{\lambda_1 + N_\sigma \sqrt{\lambda_1} - \lambda_1}{\sqrt{2\lambda_1}}\right) - erf\left(\frac{\lambda_1 - N_\sigma \sqrt{\lambda_1} - \lambda_1}{\sqrt{2\lambda_1}}\right)\right) = erf\left(\frac{N}{\sqrt{2}}\right) \tag{20}$$

The probability that a value from the bright pixels distribution is misclassified as belonging to the dark pixel distribution, i.e. the False Positive rate (FP) is given by (21):

$$FP = P\left(\lambda_1 + N_\sigma \sqrt{\lambda_1} \leq X_2 \leq \lambda_1 + N_\sigma \sqrt{\lambda_1}\right) = F_{X_2}\left(\lambda_1 + N_\sigma \sqrt{\lambda_1}\right) - F_{X_2}\left(\lambda_1 - N_\sigma \sqrt{\lambda_1}\right) =$$

$$= \frac{1}{2}\left(erf\left(\frac{\lambda_1 + N_\sigma \sqrt{\lambda_1} - \lambda_2}{\sqrt{2\lambda_2}}\right) - erf\left(\frac{\lambda_1 - N_\sigma \sqrt{\lambda_1} - \lambda_2}{\sqrt{2\lambda_2}}\right)\right) \tag{21}$$

Defining the contrast as $C = \lambda_1 - \lambda_2$, (21) can be expressed as a function of C:

$$FP = \frac{1}{2}\left(erf\left(\frac{N_\sigma \sqrt{\lambda_1} - C}{\sqrt{2\lambda_2}}\right) - erf\left(\frac{-N_\sigma \sqrt{\lambda_1} - C}{\sqrt{2\lambda_2}}\right)\right) \tag{22}$$

The contrast C, in turn, can be expressed as a function of the Contrast-to-Noise Ratio (CNR):

$$CNR = \frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}} = \frac{\lambda_2 - \lambda_1}{\sqrt{\lambda_1 + \lambda_2}} = \frac{C}{\sqrt{2\lambda_1 + C}} \tag{23}$$

$$C = \frac{CNR^2 + CNR\sqrt{CNR^2 + 8\lambda_1}}{2} \tag{24}$$

As an example, Figure 2 depicts TP and FP as functions of N_σ and CNR, which confirm that by increasing N_σ both TP and FP increase. However, the local accuracy would depend also on the percentages of dark (positives) and bright (negatives) pixels within the considered spatio – temporal

16

neighbourhood. In particular, Figure 2.(c) gives a practical demonstration of a concept expressed in the Introduction: the decrease in the original image quality of the noisy images to be restored, here quantified via the CNR, determines a progressive increase in the failure rate of the statistical assumptions at the core of denoising strategies, which obviously leads to a decrease of the denoising power and, therefore, of the gain in image quality that could be provided.

By neglecting the $F_{X_2}\left(\lambda_1 - N_\sigma\sqrt{\lambda_1}\right)$ term in (21) (thus overestimating FP) to simplify the relationship between FP and N_σ, it is also possible to derive the expression of the N_σ value which should be chosen to obtain a determined FP, given the local CNR in the spatio – temporal neighbourhood of the original noisy sequence, which is outlined in (25).

(a)

(b)

(c)

(d)
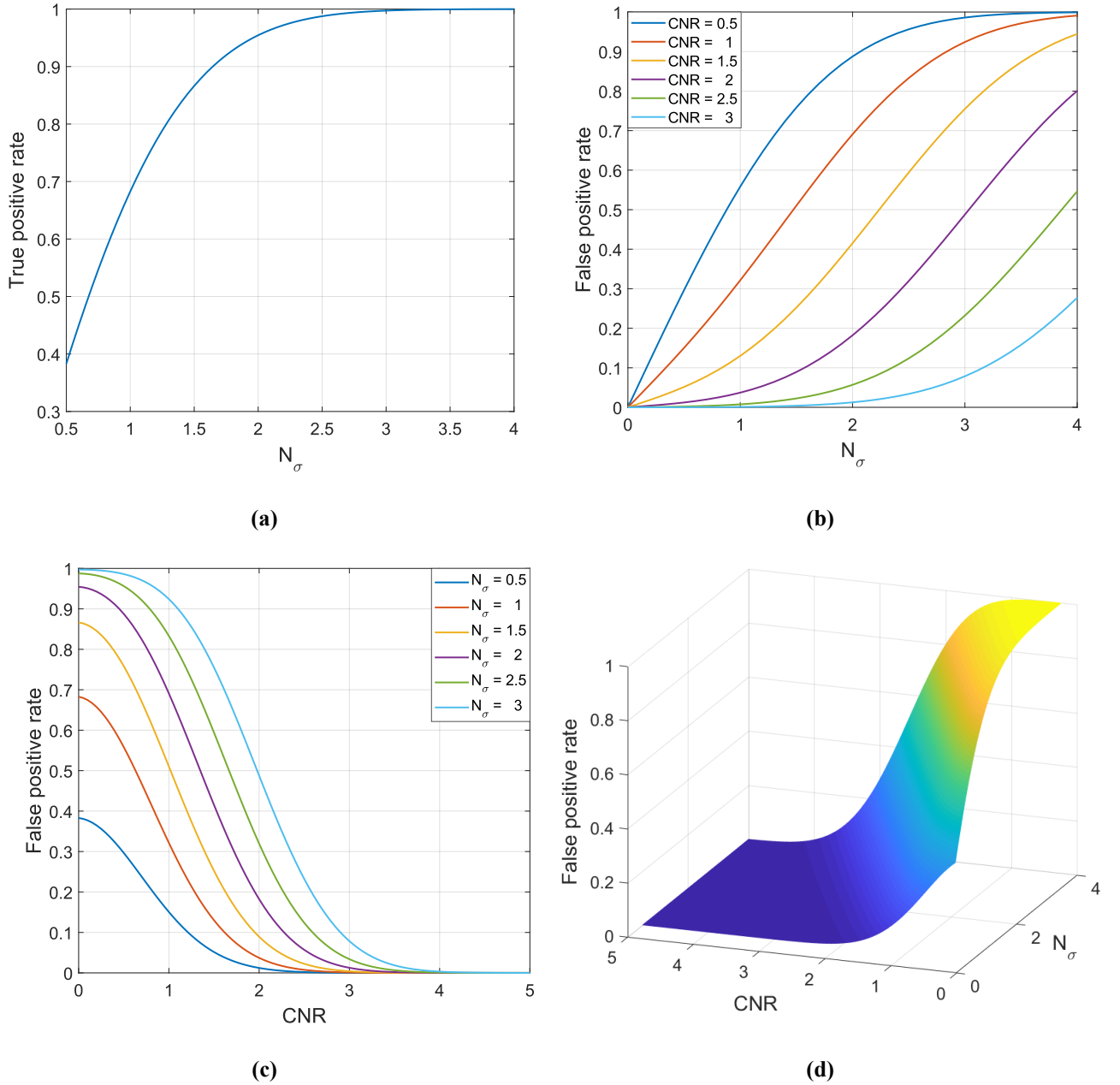
Figure 2. (a) Rate of correctly classified dark object pixels (true positive rate) as a function of N_σ ; (b) Rate of bright misclassified background pixels (false positive rate) as a function of the N_σ for six CNR values; (c) False positive rate as a function of the CNR for six N_σ values; (d) False positive rate as a 2D function of N_σ and CNR.

$$N_\sigma = \frac{\sqrt{2(\lambda_1 + C(CNR))}\ erf^{-1}(2FP - 1) + C(CNR)}{\sqrt{\lambda_1}} \tag{25}$$

where *C(CNR)* refers to the relationship described in (24). Figure 3 shows 2D and 3D plots of N_σ as a function of CNR and FP. The negative values of N_σ suggest that a left one-sided confidence interval should be adopted to obtain the desired FP for lower CNR values. However, recalling that (25) has been derived from an overestimate of FP, it is reasonable that practical values of N_σ can be chosen to ensure a determined FP also at lower CNR values, as confirmed by Figure 2.(b)

(a) (b)
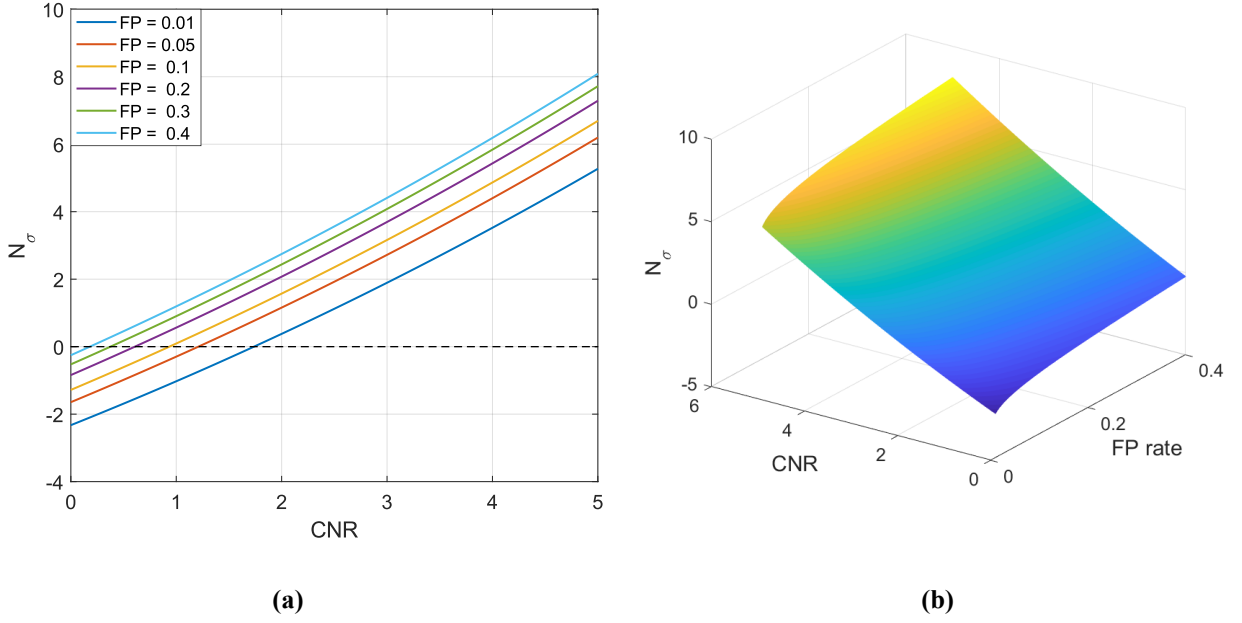
Figure 3. (a) N_σ as a function of CNR for six values of FP; (b) N_σ as a 2D function of CNR and FP.

It is worth noting that the results of the present analysis were derived based on theoretical statistics, but, in practice, the NVCA algorithm must deal with a quite limited number of noisy observations, which would result in a certain variability of the actual results with respect to the theoretical predictions. In these cases, simulation studies with various levels of accuracy and complexity could be of help to assess the reliability of the theoretical predictions and their effects in terms of image quality. A first attempt to quantitively analyse the performance of NVCA as a function of its free parameters (S, T, N_σ) is described in Chapter 6, which presents a preliminary study on a novel IQA index for X-ray images that could be used precisely to guide the identification of NVCA parameters ensuring the best trade – off between noise reduction and edge preservation.

Chapter 4

Application of NVCA to spine kinematics analysis

In this chapter, a demonstration of the effectiveness of the NVCA denoising algorithm in low-dose fluoroscopic sequences is presented within the context of a fluoroscopy-based diagnostic procedure for the assessment of spine kinematics, which will be considered as a case study for further investigations in the following chapters. In particular, the examined application is aimed at providing an automated, quantitative assessment of intervertebral kinematics from a guided flexion – extension motion sequence, imaged via a low – dose fluoroscopy protocol. The automated extraction of intervertebral kinematics is performed by means of a template matching method described in [3 – 4], which relies on a normalised cross-correlation index computed from horizontal and vertical image derivatives. As the derivative operation is highly sensitive to noise, a smoothing operation is generally included in the derivative operators kernels, but in low – dose fluoroscopy images this may be not enough to avoid the distortions caused by the noise, thus requiring additional pre – processing with smoothing filters. However, the edge blurring effects introduced by common smoothing filters significantly hinder the ability of cross-correlation to match the contours of the selected template when operating on image derivatives, hence an edge-aware noise reduction algorithm is required to provide a substantial noise reduction while preserving the edge sharpness. The study here presented assessed the effectiveness of NVCA filtering in improving image quality to allow obtaining reliable estimations of spine kinematics, and also focused on a quantitative performance comparison between the results achieved by using different derivative operators.

4.1   Spine kinematics

Spine kinematics aims to detect abnormalities in the motion of the spine segments, in order to diagnose spine pathologies. To this aim, spatial relationships between the vertebrae are measured during various activities that involve a specific spine motion, to determine whether those are in a normal range of motion (ROM) [91 – 92]. However, many studies have shown that a number of both cervical and lumbar spine pathologies do not show up with abnormal ROM, and could be rather revealed by detecting abnormal patterns of intervertebral motion. A well-known example is the cervical spine instability, which arises from the ligamentous disruption with subsequent atlantoaxial instability, often caused by traumatic flexion-extension movements exerted on the spine, e.g. in car accidents or heavy impacts on the neck [93]. It is also related to other pathologies, such as rheumatoid arthritis [94], or to congenital deviation, like those involved in the down syndrome [95]. Such a pathological condition prevents the vertebrae from maintaining their normal pattern of displacements

under physiological loads, which precisely ensures the spine stability in healthy subjects [91]. Low back pain (LBP) is another well-known pathology of the spine that can lead to substantial disability. It is considered as a consequence of degenerations of lumbar vertebral segments, often caused by degenerative disease, facet joint hypertrophy, postoperative status, postoperative spinal fusion, trauma, scoliosis or other congenital spine lesions, which can significantly affect spinal load-bearing capability [96 – 97]. The increment of the intervertebral joints that occur in these spine disorders, due to a combined action of disks, ligaments, vertebral joints and muscle degenerations, can lead to abnormal kinematics of the intervertebral segments [98 – 99].

Currently, functional flexion – extension radiography is the most widespread method for the clinical diagnosis of spine instability. However, it is limited to few end-of-range spinal positions, usually one radiograph in a neutral position and two radiographs in the final flexion and extension positions. Indeed, motion capture is a form of time sampling, as it is ultimately aimed at deriving kinematic signals; hence, using N radiographs to describe the spine motion is equivalent to acquire N samples of the time evolution of the segmental displacements pattern, which means that the continuous motion between two time samples would not be captured and should be reconstructed via interpolation methods. However, the use of a limited number of samples undoubtedly places a limitation on the information that could be extracted from the kinematic signals thus sampled, as it is well-known from the Nyquist – Shannon sampling theorem. In particular, the derivative operation for discrete signals exhibits an high sensitivity to the sampling period, showing substantial errors for poorly sampled signals.

Essentially, flexion – extension radiography is suitable for finite motion measurements (e.g. ROM), but not for measurements of velocities and accelerations associated to relative movements of intervertebral segments, as they require the computation of first and second derivatives of displacement signals. Nonetheless, many authors suggested that the instantaneous centre of rotation (ICR), which is determined from measurements of linear and angular displacement and velocities, is the most sensitive and specific measurement to detect the damage of intervertebral disk and facet joints, also at early stages. The ICR stands also as a reliable, stable parameter of the quality of vertebral motion, which allows the detection and analysis of abnormal motion patterns. Indeed, according to different studies, the ICRs were found to exhibit a wider scatter on patients with neck pain as compared to healthy subjects, while studies on patients with cervical headache report that abnormal ICRs were found in the upper cervical segments. Other studies, also on LBP, found that disk degeneration may lead to abnormal location of the related intervertebral ICR, while maintaining finite intervertebral translation and rotation within a physiological ROM [3, 91, 100 – 102].

The estimation of spine kinematics thus requires a continuous screening of specific spinal tracts during patients' motion, which cannot be provided by radiography. Instead, this task can be rather accomplished via fluoroscopy, also with a reasonably low X-ray dose [3 – 4, 99]. Certainly, the high number of frames to be processed prevents performing a manual tracking of vertebral bodies, as it would be undoubtedly burdensome and unreliable. Indeed, the vertebrae tracking process must be very accurate and precise, as even very little errors in the estimation of displacements and rotations of vertebrae generally result in much higher errors in the reconstruction of intervertebral kinematics, especially of ICR trajectories. An automated method for spine kinematics estimation from fluoroscopic sequences was proposed in [3 – 4]. It is based on a template matching operation that computes a normalized cross-correlation index between the image derivatives of a vertebra template

and of multiple rotated versions of the current frame, in order to match their contours and, thus, estimate the horizontal and vertical displacements as well as the rotation angle of the selected vertebra in each frame of the acquired fluoroscopic sequence. The computation of image derivatives requires to filter the images with derivative operators, which are basically approximations of continuous derivatives in the discrete domain.

4.2   Application and performances of derivative operators

Many different derivative operators have been proposed in literature [78, 103]. However, there is no general consensus on which one provides the best performances, because there is no unequivocal definition of edges, which is rather dependent on the specific task. One of the most common applications of derivative operators is certainly edge detection, which is considered by far one of the key steps to analyse medical images, with the aim to extract reliable diagnostic information from them. After computing the image derivatives, the edge detectors perform further processing to extract the binary edge maps. Since the noise is known to introduce spurious edges, by causing abrupt luminance variations, common edge detectors usually apply thresholding operations on the image derivatives with the aim to maintain the strongest edges, considered as the "true" edges, while completely disregarding the weaker ones, which, ideally, should only be the noisy ones. Indeed, the quality of the detected edges is highly dependent on the validity of the assumptions made by the edge detector, which could fail because of image noise, different lighting conditions, local contrast variations and edge density [104]. This unavoidably results in a loss of part of the information, depending also on the selected threshold, which plays a key role in the performances of edge detection. However, general-purpose methods for self-adaptation of the threshold have not been proposed yet, and the optimal threshold is very often identified by means of a trials and errors approach [105 – 106]. Similarly, in literature there is a lack of quantitative studies on derivative operators performances for image processing applications, even for specific tasks, and the common trend is to select an operator and heuristically find the threshold which yields the best performance. In medical imaging, it is even worse, as in most cases the evaluation is generally limited to qualitative aspects and the measurement of edge detection performances is merely based on how well edge detector markings match with the visual perception of object boundaries of the human operator [107]. Although the literature is dominated by edge detection applications, the derivative operators have also many other applications, such as the template matching method for spine kinematics described above that achieves higher performances by directly operating on image derivatives rather than binary edge maps. In this study the performances of different derivative operators were evaluated and compared by analysing the results yielded in the estimation of ICR trajectories from a low-dose fluoroscopic sequence of a guided flexion-extension motion of the lumbar spine of a healthy subject.

4.3   Lumbar spine motion protocol

The data used in this study have already been presented in a previous work [108] and consisted of a sagittal fluoroscopic image sequence of a healthy subject (see Figure 4), performing a flexion-extension lumbar spine motion while lying on his side. In order to ensure a controlled and passive motion, the subject was tied with the $L4 - L5$ segment bind to the hinge of a motorized table, which was programmed to perform a smooth arc to 40° left from the neutral position, then to 40° right and, finally, return back to the neutral position. The X-ray tube was set at 73 kV, 2 mA and the duration

of the whole screening was 30 s, with an overall measured dose of 0.9 Gy·cm^2 [109]. The X-ray images were acquired with a frame rate of 5 frames per second and a spatial resolution of 0.43 mm × 0.43 mm (pixel dimension), and an 8-bit grayscale resolution.

Figure 4. A frame of the raw fluoroscopic sequence.

4.4 Noise reduction with NVCA

Although cross-correlation is generally quite robust to noise, low-dose fluoroscopic images are usually characterised by a very poor image quality, which may invalidate the results of template matching methods. For this reason, the raw fluoroscopic sequence was pre-processed via NVCA denoising, which required a preliminary estimation of noise parameters. To this aim, a static scene at the end of the raw sequence, during which the table did not move and kept the subject fixed, was extracted by considering the last 30 frames. The experimental EVaR was obtained by computing the sample mean and variance for each pixel along the temporal dimension (i.e. by considering the grey levels assumed by the same pixel in subsequent frames) and reported in Figure 5. Its scattered appearance is due to the variability of the sample variance for such a small number of observations, which grows with the mean, as expected from the theoretical EVaR of the Poisson noise. The noise parameters, i.e. the slope a and intercept b of the straight line representing the theoretical EVaR, were obtained by performing a linear regression on the measured data. The linear regression and related parameters are reported in Figure 5 as well.
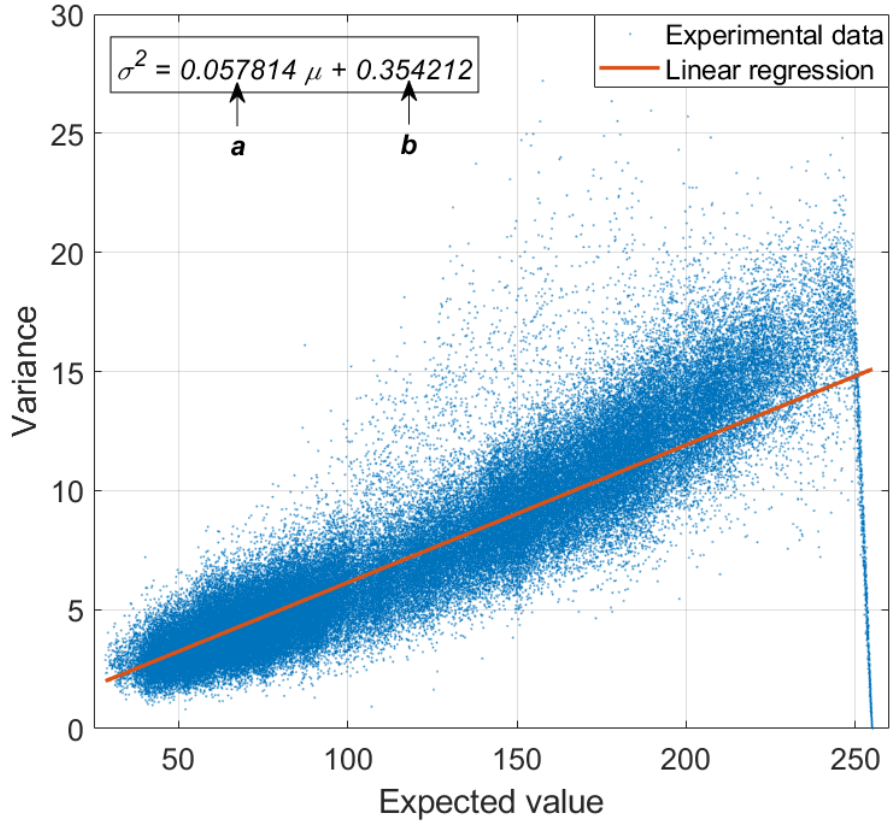
Figure 5. 2D plot of measured EVaR with linear regression and estimated noise parameters

Finally, noise reduction was accomplished by applying NVCA with a 7×7×3 spatio – temporal mask and a noise SD threshold multiplier $N_\sigma = 3$. The effectiveness of NVCA denoising can be assessed by visual comparison of Figure 6.(a-b), which show the same frame from the raw noisy sequence and the denoised sequence. It is worth noting how the NVCA algorithm substantially removed the so-called "quantum mottles", while preserving the edge sharpness also in areas with very low local contrast, such as the L4 – L5 segment at the bottom of the frame. This can be probably better appreciated in Figure 6.(c-d) that show the corresponding edges extracted from the considered frames.

4.5   Quantitative comparison of ICR trajectories

The horizontal and vertical image derivatives to be fed to the template matching block of the spine kinematics analysis were computed from the sequence denoised via the NVCA algorithm. The first-order derivative operators used to produce the different image derivatives, were proposed by Roberts, Sobel, Prewitt and Cerciello *et al.* [3], and their kernels are reported in (26), (27), (28) and (29), respectively.

(a)

(b)
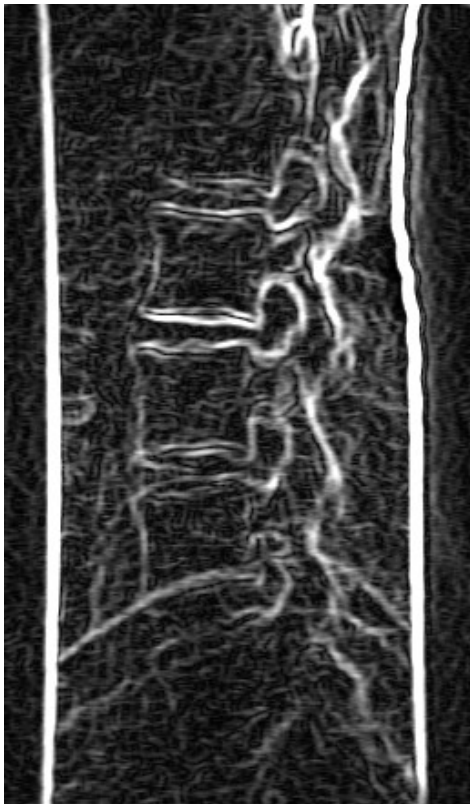
(c)
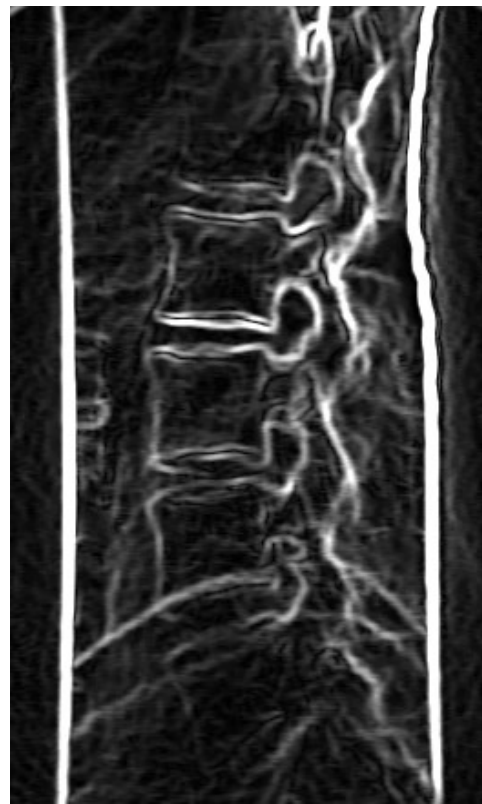
(d)

Figure 6. (a) Frame from the original noisy sequence; (b) corresponding frame from the sequence denoised with NVCA; (c-d) edges extracted with Cerciello et al. operators from the frames depicted in (a-b), respectively.

24

$$H1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad H2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \qquad (26)$$

$$H1 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad H2 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \qquad (27)$$

$$H1 = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad H2 = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad (28)$$

$$H1 = \begin{bmatrix} -1 & -1 & 0 & 0 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & 1 \end{bmatrix}$$

$$(29)$$

$$H2 = \begin{bmatrix} -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

The method for the extraction of intervertebral kinematics from fluoroscopic sequences described in [3], first requires the used to select a the template of the vertebra to be tracked from the first motionless frame (see Figure 7.(a)). Then, the template matching block computes a normalized cross-correlation index (NCC) [4] between the derivatives of the vertebra template and of several rotated versions of each considered frame. The NCC formula is reported in (30):

$$NCC_x(i,j) = \frac{\sum_{x=-N/2}^{N/2} \sum_{y=-M/2}^{M/2} g_x(i+x, j+y) \cdot t_x(x,y)}{\sqrt{\sum_{x=-N/2}^{N/2} \sum_{y=-M/2}^{M/2} g_x^2(i+x, j+y)} \cdot \sqrt{\sum_{x=-N/2}^{N/2} \sum_{y=-M/2}^{M/2} t_x^2(i+x, j+y)}}$$

$$NCC_y(i,j) = \frac{\sum_{x=-N/2}^{N/2} \sum_{y=-M/2}^{M/2} g_y(i+x, j+y) \cdot t_y(x,y)}{\sqrt{\sum_{x=-N/2}^{N/2} \sum_{y=-M/2}^{M/2} g_y^2(i+x, j+y)} \cdot \sqrt{\sum_{x=-N/2}^{N/2} \sum_{y=-M/2}^{M/2} t_y^2(i+x, j+y)}} \qquad (30)$$

$$NCC(i,j) = \frac{1}{2} NCC_x(i,j) + \frac{1}{2} NCC_y(i,j)$$

where M, N are the dimension of the vertebra template, t_x, t_y are the horizontal and vertical components of the image gradient of the vertebra template, and g_x, g_y are the corresponding components of the image gradient of the rotated version of the current frame. The rotation angle, as well as the horizontal and vertical displacements of the tracked vertebra are estimated by identifying the rotated version of the current frame which scores the highest NCC and, then, locating the spatial position of its global maximum. Finally, an interpolation of intervertebral kinematic data (i.e. the relative displacements and rotation angle of the upper vertebra with respect to the lower vertebra) is carried out with quintic smoothing splines to obtain a continuous description of the time evolution of

intervertebral displacements and rotation angles, which allows to derive also the linear and angular velocities and, eventually, the trajectory of the intervertebral ICR.

The described method was applied to extract four distinct descriptions of the kinematics of the L2 – L3 intervertebral segment, by processing the image derivatives provided by each of the four considered derivative operators. Figure 7.a depicts the selected templates for the L2 and L3 vertebrae with boundaries highlighted in red and with two red squares of 5×5 pixels located at their centres. Figure 7.b shows a subsequent frame of the sequence, where the two aforementioned red squares were located at the estimated centre coordinates of the highest template matchings for the two vertebrae.

(a)   (b)
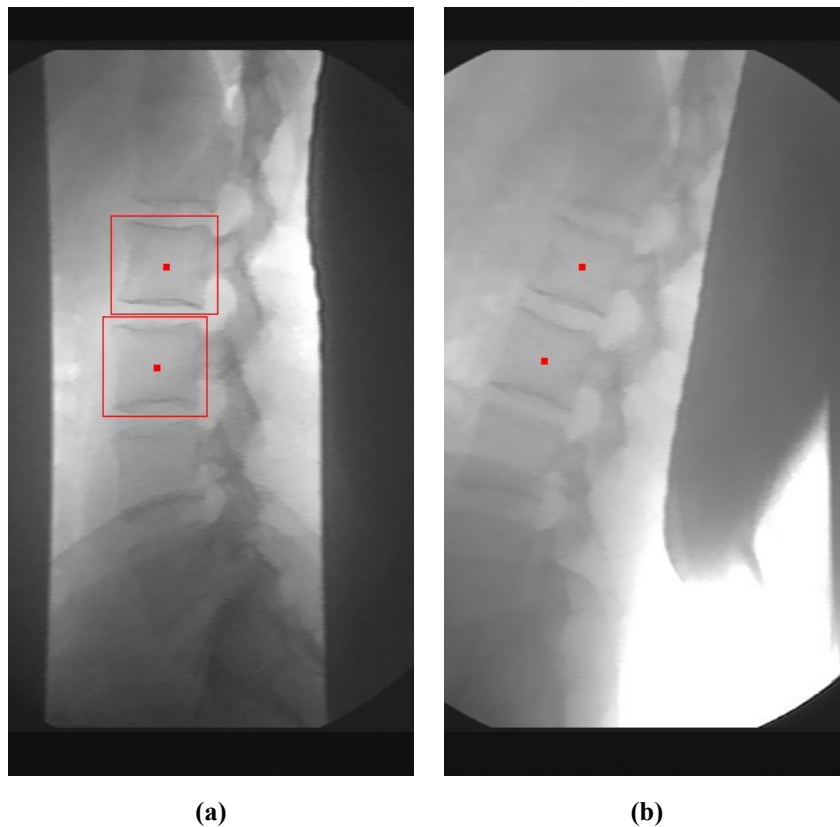
Figure 7. Initial frame with centres and borders of upper and lower vertebrae templates highlighted in red; (b) subsequent frame with estimated centres of upper and lower vertebrae highlighted in red.

The raw kinematic data of the L2 – L3 intervertebral segment estimated from the image derivatives of each frame were interpolated with quintic smoothing splines by using an equivalent cut-off frequency of 0.250 Hz (corresponding smoothing factor of 0.936), and then linear and angular velocities were computed as first analytical derivatives of the continuous splines descriptions. For the sake of comparison, the velocities were also derived directly from the raw data, by means of first order differences, to demonstrate the amount of errors obtained even with poorly time resolved kinematic signals extracted from continuous X-ray screening of body motion. To this aim, raw and spline interpolated data, obtained by using Cerciello *et al*. derivative operators, are reported in Figure 8. While the displacements exhibited a pretty good accordance, the velocities computed from raw data turned out to be very noisy, as opposed to those obtained from interpolated data, which showed

up as quite smooth. Considering well-known ICR sensitivity to errors of kinematic data, the velocities computed from raw data could not be used for ICR coordinates computation.
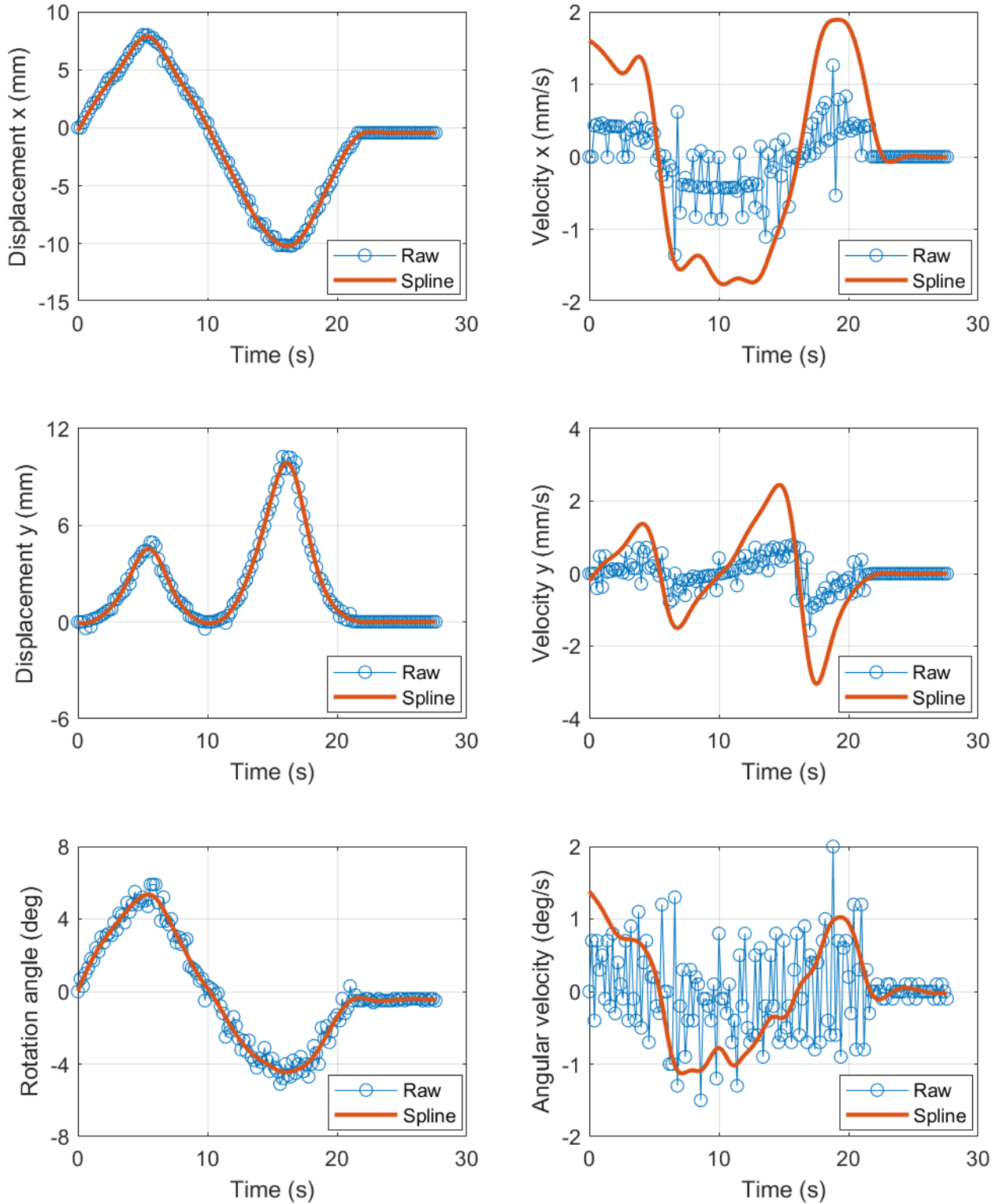
Figure 8. Comparison between raw and spline interpolated data of intervertebral kinematics obtained with Cerciello et al. operators

The ICR trajectories, i.e. the time sequence of ICR spatial coordinates, were computed from the splines descriptions of kinematic signals via the following formulas (31):

$$\begin{cases} ICR_x(t) = -\dfrac{v_y(t)}{\omega(t)} + r_x(t) \\ ICR_y(t) = \dfrac{v_x(t)}{\omega(t)} + r_y(t) \end{cases} \tag{31}$$

The ICR trajectories computed from the different image derivatives provided by the four derivative operators were quantitatively compared in pairs, by computing the mean and SD of the Euclidean distances of their corresponding points, as well as the trajectory of their point-wise differences (referred to as "difference trajectory") and the parameters of the related minimum bounding circle, i.e. the radius and centre coordinates. The results obtained for the six pairs of ICR trajectories were organized in Table 1 and Table 2.

Table 1. Euclidean distance and minimum bounding circle parameters of ICR trajectories pairs

Derivative Operator	Euclidean distance		Minimum bounding circle	
	Mean (mm)	*SD (mm)*	*Centre-Origin distance (mm)*	*Radius (mm)*
Cerciello *et al.* - Prewitt	0.37	0.09	0.47	0.29
Cerciello *et al.* - Sobel	0.39	0.10	0.52	0.31
Cerciello *et al.* - Roberts	0.59	0.14	0.72	0.46
Prewitt - Sobel	0.03	0.02	0.03	0.06
Prewitt - Roberts	0.24	0.12	0.25	0.27
Sobel - Roberts	0.23	0.11	0.22	0.25

Figure 9 shows the six pairs of compared ICR trajectories (left column), along with their corresponding difference trajectories with the minimum bounding circles (right column). Finally, in Figure 10 the ICR trajectory obtained with Cerciello *et al.* operators was superimposed on the first frame of the filtered sequence (assuming the centre of the lower vertebra's template as the origin of the reference system), in order to provide visual feedback and to allow the contextualization of the ICR coordinates locus into the actual intervertebral segment.

Prewitt and Sobel operators achieved the most similar results: indeed their related ICR trajectories scored both the minimum mean and SD of Euclidean distances (0.03 mm ± 0.02 mm) as well as the minimum radius (0.06 mm) and centre – origin distance (0.03 mm) of the minimum bounding circle of their difference trajectory. On the other hand, Cerciello et al. and Roberts

operators, provided the most different results, as their related ICR trajectories scored both the maximum mean and SD of Euclidean distances (0.59 mm ± 0.14 mm) together with the maximum radius (0.46 mm) and centre – origin distance (0.72 mm). Moreover, the trajectory obtained with Cerciello et al. operators resulted to be, on average, the most distant from the others, as, according to Table 2, the trajectory achieved the maximum average values for the mean Euclidean distance from other trajectories (0.45 mm), and also for the centre-origin distance (0.57 mm) and radius (0.35 mm) of the minimum bounding circle of the difference trajectories.

Table 2. Average parameters of each ICR trajectory with respect to the others.

Derivative Operator	Euclidean distance		Minimum bounding circle	
	Mean (mm)	*SD (mm)*	*Centre-Origin Distance (mm)*	*Radius (mm)*
Cerciello *et al.*	0.45	0.11	0.57	0.35
Prewitt	0.21	0.08	0.25	0.21
Sobel	0.22	0.08	0.26	0.21
Roberts	0.35	0.12	0.40	0.33

Detectable differences were found in the obtained ICR trajectories, but their significance level for clinical applications should be further evaluated, in order to assess if the derivative operators under discussion could be considered as equivalent or not. However, it is not possible at the moment, because there is no consensus and sufficient data on kinematic parameters features associated with specific spinal pathologies. Anyway, the results highlighted the need for a quantitative analysis of derivative operators performances in applications aimed at providing quantitative results, which in this case was carried out by quantitatively analysing the effects of different derivative operators on the eventual ICR trajectories provided to researchers/doctors. Surprisingly, this issue is rarely addressed in medical imaging studies and definitely deserves much more attention, considering the large use of derivative operators in the plethora of medical image processing applications proposed in literature.

(a)

(b)

(c)

(d)

(e)

(f)
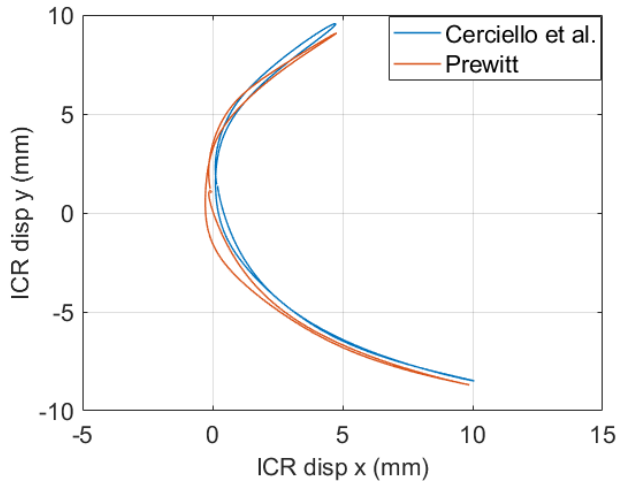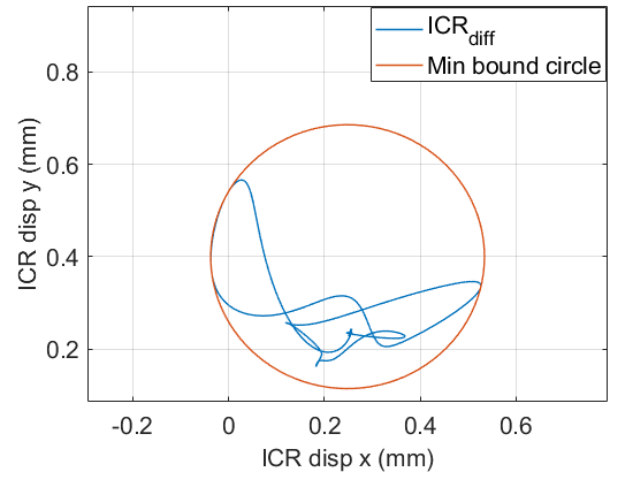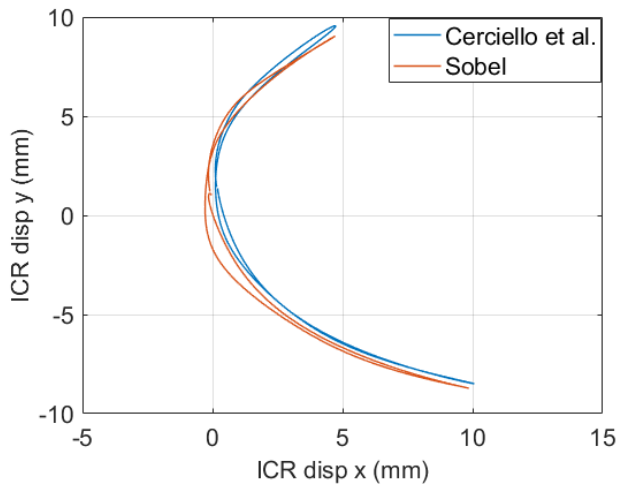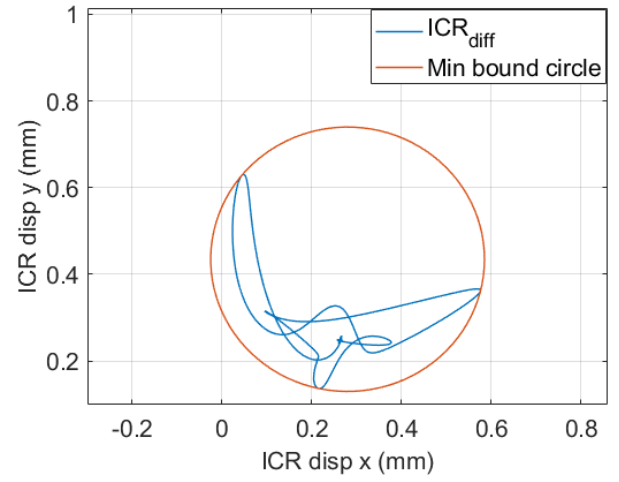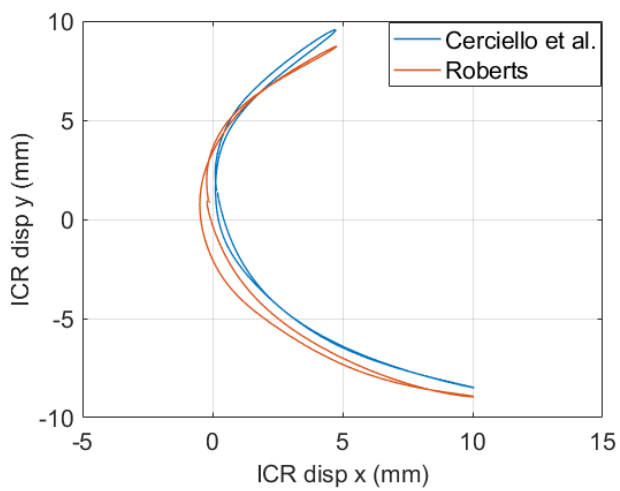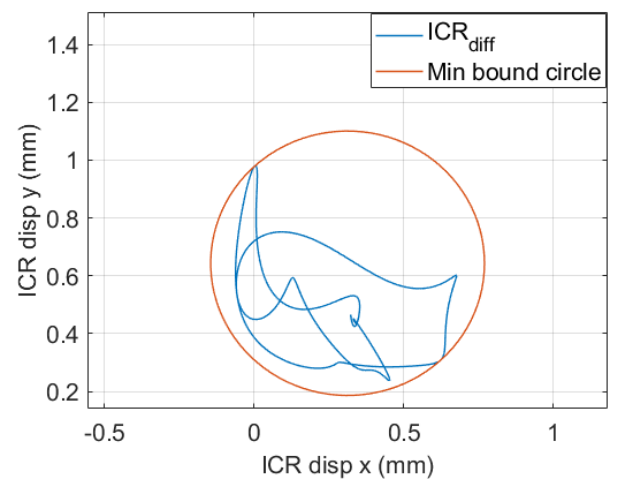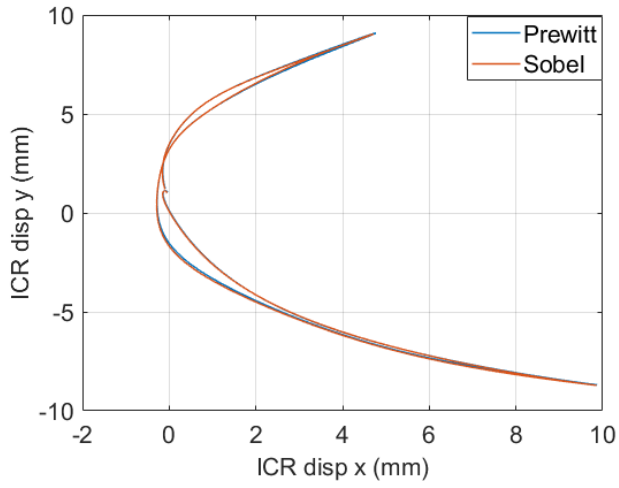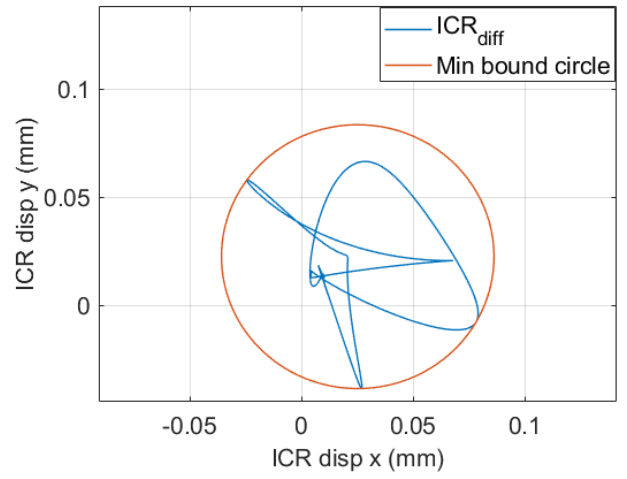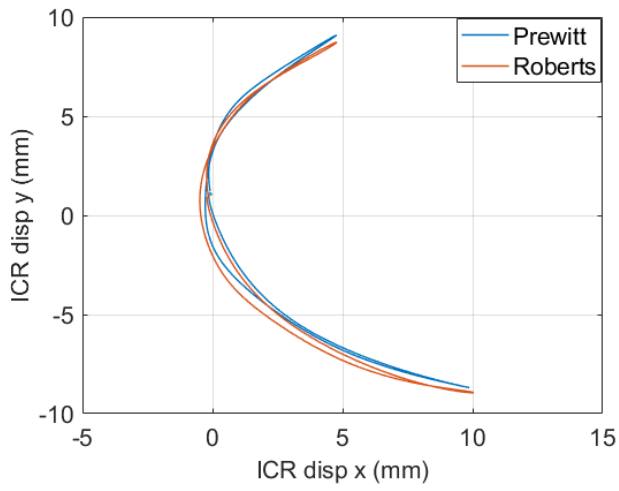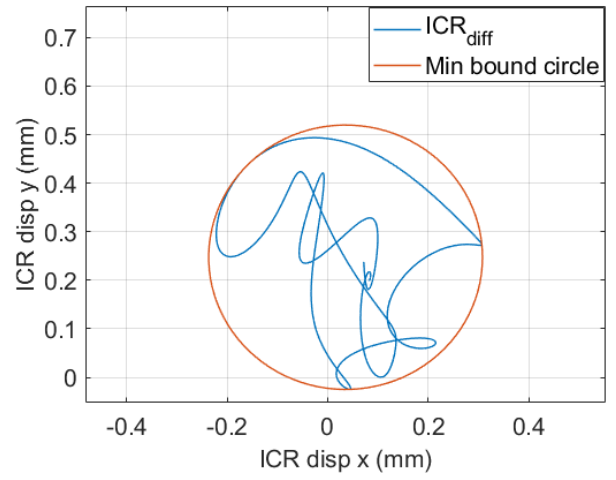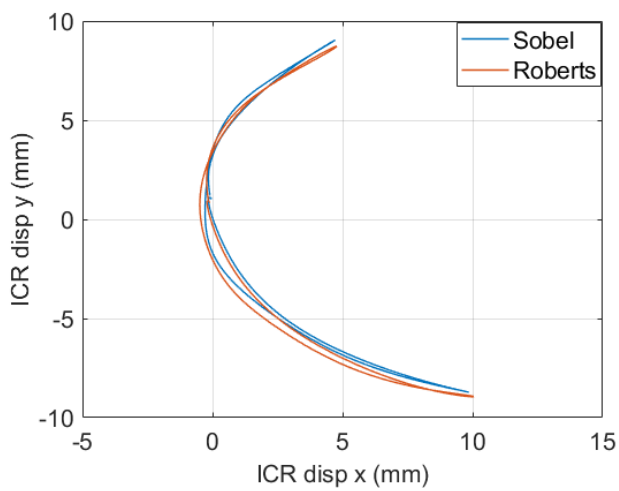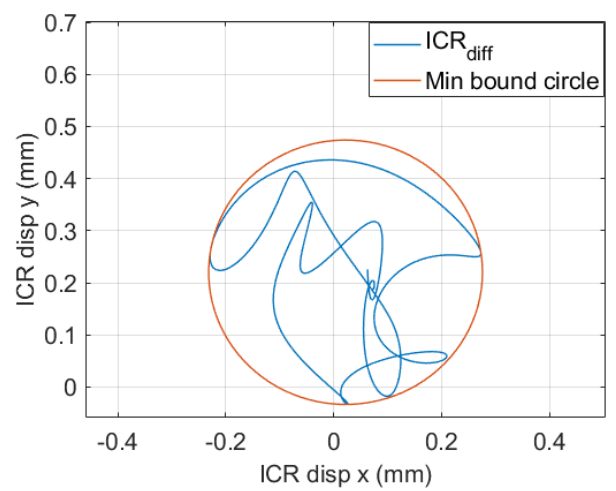
Figure 9. Comparison of ICR trajectories obtained with different derivative operators. Panels on the left column show the pairs of compared trajectories, while panels on the right column show their corresponding difference trajectory along with the minimum bounding circle.
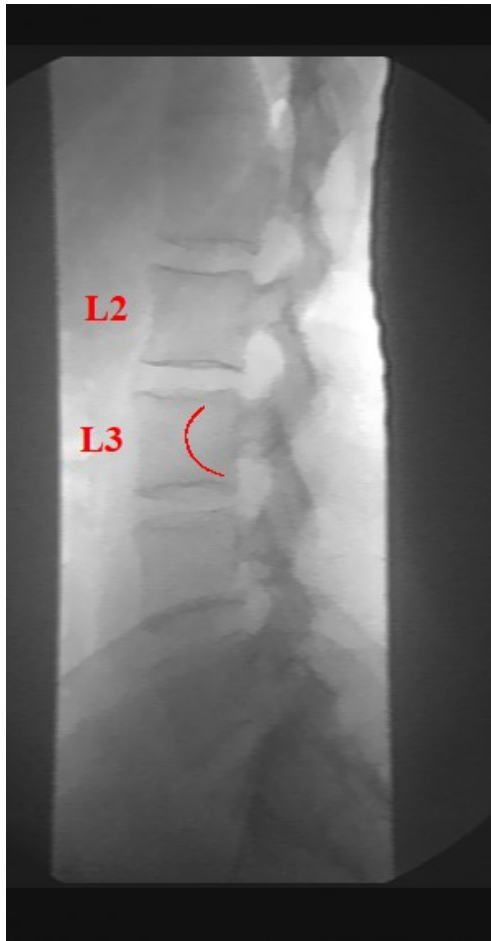
Figure 10. Superimposition of the ICR trajectory, obtained with Cerciello et al. operators, on the initial frame, with origin of reference system set at the centre of the lower vertebra template.

Chapter 5

Performance comparison with state-of-the-art algorithms

5.1   State of the art of denoising algorithms

To date, a plethora of algorithms have been proposed for noise removal in images and videos, which are based on very different principles as well as assumptions on distinctive features of images and noise [78, 103]. Nonetheless, while scientific literature is rich in approaches devised for space-invariant additive white gaussian noise (AWGN), undoubtedly much less effort has been devoted to directly address the problem of noise suppression in Poisson-corrupted data. Since the effectiveness of denoising algorithms is strongly dependent on the validity of the considered noise model, it follows that AWGN denoising strategies cannot be successfully applied to Poisson noise removal [4, 78, 103, 111]. Indeed in [40] NVCA was tested on single natural images corrupted by simulated Poisson noise and outperformed several state-of-the-art denoising algorithms (e.g. Adaptive Variational denoising [112], 4-PDE denoising [113], Total Least Squares denoising [114], K-SVD [115], Wavelet-based Hidden Markov Models [116]), especially those designed for AWGN, while also entailing a much lower computational burden. In spite of its simplicity, NVCA was so efficient in reducing noise power that it scored a Peak Signal-to-Noise Ratio (PSNR) only about 2 dB lower than the much more complex BM3Dc algorithm [46], which is an adaptation for Poisson-Gaussian noise of the most famous Block Matching and 3D filtering (BM3D [117]), still considered a state-of-the-art algorithm for single image Gaussian denoising [118 – 121].

BM3Dc relies on a variance stabilising transformation (VST), i.e. a point-wise operation that transforms an heteroscedastic random distribution into an homoscedastic one (e.g. conversion from signal-dependent to signal-independent noise). The VST for Poisson-distributed data is known as Anscombe transform (AT) [122], which provides a practically Gaussian distribution with unit variance.

$$I_{AT}(m,n) = 2\sqrt{I(m,n) + \frac{3}{8}} \tag{32}$$

$$\begin{cases} E[I_{AT}(m,n)] = 2\sqrt{E[I(m,n)] + \frac{3}{8} - \frac{1}{4E[I(m,n)]^{1/2}} + O\left(\frac{1}{mE[I(m,n)]^{3/2}}\right)} \\ Var[I_{AT}(m,n)] = 1 + O\left(\frac{1}{E[I(m,n)]^2}\right) \end{cases} \tag{33}$$

Therefore, AT should allow applying all the denoising methods devised for AWGN to the suppression of Poisson noise. Unfortunately, the AT cannot be directly applied to X-ray images, because,

according to the Poisson – Gaussian noise model presented in Chapter 2, the actual distributions of pixels grey levels result from the transformation of the pure Poisson distributions of photons arrivals, due to the detector gain and the additional electronic noise (zero-mean AWGN). The proper VST for Poisson-Gaussian distributions is the Generalised Anscombe Transform (GAT) [50 – 51, 64], which requires the knowledge or the estimation of the noise parameters *(a, b)*:

$$I_{GAT}(m,n) = \frac{2}{a}\sqrt{a \cdot I(m,n) + \frac{3}{8}a^2 + b} \qquad (34)$$

However, the inverse GAT, which is required to bring the processed data back in the original domain, is known to suffer from bias errors [50 – 51]. It is also worth underlining that, according to (33), the variance stabilization provided by both AT and GAT is fully effective only for sufficiently high expected luminance values, which means that for images obtained with lower X-ray doses (i.e. lower luminance and number of photons per pixel) it would introduce increasingly higher errors. Moreover, real fluoroscopic data are usually modified by sensor non-linearities (e.g., clipping effects) and by white compression operations, which are commonly implemented in commercial fluoroscopic devices (e.g. log-mapping, gamma-correction) [40]. Inverting these non-linearities would require proper strategies to detect transformed data, infer the adopted transformation and estimate its parameters: these are not straightforward task and would result, anyway, in a further increase of the overall computational burden, which is usually not so low. Indeed, the state-of-the-art algorithm for AWGN denoising usually exhibit high computational complexities and, sometimes, unpredictable computational times, when they are based on iterative solving of minimization problems. These aspects usually prevent them from achieving real-time operation, thus making them suitable for post-processing only.

Other recent approaches mainly concentrate on edge enhancement, to recover from blurring effects caused by noise reduction [72 – 74] and on the recognition of curvilinear structures (e.g. guidewires, catheters) [68], but they do not provide a global approach to noise reduction. Surprisingly, although it has long been known that the majority of fluoroscopy screenings is employed in interventional procedures, there is still a lack of effort in designing computationally efficient approaches for real-time processing of low-dose fluoroscopic sequences, focusing rather on decidedly more complex approaches (e.g. [69 – 70, 72 – 75]) that have not been demonstrated yet to outperform the simpler denoising algorithms devised for real-time fluoroscopic noise suppression, such as NVCA. Nevertheless, some of the authors of the BM3D algorithm, which has originally been devised for single image denoising, have more recently proposed two extensions for video processing, which are likewise widely regarded as the state-of-the-art for video denoising, and, indeed, they have been used as a reference for the performance assessment of new algorithms in many recent studies [43, 54, 76]. The most recent of them is the Video Block-Matching and 4D joint filtering (VBM4D) algorithm [123], which undoubtedly offers improved performances by explicitly taking advantage of the temporal dimension through an approach based on motion estimation. The operating principle at the core of VBM4D, however, is the same of BM3D (i.e. block matching and collaborative filtering in a transform domain), which has been proved to be efficient for images corrupted by AWGN. Considering the promising performances that this state-of-the-art video denoiser could achieve in Poisson – Gaussian noise reduction when combined with the GAT, a performance comparison with the NVCA algorithm was carried out on synthetic and real low-dose fluoroscopic sequences, both static and dynamic, also with a focus on computational performances. The performances of a simple

spatio – temporal moving average (MA) filter were also assessed to provide a reference for the effectiveness of filtering operations that are usually implemented in commercial fluoroscopic devices.

5.2 VBM4D

In VBM4D spatio-temporal volumes are constructed by tracking blocks along trajectories defined by the motion vectors and mutually similar volumes are stacked along an additional fourth dimension, thus producing a 4-D structure, termed group [123]. The VBM4D extends the paradigm of non-local grouping and collaborative filtering of BM3D, to exploit not only the spatial redundancy, but also the temporal redundancy that characterizes the video sequences. The spatio – temporal block-matching operation allows for tracking and reallocation of moving objects as well, which ensures a deep noise reduction with a good edge preservation, at least for the natural images it was devised for. These advantages made VBM4D a state-of-the-art algorithm for effective video denoising and, thus, a reference for performance assessment of novel algorithms [43, 54, 70, 74 – 77, 124 – 125]. On the other hand, VBM4D needs the entire image sequence to be processed and it is also very time consuming, thus not being suitable for a real-time implementation.

5.3 Computational complexity and computational time

The computational burden of the NVCA and VBM4D algorithms can be compared in terms of the number of required multiplications/divisions for an input of size N, which is generally expressed by the big-O notation [89 – 90]. Let consider a single frame of $\sqrt{N} \times \sqrt{N}$ pixels to be processed. The BM3D algorithm can be mainly decomposed in two phases: the block matching (grouping) and the 3D collaborative filtering. The block matching phase has a computational complexity order of $O(N^2 K)$, with a $\sqrt{K} \times \sqrt{K}$ mask size [126 – 127]. The 3D collaborative filtering phase provides for the processing of the 3D arrays identified by the block matching operation, and consists of 3D transformation, spectrum shrinkage, and 3D inverse transformation. Considering the case of a 3D-FFT transform [128] (which has a computational complexity of $O(N \log N)$, with N being the total number of data to be processed) applied to a single 3D array composed by S similar blocks of $\sqrt{K} \times \sqrt{K}$ pixels, the computational complexity order can be expressed as $O((K \cdot S) \log (K \cdot S))$. Indicating with R and G, respectively, the total number of $\sqrt{K} \times \sqrt{K}$ blocks in the image, and the number of 3D arrays identified by the block matching operation, and assuming every single 3D array to be composed by S similar blocks, it is possible to express the computational complexity order of the whole 3D filtering operation as $O(G \cdot ((K \cdot S) \log(K \cdot S)))$. Since $G = R/S$ and $R = N/K$, the computational complexity can be expressed as $O(N \log(K \cdot S))$. If the block matching does not find any similar blocks (i.e. $S = 1$), the computational complexity order is $O(N \log K)$. Alternatively, if the block matching provides a single 3D array composed by all the R blocks (i.e. $S = R$), the computational complexity order is $O(N \log N)$. The computational complexity of the block matching is of some orders of magnitude higher than that of the 3D filtering, and so it dominates the computational complexity of the whole BM3D algorithm. Since VBM4D works along an additional dimension (i.e. time), thus processing a larger amount of data, its computational complexity is certainly higher than that of BM3D. NVCA, instead, provides for a single division operation per pixel [40, 60 – 61] and, therefore, considering same frame of $\sqrt{N} \times \sqrt{N}$ pixels, it has a computational complexity order of $O(N)$, which is several orders of magnitude lower than VBM4D's one. Computational times were estimated for MA, NVCA and

VBM4D filters to carry out a quantitative comparison of their actual computational burden. The evaluation of computational times was performed on a real fluoroscopic sequence acquired by imaging a static scene of a custom step phantom, which will be detailed in the following paragraphs.

5.4 Image quality assessment indices

A quantitative comparison of the improvements in visibility provided by the analysed algorithms requires objective measures of image quality. To this aim, three image quality assessment (IQA) indices were considered, namely the Full-Width at Half-Maximum of the line spread function (referred to as FWHM), the CNR and the Feature Similarity Index [129].

FWHM is a No-Reference IQA index that quantifies the blurring introduced by a filtering operation, by measuring the width of the line spread function obtained as the derivative of the edge spread function (i.e. the edge "step" profile). Essentially, FWHM measures the width of the edge region where there is a significant spatial luminance variation, so, the more the edge is blurred (i.e. its luminance is literally spread over its neighbourhood), the more the edge is large, scoring an higher FWHM. Usually, an horizontal or vertical edge is manually selected, which is considered as very sharp (at least before the potential edge blurring had occurred); then a line of pixels in the orthogonal direction is extracted and fitted with a sigmoidal function representing the edge profile, in order to reduce the influence of spurious noisy luminance fluctuations on the FWHM estimates. In this analysis the following function was used:

$$\psi(x) = \frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{x - c}{\sqrt{2}d}\right)\right) \tag{35}$$

where *erf(·)* denotes the error function, *x* is the spatial variable (pixel position along the selected line), and *c*, *d* are fitting coefficients, which basically determine the location and steepness of the edge spread function. The FWHM was computed as FWHM = 2.355 · *d*. It was estimated as the mean value over consecutive edge profiles and its uncertainty was estimated as the standard deviation over these estimates.

The CNR is a measure of details visibility and was adopted to quantify the local trade-off between noise reduction and details preservation. It was computed with the following formula:

$$CNR = \frac{\mu_{ROIA} - \mu_{ROIB}}{\sqrt{\sigma^2_{ROIA} + \sigma^2_{ROIB}}} \tag{36}$$

Where ROIA and ROIB refer to the regions of interest (ROI) shown in Figure 12, i.e. two smooth regions located at both sides of an edge (without including the edge region).

FSIM [129] is a Full-Reference IQA index, which requires a reference (ideally noiseless) image to be computed. It was developed within the framework of the well-known Structural Similarity index (SSIM), which quantifies image quality as the sum of three measured components, namely luminance, contrast and structure. SSIM is claimed to be based on a perceptual model that takes into account important predictors of the response to visual stimuli of the human vision system (HVS), thus providing a score that is highly correlated with the perceived image quality. However, in a very recent publication [130], Nilson & Akenine-Möller from NVIDIA have reported that this feature of SSIM lacks a scientific foundation and have also shown some paradoxes in its actual

behaviour. FSIM maintains the same expression for image quality, but replaces contrast and structure components with two novel measures, namely the phase congruency and the gradient magnitude, which take more into account the edge information and have been shown to represent low-level features employed by the HVS in the first interpretation of the scene, according to neurobiological studies [129, 131].

5.5 Static fluoroscopic sequence of a step phantom

The first analysed sequence was acquired with a commercial fluoroscopic device (GE 9900 Elite C-arm [132]), by imaging a static scene with the custom-made stop phantom that is schematically depicted in Figure 11. The round field of view at the patient table had a diameter of 9 inch; the X-ray tube peak voltage and current were set to 57 kVp and 1 mA, respectively, to simulate a real interventional cardiac procedure; a sequence of 712 consecutive frames of 328 x 333 pixels were acquired at a frame rate of 25 fps and the grey values were digitized in a 16-bit scale. The step phantom is composed of seven superimposed, square, 1 mm – thick aluminium sheets with edges of 30 cm, 26 cm, 22 cm, 18cm, 14 cm, 10 cm and 6 cm. They were piled-up with the centres placed on the same axis in order to produce a step phantom whose projected images presented flat regions with different average values. The sheets were fixed together by two metal screws and bolts, which are also visible in the object Figure 12, along with the ROIs defined for IQA.
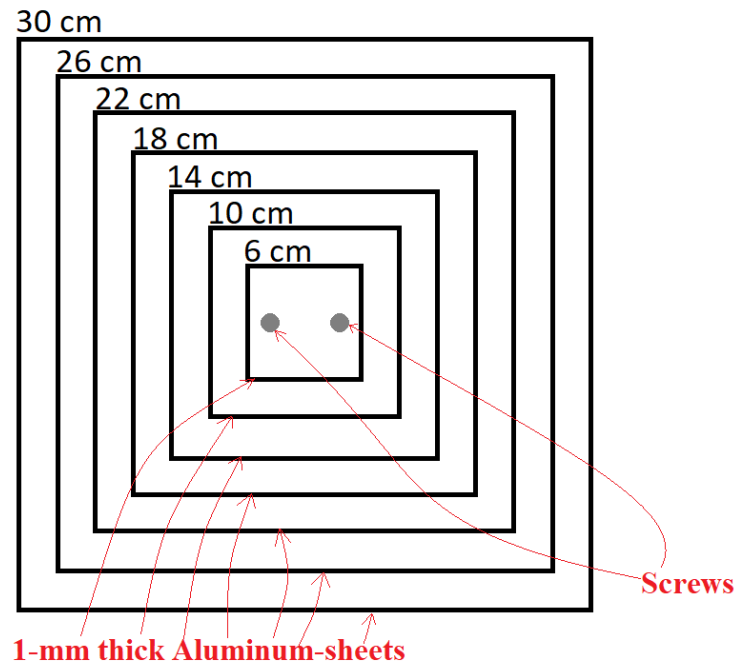
Figure 11. Schematic of the home-made test object used for the noise characterization.
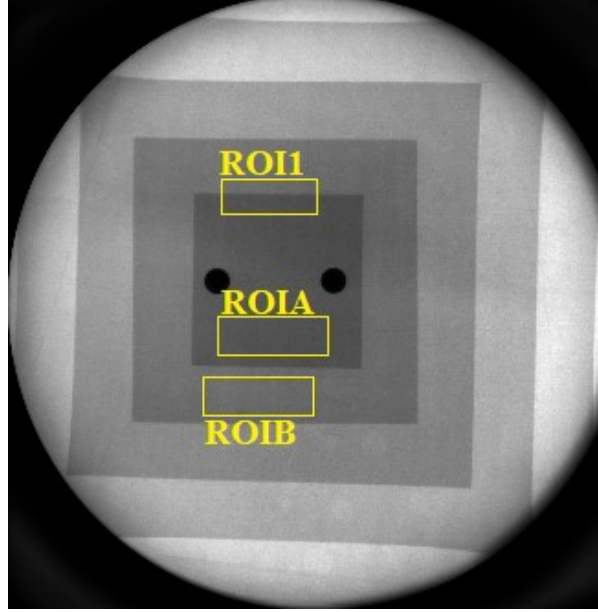
Figure 12. A frame of the aluminum step phantom fluoroscopic sequence. The manually selected regions of interest are shown in yellow: the ROI1 was chosen for the FWHM estimation and the ROIA and ROIB were chosen for CNR estimation. The two dark circles in the middle correspond to the two metallic screws and bolts that hold the aluminum sheets together.

Since both NVCA and GAT require the knowledge of noise parameters, the EVaR was measured from the acquired fluoroscopic sequence. The method presented in Chapter 4 was adopted, as well as the one presented in [47 – 48, 52], which takes also into account the data clipping due to sensor response, and provides an automatic segmentation of uniform regions in a single image, from which the noise parameters estimates are extracted. The former estimated a and b to be 43.90×0-4 a.u. and -0.99×10-4 a.u.^{2}, respectively, with an high goodness of fit (R^2 = 0.9196). The latter was performed in each of the 712 frames and the estimated a and b parameters were reported in distinct plots in Figure 13, by considering the frame numbers as abscissas. The median and SD values were 37.91×10-4 a.u. (std dev = 2.25 ×10-4) and 0.05×10-4 a.u.2 (std dev = 0.75 ×10-4) for a and b res. Figure 14, depicts the straight lines representing the EVaRs with the parameters estimated from the temporal sequences of pixels (dash dotted lines) and with the method described in [48].

The Poisson – Gaussian noise model described in Chapter 2 assumes the pixels grey values to be uncorrelated. However, the limited spatial resolution of the imaging system determines correlation between adjacent pixels in the spatial domain. In order to quantify such a correlation, the 2D autocorrelation matrix of ROIA (see Figure 12) was computed. This ROI was selected on a flat, uniform region (i.e. not including edges) at the centre of the field of view. The 2D autocorrelation matrix of ROIA was calculated for each frame and averaged across the computed matrices in order to obtain a single matrix. Finally, the 1D autocorrelation curve was calculated as the radial profile of the 2D matrix. The temporal correlation of consecutive pixel values was also studied by computing the 1D autocorrelation of the time sequence of each image pixel and averaging over all pixels. Figure 15 show a comparison between spatial and temporal 1D autocorrelation curves, where it can be observed that the temporal autocorrelation scored a value of 0.09 a.u.^{2} for a single inter-frame time interval, and, similarly, the spatial autocorrelation, scored a value of 0.15 a.u.^{2} for a single pixel size. These results confirm the hypothesis of weak, practically negligible temporal correlation, due to the reduced response time of modern flat panels (less than 1 ms versus the 33 ms sampling time,

corresponding to the maximum frame rate of 30 fps), but also the spatial correlation scored low values.

a)

b)
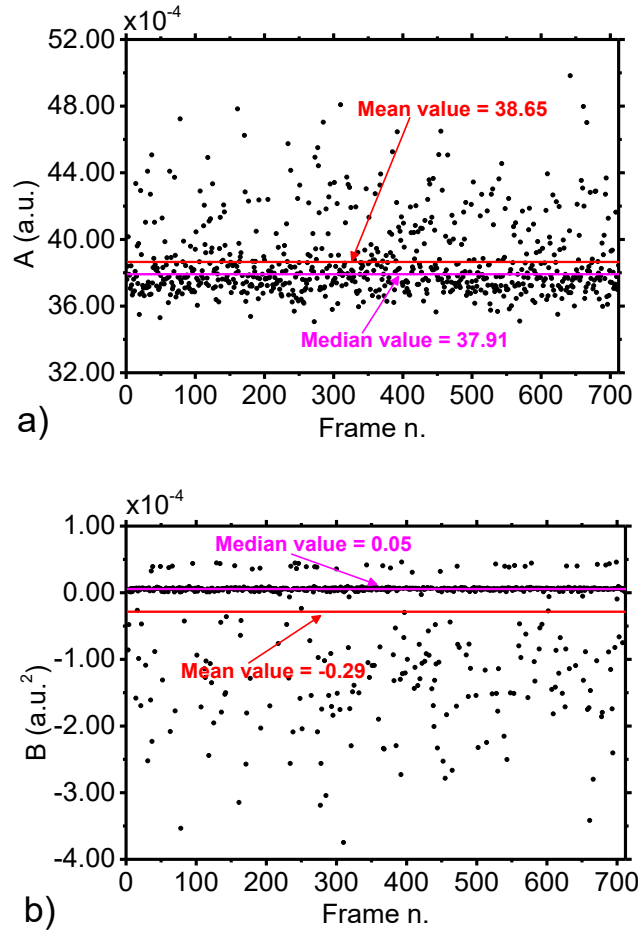
Figure 13. a) A and b) B coefficients estimated via the algorithm provided in [48]. Since A and B are estimated for each frame, the image frame number is reported on the x-axis.
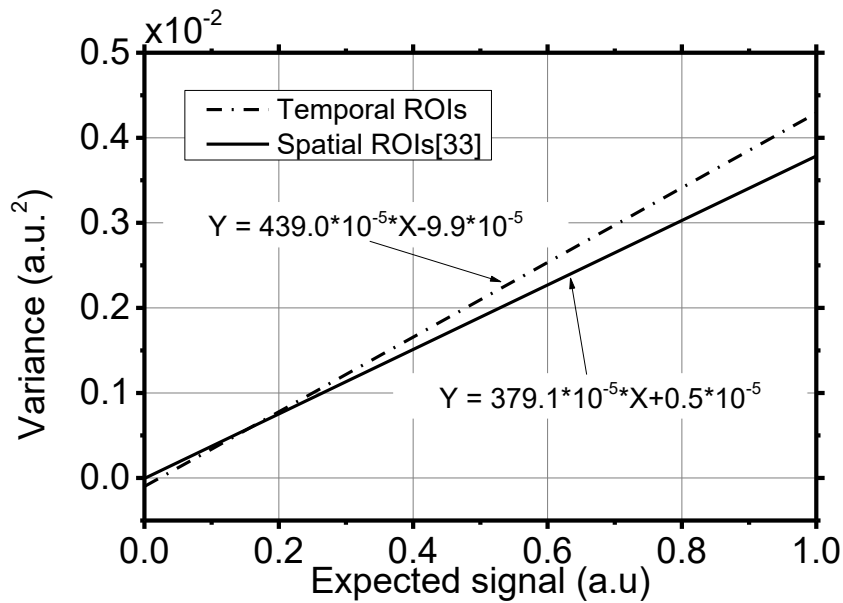
Figure 14. Expected value – variance relationship as estimated in space according to [48] and in time using the fluoroscopy sequence of the aluminum phantom.
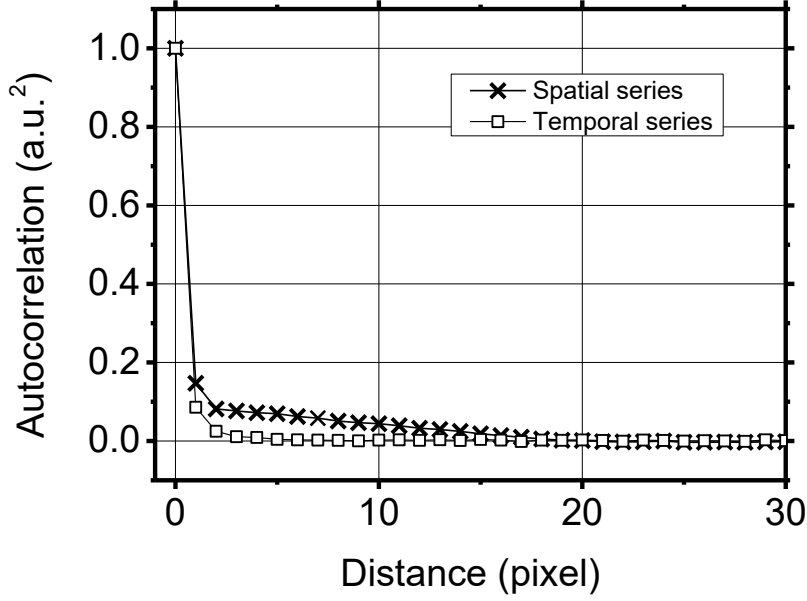
Figure 15. Average autocorrelation function evaluated in space (black crosses) and over time (white squares). The fluoroscopy sequence of the aluminum phantom was used.

After the noise parameters estimation, the step phantom sequence was filtered with VBM4D, NVCA with 3x3x3, 5x5x5 and 7x7x7 mask sizes and N_σ = 1, 1.5, 2, 2.5, 3, and spatio – temporal moving average with the same mask sizes used for NVCA. Figure 16 shows a comparison of the same frame extracted from the original noisy sequence and the sequences filtered via VBM4D, NVCA with 3x3x3 and 5x5x5 pixels mask sizes and N_σ = 2, spatio – temporal moving average with 3x3x3 and 5x5x5 pixels mask size. Figure 17 reports the profile across the edge outlined in yellow in Figure 16.a) for the original noisy sequence, as well as for the sequence processed via NVCA with a 7×7×7 pixels mask size and N_σ = 2, and for that processed via a moving average with a 7×7×7 pixels mask size. Figure 18 outlines the FWHM values, measured for the raw and processed data on 40 consecutive vertical profiles sampled in the ROI1 in Figure 12. As expected, it can be seen that by increasing N_σ, with a fixed mask size, also the FWHM achieved by NVCA filtering increased, and that happened also for fixed N_σ for increasing mask sizes (1.8 pixels for N_σ = 1, 3.1 pixels and 4.4 pixels for N_σ = 2 and N_σ = 3, respectively) Moreover, the FWHM achieved by the moving average turned out to be always higher than the NVCA ones for corresponding mask sizes, which indicated an higher smoothing. VBM4D, instead, caused a decidedly smaller increase of the FWHM, which means that it has introduced a much lower smoothing. Indeed, only by setting N_σ = 1 it was possible to obtain comparable FWHM with the NVCA filter.

The CNR values are depicted in Figure 19 where it can be assessed that VBM4D achieved the highest CNR and that, despite the strong dependency of the CNR on NVCA threshold, it was not possible to make NVCA reach the same CNR levels of VBM4D by increasing N_σ. Surprisingly, all the sequences processed via the moving average filter scored CNR values that turned out to be always higher than those scored by NVCA, although the moving average had produced more edge blurring. This result highlighted that the CNR is not as edge-aware as it could seem and tends to weight noise reduction more than edge-preservation. With respect to the original noisy sequence, VBM4D, NVCA and moving average filter provided an increase in CNR of 14%, 10% and 13%, respectively.
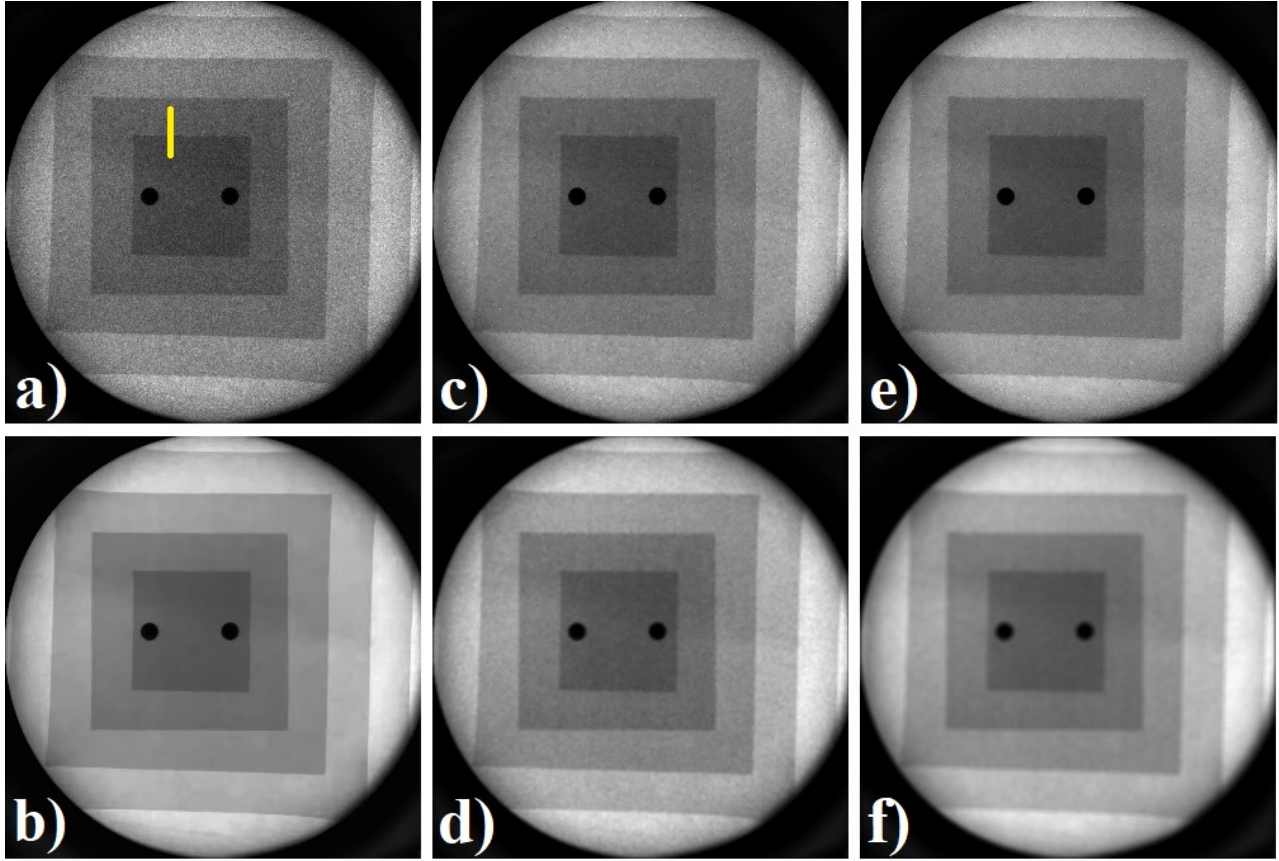
Figure 16. a) Raw image; b) VBM4D with algebraic inversion; c) NVCA, 3x3x3 mask N_σ = 2 ; d) moving average 3x3x3 mask; e) NVCA, 5x5x5 mask N_σ = 2; f) moving average 5x5x5 mask. The yellow vertical line in a) indicates the profile selected for the spatial resolution
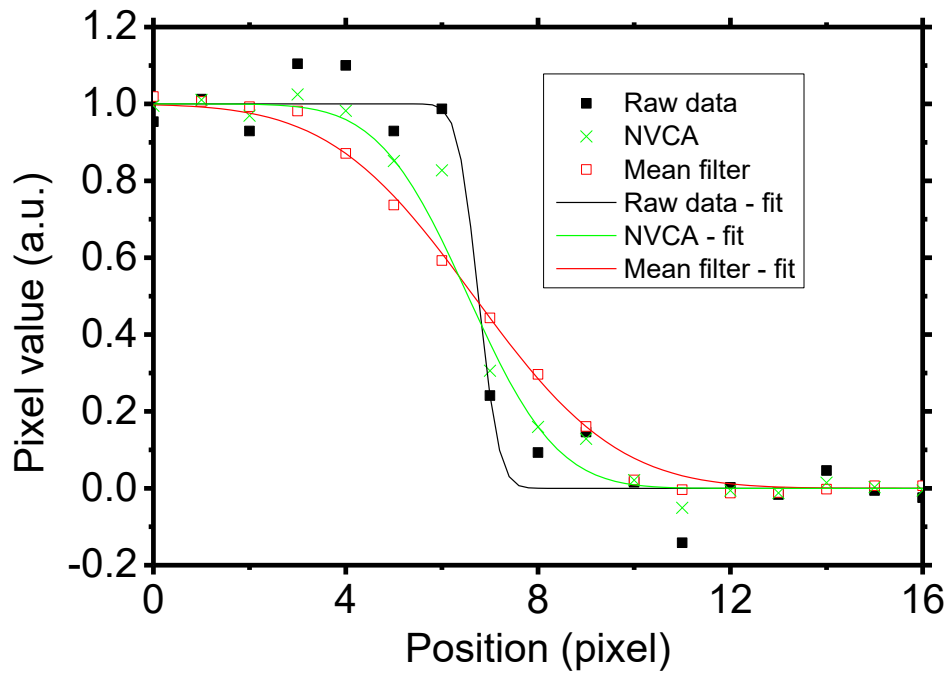
Figure 17. Profiles across the edge outlined in Figure 16.a) for the raw image, the image filtered via NVCA algorithm (7x7x7 mask N_σ =2) and that filtered with a 7x7x7 average filter. Continuous lines represent the fitting curves adopted. The R^2 was 0.9782 for the raw data, 0.9970 in the case of the average filter and 0.9977 in the case of the NVCA filter.
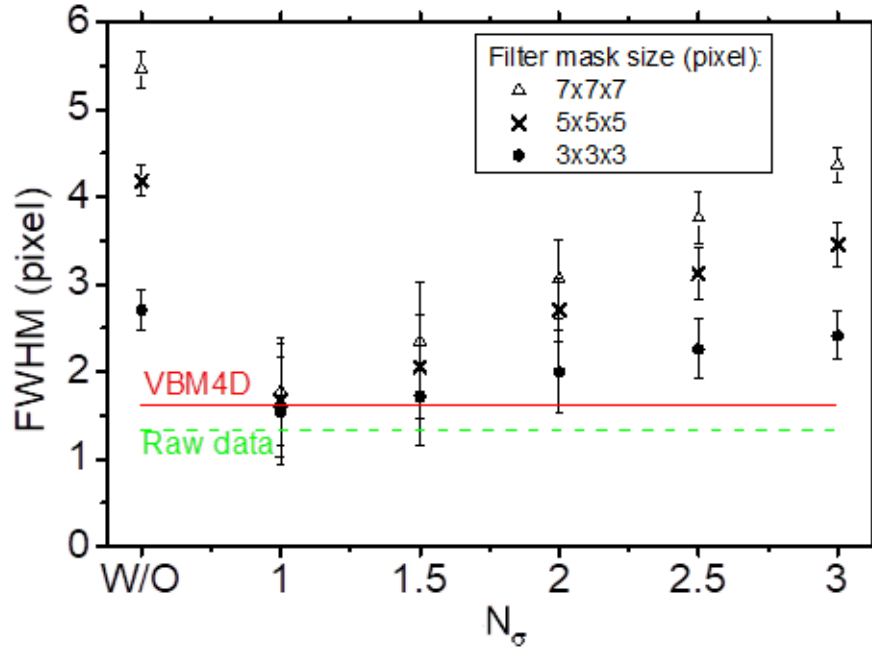
Figure 18. FWHM values for the evaluated filters and the raw data. The W/O label corresponds to the moving average filters.
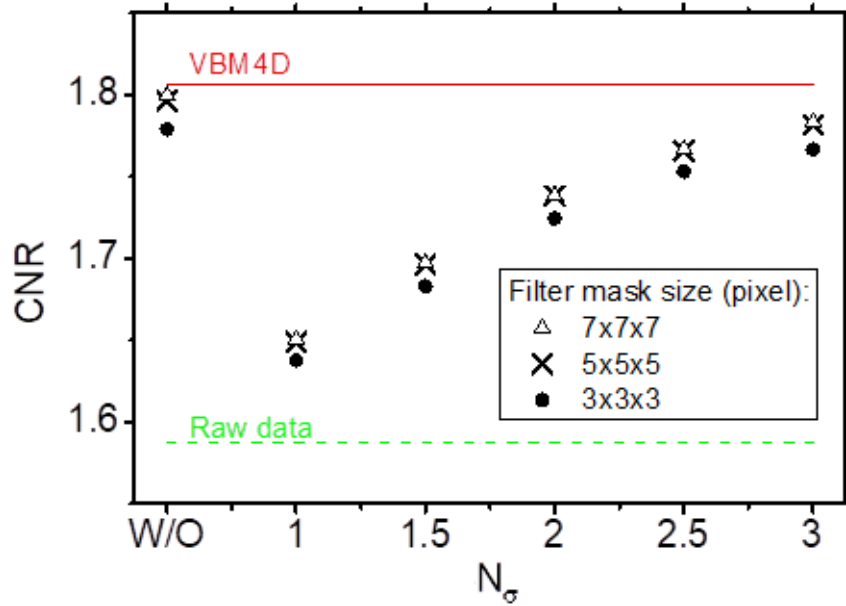
Figure 19. CNR values for the NVCA filters, VBM4D and raw data. The W/O label corresponds to the moving average filters.

Figure 20 reports the FSIM global quality index, computed for the NVCA filtered images by using the time average of the whole sequence as the reference image, considering different mask sizes and different thresholds.

For threshold levels of $N_\sigma = 2.5$ and $N_\sigma = 3$, the NVCA led to FSIM values comparable or larger than those of the VBM4D. However, it is important to underline that for such values of N_σ, the FWHM of NVCA almost doubled that of VBM4D. This result also shed a light on the claimed edge-awareness of the FSIM index, which would have "guided" the NVCA algorithm toward an increased blurring instead of suggesting the result of most edge-aware denoising.
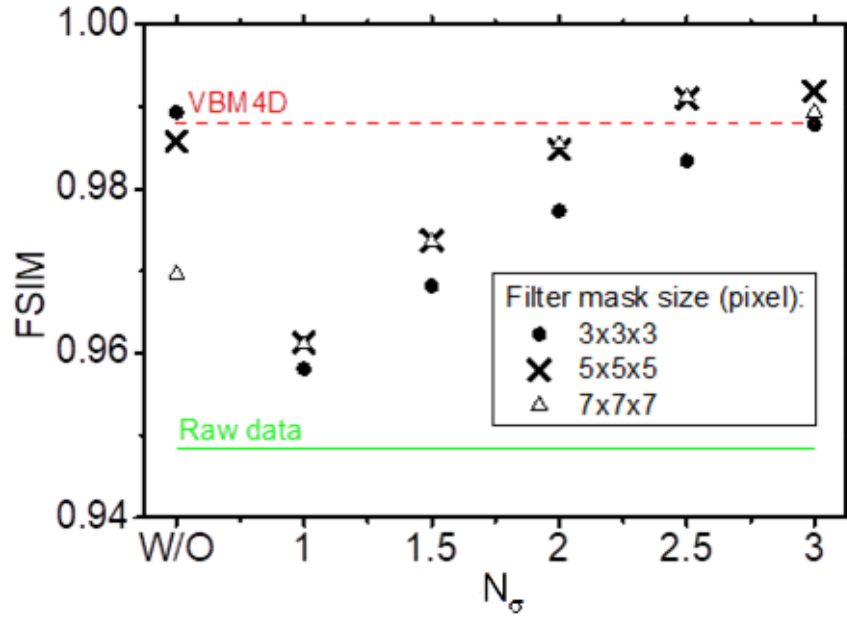
Figure 20. FSIM indices obtained by using the NVCA, VBM4D and the average filter (labelled W/O) denoisers. It was evaluated on the static image of the aluminium step phantom; the reference image was obtained by averaging over time all the raw fluoroscopic frames.

Finally, the computational times required by MA, NVCA and VBM4D to process the 712 frames of the step phantom sequence were determined. All the tests were conducted in MATLAB® R2016b, run on an Intel Core i7-3770 CPU at 3.40 GHz, and all compared filters were implemented via MEX files (i.e. C/C++ subroutines created in MATLAB®). In Table 3 the computational times required by MA and NVCA are reported as percentages of the computational time required by VBM4D. As expected, computational times required by the NVCA filters with different mask sizes turned out to be just few percent of that required by the VBM4D. Hardware implementations of NVCA, such as those presented in [61], can further reduce the computational times by computing the local noise SD threshold from a lookup table instead of directly implementing the formula reported in (14), which involves the computation of square root and multiplication operations.

Table 3. Relative computational times of the NVCA and moving average (MA) filters calculated with respect to the computational time of VBM4D, taken as reference.

Spatial mask dimension	3×3						5×5						7×7					
Temporal mask dimension	**1**	**2**	**3**	**5**	**7**	**9**	**1**	**2**	**3**	**5**	**7**	**9**	**1**	**2**	**3**	**5**	**7**	**9**
NVCA-to-VBM4D computation time (%)	1.5	2.1	2.4	3.1	3.7	4.3	2.7	3.9	4.7	6.6	8.5	10.0	4.7	6.4	8.1	11.7	15.2	18.3
MA-to-VBM4D computation time (%)	1.1	1.2	1.2	1.3	1.4	1.5	1.1	1.4	1.7	2.1	2.3	2.7	1.6	2.0	2.5	3.1	3.7	4.3

5.6 Synthetic sequence with moving object

In order to evaluate the blurring effects introduced when imaging moving objects, three synthetic, noiseless sequences were designed (Figure 21), starting from a fully digital contrast-detail phantom, set as a static background, and then including a rectangular, high-contrast object moving horizontally from left to right at a speed of 1, 2 and 3 pixels per frame. All the objects presented ideal edges and the contrast between the moving object and the image background was 46% (evaluated as the percent ratio between the mean pixel value within the rectangle and the mean pixel value of the image background). Afterward, three noisy sequences were devised by adding a simulated Poisson noise to the noiseless sequences just created. The yellow ROI in Figure 21 shows the edge portion selected for the FWHM evaluation (the ROI includes as much vertical edges as possible, excluding the upper and lower corners). The overall FWHM was estimated as the average FWHM over 10 consecutive profiles. The ROI dimension was chosen in order to include as many lines as possible with the aim of reducing the noise influence on the FWHM estimates. The FWHM evaluated across the selected edge are reported in Figure 23.
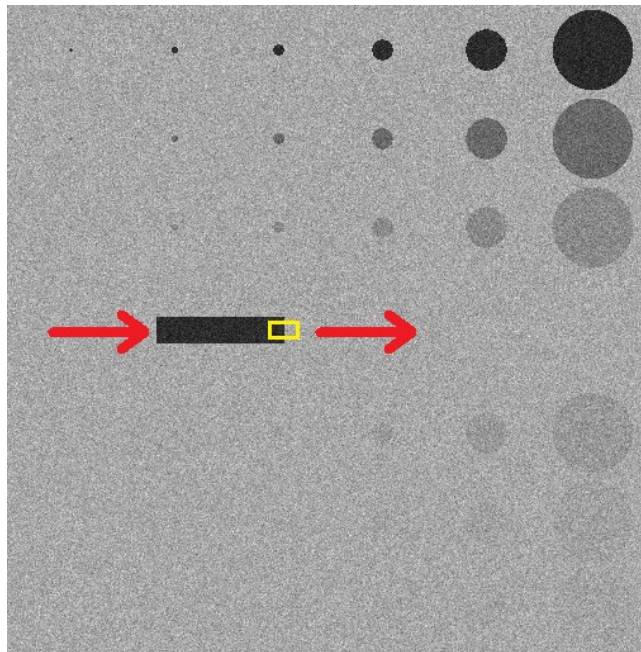
Figure 21. Digital phantom with a moving rectangular insert. The red arrows indicate the moving direction of the insert over the consecutive frames. The yellow ROI outlines the region for the evaluation of the FWHM on the insert edge.

Figure 22 shows a profile across the ideal edge of the object moving at 1 pixel/frame after the application of VBM4D, NVCA and moving average filter with 5×5×5 pixels mask size. It can be noted that while the simple average filter introduced a significant motion blur (FWHM values more than 20 times higher than those of the other filters), both VBM4D and NVCA well preserved the edge sharpness, but VBM4D provided a higher noise reduction.

For the moving object speed of 1 pixel/frame, differences in motion blur introduced by VBM4D and NVCA resulted negligible. However, the spatial resolution of the sequence denoised by VBM4D degraded when increasing object speed (2.2 times when increasing from 1 to 2 pixels/frame), while the NVCA spatial resolution remained unaltered. Also in the case of moving average filter; the FWHM increased as the object speed increased.
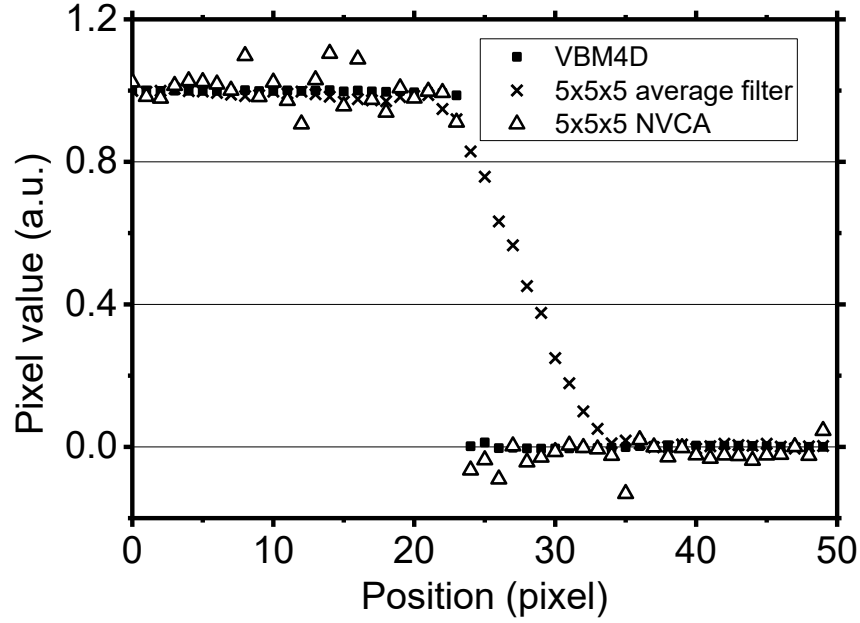
Figure 22. Edge profiles across the edge of the moving object in the sequence of the digital phantom for 5×5×5 average filter, VBM4D and 5×5×5 NVCA filter (N_σ = 2). Insert speed = 1 pixel/frame.
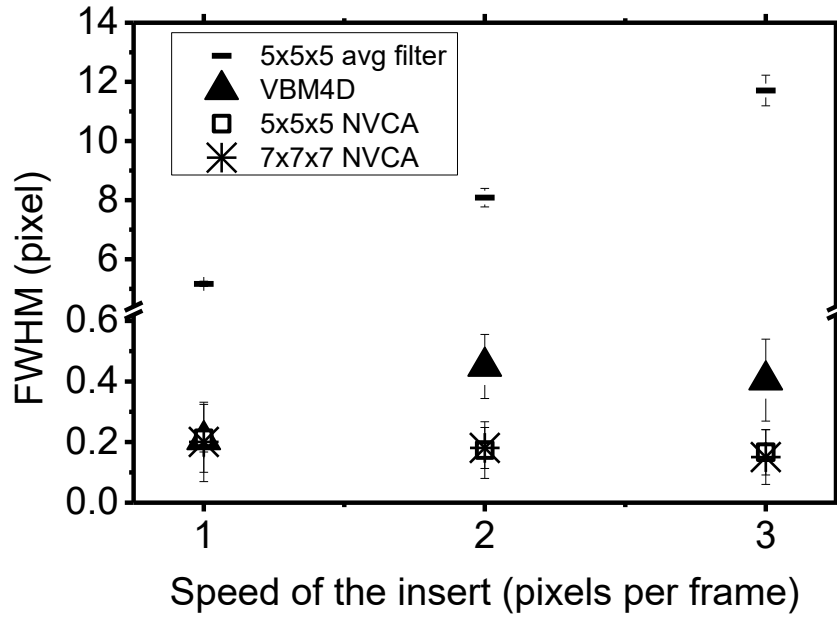
Figure 23. FWHM evaluated across a moving edge in the digital phantom for 5×5×5 average filter, VBM4D and 5×5×5 and 7×7×7 NVCA filter (N_σ = 2).

5.7 Interventional fluoroscopy sequence

NVCA and VBM4D performances were evaluated on a further real fluoroscopic sequence (Figure 24), acquired during a surgical procedure, which shows a moving radiopaque needle, as well as a small implanted catheter. The sequence was acquired with the same device and settings used for the aluminium step phantom (GE 9900 Elite C-arm [132]; tube voltage = 57 kVp, anode current = 1 mA; detector frame rate = 25 fps, frames dimension = 328×333 pixels and grey levels digitalized in a 16-bit scale). Figure 25 shows a ROI of the original surgical sequence and of its versions denoised via

45
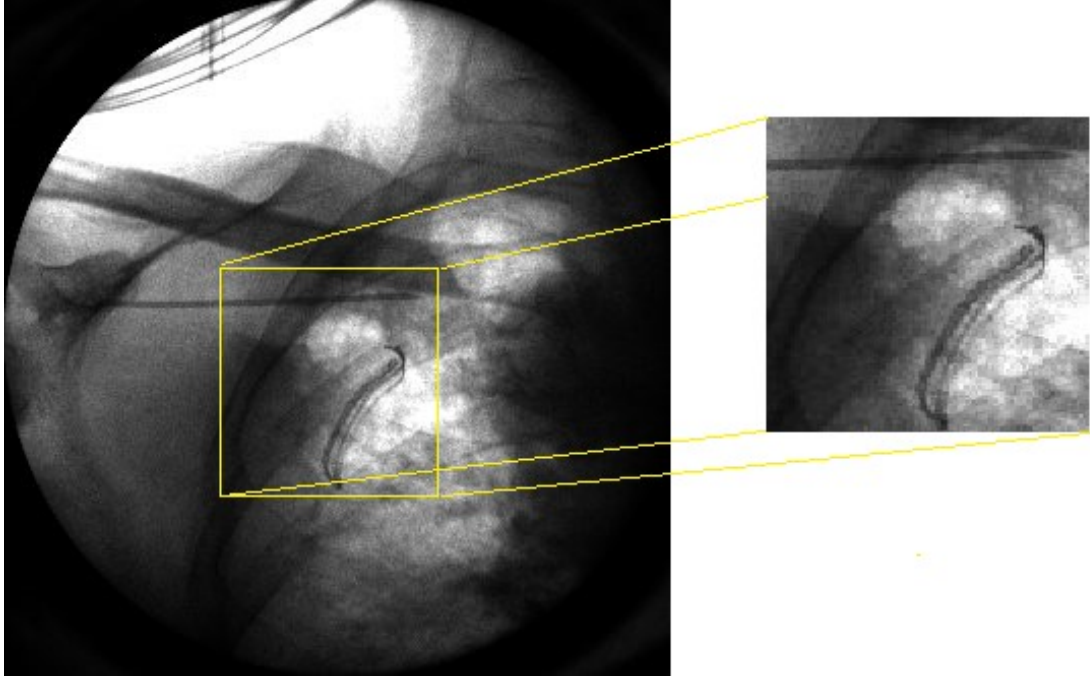
Figure 24. A frame of the surgical fluoroscopy sequence (left) and an enlargement of its central region (right). The radiopaque needle appears at the top of the image enlargement.
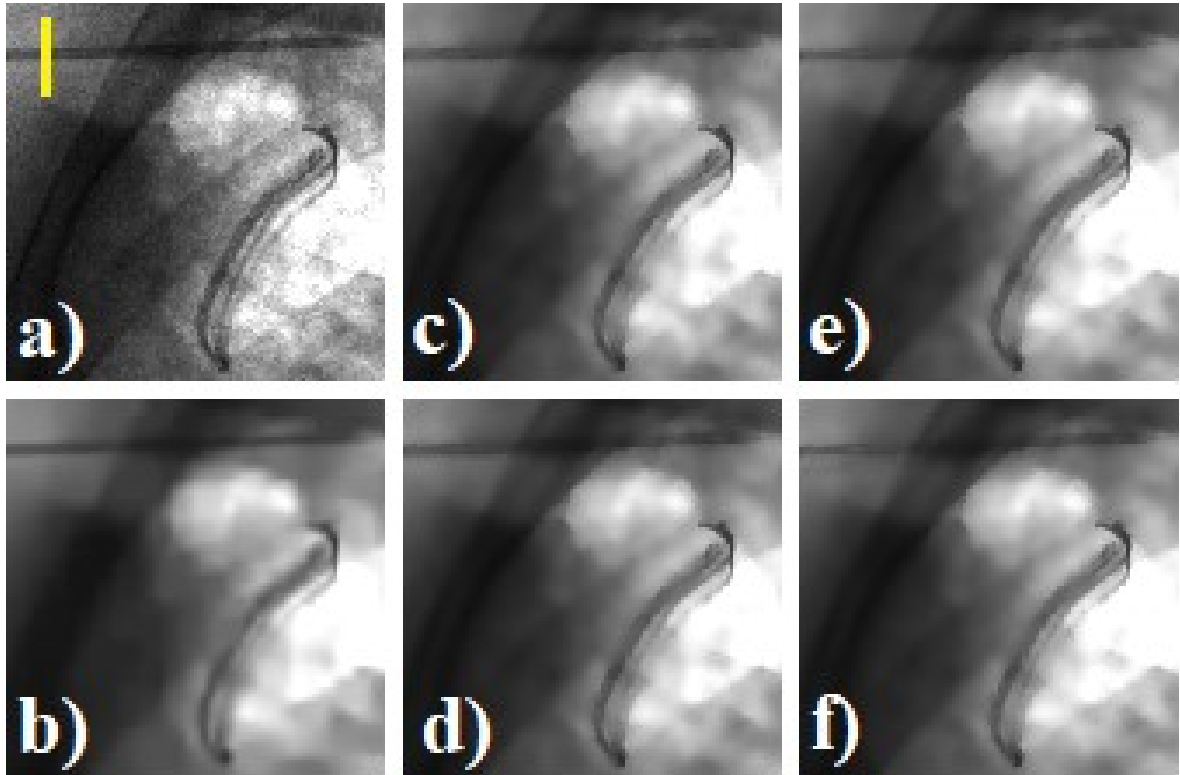
Figure 25. The enlargement of the real fluoroscopic image including the radiopaque needle: a) raw image, the yellow vertical line was manually placed across the needle to evaluate blur, b) image filtered with VBM4D and algebraic inversion of the Anscombe transform; image filtered with NVCA (N_σ = 1.5) with mask size of c) 5x5x1 pixels, d) 5x5x3 pixels, e) 5x5x5 pixels and f) 5x5x7 pixels.

VBM4D and NVCA filter with N_σ =1.5 and mask size of 5×5×1 pixels (spatial filter), 5×5×3 pixels, 5×5×5 pixels and 5×5×7 pixels. From Figure 26 it can be seen that no substantial differences in the profile blurring were caused by passing from a pure spatial filter (5×5×1 pixels mask) to a mask size

of 5×5×3 pixels. However, increasing the temporal extent of NVCA mask to 5 pixels led to a reduction in the signal peak level along with a profile spread.
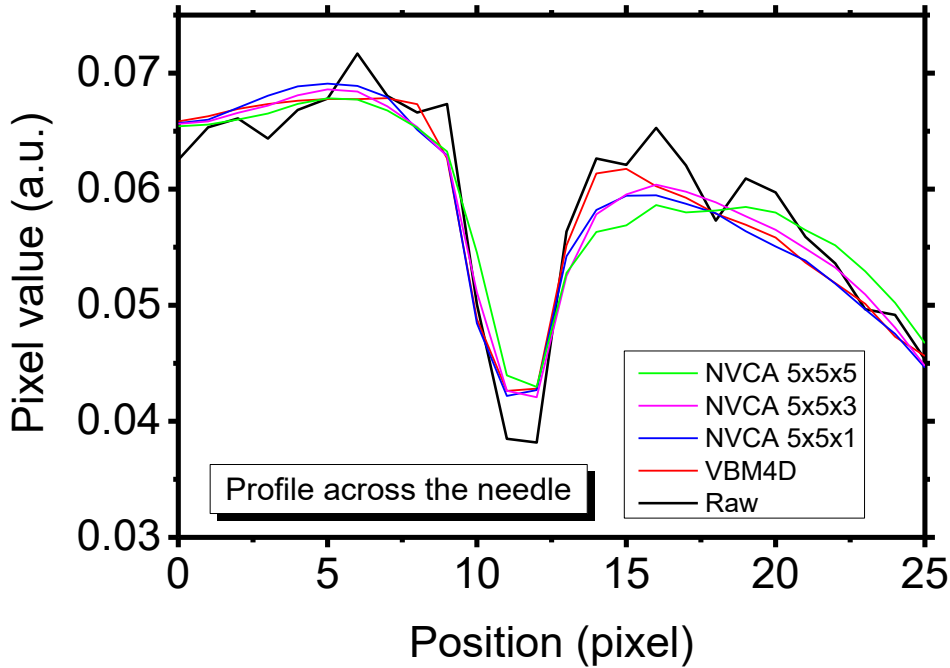
Figure 26. Profile across the needle outlined in Figure 25.

However, by visually comparing the denoising performances of VBM4D and NVCA, a number of undesired effects could be noted. As an example, VBM4D caused losses of contrast and edge sharpness in originally low-contrast areas, such as the ribs and the portions of the needle in the vicinity of the ribs boundaries and over the ribs. Moreover, the small dark catheter appears as extremely blurred and also fragmented in many points. In general, VBM4D introduced a "cartoon-like" effect [133] in almost all areas of the image, probably caused by the patch-based operation and the final aggregation. Indeed, medical images are much more complex, e.g. they do not exhibit large smooth areas, like natural images, therefore, patch-based methods tend to fail more easily [70] by both grouping together patches that are actually not so similar, thus exceeding in smoothing operation, or not being able to find enough patches, thus providing poor noise reduction. NVCA preserved much more the edge sharpness by reducing the amount of smoothing in the neighbourhood of edges (because of the exclusion of more outliers belonging to other objects), especially of larger moving objects which indeed represent the worst condition for the spatio-temporal edge-aware filtering.

5.8   Image derivatives in flexion – extension fluoroscopy sequence

NVCA and VBM4D were also tested on the flexion – extension fluoroscopy sequence of the lumbar spine presented in Chapter 4. First, the noise parameters were extracted from the static scene depicted in the last 30 frames of the sequence, by using the EVaR estimation approach based on time sequences of pixels grey levels. Then the noise parameters just retrieved were used to apply the GAT to the whole sequence, in order to obtain the transformed sequence to be processed via VBM4D. For demonstration purposes, the EVaR estimation was repeated on the transformed data, to show the effectiveness of the variance stabilisation provided by the GAT. The results of EVaR estimation from original and transformed data are depicted in Figure 27.(a-b), respectively, where it can be seen that

the GAT allowed obtaining a distribution with approximately unit variance and absolutely negligible signal-dependence.
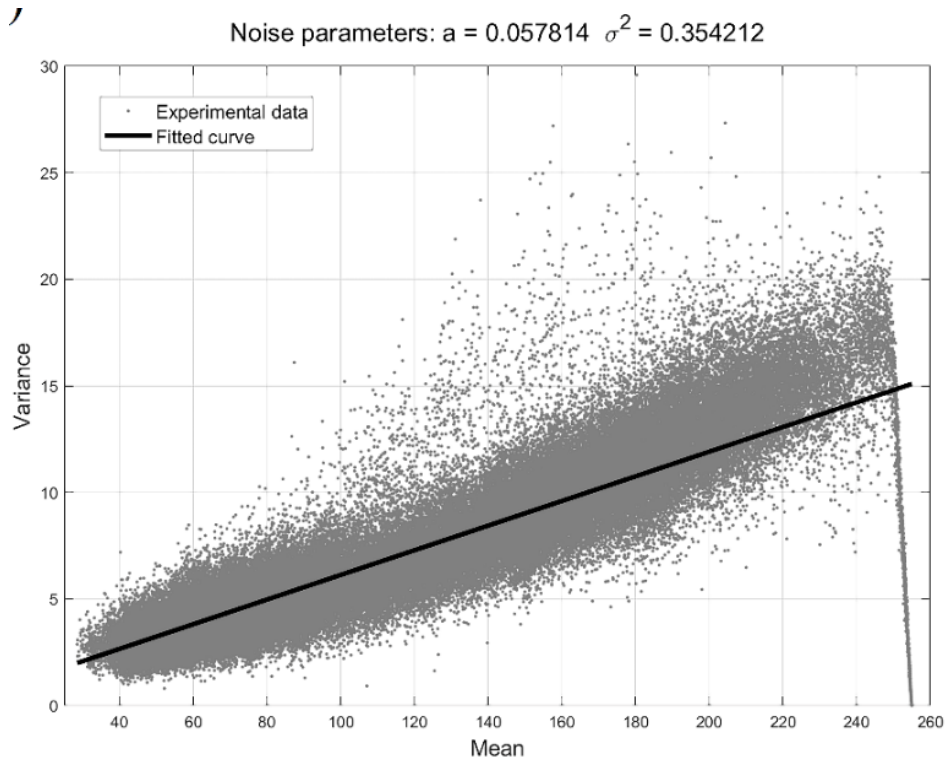
(a)
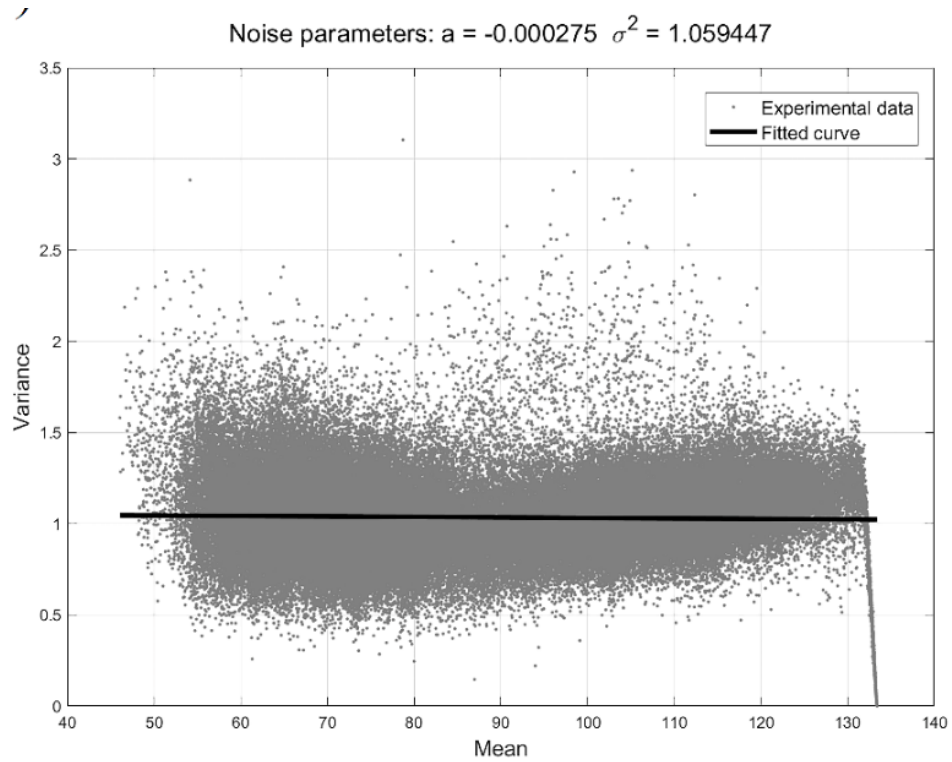
(b)

Figure 27. Experimental EVaRs (scatter points), linear regressions (straight lines) and noise parameters for: a) original noisy sequence; b) noisy sequence after generalised Anscombe transformation.

48

Afterward, an NVCA filter with 11×11×7 mask size and $N_\sigma = 3$ was applied to the raw fluoroscopic sequence, while VBM4D was applied to the transformed data, which were eventually processed with the inverse transform, to bring them back in their original domain. The results are depicted in Figure 28, where, surprisingly, it is quite evident by visual inspection that the VBM4D provided a lesser noise power reduction as compared to NVCA. Nonetheless, as already demonstrated in Chapter 4 for a smaller mask size, NVCA provided a very accurate edge preservation. This can be assessed also by visual inspection in Figure 29, where a ROI of the original noisy sequence including the L2 vertebral body is depicted (Figure 29.a) along with the related image gradients of the noisy and filtered sequences, which were obtained via the Sobel operators described in (27) in Chapter 4. It is very clear from Figure 29, that NVCA outperformed VBM4D by simultaneously providing a deeper noise reduction (in Figure 29.d many spurious edges due to the residual noise are still visible) and a decidedly more powerful preservation of edge sharpness that allowed accurate delineation of vertebral bodies contours , while VBM4D caused a significant blurring also for high-contrast edges.
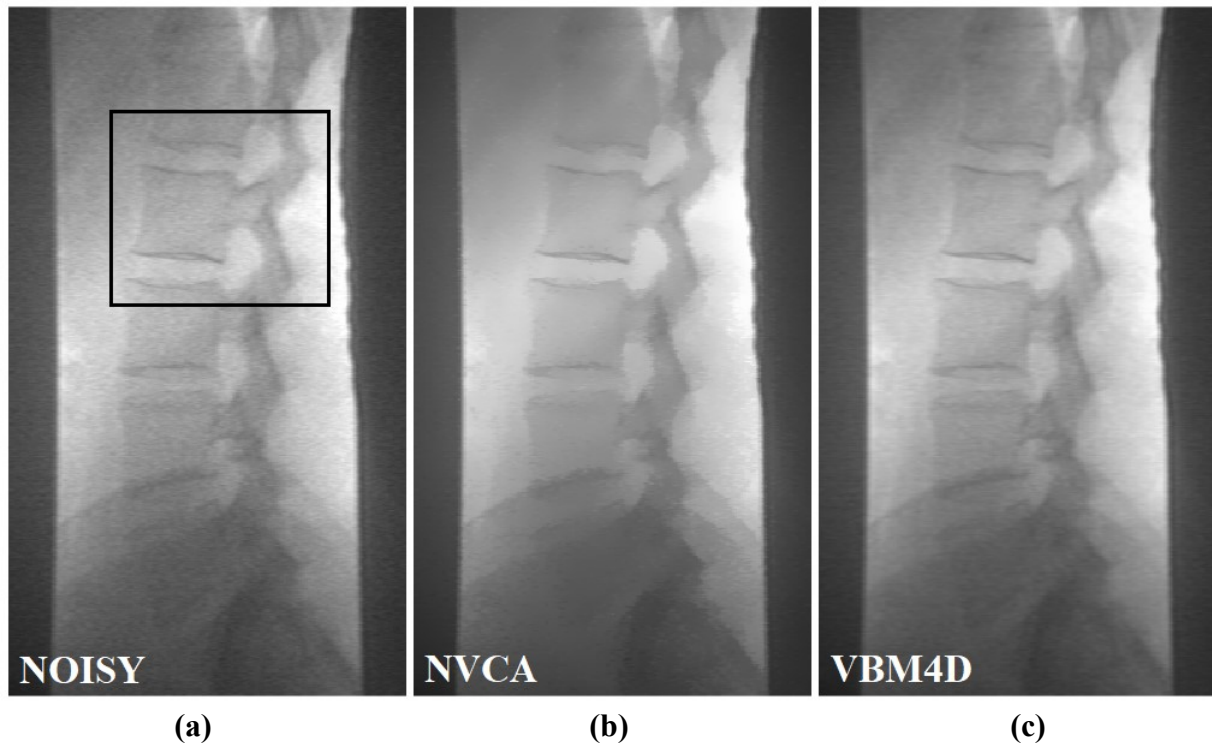
Figure 28. Comparison of a single frame from: (a) noisy sequence with selected ROI; (b-c) sequences filtered with NVCA and VBM4D.

The analyses presented in this chapter focused on the performance assessment of NVCA, VBM4D and MA filters, in terms of details visibility and edge blurring provided in the filtered sequences. The denoising algorithms were tested on sequences with both static and moving scenes, simulated and real fluoroscopic noise, with and without clinically relevant structures.

VBM4D provided the best performances in the step-phantom sequence, which represented a very long static scene, with real fluoroscopic noise but without any clinically relevant structure. Considering that VBM4D is designed as a post-processing algorithm and takes into account the whole spatio – temporal volume of the sequence, and that the step-phantom scene undoubtedly shares many
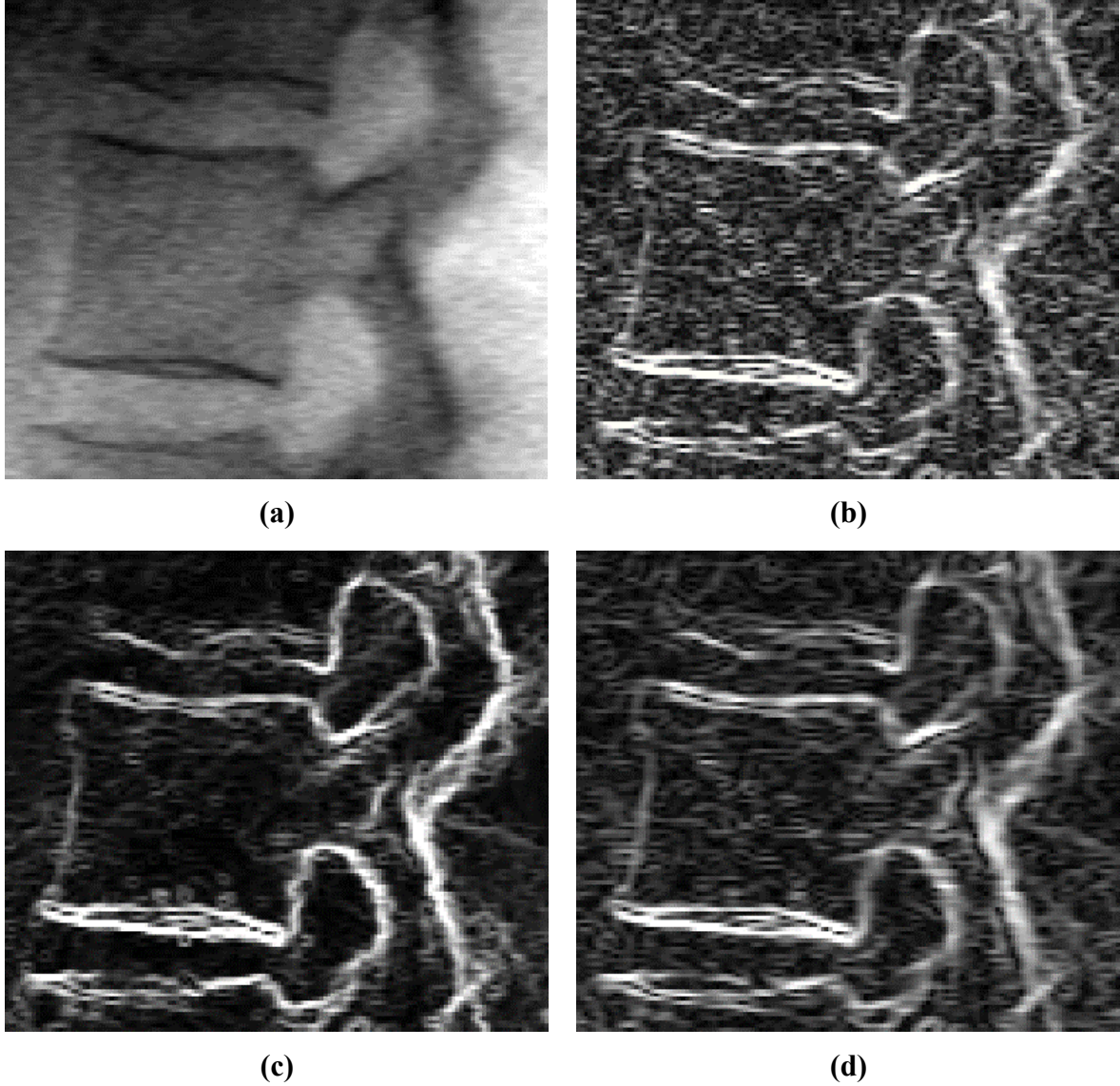
Figure 29. (a) ROI extracted from the noisy sequence, shown in Figure 28.a; the same ROI extracted from the image gradients of the (b) noisy, (c) NVCA-filtered and (d) VBM4D-filtered sequences, respectively. The luminance of the image gradients was stretched in the same grey levels interval, so as to allow a fair comparison.

features with natural images, such as being dominated by large smooth areas and including sharp, high-contrast edges, this result was not unexpected.

Undoubtedly, this scene is not very challenging as compared to those the VBM4D has already been tested on: indeed, the absence of any kind of motion obviously prevented motion blur to show up, even when applying the simple MA filter; moreover the presence of sharp, high-contrast, geometric objects contours assuredly increased the amount of similar patches, within the same frame and even more so in the whole sequence, as the scene is absolutely motionless, and eased their recognition, thus supporting a significant increase in the VBM4D denoising power. Nonetheless, the VBM4D provided a 14 % increase in CNR, which was not so higher than the 13 % and 10 % increases provided, respectively, by MA and NVCA, which instead required computational times that turned out to be few percent of the time required by VBM4D.

The first issues arose when processing the synthetic sequences with the moving object, which did not include real fluoroscopic noise, nor clinically relevant structures. Indeed, the VBM4D

provided reasonable edge preservation only for the lowest object speed of 1 pixels/frame, causing an increased blurring of originally sharp and high-contrast edges for increasing speeds, as proved by the increase in the estimated FWHM of the edge spread function. Instead, NVCA provided an accurate edge preservation, regardless of the object speed, which suggests a kind of stability of its performances with respect to very different conditions of the imaged scene. It is worth noting that, as in the case of the step-phantom sequence, the absence of complex, clinically relevant structures, as well as the presence of high-contrast objects contours, certainly eased the grouping of similar patches within the whole sequence; moreover, the simulated Poisson noise undoubtedly ensured better performances in the variance stabilisation provided by the GAT, as compared to the case of real fluoroscopic noise, which has actually more complex features as already explained in Chapter 2. Nevertheless, the performances of VBM4D turned out to be significantly degraded, thus suggesting that the presence of objects moving at various speeds is a critical factor for VBM4D performances, regardless of the details visibility and the structures complexity, which is quite surprising for a video denoising algorithm that is claimed to provide for a motion estimation process to avoid motion blur and boost denoising performances.

These aspects were heightened when processing the interventional fluoroscopy sequence that represented a moving scene of a cardiac surgical procedure, which included both real fluoroscopic noise and clinically relevant structures. The denoising results provided by all the NVCA filters with different mask sizes clearly outperformed the result of VBM4D, which presented many critical quality issues, among which a general cartoon-like appearance; the loss of contrast and edge sharpness of low-contrast and/or moving edges of clinical relevance, i.e. the contours of organs (e.g. ribs, lungs) as well as the contours of surgical instruments, such as an high-contrast, slowly moving, implanted catheter, which appeared completely blurred and fragmented in his thinnest parts, and an inserting needle, which appeared blurred when approaching or superimposing on ribs areas in the X-ray projection.

While VBM4D remains undoubtedly a state-of-the-art denoising algorithm for natural image sequences corrupted by AWGN, this study demonstrated that, even if adapted to the Poisson – Gaussian denoising problem via the GAT, it is not able to ensure an efficient, edge-aware denoising of real low-dose fluoroscopic sequences, beside also exhibiting a very high computational burden as compared to simpler, yet effective approaches like NVCA. This result, not only prevents its application to the real-time processing of fluoroscopic sequences during interventional procedures, but also discourage its potential use as a post-processing algorithm for previously acquired, low-dose fluoroscopic sequences.

The use of "automated" IQA indices like CNR and FSIM, highlighted some of the critical issues that arise when applying this indices to medical images. Indeed, both CNR and FSIM suggested that the best denoising results were achieved by NVCA when using the highest N_σ values, which also yielded that highest FWHM (i.e. the worst edge blurring) though. These results demonstrate that CNR and FSIM are not as edge-aware as they were thought to be, and highlight the urge of conceiving an IQA index that is well-suited to describe the most important features that influence the actual quality of X-ray medical images as perceived by a human operator. Such an IQA index, indeed, would also be of help to guide the analysis and identification of the best settings for parameterised algorithms like NVCA, for which the final image quality (at least from a qualitative visual point of view) is known to be heavily dependent on parameters selection in a not straightforward fashion. To this aim,

a novel edge-aware IQA index for X-ray images is proposed in Chapter 6, along with its preliminary assessment against CNR and FSIM indices.

Chapter 6

Edge-aware image quality assessment for X-ray image denoising optimization

This chapter addresses two issues that emerged in the previous chapters, namely the reliable quality assessment of medical X-ray images and the optimisation of NVCA performances through identification of the optimal parameters settings. Indeed, it has been shown that NVCA performances have a significant dependency on the particular choice of its parameters and on various image features, which has not been fully understood yet. Moreover, both CNR and FSIM proved unable to give a good measure of edge perceptibility (even though they are commonly considered to be quite sensitive to edge sharpness), and to give it enough weight in the evaluation of the image quality of fluoroscopic images, where edges play a key role in the ability of human eye to understand the scene. It follows that also the objective comparison between different denoising strategies would suffer from this flaw, thus giving potentially unreliable results. Both issues may share a common solution: the design of an edge-aware IQA index that could give a more accurate evaluation of X-ray image quality and, therefore, guide the identification of the optimal parameters settings that yield the highest image quality. In this chapter a novel IQA index is presented and a preliminary performance comparison against FSIM and CNR is carried out by simultaneously assessing their ability to identify the NVCA denoising result with the highest image quality and, as a consequence, the related optimal parameters choice for NVCA in a synthetic fluoroscopic sequence with definite features.

6.1 Image quality assessment

Image quality assessment is a fundamental task for the analysis of image quality enhancement [55, 78 − 80]. Enhancement algorithms are aimed at improving different features, such as mean luminance, contrast, spatial resolution, edge sharpness, which influence the image quality as perceived by the human vision system. However, the HVS has a very complex behaviour, which cannot be accounted for by a mere recipe of simple image parameters. As an example, the HVS, just like other human senses, is subject to masking effects, which alter in various ways the perceptibility of certain details or components of the visual informational content, depending on several conditions [134]. In addition, the HVS exhibits astonishing, and in some way still unclear, statistical capabilities, which result in a significant robustness to various kinds of distortions and noises [12]. Nonetheless, the possibility to predict the quality perceived by the HVS with reasonable accuracy is still desirable and required in many applications, e.g. real-time quality check during images acquisition and transmission, objective comparison of methods for image quality enhancement, parameters optimization for image processing algorithms. To this aim, several IQA metrics and indices have been proposed in literature, which can be divided in three main categories, namely Full-Reference IQA (FR-IQA), Reduced-Reference IQA (RR-IQA) and No-Reference IQA (NR-IQA) [55, 78 − 80]. FR-IQA indices are aimed

at quantifying the loss in image quality by measuring the discrepancies between a reference image and a distorted version, and require the availability of the full reference image. RR-IQA does not require the full reference image but only synthetic information. Since in most practical applications the reference image is usually unavailable, a number of NR–IQA indices have been proposed, which attempt to quantify the quality of a distorted image without directly comparing it with the original one. However, NR–IQA indices are usually tailored to natural images, which differ substantially from medical images in terms of structural content, statistical distributions of noise, and image artifacts [79], and are generally not competitive with FR-IQA metrics, as they need to deal with multiple kinds of distortion [55]. FR-IQA metrics are then preferable to evaluate the ability of an algorithm to recover the original content of a medical image from a distorted version. Examples of classic, widespread FR-IQA metrics are the Mean Squared Error (MSE) and the PSNR, which are both based on a measure of the error between the original reference image and the distorted image and have no relation with the human visual perception. SSIM has been one of the first IQA indices to be claimed to rely on a perception-related model that could better describe the image quality as perceived by the human eye, even though Nilson & Akenine- Möller from NVIDIA recently argued that there is no evidence of such relationship with human perception and demonstrated that SSIM produces the largest quality estimation errors precisely on the edges [130], which play a key role in the recognition and tracking of objects and organs in fluoroscopy. The HVS relies on edges much more than smooth areas to interpret the scene of an image, and it is very robust to noise [135 – 137]. While smooth areas are mainly affected by noise, the edges are corrupted by blurring effects as well, which should be considered at least as much as the noise when evaluating image quality. However, since the number of edge pixels is statistically much lower than that of pixels in smooth areas, the blurring effects are often disregarded in favour of noise, especially in IQA indices that rely on pooling a single quality score from pixel-wise distortion measures which are given the same importance, regardless of their spatial positions [129]. Indeed, in [138 – 140] it has been shown that by using space-variant weights that take into account the kind and amount of contribution of each pixel to the perceived image quality, the predictive performances of various IQA indices, even as simple as PSNR, could be improved. This supports the intuition that pixels can provide very different contributions to the HVS perception, based on the role they play in the scene interpretation, and suggests that common pooling strategies are not so favourable.

6.2   Sensitivity of Edge Detection

In literature, many edge-aware FR-IQA indices have been proposed (e.g. [129, 141 – 142]), among which FSIM is currently the most recent and referenced index and, therefore, could be considered as the state of the art. Nevertheless, in the previous chapters it proved uncapable of detecting the loss in edge sharpness due to blurring effects, which was instead measured via the FWHM of the edge spread function. A novel edge-aware IQA index, named Sensitivity of Edge Detection (SED), is then proposed to provide a reliable measure of image quality in fluoroscopic sequences. The index is based on the assumption that the sharper the edges, the better the edge detection. Therefore, edge detection is adopted to quantify the differences in sharpness and visibility of edges between the original and distorted versions of an image. In particular, the index measures the sensitivity of edge detection, defined as the ratio of true edge pixels in the distorted image to the total of edge pixels in the original image, both identified by using the same parameters of the Canny edge detector [143]. The map of true edge pixels is obtained by computing a pixel-wise logical AND operation between the edge maps

54

of the distorted and original versions of the image. Then, the number of edge pixels is derived both in the true edges map and in the edges map of the original image, and their ratio is computed to obtain the SED score, as reported in (37):

$$E_{test} = EdgeMap(I_{test})$$
$$E_{ref} = EdgeMap(I_{ref})$$

$$SED = \frac{|E_{test} \wedge E_{ref}|_1}{|E_{ref}|_1}$$

(37)

6.3 Synthetic test sequence design

A synthetic fluoroscopic sequence was used to assess the capabilities of FSIM and SED to identify the NVCA parameters values that yield the best trade-off between noise reduction and edge preservation in moving scenes. The building block of the sequence was a sequence of binary maps representing a full circle with a radius of 7 pixels, moving along a square trajectory at a speed of 1 pixel/frame. The final sequence of binary maps (50 frames of 100×100 pixels), referred to as binary sequence, was devised via superimposition of several building blocks at randomly distributed starting points. Then, an 8-bit grayscale, noiseless sequence (referred to as reference sequence) was obtained by assigning grey levels of 127 and 86 to the background and the moving circles, respectively. Finally, the reference sequence was corrupted with simulated Poisson noise, so as to obtain a noisy sequence with a CNR of 2 (defined as in (36)). The same frame extracted from binary, reference and noisy sequences are depicted in Figure 30. The actual CNR obtained by adding Poisson noise to the reference sequence, computed by applying the formula reported in (36), was equal to 1.99, thus confirming that the condition on CNR was matched.

(a) (b) (c)
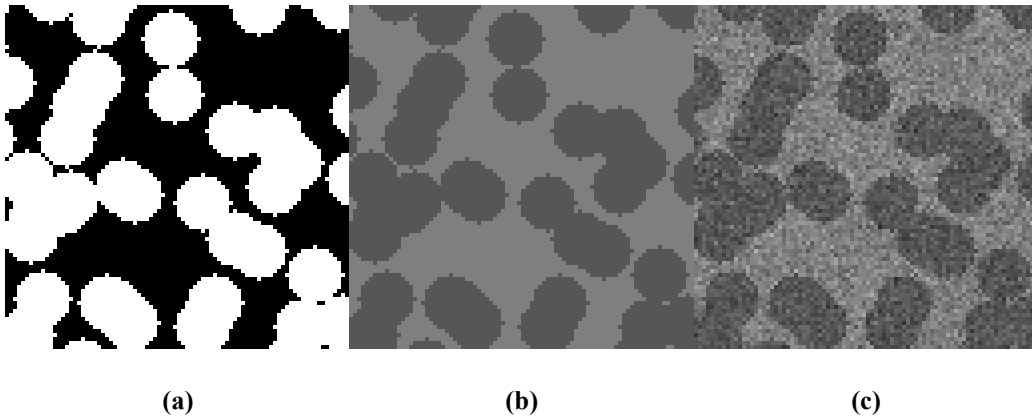
Figure 30. Comparison of the 20th frame of (a) binary sequence; (b) noiseless sequence; (c) noisy sequence.

6.4 Identification of optimal NVCA parameters based on FSIM and SED scores

The noisy sequence was filtered with NVCA by considering all the combinations of parameters values, within certain ranges, thus producing several filtered versions of the noisy sequence. In particular, the ranges considered for each parameter were: 1 to 15 in steps of 2 for S (spatial mask

size), 1 to 7 in steps of 1 for T (temporal mask size), 1 to 4 in steps of 0.2 for N_σ. Therefore, a total of 1680 filtered sequences were obtained.

The FSIM and SED scores were computed for the noisy sequence and for all the 1680 filtered versions obtained with NVCA filtering, by considering the noiseless reference sequence as the ground truth. The FSIM scores were computed via the MATLAB function "FSIM Index with automatic downsampling" (Version 1.0, Copyright© 2010 Lin Zhang, Lei Zhang, Xuanqin Mou and David Zhang) [129]. The Canny edge detection involved in the computation of the SED scores was performed via the MATLAB "edge" function (Version 10.1, R2017b). The default parameters set in the "edge" function were adopted, as they provided good results in the reference sequence. The FSIM and SED scores were organized in 3D matrices (15×7×16) and for each of them the overall maximum value was located, in order to identify the related NVCA parameters.

The optimal parameters values identified by FSIM and SED indices are reported in Table 4. The best results, according to both IQA indices, were obtained with the same filter mask dimensions (i.e. 5×5×2), but with two different values for N_σ (higher values produce higher noise reduction but also heavier edge blurring). Moreover, other image quality related parameters were also computed for the noisy and filtered sequences. In particular, the CNR, as well as contrast and noise SD were calculated and reported in Table 5.

Table 4. Comparison of IQA results: the first three columns are the optimal parameters identified with FSIM and SED indices, the last two columns are the IQA scores achieved by the noisy sequence and the best sequences identified with FSIM and SED indices

Sequence	S_{opt}	T_{opt}	$N_{\sigma\ opt}$	FSIM	SED
V_{noise}	-	-	-	0.60	0.82
V_{FSIM}	5	2	3	0.88	0.82
V_{SED}	5	2	2	0.82	0.85

Table 5. Comparison of contrast, noise SD and CNR of the noisy and filtered sequences with highest IQA scores.

Sequence	Contrast	Noise SD	CNR
V_{noise}	41.0	20.6	1.99
V_{FSIM}	36.2	10.4	3.50
V_{SED}	38.8	12.1	3.21

The filtered sequences with the highest image quality scores can be visually compared in Figure 31, where the same frames extracted from the reference and noisy sequences are reported in the upper panels, while those extracted from the best results identified by FSIM and SED are reported in the lower panels. It can be clearly observed that the sequence that achieved the highest SED score (Figure 31.(d)) is characterized by sharper edges with respect to the sequence that obtained the highest

FSIM score (Figure 31.(c)), which instead provided slightly better noise reduction. Finally, for the sake of completeness, the edges revealed by the Canny edge detector, adopted in the SED index computation, are reported in Figure 32. For the reference sequence the detected edges are highlighted in purple, while for the noisy and filtered sequences, the correctly detected edges (true edges) and the missed ones are highlighted in red and blue, respectively.

(a) (b)

(c) (d)
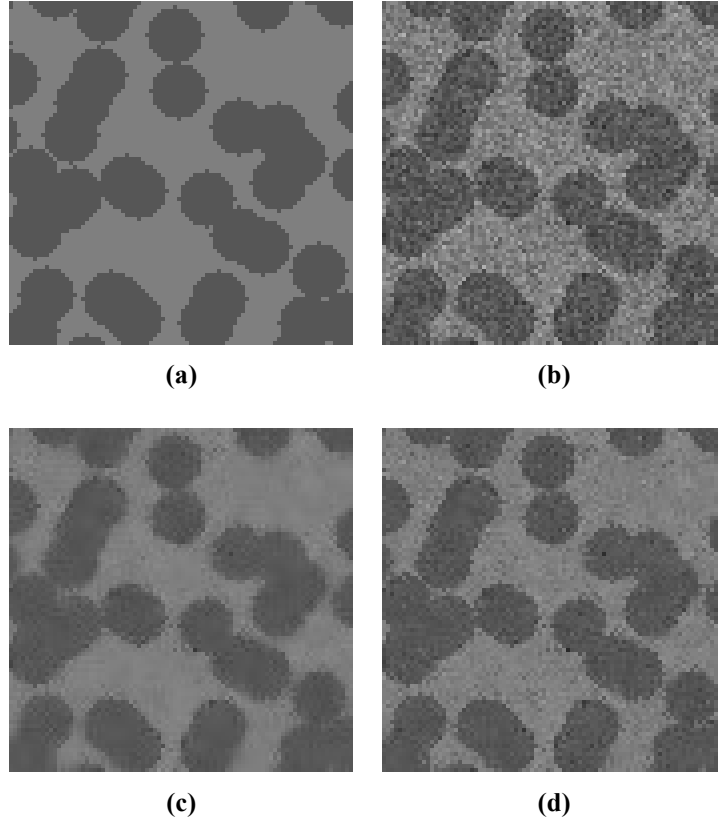
Figure 31. Comparison of a frame from (a) original noiseless sequence; (b) noisy sequence; (c) filtered sequence with higher FSIM score; (d) filtered sequence with higher SED score

According to the results, the best sequence identified by SED was characterized by an overall higher edge sharpness with respect to the best sequence suggested by FSIM, as it can be assessed by visual comparison of Figure 31.(c-d) and Figure 32.(c-d). This could also be inferred from a quantitative comparison of their SED scores, as they indicate the percentage of correctly detected edge pixels, which is surely dependent on the edge sharpness and the local contrast. This seems to be confirmed also by the values reported in Table 5, where it can be noticed that the NVCA filter which produced the best quality sequence according to FSIM, provided poor edges preservation, as it achieved the lowest contrast value.

However, it provided a greater noise reduction (also at the cost of smearing edges) and so it achieved a slightly better CNR. It is worth mentioning that the sequence with the highest FSIM score achieved almost the same SED score of the noisy sequence (no significant improvements in percentage of correctly detected edges), while the sequence with the best SED score achieved a much higher FSIM score with respect to the noisy sequence, and also an higher SED score (significant improvements in both noise reduction and percentage of correctly detected edges).
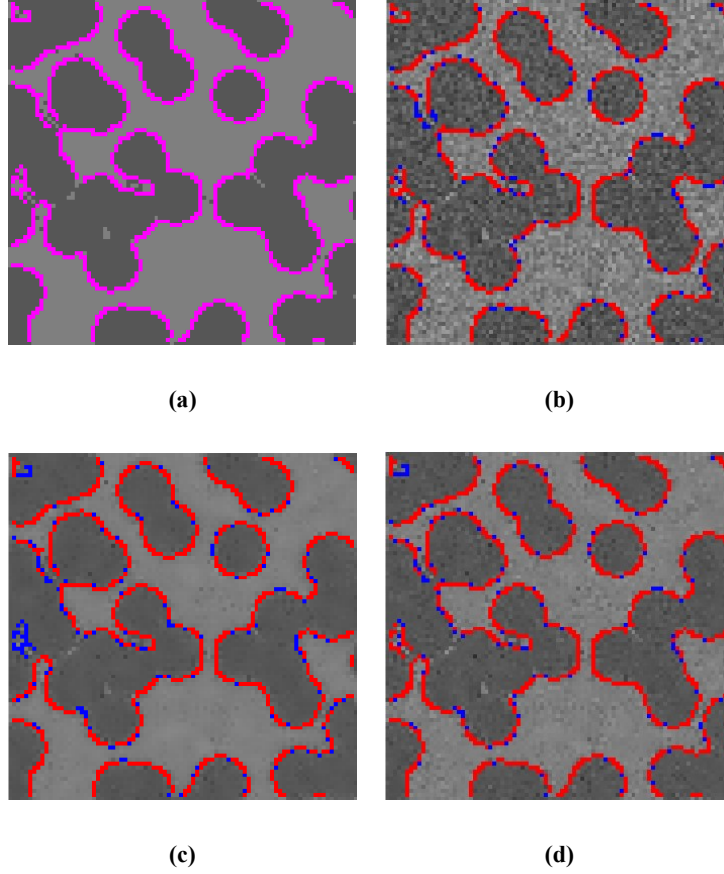
Figure 32. Comparison of the edges detected in a frame of the noiseless, the noisy and the filtered sequences with higher IQA scores. In the original noiseless sequence (a) the edges are highlighted in purple, while in the noisy and filtered images (b-c) the correctly detected and the missed edges are highlighted in red and blue, respectively.

A more in-depth analysis of the performances of NVCA and the behaviour of SED and FSIM could be carried out by analysing the variations of the optimal results that could be obtained by varying N_σ. Indeed, it is reasonable that the optimal settings for NVCA would vary according to different features that influence the quality of the original noisy sequence, thus preventing the identification of a unique optimal setting. Hence, unless a robust method is conceived to determine those features and to automatically set the optimal NVCA parameters accordingly, a manual setting of these parameters should be performed, so it would be interesting to figure out how the mask sizes should be varied when increasing or decreasing N_σ. Figure 33 depicts the results of this analysis.

In Figure 33.(a) it could be observed that, in practice, by increasing N_σ, nothing changes in the optimal sizes of the mask according to FSIM ($5\times5\times2$), apart from a slight increase of the optimal T value for $N_\sigma < 1.5$. Instead, in Figure 33.(b) it is clear that, according to SED, the mask sizes should be reduced when increasing N_σ. This is understandable, as an increase in N_σ results in an increased inferential error probability, whose undesired effects could be amplified by relatively large mask sizes, simply because of the higher absolute number of pixels that could be misclassified during the NVCA filtering. Looking at the lower panels in Figure 33, it can be noticed that the trend of SED with N_σ has a local maximum that is much more emphasised, as compared to the trend of FSIM, which, instead, decreases very slowly as N_σ increases. These trends explain the inability of FSIM to account for the edge blurring effects, due to large N_σ values, as well as the edge-awareness of SED, which penalizes both excessively blurred edges, due to large N_σ values, and excessively noisy edges, due to small N_σ values.

58
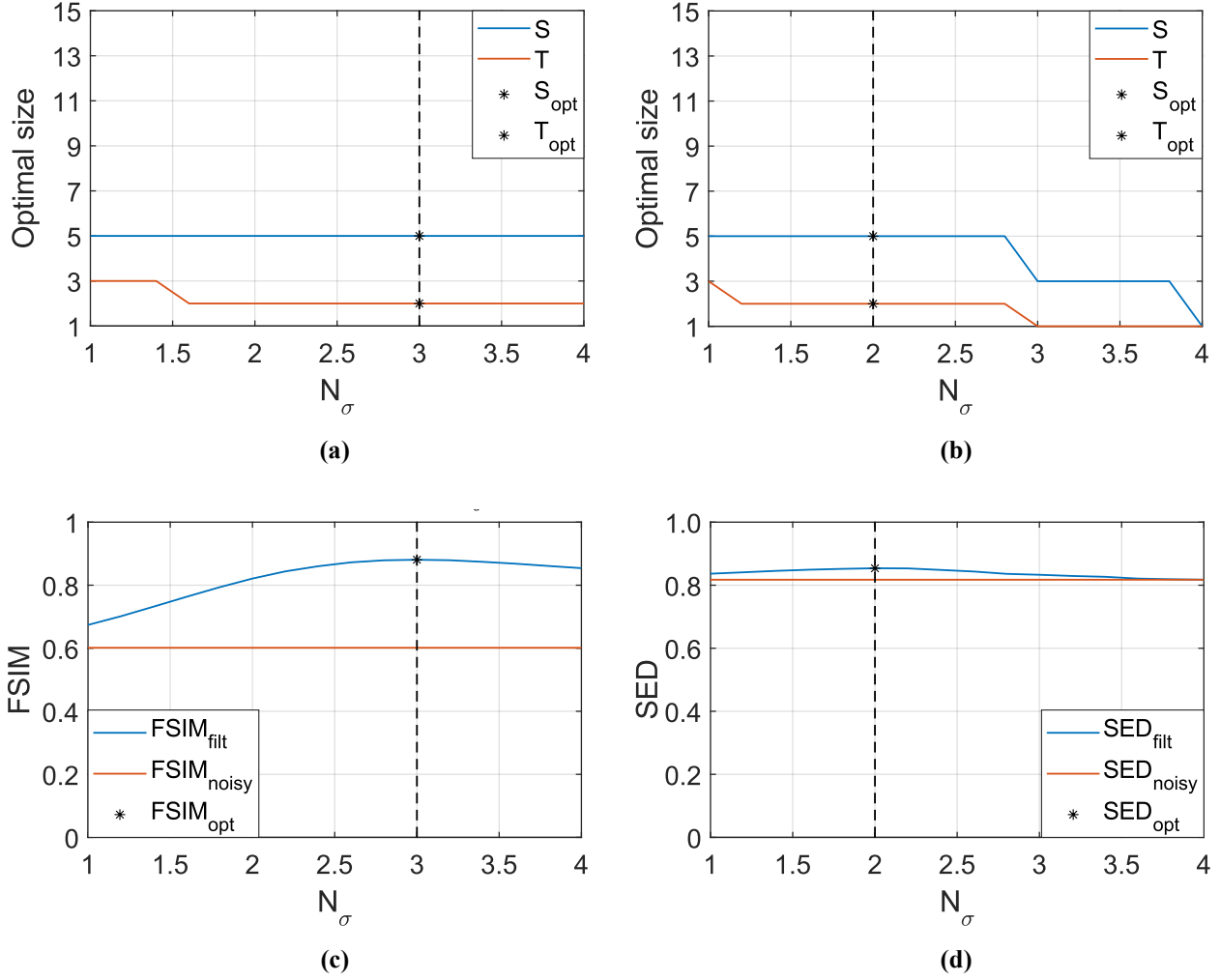
Figure 33. Optimal mask sizes and IQA scores as functions of N_σ. First row shows the optimal mask sizes as functions of NSIGMA, according to (a) FSIM and (b) SED. Second row shows the IQA scores as functions of NSIGMA, obtained by using the corresponding optimal mask sizes depicted in the first row.

These behaviours of the FSIM and SED indices would be even more stressed in noisy conditions with lower CNR, because the noise amplitude and the inferential error probability would be higher and, as a consequence, also the weights that they would have on the image quality, according to FSIM and SED, respectively. For this reason, the above described analyses were carried out also on a noisy sequence with a further reduced CNR, which was set to be equal to 1. It is worth underlining that this condition is quite challenging for the statistical validity of the basic principle of NVCA, as already shown in Chapter 3. Nonetheless, it is an interesting condition to assess the denoising performances that NVCA would provide for low-contrast details in very noisy conditions, which could be encountered in very low-dose fluoroscopy procedures. Table 6 and Table 7 outline the results of the analysis for the noisy sequence with CNR = 1. It can be noted that, besides suggesting a lower $N_{\sigma\,opt}$ as in the case with CNR = 2, SED also suggested smaller mask sizes as compared to FSIM, probably to deal with the increased inferential error probability and the related risk to produce blurring effects. As expected, the best result according to FSIM did not increase the SED score with respect to the noisy data and caused a significant reduction of contrast and noise SD, while the best result according to SED increased both FSIM and SED scores with respect to the noisy data and preserved much more the contrast, at the cost of an higher residual noise SD. Also in this case, the highest CNR was achieved by the best result according to FSIM, even though it provided

the highest contrast loss, thus proving that the noise power has more weight in the image quality measured by both FSIM and CNR.

Table 6. Results for noisy sequence with CNR = 1. Comparison of IQA results: the first three columns are the optimal parameters identified with FSIM and SED indices, the last two columns are the IQA scores achieved by the noisy sequence and the best sequences identified with FSIM and SED indices.

Sequence	S_{opt}	T_{opt}	$N_{\sigma\ opt}$	FSIM	SED
V_{noise}	-	-	-	0.54	0.65
V_{FSIM}	5	3	3.4	0.87	0.57
V_{SED}	3	2	2.2	0.76	0.69

Table 7. Results for noisy sequence with CNR = 1. Comparison of contrast, noise SD and CNR of the noisy and filtered sequences with highest IQA scores.

Sequence	Contrast	Noise SD	CNR
V_{noise}	22.0	21.5	1.02
V_{FSIM}	16.8	6.41	2.62
V_{SED}	20.1	10.3	1.96

A visual comparison of the filtered sequences with the highest image quality scores can be carried out in Figure 34 and Figure 35. The best sequence according to FSIM appears as deeply blurred and poor contrasted (Figure 34.(c)), and indeed exhibits a conspicuous number of missed edge pixels (the blue ones in Figure 35.(c)), even higher than those of the noisy sequence (Figure 35.(b)). The best sequence according to SED, instead, features significantly softer edge blurring effects and maintains the contrast almost unaltered, resulting with no doubts in a decidedly higher visual image quality (Figure 34.(d)). Indeed, it provided better edge detection performances with respect to the noisy sequence, as proven by the higher number of true edge pixels (highlighted in red) and the lower number of missed edge pixels (highlighted in blue) in Figure 35.(d).

Figure 36 shows the histograms of pixels grey values computed for the noisy and the filtered sequences. In Figure 36.(a) the two Poisson distributions related to the circles and the background cannot be distinguished and the distribution appears as unimodal, which is compliant with the condition of CNR = 1. This situation seems impossible to deal with by employing the core criterion of NVCA, because the distributions of the background and the circles are almost overlapping. Nonetheless, NVCA succeeded in reducing the noise power so as to allow the two distributions to be distinguished again in Figure 36.(b-c). It is worth noting that the distributions in Figure 36.(b), albeit exhibiting a reduced SD as compared to those showed in Figure 36.(c), turn out to be also heavily skewed, which suggests the occurrence of some kind of distortion in the filtering operation.
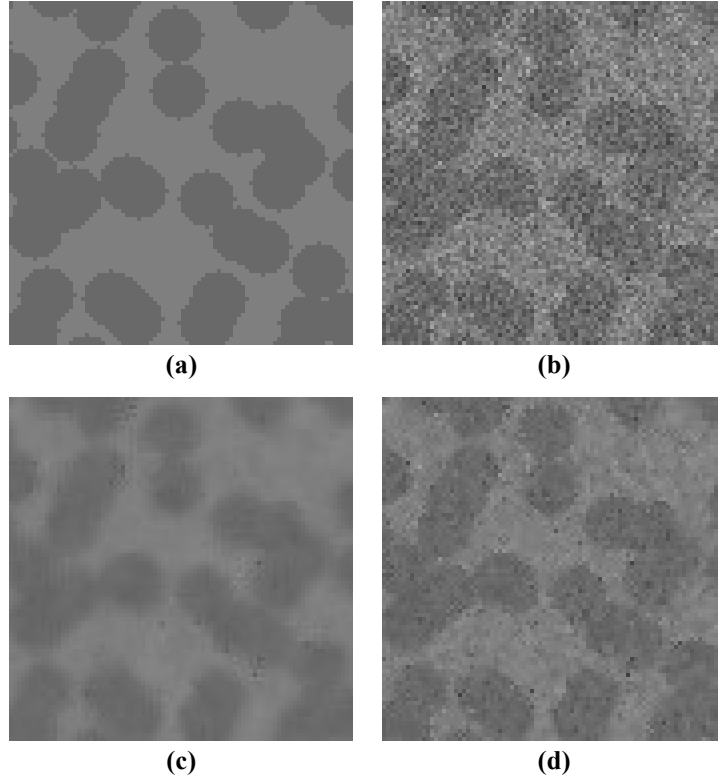
Figure 34. Results for noisy sequence with CNR = 1. Comparison of a frame from (a) original noiseless sequence; (b) noisy sequence; (c) filtered sequence with higher FSIM score; (d) filtered sequence with higher SED score
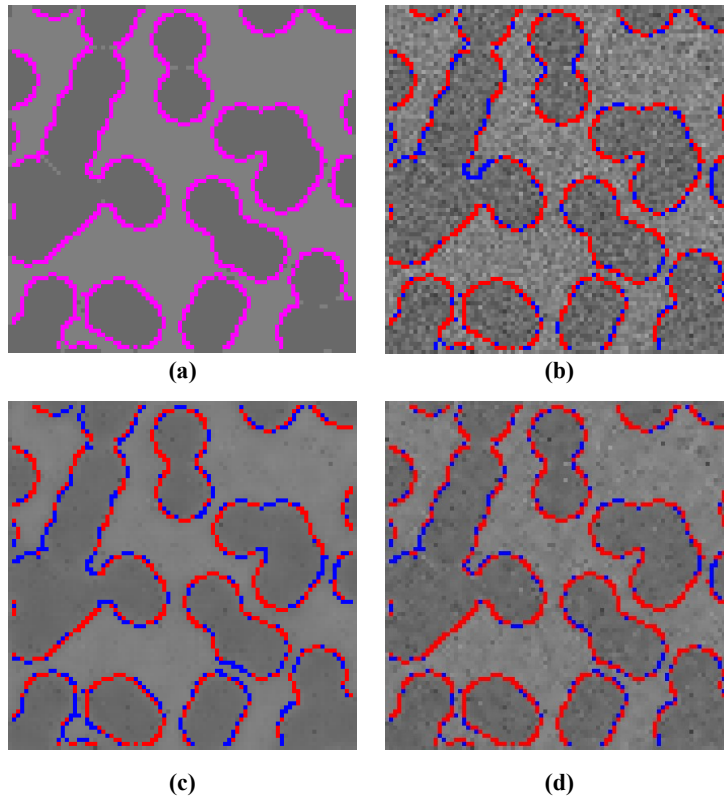
Figure 35. Results for noisy sequence with CNR = 1. Comparison of the edges detected in a frame of the noiseless, the noisy and the filtered sequences with higher IQA scores. In the original noiseless sequence (a) the edges are highlighted in purple, while in the noisy and filtered images (b-c) the correctly detected and the missed edges are highlighted in red and blue, respectively.

61
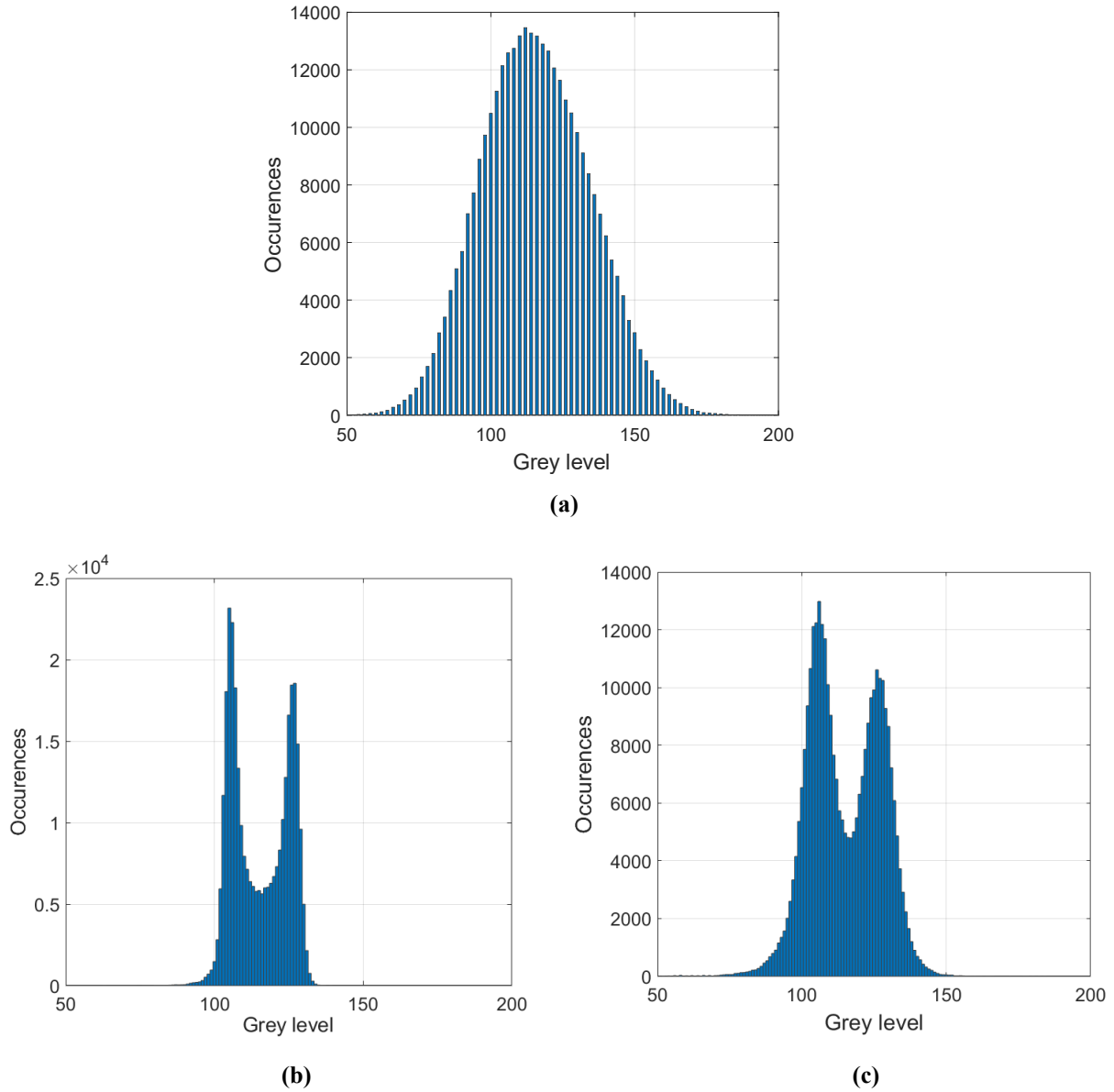
Figure 36. Results for noisy sequence with CNR = 1. Comparison of histograms of grey level occurrences in (a) noisy sequence; (b) filtered sequence with higher FSIM score; (c) filtered sequence with higher SED score.

In Figure 37.(a) the optimal mask sizes, along with the FSIM and SED scores are depicted as a function of N_σ. As in Figure 33, it could be observed that by increasing N_σ the optimal sizes of the mask (5×5×3) remain practically unaltered, according to FSIM, and, this time, this is also true for SED, which, instead, selected different optimal mask sizes (3×3×2) . Looking at the lower panels in Figure 37, it can be observed that the trend of FSIM practically does not show a local maximum, while the trend of SED with N_σ is much more emphasised, with a clear maximum at $N_\sigma = 2.2$.

The preliminary results obtained in this study suggest that the novel SED index could be more efficient than FSIM in identifying the filter parameters which yield the best trade-off between noise reduction and edge preservation for the processing of low-dose fluoroscopy sequences. However, more extensive analysis of SED performances should be carried out in different conditions, e.g. different objects shape, dimension and speed, also comparing SED against other well-established IQA indices.
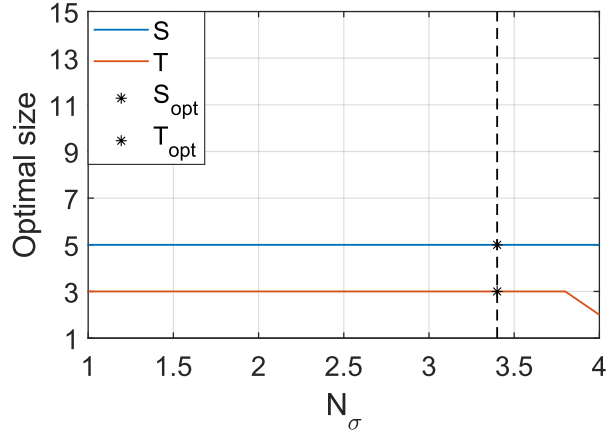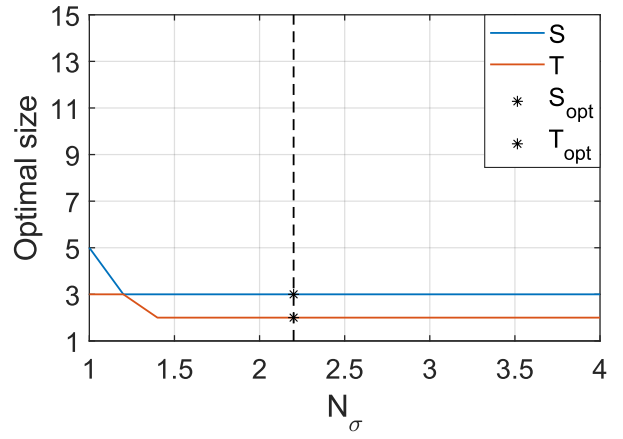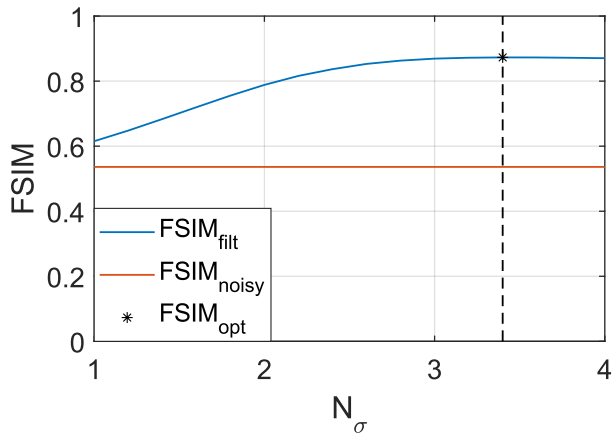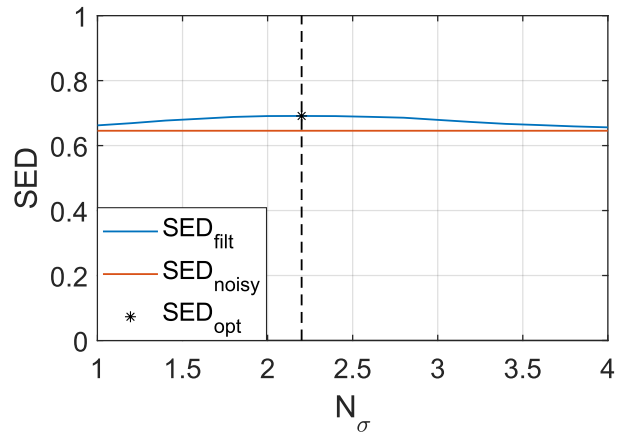
Figure 37. Results for noisy sequence with CNR = 1. Optimal mask sizes and IQA scores as functions of N_σ. First row shows the optimal mask sizes as functions of NSIGMA, according to (a) FSIM and (b) SED. Second row shows the IQA scores as functions of NSIGMA, obtained by using the corresponding optimal mask sizes depicted in the first row.

Chapter 7

Improved NVCA with spatio-temporal domain separation

In this chapter, an improved version of the NVCA algorithm is presented, which outperforms in many cases the state of the art of Poisson denoising in video sequences, always ensuring at least comparable performances. As in the original NVCA algorithm, the computational burden is kept sufficiently low, therefore, unlike many state-of-the-art algorithms characterized by a very high complexity that prevents their actual real-time implementation, the proposed algorithm is ideally suited for real-time, hardware implementation on small Field Programmable Gate Arrays (FPGA), also requiring very limited percentages of their hardware resources.

The main improvements can be summarized as follows:

a) The joint spatio – temporal filtering is split in cascaded temporal and spatial filtering operations, in order to increase the ability to discriminate and restore low-contrast details in very noisy conditions;

b) IIR filters are used for time filtering, which allows for equivalent long time averaging while reducing the hardware resource burden, in particular the memory that would be required to store many high-resolution frames to filter with long time windows;

c) local adaptive thresholds are introduced in both temporal and spatial domains to improve the recognition of noisy and edge pixels, depending on their amount of motion activity, in order to jointly avoid motion blur and increase the ability to denoise and preserve static or slow moving low-contrast edges;

d) isolated pixels removal is also included in spatial filtering to deal with the tails of noise distributions that are too far from their expected value and would be hardly filtered by the core criterion of NVCA;

e) real-time filtering of 1024×1024 resolution video at 49 fps is demonstrated with an hardware implementation on the smallest FPGA of the Altera StratixIV family, by using, at most, the 22% of its hardware resources.

The improved NVCA algorithm is tested against several state-of-the-art denoising algorithms and is compared with the hardware implementation of the original NVCA algorithm presented in [61].

7.1 Spatio-temporal domain separation filtering algorithm

The main idea at the core of this new algorithm comes from the observation that the scenes imaged in fluoroscopic procedures are characterised by both quasi-static and dynamic areas. While the latter pose unavoidable limitations to the denoising operation due to the motion of objects, which prevents the possibility to average a significant number of noisy observations and to provide substantial noise power reduction (otherwise motion blur effects could occur), the former allow exploiting much more the temporal redundancy of image sequences, thus providing stronger edge-aware noise reduction in the time domain, which, in turn, allows for a more accurate discrimination of low-contrast edges in the spatial domain. Hence, this observation supports the development of an approach based on the separation of time and spatial domain filtering.

Indeed, when performing a joint spatio – temporal conditioned average, as in the original NVCA algorithm, the same threshold is used to discriminate between noisy and edge pixels both in space and time within the same spatio – temporal neighbourhood. This principle of operation determines a trade-off between a deep noise reduction, which could be achieved by selecting a relatively high threshold (i.e. high values for N_σ) at the risk of causing edge blurring, and the preservation of sharpness of low-contrast edges, which requires the selection of a lower threshold at the cost of achieving a poor noise reduction. However, if pixels in a certain area are characterised by a static luminance, pure temporal filtering provides absolutely edge-preserving noise reduction, thus improving the CNR of low-contrast edges without affecting their sharpness. The higher CNR thus obtained brings two advantages: an improved edge-awareness, i.e. the ability to distinguish the luminance variations due to low-contrast edges from the spurious fluctuations caused by noise, and a lower level of noise reduction required to accomplish a reasonable structures perceptibility. The latter complies with the selection of lower values for N_σ, which unavoidably entails a lower number of pixels included in the filtering operation, i.e. a lower noise reduction ratio (NRR), but is also essential to achieve the improved edge-awareness. For a given edge contrast, the possibility to decrease N_σ depends on the noise level of each time-filtered pixel, that is on the NRR, which, in turn, depends on the number of averaged pixels involved in its computation. By exploiting this information, an equivalent space-variant N_σ can be determined, which defines an adaptive spatial threshold allowing for the improved performances just described.

Figure 38 shows the conceptual overview of the filtering algorithm. The filter operates on a video with a $L{\times}K$ pixels frame size. The input luminance of the pixel at the position *(i,j)* of the *n*-th frame is referred to as *p(i,j,n)*. The algorithm consists of a first stage composed by temporal filters, referred to as *T-filters*, and a second stage composed by a spatial filter, referred to as *S-filter*. Each *T-filter(i,j)* performs an NVCA filtering along the temporal dimension on the luminance of the pixels at the position *(i,j)* in subsequent frames. In practice, the ensemble of the *T-filters* can be regarded as a NVCA filter with a purely temporal mask (i.e. with spatial mask size $S = 1$). The output of the *T-filters*, referred to as *p_T(i,j,n)*, is then fed at the input of the *S-filter*, that, for each frame *n*, performs an NVCA filtering in the spatial domain on pixels within a $(2X{+}1){\times}(2Y{+}1)$ neighbourhood of the position *(i,j)*. Indeed, the *S-filter* can be regarded as a NVCA filter with a purely spatial mask (i.e. with temporal mask size $T = 1$). The output of the *S-filter* is referred to as *p_S(i,j,n)*.

A synthetic sequence with a 512×512 frame size, corrupted by simulated Poisson noise, is used to show the performance of the subsequent steps of the algorithm. The sequence includes a static background, resembling a contrast – detail phantom used in quality assessment of X-ray devices (flat background with circles of different sizes and contrast), and a dark rectangle moving from left to right. A frame of the noisy reference video is shown in Figure 39.(a).
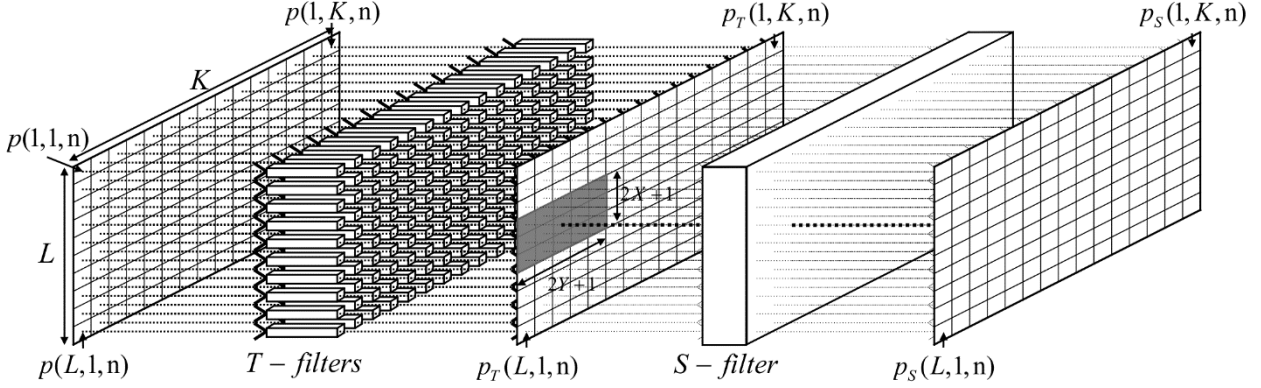
Figure 38. Schematic representation of the improved NVCA algorithm operation with the cascaded temporal and spatial denoising stages.

7.2 Temporal filters design

Let us consider a pixel at position *(i,j)*, which is initially supposed to belong to a static object. The pixels *p(i,j,n)* are random variables with Poisson distribution, potentially modified by non-linear point-wise luminance transformations (e.g. log-mapping, gamma-correction), which can be locally approximated [85, 86] as a normal distribution $N(\mu,\sigma^2)$, if the number of photons per pixel is greater than 10 or 20 (hypothesis widely verified also in fluoroscopy), where the variance σ^2 is a function of the mean μ:

$$\sigma^2(i,j,n) = G(\mu(i,j,n)) \tag{38}$$

If the object is static, the input pixels of the *T-filter* at the position *(i,j)* have a constant mean, and therefore a constant variance as well. In addition, the reduced response time of modern X-ray detectors (~1 ms [41], much lower than the minimum frame period of 33 ms, corresponding to the maximum frame rate of 30 fps) allows assuming that pixels *p(i,j,n)* at the same spatial position *(i,j)* and in different frames *n* are mutually independent and normally-distributed. In this hypothesis, it is well known that the best filtering technique is represented by the moving average, which equally weights all the samples in a temporal window of length *M*, [144]:

$$p_T(i,j,n) = \frac{1}{M}\sum_{m=0}^{M-1} p(i,j,n-m) \tag{39}$$

Obviously, the average in (39) results in motion blur effects for moving objects, which worsen for increasing M, thus determining a trade-off between noise reduction and edge preservation. Moreover, it is worth noting that (39) is poorly suited for an hardware implementation if a large number of frames needs to be filtered, because a great amount of memory would be required for them to be stored. As an example, in a finite impulse response (FIR) filter implementation (e.g. [40, 61]), storing 128 frames of a video with 1024×1024 frame size an 8 bits/pixel resolution would require 128 MB

of memory, and, at a frame rate of 25 fps, a memory bandwidth of 3.1 GB/s. For this reason, FIR filter based implementations generally employ much lower frame memory lengths (*M*), which unavoidably limits the NRR the circuit can provide.
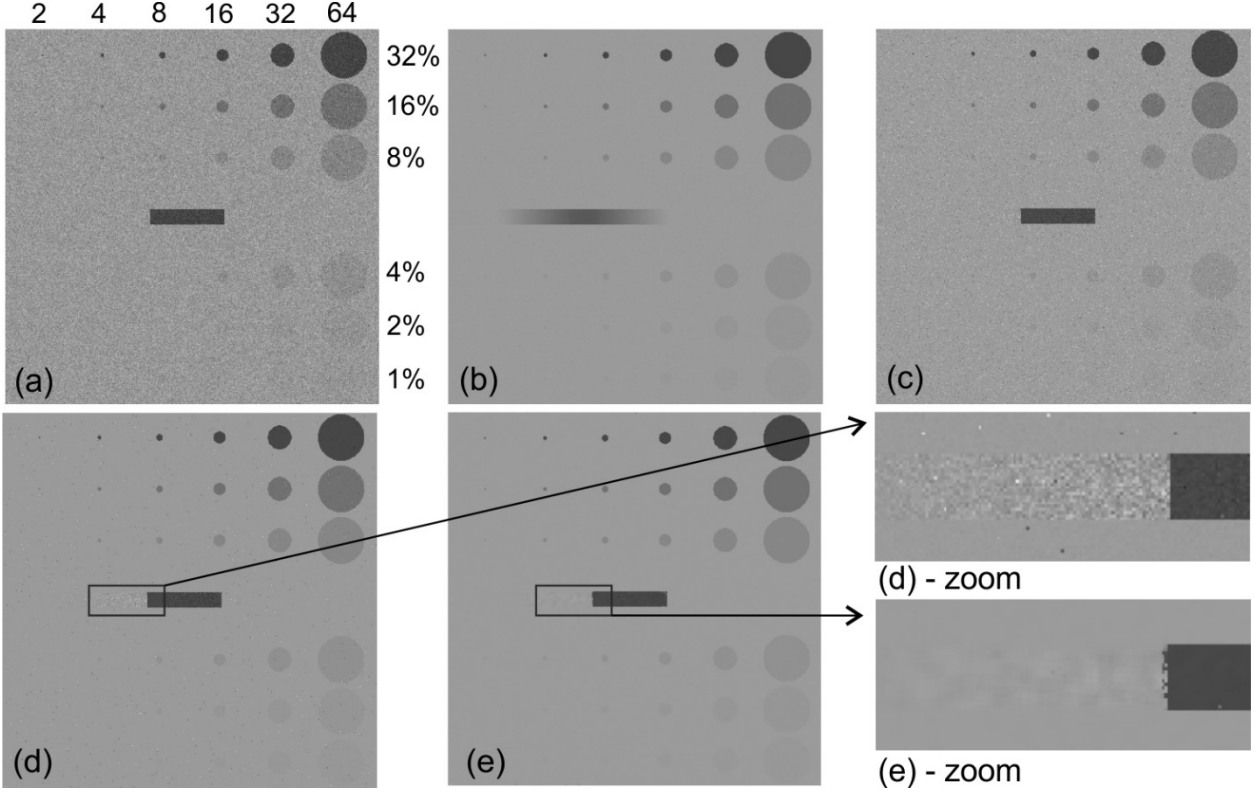
Figure 39. Filtered frame of a test video (filter parameters: M=128, X=Y=1) that presents in the background scene circles with different sizes (diameters on top of (a)) and contrast (contrast at right of (a)) and a rectangular object that moves from left to right: (a) input noisy video; (b) effect of the temporal filtering algorithm realized by using a simple IIR filter. The motion blur effect is relevant; (c) output of the temporal filtering algorithm realized by using the conditioned IIR approach. The motion blur effect is removed but the noise on the background scene is increased (compare (b) with (c)) due to the phenomenon of false reset of the filter; (d) output of the proposed T-filter temporal algorithm. The problem of false reset is almost completely resolved; (e) output of the complete spatio-temporal denoising algorithm (S filter cascaded to T-filters). Note the reduction of noise on the moving object and the complete elimination of the false reset on the background scene

To overcome this issue, infinite impulse response (IIR) filters could be employed for temporal filtering, so the *T-filters* shown in Figure 38 are designed as *N*-th order IIR filters that approximate the impulse response of a moving average FIR filter with $M \gg N$. Therefore, the IIR *T-filters* can be described by the following equation:

$$p_T(i,j,n) = \sum_{m=0}^{N} b_m p(i,j,n-m) - \sum_{m=1}^{N} a_m p_T(i,j,n-m) \qquad (40)$$

Using (40) instead of (39) results in a significantly reduced amount of the memory needed to implement the *T-filter* (e.g. one IIR filter with $N = 10$ can be used to approximate the behaviour of a FIR filter with $M = 128$). Indeed, a method to determine the filter coefficients (b_m, a_m) that allow approximating (39) with (40) should be devised, and would be presented in the next paragraphs. Figure 40 shows the signal flow-graph of the IIR filter. In order to simplify algorithm description, the filtering operation (40) is expressed as follows:

$$p_T(i,j,n) = b_o p(i,j,n) + \sum_{m=1}^{N} b_m si_m(i,j,n) - \sum_{m=1}^{N} a_m so_m(i,j,n) \qquad (41)$$

$$
\begin{aligned}
si_m(i,j,n) &= p(i,j,n-m) \quad for: \quad m = 1..N \\
so_m(i,j,n) &= p_T(i,j,n-m) \quad for: \quad m = 1..N
\end{aligned}
\qquad (42)
$$

where the delayed inputs and outputs, [$si_m(i,j,n)$, $so_m(i,j,n)$], i.e. the outputs of the delay elements, are grouped together and referred to as *filter state*. For a static pixel, whose luminance fluctuations are solely caused by noise, the elaboration proceeds through the IIR approach described in (41)-(42).
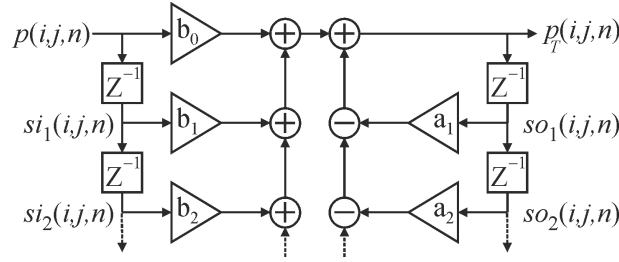
Figure 40. Signal flow-graph of the IIR filter. The output of the delay elements [si_m(i,j,n), so_m(i,j,n)] with m = 1..N is the filter state

Figure 39.(b) shows a frame of a test video processed with the IIR filter (41)-(42) that approximates a temporal filter of length $M = 128$. A significant motion blur effect on the central rectangle can be observed in Figure 39.(b), which was caused by the transient of the temporal filter. The designed *T-filters* actually implement a simple moving average in the time domain, which is known to be unable to provide an edge-preserving noise reduction for moving objects. Hence, the edge-preserving criterion at the core of NVCA denoising should be embedded in the IIR *T-filters*, in order to perform the average computation across the frames only on the pixels at the same spatial position that belong to the same object. However, considering that the average computation on the M frames is not explicit as in a FIR filter, it follows that the original NVCA criterion based on the determination of a binary mask obtained by comparing the differences between all mask pixels with the current pixel could not be implemented, hence a different way of embedding this criterion was conceived, which is based on a *filter reset* approach.

Let us assume that pixel *p(i,j,n)* has been classified as belonging to a different moving object with respect to the pixels *p(i,j,n – 1)* previously processed by the *T-filter(i,j)*. In this hypothesis, the filter state is completely reset to impose a novel steady state condition for the IIR filter, thus the temporal filtering of the pixel and the update of the filter state proceed as follows:

$$p_T(i,j,n) = G_{DC} \cdot p(i,j,n) \qquad (43)$$

$$
\begin{aligned}
si_m(i,j,n+1) &= p(i,j,n) \quad for: \quad m = 1..N \\
so_m(i,j,n+1) &= p_T(i,j,n) \quad for: \quad m = 1..N
\end{aligned}
\qquad (44)
$$

where:

$$G_{DC} = \frac{\sum_{m=1}^{N} b_m}{1 + \sum_{m=1}^{N} a_m} \qquad (45)$$

This event, named *filter reset*, completely discards the actual history of the pixel, which is replaced according to the assumption that the past pixel inputs in the filter memory had all been equal to *p(i,j,n)* precisely. Obviously, considering that all past pixels inputs are equal, thus representing a static input signal, the output pixel *pT(i,j,n)* is equal to the input pixel *p(i,j,n)* multiplied by the DC gain of the IIR filter, as reported in (45). Accordingly, also all the past pixel outputs in the filter memory are set to *pT(i,j,n)*, thus completing the reset of filter state. For the following frames, the processing continues by using (41)-(42) if no new filter reset occurs at position *(i,j)*.

To ascertain that a pixel *p(i,j,n)* corresponds to a different moving object, an adaptive temporal threshold *TT(i,j,n)* is introduced, so that a filter reset in pixel *(i,j)* is performed when the following condition holds:

$$|p(i, j, n) - p_T(i, j, n - 1)| > T_T(i, j, n) \qquad (46)$$

where the threshold *TT(i,j,n)* can be derived as follows. In order to identify whether *p(i,j,n)* belongs to a different moving object, *p(i,j,n)* can be compared with the previous output of the *T-filter*, i.e. *pT(i,j,n−1)*, which is a weighted sum of *p(i,j,n−1)*, mutually independent and assumed as belonging to the same object, and indeed represents the most accurate information about the expected value of pixel luminance at position *(i,j)* in the previous frames. Recalling that $p(i,j,n) \sim N[\mu(i,j,n), \sigma^2(i,j,n)]$, where, according to (38), $\sigma^2(i,j,n)$ is function of $\mu(i,j,n)$, it follows that $p_T(i,j,n-1) \sim N(\mu_T(i,j,n-1), \sigma_T^2(i,j,n-1))$.

The problem of defining whether *p(i,j,n)* belongs to the same object of *pT(i,j,n−1)* by using (46) is the statistical hypothesis testing problem defined by (47):

$$\boldsymbol{H_0}: \quad \mu(i, j, n) = \mu_T(i, j, n - 1) \qquad (47)$$

In practice, (46) assumes that if the difference in luminance between the current input pixel *p(i,j,n)* and the last output pixel *pT(i,j,n)* (i.e. the closest approximation of the expected luminance value of previously processed pixels) is higher than the threshold, then *p(i,j,n)* and *pT(i,j,n)* do not belong to the same distribution and a *filter reset* is required. Obviously this condition is not always verified and, given the selected threshold *TT(i,j,n)*, a probability of error α, referred to as *false reset* probability, can be determined, which corresponds to the probability that, although the two pixels belong to the same distribution (i.e. $\boldsymbol{H_0}$ is verified), their difference exceeds the selected threshold:

$$P[|p(i, j, n) - p_T(i, j, n - 1)| > T_T(i, j, n) \mid \boldsymbol{H_0}] = \alpha \qquad (48)$$

where 1−α represents the confidence level of the test. Recalling that the two random variables *p(i,j,n)* and *pT(i,j,n−1)* are mutually independent, it follows that, under the hypothesis $\boldsymbol{H_0}$, the random variable corresponding to their difference can be expressed as in (49):

$$p(i, j, n) - p_T(i, j, n - 1) \quad \sim \quad N[0, \sigma_D^2(i, j, n - 1)] \qquad (49)$$

$$\sigma_D^2(i, j, n - 1) = \sigma^2(i, j, n) + \sigma_T^2(i, j, n - 1) \qquad (50)$$

The variance $\sigma_T^2(i,j,n-1)$ of the IIR filter response can be computed by approximating the filter impulse response with that of the emulated moving average filter (i.e. a rectangular window of length *M)*:

$$p_T(i,j,n-1) = \sum_{k=1}^{\infty} h_m(k-1)p'(i,j,n-k) \cong \frac{1}{M}\sum_{k=1}^{M} p'(i,j,n-k) \tag{51}$$

with $p'(i,j,n-k)$ defined as follows:

$$p'(i,j,n-k) = \begin{cases} p(i,j,n-k) & k \leq m(i,j,n-1) \\ p(i,j,n-m(i,j,n-1)) & k > m(i,j,n-1) \end{cases} \tag{52}$$

where $m(i,j,n-1)\in[1,M]$ is the number of frames from which the pixel $p_T(i,j,n-1)$ can be considered static. As an example, if the last *filter reset* occurred at the frame $n-k^*$, then $m(i,j,n-1) = k^*$. In (52) $p(i,j,n-k)$ are mutually independent and belong to the same distribution, that is they have the same expected value and variance ($\mu_T(i,j,n-1)$, $\sigma_T^2(i,j,n-1)$). In addition, under hypothesis $\mathbf{H_0}$, their expected value and variance are equal to those of $p(i,j,n)$, that is $p(i,j,n-k) \sim N(\mu(i,j,n), \sigma^2(i,j,n))$. From (51)-(52) it is possible to derive the expression of $\sigma_T^2(i,j,n-1)$:

$$\sigma_T^2(i,j,n-1) = \sigma^2(i,j,n) \cdot \left(\frac{m(i,j,n-1)^2}{M^2} - \frac{m(i,j,n-1)}{M}\left(2+\frac{1}{M}\right) + \left(1+\frac{2}{M}\right) \right) \tag{53}$$

Then, the expression of $\sigma_D^2(i,j,n-1)$ can be computed by substituting (53) in (50):

$$\sigma_D^2(i,j,n-1) = \sigma^2(i,j,n) \cdot g(m(i,j,n-1)) \tag{54}$$

$$g(m) = \frac{m^2}{M^2} - \frac{m}{M}\left(2+\frac{1}{M}\right) + \left(2+\frac{2}{M}\right) \tag{55}$$

Finally, the solution of (48) can be obtained from (49) as follows:

$$T_{T(i,j,n)} = CDF_N^{-1}\left(1-\frac{\alpha}{2}\right) \cdot \sqrt{\sigma^2(i,j,n) \cdot g(m(i,j,n-1))} \tag{56}$$

where $CDF_N^{-1}(p)$ is the inverse normal cumulative distribution function, α is the *false reset* probability, $m(i,j,n-1)\in[1,M]$ is a counter which counts, at the position (i,j), the number of frames from a previous filter reset, and $g(m)$ is the non-linear function defined in (55). It is worth noting that the multiplicative coefficient of the square root in (56) corresponds to the N_σ coefficient of the original NVCA formulation, which here has been explicitly related to the probability of erroneous classification of pixels as belonging or not to the same distribution. Moreover The *false reset* probability was set to $\alpha = 2.7\times10^{-3}$, which gives $CDF_N^{-1}(1-\alpha/2) = 3$.

According to (38), the noise variance $\sigma^2(i,j,n)$ depends on the expected luminance $\mu(i,j,n)$, which is obviously unknown, and can be approximated as the previous filtered output $p_T(i,j,n-1)$:

$$\sigma^2(i,j,n) \simeq G(p_T(i,j,n-1)) \tag{57}$$

From (56)-(57), the threshold $T_T(i,j,n)$ becomes a function of $p_T(i,j,n-1)$ and $m(i,j,n-1)$, and with the chosen α value can be expressed as follows:

$$T_T(i,j,n) = 3 \cdot \sqrt{G(p_T(i,j,n-1)) \cdot g(m(i,j,n-1))} \tag{58}$$

The expression in (58) can be easily computed in hardware, as $p_T(i,j,n\text{-}1)$ is part of the filter state, $m(i,j,n\text{-}1)$ is the count value of a 7-bits counter and the two non-linear functions $G(\cdot)$ and $g(\cdot)$ can be implemented via lookup tables.

Figure 41 shows the plot of threshold $T_T(i,j,n)$ as a function of $m(i,j,n\text{-}1)$. The proposed approach is able to adaptively reduce the threshold $T_T(i,j,n)$ when the time from the previous filter reset increases ($m(i,j,n\text{-}1)$ increases), thus improving the ability of the algorithm to restore low-contrast edges which remain static for a sufficient number of frames. Figure 39.(c) shows the filtering result obtained by using the *T-filter* algorithm with filter reset. The conditioned filtering operation avoids motion blur on moving objects. However, a significant noise appears on the background scene, which is due to the occurrence of *false resets* in the considered frame, as well as to the occurrence of *false resets* in previous frames that propagate to the current frame.

This effect can be mitigated by reducing the *false reset* probability α, which however would increase the threshold $T_T(i,j,n)$ and, consequently, the motion blur. Instead, a *false reset* propagation avoidance could be achieved by modifying the *T-filter* algorithm to avoid that a *false reset* from the previous frames propagates to the current frame.
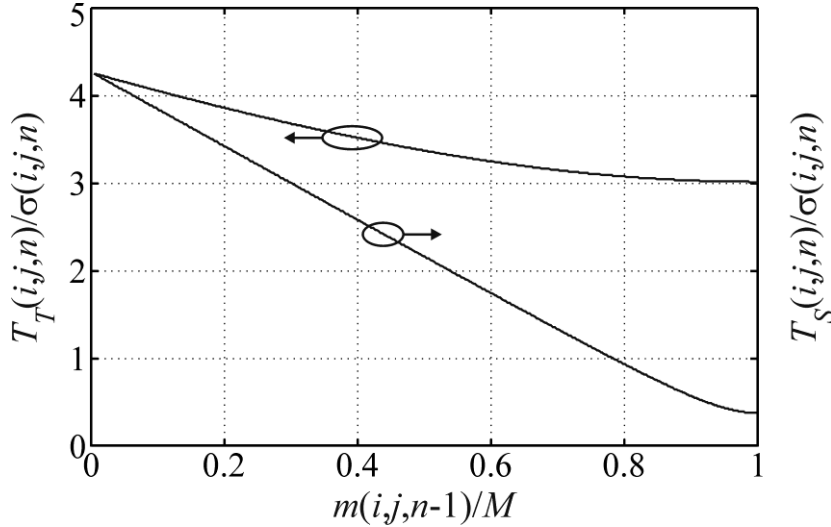
Figure 41. Plot of the thresholds $T_T(i,j,n)$ and $T_S(i,j,n)$ as a function of $m(i,j,n\text{-}1)$. In the plot $M = 128$ and $CDF_N^{-1}(1\text{-}\alpha/2)=3$ are assumed

Assuming that a pixel $p(i,j,n)$ of the n-th frame is being processed, and that a filter reset occurred in frame $n\!-\!1$ in the same position (i,j) (that is $m(i,j,n\text{-}1) = 1$), to verify whether a *false reset* occurred at frame $n\!-\!1$, the current pixel value $p(i,j,n)$ should be compared with $p_T(i,j,n\!-\!2)$ (i.e. the last output of the pixel distribution preceding the *false reset* event). A *false reset* is detected, with probability $1\!-\!\alpha^2$, if the current pixel $p(i,j,n)$ turns out to belong to the same distribution of $p_T(i,j,n\!-\!2)$, i.e. if the following condition holds:

$$|p(i,j,n) - p_T(i,j,n-2)| < T_T(i,j,n) \tag{59}$$

Once a false reset is detected, a corrective action can be performed, which consists in restoring the filter state at the instant $n-1$, as follows:

$$p_T(i,j,n) = b_o p(i,j,n) + \sum_{m=1}^{N} b_m s i_m(i,j,n-1) - \sum_{m=1}^{N} a_m s o_m(i,j,n-1) \tag{60}$$

72

This approach reduces the probability of propagation of the *false reset* events for more than one frame from α to α^2, that is, from $2.7 \cdot 10^{-3}$ to $7.3 \cdot 10^{-6}$, for the particular choice of α.

If the condition (59) is not verified, the elaboration proceeds by using (41)-(42). The flowchart in Figure 42 summarizes the complete *T-filter* algorithm. Figure 39.(d) shows the filtering result obtained by using the complete *T-filter* algorithm of Figure 42. The algorithm showed very good capabilities to track moving objects while being able to reduce the noise due to *false reset* events. Moreover, it could be noted that some background noise due to residual *false reset* events still shows up within the frames (compare the background scene in Figure 39.(b) with Figure 39.(d)). This noise is the one generated within the current frame (which does not propagate from previous frames) and will be further attenuated by the spatial filter described in the following paragraph.
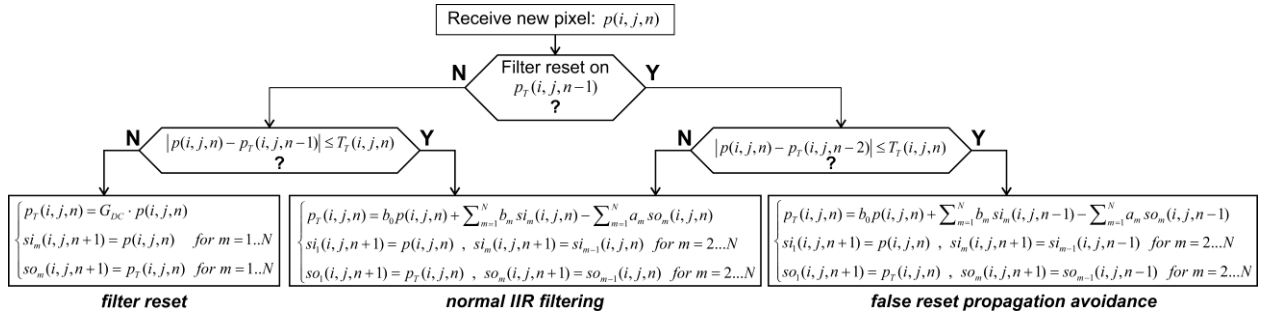
Figure 42. Flowchart of the T-filter algorithm, including filter reset and false reset propagation avoidance

As previously mentioned, a method to design the IIR T-filters coefficients should be devised. Several numerical methods exist to design IIR filters [145 – 148] . However, most techniques (e.g. [145 – 146]) aims to minimize error metrics in the frequency domain, while in the present application the error in the time domain is the actual target. An IIR filter design technique was conceived based on the Steiglitz – McBride iterative method [148], which attempts to minimize the squared error in the time domain between the actual impulse response of the IIR filter and the desired impulse response $h_d(n)$, that in this case is represented by a rectangular window of length M. More in-depth information on the adopted design procedure can be found in [149].

7.3 Spatial filters design

As shown in Figure 38, the output frame of the *T-filters* stage is fed to an *S-filter* that performs a conditioned spatial average on the pixels $p_T(i',j',n)$ of the $(2X+1) \times (2Y+1)$ neighbourhood of $p_T(i,j,n)$, by excluding all pixels that do not belong to the same object of the central pixel, that is all the $p_T(i',j',n)$ whose difference in luminance with $p_T(i,j,n)$ exceeds a determined spatial threshold $T_S(i,j,n)$, which can be derived as follows.

By following the same approach used for *T-filters*, the generic pixel $p_T(i',j',n)$ is considered in the average computation only if it is statistically inferred that it belongs to the same object of pixel $p_T(i,j,n)$ by using the following criterion:

$$\left| p_T(i',j',n) - p_T(i,j,n) \right| \leq T_S(i,j,n) \tag{61}$$

This results in the following statistical hypothesis testing problem:

$$\boldsymbol{H_0}: \quad \mu_T(i',j',n) = \mu_T(i,j,n) \tag{62}$$

Essentially, as in the case of temporal filtering, (61) assumes that if the difference in luminance between the central pixel $p_T(i,j,n)$ and its neighbours pixels $p_T(i',j',n)$ is lower than the threshold $T_S(i,j,n)$, then they belong to the same distribution and, therefore, are included in the filtering operation. The problem is solved by defining a threshold $T_S(i,j,i',j',n)$, for which it is possible to determine the error probability of the criterion described in (61):

$$P\left[\left|p_T(i',j',n) - p_T(i,j,n)\right| > T_S(i,j,i',j',n) \mid \boldsymbol{H_0}\right] = \alpha \tag{63}$$

It is worth noting that, according to (63), the threshold T_S is not only a function of (i,j) but also of (i',j'), as, indeed, the variance of the random variable corresponding the pixels difference $p_T(i,j,n) - p_T(i',j',n)$ depends on the variances of both $p_T(i,j,n)$ and $p_T(i',j',n)$, which, in turn, according to (53), depend on $m(i,j,n)$ and $m(i',j',n)$, respectively. However, with the help of a large number of simulations, it has been verified that it is possible to use the same threshold for all pixels $p_T(i',j',n)$ in the considered neighbourhood of $p_T(i,j,n)$, by assuming $\sigma_T^2(i',j',n) = \sigma_T^2(i,j,n)$, which leads to use a threshold $T_S(i,j,n)$ defined as:

$$T_S(i,j,n) \triangleq T_S(i,j,i',j',n)\Big|_{\sigma_T^2(i',j',n)=\sigma_T^2(i,j,n)} \tag{64}$$

which allows deriving the following expression for the threshold:

$$T_S(i,j,n) = CDF_N^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{2 \cdot \sigma_T^2(i,j,n))} \tag{65}$$

Considering the expression of $\sigma_T^2(i,j,n)$ given in (53), and also (54) and (50), the threshold can be expressed as:

$$T_S(i,j,n) = CDF_N^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{2 \cdot \sigma^2(i,j,n) \cdot (g(m(i,j,n)) - 1)} \tag{66}$$

In (66) the variance $\sigma^2(i,j,n)$ of the noise depends on $\mu(i,j,n)$, which is obviously unknown and can be approximated as $p_T(i,j,n)$. The threshold $T_S(i,j,n)$ can be expressed as in (67) by setting the same error probability considered in the temporal filtering $\alpha = 2.7\times10^{-3}$, which yields $CDF_N^{-1}(1-\alpha/2) = 3$:

$$T_S(i,j,n) = 3\sqrt{2 \cdot G(p_T(i,j,n)) \cdot (g(m(i,j,n-1)) - 1)} \tag{67}$$

Figure 41 shows the threshold $T_S(i,j,n)$ as a function of $m(i,j,n)$ for $M = 128$. Note that, by increasing $m(i,j,n)$, the noise of pixels $p_T(i',j',n)$ is reduced, which allows using a very low threshold $T_S(i,j,n)$, while maintaining a good confidence level for the statistical criterion in (61). As an example, Figure 41 shows that $T_S(i,j,n)$ reduces to $0.375 \cdot \sigma(i,j,n)$ for $m(i,j,n) = M = 128$. The automatic adaptation of the thresholds to match the local estimated noise level, grants to the improved NVCA algorithm the ability to recognize and restore very low-contrast object contours.

By following the averaging approach of the original NVCA algorithm, the pixels $p_T(i',j',n)$ classified as belonging to the same distribution of the current pixel $p_T(i,j,n)$ could be simply averaged together to produce the output pixel $p_S(i,j,n)$. However, it should be observed that the pixels $p_T(i',j',n)$ are not equally reliable in terms of the information they bring about the true expected luminance of the current pixel $p_T(i,j,n)$. Indeed, according to (53), pixels with larger associated m values, i.e. pixels which have been averaged for a larger number of frames, are characterized by lower noise and, therefore, provide more reliable information on the true (noiseless) luminance of the current pixel neighbourhood, which is used in practice to infer the current pixel luminance. Hence, when estimating the average neighbourhood luminance, to infer the true current pixel luminance, pixels with larger m

values should be given more weight. This was attained by implementing the *weighted conditioned average* of pixels $p_T(i',j',n)$ described in (68):

$$p_S(i,j,n) = \frac{\sum_{i'=i-X}^{i+X}\sum_{j'=j-Y}^{j+Y} p_T(i',j',n) \cdot m_{th}(i',j',n)}{\sum_{i'=i-X}^{i+X}\sum_{j'=j-Y}^{j+Y} m_{th}(i',j',n)} \tag{68}$$

where the weights $m_{th}(i',j',n)$ given to each pixel are equal to $m(i',j',n)$, for the pixels belonging to the distribution of the current pixel, and forced to zero for those violating (61):

$$m_{th}(i',j',n) = \begin{cases} m(i',j',n) & if \ \left| p_T(i',j',n) - p_T(i,j,n) \right| \le T_S(i,j,n) \\ 0 & otherwise \end{cases} \tag{69}$$

An additional improvement of the *S-filter* performance can be achieved by separately processing the *isolated pixels*. A pixel $p_T(i,j,n)$ is considered isolated if there are no pixels $p_T(i',j',n)$ in its 3×3 spatial neighbourhood that belong to the same object, that is:

$$\left| p_T(i',j',n) - p_T(i,j,n) \right| > T_S(i,j,n) , \qquad \forall i' \in [i-1;i+1], \forall j' \in [j-1;j+1] \tag{70}$$

Isolated pixels may be realizations of the tails of noise distributions, but they can also be caused by either the residual *false reset* noise introduced by the *T-filters* or the speckle noise of the sensor. Therefore, a processing stage dedicated to *isolated pixels* was included in the *S-filter* to detect the isolated pixels by means of (70) and replace each of them with the unconditioned weighted average of the pixels in its 3×3 spatial neighbourhood, computed by excluding the isolated pixel itself from the average computation. In conclusion, the complete description of *S-filter* operation is given in (71):

$$p_S(i,j,n) = \begin{cases} \dfrac{\sum_{i'=i-X}^{i+X}\sum_{j'=j-Y}^{j+Y} p_T(i',j',n) \cdot m_{th}(i',j',n)}{\sum_{i'=i-X}^{i+X}\sum_{j'=j-Y}^{j+Y} m_{th}(i',j',n)} & \begin{array}{l} if \exists p_T(i',j',n)\big|_{\substack{i'\ne i \\ j'\ne j}}: \\ \left| p_T(i',j',n) - p_T(i,j,n) \right| \le T_S(i,j,n) \end{array} \\ \\ \dfrac{\sum_{\substack{i'=i-1 \\ i'\ne i}}^{i+1}\sum_{\substack{j'=j-1 \\ j'\ne j}}^{j+1} p_T(i',j',n) \cdot m(i',j',n)}{\sum_{\substack{i'=i-1 \\ i'\ne i}}^{i+1}\sum_{\substack{j'=j-1 \\ j'\ne j}}^{j+1} m(i',j',n)} & \begin{array}{c} otherwise \\ (isolated \ pixel) \end{array} \end{cases} \tag{71}$$

Figure 39.(e) shows the result provided by the complete cascaded temporal and spatial filtering scheme depicted in Figure 38, where it can be observed that the spatial conditioned average helped in further reducing both the residual noise on the moving object and the *false reset* noise of the *T-filters*, while ensuring a very accurate edge delineation.

7.4 Performance comparison with state of the art

The performances of the improved NVCA algorithm just described were compared against several state-of-the-art denoising methods, namely 4-PDE [113], BM3D [117], BM3Dc [46], UINTA [150], NLM [151], VBM3D [152], STGSM [153], as well as the original NVCA algorithm [40, 61]. Since some of these algorithms feature a number of parameters to be set, a parameter tuning procedure was carried out to make those algorithms achieve the finest image restoration. Two image quality

metrics were considered, namely PSNR and SSIM [134]. The first test was performed on a synthetic, noiseless, static sequence (assumed as reference) composed by 500 frames of 512×512 pixels, which was then corrupted with a simulated Poisson noise (PSNR$_{in}$ = 20.1 dB, SSIM$_{in}$ = 0.234). The results are reported in Table 8, where it can be verified that VBM3D obtained the best performances by achieving a PSNR$_{out}$ of 45.3 dB (i.e. a noise reduction of 25.2 dB) and a SSIM$_{out}$ of 0.995. Indeed, the original NVCA algorithm [40] achieved a SSIM$_{out}$ of 0.968 and a PSNR$_{out}$ of 30.8 dB, which was about 14.5 dB lower than VBM3D. However, it should also be noted that, while NVCA has already attained real-time operation via hardware implementation, as demonstrated in [61], this could not be accomplished by VBM3D, due to its very high computational burden. The improved NVCA filter here proposed, surprisingly, outperformed VBM3D and also achieved comparable performances for small mask sizes (e.g. with $X = Y = 1$ and $M = 32$, i.e. an equivalent 3x3x32 spatio – temporal mask, it achieved a PSNR$_{out}$ of 44.3 dB while VBM3D achieved a PSNR$_{out}$ of 45.3 dB).

To assess the effectiveness of the improved NVCA algorithm in an experimental scenario, a real fluoroscopic sequence was used, which represented a static scene of an X-ray step phantom. The phantom was made by seven square aluminium sheets (side lengths from 6 cm to 30 cm), each 1 mm thick, superimposed one to each other in a ladder structure (see Figure 11 in Chapter 5). The scene was imaged with a GE OEC 9900 [132] fluoroscope and RAW video data have been processed. Figure 43.(a) shows an unfiltered frame of the scene (Figure 43.(b) is a 100×300 pixels detail).

Table 8. Denoising methods comparison. As input a test video was used, which was composed by 500 frames of a static image (assumed as reference) with resolution 512×512 pixel, corrupted by a simulated Poisson noise (PSNRin = 20.1 dB SSIMin = 0.234)

Denoising method	4-PDE	BM3D	VBM3D	BM3Dc	NLM	UINTA	STGSM	NVCA X=Y=3 K=5
PSNR$_{out}$ (dB)	28.4	39.9	**45.3**	40.1	37.7	28.8	40.2	30.8
SSIM$_{out}$	0.740	0.990	**0.995**	0.991	0.980	0.840	0.993	0.968
Improved NVCA (various temporal and spatial mask sizes)								
Denoising method	X=Y=1 M=32	X=Y=2 M=32	X=Y=3 M=32	X=Y=1 M=64	X=Y=2 M=64	X=Y=3 M=64	X=Y=1 M=128	X=Y=2 M=128	X=Y=3 M=128
PSNR$_{out}$ (dB)	44.3	48.1	49.8	47.3	50.9	52.5	49.9	53.1	**54.4**
SSIM$_{out}$	0.989	0.996	0.998	0.995	0.998	0.999	0.997	0.999	**0.999**

| (a) | (b) | (c) |
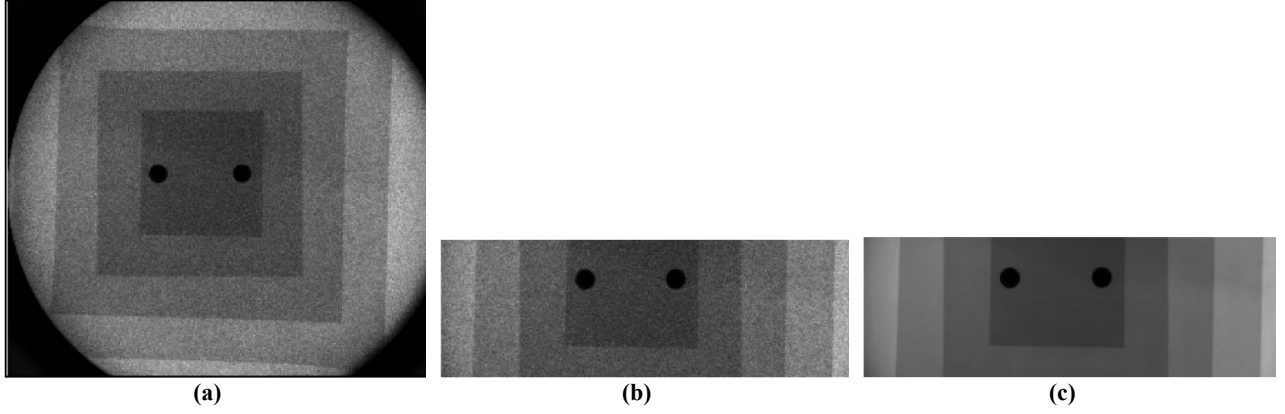
Figure 43. (a) Fluoroscopic image from the test video sequence; (b) 100x300 pixels zoom of (a); (c) result of the proposed filter (X=Y=1, M=32)

Table 9. Comparison of denoising methods for a 100x300 pixels region taken from a real fluoroscopic sequence. Image characteristics: $PSNR_{in} = 24.5dB$ $SSIM_{in} = 0.462$.

Denoising method	$PSNR_{out}$ (dB)	$SSIM_{out}$
4-PDE	31.7	0.855
BM3D	36.4	0.963
VBM3D	38.2	0.973
BM3Dc	37.5	0.975
NLM	36.2	0.980
UINTA	28.9	0.879
STGSM	36.9	0.971
NVCA X=Y=3 K=5	31.1	0.931
Improved NVCA X=Y=1 M=32	**40.0**	**0.981**

The zoomed image processed by using the proposed spatio-temporal filter ($X = Y = 1$, $M = 32$) is shown in Figure 43.(c) and confirms that the proposed algorithm provides an effective, edge-aware denoising. Table 9 quantifies the performances achieved on the experimental video and compares again the proposed approach with the state of the art. Table 9 confirms that the proposed technique, also with $M = 32$, achieves performances comparable to VBM3D.

Two publicly available video sequences [112], also corrupted with simulated Poisson noise, were considered to further test the proposed algorithm. The single frames obtained in the case of bridge close video with a $PSNR_{in}$ of 20 dB are depicted in Figure 44. Table 10 outlines the results for two input noise levels, which confirm that the improved NVCA provided similar performances as compared to the best performing approaches (VBM3D, STGSM). Indeed, the new algorithm achieved a $PSNR_{out}$ of only about 2 dB less than VBM3D and an higher $SSIM_{out}$ (0.915 vs. 0.896), but was slightly outperformed by STGSM, which achieved both an higher PSNRout (35.4 dB vs. 31.2 dB) and an higher $SSIM_{out}$ (0.945 vs. 0.915). Nonetheless, STGSM requires forward and inverse wavelet transforms, as well as motion estimation/compensation algorithms, which make it unsuitable for real-time hardware implementation, similarly to VBM3D.
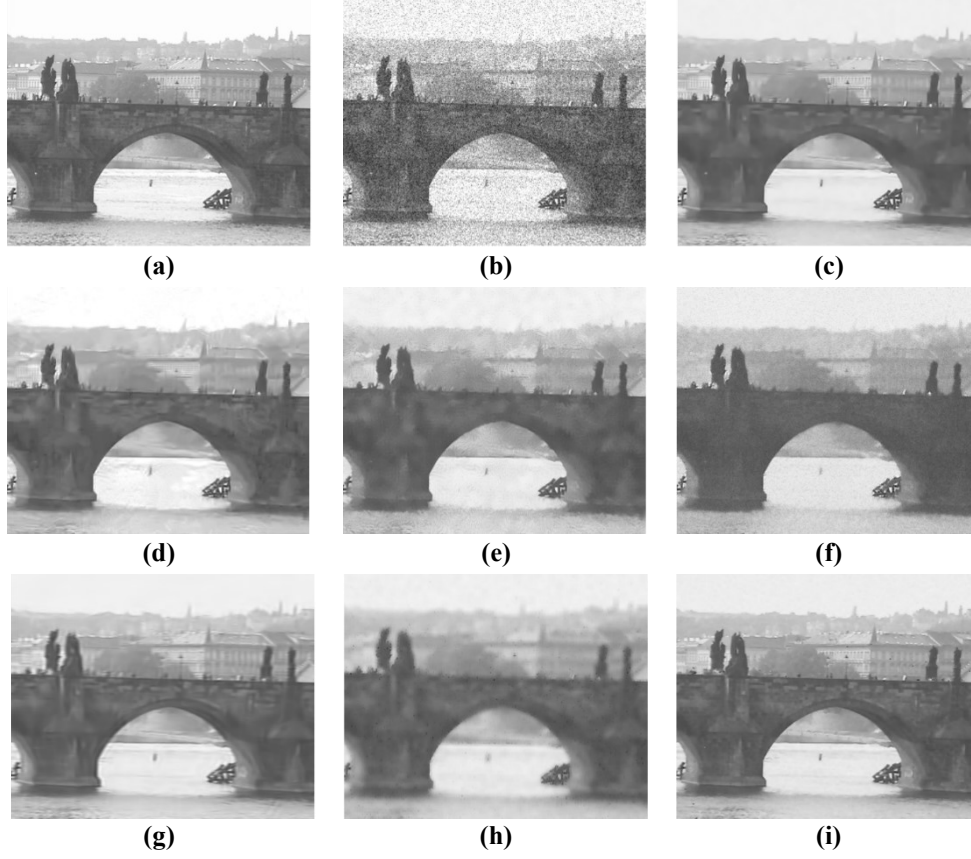
Figure 44. Single frame result images obtained filtering the "bridge close" publicly available video [112]. (a) Noiseless reference; (b) Noisy (PSNR$_{in}$=20dB); (c) VBM3D; (d) BM3Dc; (e) NLM; (f) UINTA; (g) STGSM; (h) NVCA (X=Y=3, K=5); (i) improved NVCA (X=Y=1, M=32).

Table 10. Comparison of denoising methods for two publicly available video sequences [112] to which Poisson noise has been applied.

Denoising method	PSNR$_{in}$ = 20dB				PSNR$_{in}$ = 15dB			
	Akiyo		Bridge close		Akiyo		Bridge close	
	PSNR$_{out}$ (dB)	SSIM$_{out}$	PSNR$_{out}$ (dB)	SSIM$_{out}$	PSNR$_{out}$ (dB)	SSIM$_{out}$	PSNR$_{out}$ (dB)	SSIM$_{out}$
VBM3D	33.5	0.896	29.0	0.831	28.9	0.814	24.9	0.740
BM3Dc	32.3	0.905	28.5	0.751	24.6	0.493	25.7	0.615
NLM	28.3	0.857	27.0	0.701	25.4	0.623	23.8	0.568
UINTA	26.8	0.690	25.2	0.628	16.2	0.307	18.1	0.310
STGSM	35.4	0.945	27.6	0.813	30.7	0.907	23.7	0.718
NVCA X=Y=3 K=5	28.9	0.785	25.8	0.676	22.7	0.354	22.3	0.519
Improved NVCA X=Y=1 M=32	31.2	0.915	28.5	0.858	30.1	0.876	25.1	0.802

7.5 Hardware Implementation

The improved NVCA algorithm, as mentioned above, was implemented in hardware to achieve a real-time operation. In particular, 10-th order IIR *T-filters* (i.e. *N* = 10) performing an equivalent moving average on 128 samples (i.e. M = 128), as well as an *S-filter* with 3×3 spatial mask size (i.e. *X* = *Y* = 1), were implemented on an Altera StratixIV EP4SGX70HF35C2 FPGA, which is the smallest FPGA of the StratixIV family, i.e. the one with the lowest amount of hardware resources. The Altera Quartus II software was used to accomplish the synthesis and the *place & route* procedures for the circuit, while the Mentor Questa Sim software was employed for behavioural and gate level simulations. Figure 45 shows the top-level architecture of the designed circuit. The system receives the stream of input pixels $Pix_{in}(i,j,n)$ in a serial fashion. The pixels, represented on 8 bits (as indicated by the number of parallel lines reported for the arrow corresponding to the stream of input data), are ordered by columns (*i* index), then by rows (*j* index) and finally by frames (*n* index), and are synchronized by the clk_{filter} signal. A frame of size 1024×1024 (*K* = *L* = 1024) was assumed, so the state of 1024×1024 = 1,048,576 IIR filters was stored in the external memory of Figure 45, which in the present implementation was a 1 GB DDR2 800 SDRAM.
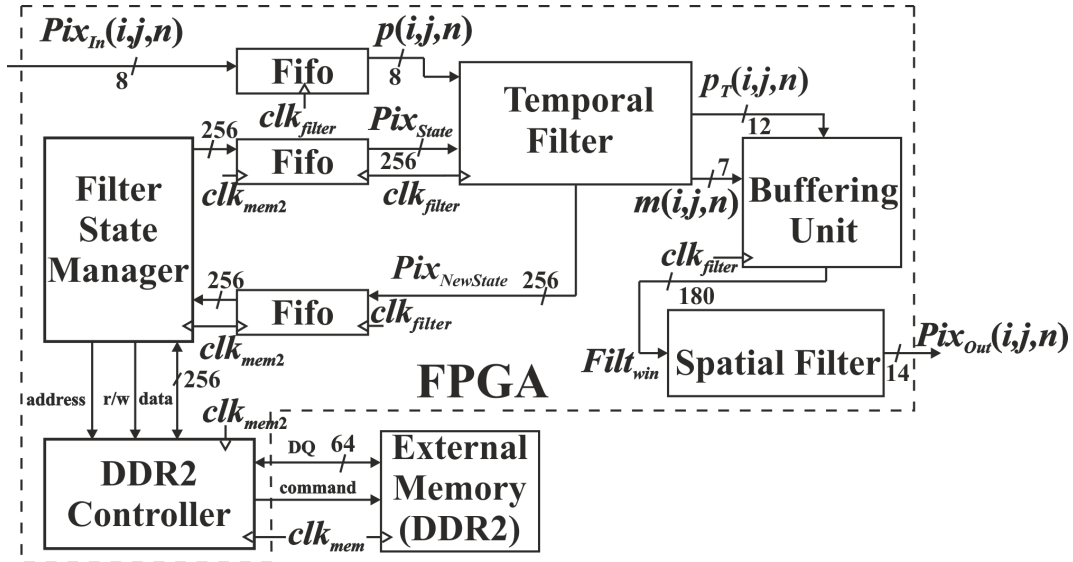
Figure 45. Top-level architecture of the proposed filtering system

The designed circuit uses only the 22% of the logic resources of the target FPGA, which means that plenty of resources is still available for additional circuitry. In particular, Table 11 details the logic resource usage for each block of the top-level architecture depicted in Figure 45, where it can be assessed that the percentage usage is well balanced for different resources, such as ALM, Flip Flops, embedded memory, and DSP elements, and never exceeds the 22% of the related total resources. As expected, the frame rate turned out to be limited by the filter block (*T-filters* + *S-filter*) to 49 fps for the considered frame size of 1024×1024 pixels. It is worth mentioning that most modern fluoroscopic devices usually provide 1024×1024 greyscale images with a frame rate up to 30 fps, for which the proposed circuit would easily achieve real-time operation.

Table 12 outlines performance and resource usage of the improved NVCA algorithm as compared to the hardware implementation of the original NVCA algorithm described in [108], which was based on a spatio – temporal FIR filter. The comparison clearly shows that the improved NVCA

clearly outperforms the original NVCA by entailing a lower percentage usage for all different resources but the DSP elements.

Table 11. Resource usage of the proposed circuit

Proposed Circuit	#ALM (%)	#FF (%)	#M9K (%)	MLAB (kb) (%)	#DSP (%)	Frequency (MHz)	Frame rate (fps)*
Filter*	3041 (10%)	2408 (4%)	0 (0%)	0 (0%)	14 (4%)	52.3	49
Filter State Manager + Buffering Unit	3057 (11%)	3404 (6%)	10 (2%)	63 (7%)	0 (0%)	200	138
FIFO	139 (<1%)	165 (<1%)	16 (3%)	0.064 (<1%)	0 (0%)	200 and 52.3	-
Total	6237 (22%)	5977 (11%)	26 (6%)	63 (7%)	14 (4%)	-	49

*Frame size 1024 × 1024

The introduction of additional multipliers in the *S-filter*, which were needed to implement the operation described by (71), determined the higher usage of DSP elements, whose percentage is in any case very low (4% only). It is worth highlighting that the improved NVCA hardware implementation entails a sensible reduction (about 75%) of the ALM and Flip flops usage with respect to the circuit proposed in [108], mainly because of the reduced number of buffering elements within the FPGA. This is attained via the separation of the temporal and spatial filtering operations, which allows reducing the hardware complexity, and the use of the IIR scheme in place of the FIR scheme, which further optimizes the implementation of the *T-filters*.

Table 12. Comparison with original NVCA hardware implementation presented in [108]

Circuit	#ALM (%)	#FF (%)	#M9K (%)	MLAB (kb) (%)	#DSP (%)	Frequency (MHz)	Frame rate (fps)*
Original NVCA [108]	23202 (80%)	22480 (39%)	116 (25%)	103 (11%)	2 (<1%)	51.4	49
Improved NVCA [112]	6237 (22%)	5977 (11%)	26 (6%)	63 (7%)	14 (4%)	52.3	49

*Frame size 1024 × 1024

7.6 Progresses and limitations

The improvements brought to the original NVCA algorithm, i.e. the separation of the filtering operation in the space and time domains, the IIR-based implementation of temporal filtering and the use of adaptive thresholds, yielded both enhanced denoising performances and reduced hardware usage. However, more extensive testing should still be carried out to assess the performances of the improved NVCA algorithm in different challenging conditions, such as low-contrast high-noise moving edges (i.e. low CNR_{in}), with different speed and size. In addition, the new algorithm features five free parameters in its most general formulation, namely the equivalent temporal window length M and the IIR order N for the *T-filters*, the spatial mask size S for the *S-filter*, and the two inferential error probabilities α_T (*false reset* probability) and α_S, for the *T-filters* and *S-filter*, respectively. Hence, for the algorithm to be effectively used in both online and offline processing, some indications should be devised to find the optimal parameters settings for different situations and conditions, which requires further investigation, as well as the assessment of an edge-aware IQA index that provides reliable estimates of image quality, as showed in Chapter 6. Moreover, the possibility to tune these parameters in real time should also be addressed in the context of the hardware implementation, as their effects on the circuital design are very different. Indeed, changing the filter mask sizes in real time could be quite challenging, as it would require the stream of data to be modified between different blocks of the top-level architecture; adapting the equivalent time window length M for the *T-filters* would either require the computation of new IIR filters coefficients through the iterative procedure described in [112], which is clearly unsuitable for real-time operation, or the predetermination of a limited set of coefficients to be selected in real time, thus providing for a fixed number of settings for the parameter M; modulating the inferential error probabilities (α_T, α_S), instead, could be easily achieved by either adding a multiplier at the output of each LUT for threshold computation (designed by considering a unit multiplier in (58) and (67) in place of the $CDF^{-1}(1-\alpha/2)$ term) or by implementing a 2D LUT that contains the threshold values corresponding to a number of fixed settings for the parameter α.

Finally, an unsolved issue still remains, which has never been addressed and would prevent NVCA and any denoising strategy based on the *a priori* knowledge of noise statistics (e.g. all the AWGN denoising schemes combined with the generalised Anscombe transform) from achieving an effective real-time denoising of fluoroscopic sequences. Indeed, the noise parameters (introduced in Chapter 2) required by these strategies may experience substantial variations due to alterations of imaging conditions and, thus, they cannot be inferred *a priori*; hence, they are usually extracted offline directly from the noisy data to be processed, which clearly does not comply with real-time operation. Ideally, the noise parameters estimation should be performed after any change in the imaging conditions, but this would be hardly feasible, especially during image-guided interventional procedures, which undoubtedly represent the most appealing and widespread applications of low-dose fluoroscopy. This awkward yet crucial issue is addressed in Chapter 8.

Chapter 8

Noise characterization in fluoroscopic devices

This chapter focuses on a very important issue that has never been addressed in literature and has serious implications on the real-time performances of algorithms for quantum noise suppression: the estimation of noise parameters in time-varying imaging conditions. Indeed, the denoising performances of algorithms that take advantage of the *a priori* knowledge of noise properties, such as NVCA, depend not only on the selection of potential tuning parameters, but also on the validity of the noise model and, of course, on the accuracy of its parameters, commonly referred to as noise parameters [4, 63, 103, 111]. However, noise parameters may vary due to alterations of imaging conditions among distinct image acquisitions, and even within the same acquisition, as in the case of video sequences. As an example, in natural images and videos acquisition, fluctuations of environmental light conditions, as well as the adjustment of settings of the imaging device, may alter the statistical properties of the noise that affects the acquired images. In X-ray imaging, the light source is regulated by acting on two controls of the imaging device itself, that is the current and the peak-voltage of the X-ray tube, which, therefore, appear as the main parameters the noise properties depend on. These parameters can be adjusted during fluoroscopically-guided interventional procedures, in order to attain the desired details visibility for different tissues, while limiting the X-ray dose delivered to the patient. Hence, these fluoroscopic sequences turn out to be corrupted by a quantum noise with time-varying statistics, which would invalidate any previously performed estimation of the noise parameters, thus limiting the effectiveness of denoising strategies based on noise statistics. A solution to this issue is investigated, which is based on the possibility of pre-characterizing the noise parameters for a single fluoroscopic device at many different X-ray tube settings, in order to be able to switch among different predetermined parameters in real time, according to the particular choice of the X-ray tube settings.

8.1 Innovative approach: *a priori* noise characterization

Denoising approaches that make direct use of Poisson statistics, as well as those based on the combination of generalized Anscombe transform and AWGN denoising schemes, both require noise parameters to be accurately estimated from noisy images prior to their actual processing, in order to achieve a reasonable trade-off between noise reduction and edge preservation, especially in images that are heavily affected by noise (e.g. low-dose X-ray images). While it is not usually a major concern in offline implementations, it could be a serious limitation in real-time operation, which undoubtedly represents the most appealing application of fluoroscopic sequence denoising. Indeed, the variations of tube settings and detector gain that occur during a fluoroscopically-guided procedure modulate the statistics of quantum noise, and this can be drawn also from (11) and (12), as the number and energies

of the incident X-ray photons depend on the current and the peak-voltage of the X-ray tube, which, together with the peculiar energy spectrum of the X-ray tube and the overall energy-to-voltage conversion factor of the detector (i.e. the detector gain), influence the expected value of the signal at the detector and, accordingly, the noise variance. Hence, the estimation of noise parameters should be repeated ideally after any change in X-ray tube settings to ensure the highest denoising performances, but this is hardly feasible in practice.

An *a priori* characterization of noise at different X-ray tube settings could obviate the need for inferring noise statics prior to each new image sequence acquisition, by providing predetermined noise parameters values to be selected in real time, according to the current tube setting. This would always provide accurate noise parameters estimates, thus enabling the effective real-time implementation of edge-aware denoising strategies that take advantage of noise statistics to improve the image quality in fluoroscopically-guided interventional procedure. However, the reliability of such an approach should be effectively assessed for real fluoroscopic images, because the actual dependency of quantum noise statistics on X-ray tube settings has never been investigated.

The algorithm considered for noise estimation (described in Chapter 4) has already been used in previous publications about the NVCA denoising algorithm [40, 60], as well as in Chapters 4, 5 and 7, but its performances have never been assessed thoroughly, and this could affect the reliability of the conclusions that could be drawn about the feasibility of the proposed approach. Therefore, the algorithm was first tested *in silico* on several synthetic fluoroscopic sequences, which were corrupted by different levels of simulated mixed Poisson – Gaussian noise. The ability of the algorithm to retrieve the noise parameters with reasonable accuracy was assessed by varying the number and distribution of grey levels within the designed sequences, as well as the number of frames exploited for noise estimation. Indeed, considering that the algorithm infers the statistics of noise by computing the sample mean and variance of each pixel along the temporal dimension, it follows that the number of frames available for noise characterization poses a limitation on the actual number of observations of the random processes that describe each pixel luminance. This results in a certain variability of the variance values corresponding to the same mean value, which affects the accuracy of noise parameters estimates.

Afterward, the matching of noise parameters was assessed between real fluoroscopic sequences acquired by imaging two different X-ray phantoms via the same commercial fluoroscopic device with corresponding X-ray tube settings, as it would support the prospect of pre-calibrating the noise parameters at many different tube settings and using them directly in real-time denoising of new fluoroscopic sequences acquired in the same conditions.

8.2   Validation on sequences with a variable number of grey levels

The estimation of noise parameters depends on the number and distribution of the EVaR points (i.e. the expected value – variance couples) on which the linear regression is performed, i.e. on the number and distribution of grey levels within the scene. For this reason, 14 synthetic sequences were designed to represent static scenes with different number and distribution of grey levels. Each sequence was composed by 100 frames of 128×128 pixels represented on 8 bits. The grey levels were assigned to the 128 columns of the scenes in a periodic fashion, from the darkest to the lightest level

and then starting again from the darkest one. The first seven sequences, depicted in Figure 46, included 2 up to 128 grey levels in powers of 2, equally spaced in the interval [64;192], which is 128 wide and centred at the half of the whole representation interval. The further seven sequences, shown in Figure 47, included 8 grey levels, equally spaced in the intervals described in (72), which are centred at the half of the representation interval and have a decreasing width from 48 down to 16.

$$[64 + 8k;\ 192 - 8k], \qquad k = 1,2,..7 \tag{72}$$
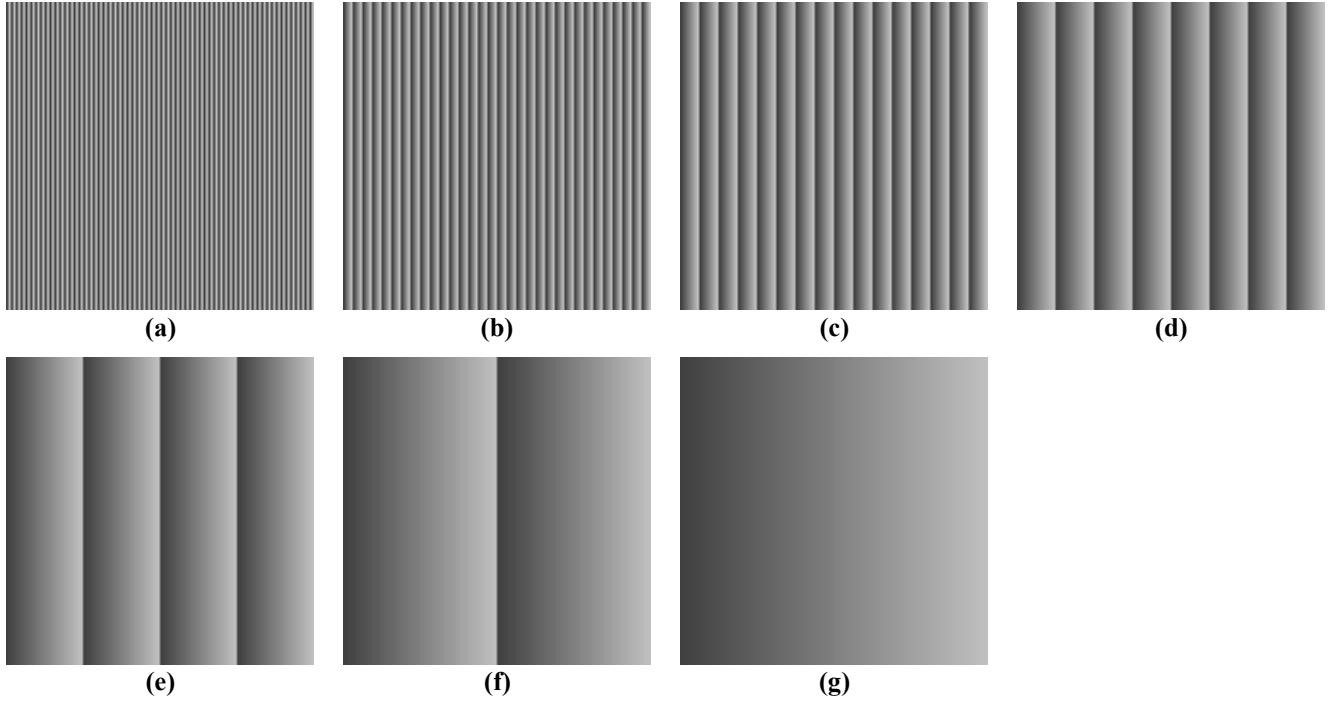
Figure 46. Static frames of the seven synthetic sequences with increasing number of grey levels (2 to 128 in power of 2) equally spaced in the range [64;192].
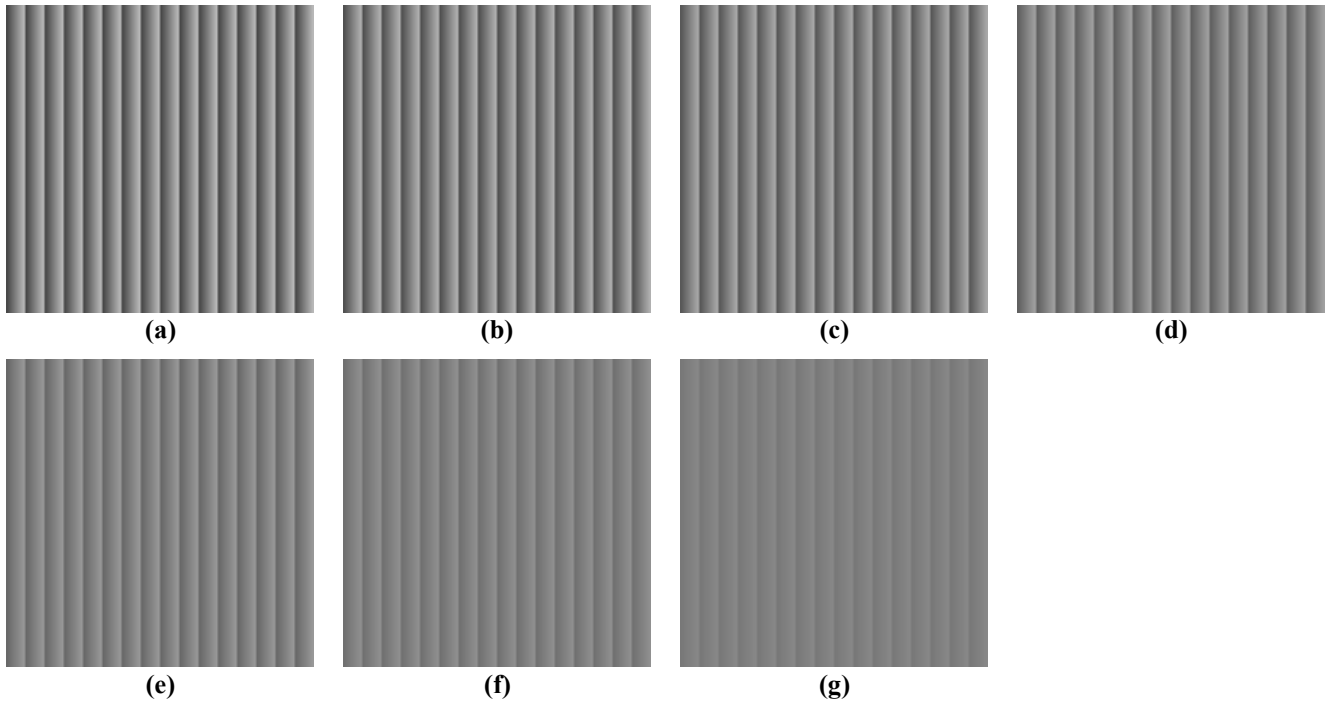
Figure 47. Static frames of the seven synthetic sequences with 8 grey levels equally spaced in narrowing ranges.

Each of the fourteen sequences with variable number of grey levels was corrupted with six different levels of simulated mixed Poisson-Gaussian noise, by using all the combinations of values considered for noise parameters (reported in Table 13). Noise estimation was performed in each of the resulting 84 noisy sequences by considering 4 different number of available frames, i.e. 10, 25, 50, 100. Therefore, a total of 336 noise estimates were actually retrieved (i.e. 56 for each noise level).

Table 13. Noise parameters of the noise levels used to corrupt the synthetic scenes

Noise level	a	b
Level 1	0.5	0
Level 2	1	0
Level 3	2	0
Level 4	0.5	144
Level 5	1	144
Level 6	2	144

In Table 14 the noise parameters estimates extracted from all the 14 synthetic sequences with variable number of grey levels are reported. For each sequence, the parameters were subdivided by the corresponding noise level and the number of frames used for noise estimation. Table 15 outlines the relative estimation errors, except for the errors related to null nominal values of parameter b (i.e. noise levels 1 to 3), which were reported as absolute errors and highlighted in blue. A substantial difference was observed in the estimation errors obtained in sequences 1-7 and sequences 8-14, which turned out to be on average consistently higher than those of the former. Mean and SD of the estimation errors are reported in Table 16.

Table 14. Noise parameters estimates extracted from the synthetic sequences with variable number of grey levels.

SEQUENCE	NOISE LEVEL	NOISE PARAMETERS ESTIMATES							
		F = 100		F = 50		F = 25		F = 10	
		a	b	a	b	a	b	a	b
# 1	Level 1	0.499	0.219	0.499	0.164	0.496	0.486	0.492	0.541
	Level 2	1.000	0.177	0.997	0.509	0.999	0.301	0.982	1.420
	Level 3	2.002	0.000	1.997	0.423	1.990	0.890	1.979	1.611
	Level 4	0.501	144.123	0.498	144.267	0.497	144.292	0.485	145.350
	Level 5	0.998	144.704	0.999	144.677	1.007	144.210	0.969	148.179
	Level 6	1.982	145.831	1.980	145.990	1.972	146.625	1.991	142.812
# 2	Level 1	0.499	0.000	0.497	0.263	0.497	0.319	0.502	0.000
	Level 2	0.996	0.644	1.000	0.296	1.004	0.017	1.005	0.000
	Level 3	1.991	1.026	1.986	1.454	1.977	1.961	1.948	5.245
	Level 4	0.499	144.388	0.492	145.639	0.483	146.561	0.484	146.777
	Level 5	1.007	143.364	1.007	142.704	1.012	142.916	0.996	145.420
	Level 6	1.972	146.909	1.966	147.741	1.956	148.271	1.951	150.376
# 3	Level 1	0.499	0.000	0.499	0.000	0.498	0.076	0.495	0.475

	Level 2	1.000	0.032	0.997	0.448	0.999	0.000	1.005	0.000
	Level 3	2.003	0.000	2.003	0.000	1.982	2.372	1.937	7.186
	Level 4	0.494	144.731	0.492	145.069	0.491	144.987	0.478	147.119
	Level 5	0.993	144.216	0.990	144.539	0.989	145.180	1.014	142.565
	Level 6	1.995	144.514	1.992	144.470	1.965	147.347	1.900	154.883
# 4	Level 1	0.500	0.000	0.501	0.000	0.499	0.183	0.498	0.290
	Level 2	0.997	0.426	0.994	0.685	0.993	0.731	0.980	2.008
	Level 3	2.003	0.000	2.001	0.000	1.996	0.427	1.967	4.308
	Level 4	0.503	143.416	0.504	143.169	0.507	142.878	0.509	142.205
	Level 5	0.999	144.480	0.985	145.829	0.998	144.244	0.991	144.283
	Level 6	2.007	143.594	1.992	145.409	1.993	145.278	1.956	149.775
# 5	Level 1	0.499	0.225	0.500	0.163	0.498	0.371	0.499	0.546
	Level 2	0.998	0.319	1.000	0.000	0.992	0.666	0.983	1.701
	Level 3	2.002	0.000	2.001	0.000	1.999	0.000	1.988	0.799
	Level 4	0.498	144.208	0.496	144.494	0.485	145.650	0.489	145.185
	Level 5	0.991	145.120	0.995	144.875	0.991	144.985	0.992	143.967
	Level 6	1.996	144.303	2.000	144.076	1.965	148.333	1.946	150.476
# 6	Level 1	0.499	0.000	0.500	0.000	0.502	0.000	0.502	0.000
	Level 2	0.999	0.245	0.998	0.114	0.989	1.001	0.998	0.493
	Level 3	1.985	2.175	1.978	3.152	1.979	3.007	1.954	5.703
	Level 4	0.500	144.032	0.493	144.862	0.504	144.187	0.475	147.657
	Level 5	1.000	144.427	1.007	143.392	1.001	143.756	1.003	143.969
	Level 6	1.967	148.324	1.962	148.507	1.948	150.805	1.939	151.995
# 7	Level 1	0.497	0.297	0.497	0.232	0.502	0.000	0.502	0.000
	Level 2	1.001	0.000	1.000	0.000	0.998	0.000	0.989	0.932
	Level 3	2.005	0.000	2.005	0.000	1.991	1.350	1.998	0.000
	Level 4	0.490	145.374	0.496	144.925	0.503	144.809	0.518	143.333
	Level 5	1.004	143.627	1.000	144.352	1.000	144.691	0.989	144.986
	Level 6	1.994	145.828	1.992	145.552	1.966	148.917	1.981	145.708
# 8	Level 1	0.499	0.092	0.496	0.427	0.493	0.958	0.488	1.766
	Level 2	0.997	0.525	1.004	0.000	1.003	0.000	1.006	0.341
	Level 3	1.984	1.572	1.980	2.232	1.974	2.874	1.956	5.122
	Level 4	0.495	144.268	0.487	145.075	0.477	147.021	0.471	148.122
	Level 5	0.988	145.929	1.004	144.260	0.997	144.697	0.998	145.961
	Level 6	1.959	149.049	1.957	149.624	1.959	149.799	1.914	154.707
# 9	Level 1	0.500	0.000	0.499	0.000	0.499	0.000	0.499	0.000
	Level 2	0.995	0.806	0.987	1.750	0.979	2.564	0.991	1.684
	Level 3	1.982	2.547	1.975	3.152	1.987	1.219	1.963	4.332
	Level 4	0.505	143.527	0.506	142.999	0.511	142.740	0.518	142.698
	Level 5	1.002	143.975	0.991	145.257	0.967	147.542	0.970	147.524
	Level 6	1.985	145.971	1.960	149.439	1.979	146.578	1.982	146.331
# 10	Level 1	0.499	0.006	0.496	0.408	0.496	0.380	0.494	0.753
	Level 2	0.996	0.263	0.996	0.243	0.991	0.840	0.976	2.465
	Level 3	1.995	0.862	1.987	1.832	1.961	5.898	1.902	13.168
	Level 4	0.488	145.676	0.500	144.077	0.519	142.098	0.518	141.881
	Level 5	0.993	145.449	0.996	144.993	0.966	148.708	0.936	152.320
	Level 6	1.997	144.294	1.981	145.680	1.957	149.669	1.818	167.430

# 11	Level 1	0.500	0.000	0.499	0.191	0.493	0.983	0.473	3.255
	Level 2	0.995	0.663	1.000	0.000	0.994	0.735	0.980	2.463
	Level 3	1.988	1.764	1.966	4.574	1.966	4.751	1.901	12.557
	Level 4	0.503	144.160	0.501	144.395	0.499	144.464	0.495	145.074
	Level 5	1.000	143.979	1.010	142.701	1.004	143.532	0.943	153.708
	Level 6	1.968	148.911	1.917	155.485	1.890	159.302	1.832	165.283
# 12	Level 1	0.500	0.000	0.499	0.088	0.498	0.369	0.501	0.000
	Level 2	1.001	0.000	1.002	0.000	1.005	0.000	0.957	5.675
	Level 3	1.979	3.193	1.956	6.100	1.944	8.350	1.805	24.896
	Level 4	0.489	145.414	0.474	146.719	0.465	147.809	0.455	149.046
	Level 5	0.974	147.411	1.002	143.844	0.988	145.695	0.875	159.413
	Level 6	1.928	153.771	1.903	157.066	1.873	161.648	1.642	190.281
# 13	Level 1	0.494	0.792	0.487	1.628	0.475	3.250	0.479	2.813
	Level 2	0.962	4.859	0.979	2.943	0.976	3.194	0.990	1.739
	Level 3	1.928	9.451	1.890	13.995	1.801	25.274	1.612	48.691
	Level 4	0.474	147.159	0.439	151.314	0.466	148.189	0.433	152.869
	Level 5	1.005	143.476	0.992	145.375	0.977	147.319	0.878	160.387
	Level 6	1.952	150.488	1.914	154.792	1.834	165.616	1.505	205.340
# 14	Level 1	0.502	0.000	0.501	0.000	0.496	0.662	0.476	3.158
	Level 2	0.901	12.725	0.886	14.400	0.832	21.182	0.635	46.330
	Level 3	1.779	28.759	1.754	31.707	1.425	74.866	0.984	131.379
	Level 4	0.491	145.234	0.509	142.511	0.460	148.487	0.267	173.398
	Level 5	0.824	166.638	0.845	163.836	0.802	169.539	0.483	210.223
	Level 6	1.652	188.843	1.540	202.970	1.183	250.552	0.622	321.903

Table 15. Errors on noise parameters estimates extracted from the synthetic sequences with variable number of grey levels. All values are expressed as relative errors, except for those reported in blue, which are expressed as absolute errors, since they refer to a null parameter (b = 0).

SEQUENCE	NOISE LEVEL	ERRORS ON NOISE PARAMETERS ESTIMATES							
		F = 100		F = 50		F = 25		F = 10	
		a (%)	b (%)	a (%)	b (%)	a (%)	b (%)	a (%)	b
# 1	Level 1	-0.162	0.219	-0.136	0.164	-0.768	0.486	-1.55	0.541
	Level 2	0.0147	0.177	-0.291	0.509	-0.0539	0.301	-1.84	1.420
	Level 3	0.0808	0	-0.170	0.423	-0.508	0.890	-1.04	1.61
	Level 4	0.280	0.0856	-0.352	0.185	-0.545	0.203	-2.91	0.937
	Level 5	-0.203	0.489	-0.0668	0.470	0.652	0.146	-3.12	2.90
	Level 6	-0.902	1.27	-1.01	1.38	-1.42	1.82	-0.466	-0.825
# 2	Level 1	-0.120	0	-0.531	0.263	-0.615	0.319	0.452	0
	Level 2	-0.371	0.644	-0.0341	0.296	0.412	0.0166	0.513	0
	Level 3	-0.462	1.03	-0.699	1.45	-1.16	1.96	-2.62	5.25
	Level 4	-0.103	0.269	-1.68	1.14	-3.33	1.78	-3.16	1.93
	Level 5	0.692	-0.441	0.652	-0.900	1.24	-0.753	-0.416	0.986
	Level 6	-1.41	2.02	-1.71	2.60	-2.22	2.97	-2.44	4.43
# 3	Level 1	-0.192	0	-0.258	0	-0.457	0.0757	-0.998	0.475

	Level 2	0.0387	0.0323	-0.317	0.448	-0.140	0	0.515	0
	Level 3	0.143	0	0.147	0	-0.897	2.37	-3.15	7.19
	Level 4	-1.19	0.507	-1.69	0.742	-1.86	0.685	-4.30	2.17
	Level 5	-0.656	0.150	-0.987	0.374	-1.15	0.819	1.43	-0.997
	Level 6	-0.251	0.357	-0.413	0.326	-1.73	2.32	-5.01	7.56
# 4	Level 1	0.0441	0	0.161	0	-0.258	0.183	-0.303	0.290
	Level 2	-0.280	0.426	-0.634	0.685	-0.734	0.731	-2.04	2.01
	Level 3	0.169	0	0.0460	0	-0.195	0.427	-1.64	4.31
	Level 4	0.665	-0.406	0.806	-0.577	1.39	-0.779	1.80	-1.25
	Level 5	-0.0787	0.333	-1.49	1.27	-0.187	0.170	-0.862	0.197
	Level 6	0.346	-0.282	-0.413	0.978	-0.335	0.888	-2.19	4.01
# 5	Level 1	-0.101	0.225	-0.0615	0.163	-0.336	0.371	-0.187	0.546
	Level 2	-0.213	0.319	-0.0221	0	-0.783	0.666	-1.73	1.70
	Level 3	0.110	0	0.0465	0	-0.0326	0	-0.614	0.799
	Level 4	-0.444	0.145	-0.762	0.343	-2.96	1.15	-2.30	0.823
	Level 5	-0.928	0.778	-0.464	0.608	-0.868	0.684	-0.839	-0.0229
	Level 6	-0.206	0.210	0.0003	0.0526	-1.74	3.01	-2.68	4.50
# 6	Level 1	-0.159	0	0.0337	0	0.417	0	0.396	0
	Level 2	-0.132	0.245	-0.187	0.114	-1.09	1.00	-0.182	0.493
	Level 3	-0.751	2.18	-1.11	3.15	-1.03	3.01	-2.30	5.70
	Level 4	-0.0030	0.0225	-1.43	0.599	0.896	0.130	-5.01	2.54
	Level 5	-0.0178	0.297	0.671	-0.422	0.0955	-0.169	0.330	-0.0215
	Level 6	-1.66	3.00	-1.88	3.13	-2.58	4.73	-3.05	5.55
# 7	Level 1	-0.506	0.297	-0.501	0.232	0.334	0	0.402	0
	Level 2	0.106	0	0.043	0	-0.225	0	-1.11	0.932
	Level 3	0.258	0	0.260	0	-0.453	1.35	-0.11	0
	Level 4	-2.04	0.954	-0.859	0.642	0.639	0.562	3.53	-0.463
	Level 5	0.393	-0.259	-0.0288	0.245	0.0427	0.480	-1.07	0.685
	Level 6	-0.315	1.27	-0.392	1.08	-1.68	3.42	-0.931	1.19
# 8	Level 1	-0.199	0.0919	-0.766	0.427	-1.40	0.958	-2.45	1.77
	Level 2	-0.329	0.525	0.362	0	0.259	0	0.588	0.341
	Level 3	-0.783	1.57	-1.02	2.23	-1.31	2.87	-2.22	5.12
	Level 4	-1.02	0.186	-2.54	0.746	-4.52	2.10	-5.85	2.86
	Level 5	-1.15	1.34	0.394	0.181	-0.304	0.484	-0.236	1.36
	Level 6	-2.04	3.51	-2.15	3.91	-2.07	4.03	-4.28	7.44
# 9	Level 1	0.0822	0	-0.141	0	-0.233	0	-0.158	0
	Level 2	-0.496	0.807	-1.32	1.750	-2.09	2.56	-0.889	1.68
	Level 3	-0.888	2.55	-1.26	3.15	-0.642	1.22	-1.87	4.33
	Level 4	1.07	-0.329	1.28	-0.695	2.17	-0.875	3.55	-0.904
	Level 5	0.156	-0.0175	-0.934	0.873	-3.30	2.46	-3.05	2.45
	Level 6	-0.735	1.37	-2.01	3.78	-1.03	1.79	-0.899	1.62
# 10	Level 1	-0.199	0.00626	-0.828	0.408	-0.712	0.380	-1.22	0.753
	Level 2	-0.355	0.263	-0.401	0.243	-0.926	0.840	-2.41	2.47
	Level 3	-0.265	0.862	-0.626	1.83	-1.97	5.90	-4.90	13.2
	Level 4	-2.32	1.16	0.0743	0.0532	3.72	-1.32	3.58	-1.47
	Level 5	-0.736	1.01	-0.375	0.689	-3.36	3.27	-6.38	5.78
	Level 6	-0.156	0.204	-0.962	1.17	-2.15	3.94	-9.10	16.3

# 11	Level 1	0.0627	0	-0.172	0.191	-1.49	0.983	-5.33	3.26
	Level 2	-0.513	0.663	-0.0014	0	-0.564	0.735	-1.98	2.46
	Level 3	-0.598	1.76	-1.68	4.57	-1.69	4.75	-4.94	12.6
	Level 4	0.672	0.111	0.223	0.275	-0.292	0.323	-0.937	0.746
	Level 5	-0.0336	-0.0143	0.985	-0.902	0.435	-0.325	-5.69	6.74
	Level 6	-1.61	3.41	-4.15	7.98	-5.51	10.6	-8.42	14.8
# 12	Level 1	-0.0121	0	-0.176	0.0885	-0.475	0.369	0.222	0
	Level 2	0.127	0	0.195	0	0.459	0	-4.29	5.68
	Level 3	-1.07	3.19	-2.20	6.10	-2.80	8.35	-9.76	24.9
	Level 4	-2.28	0.982	-5.29	1.89	-6.92	2.65	-8.94	3.50
	Level 5	-2.59	2.37	0.237	-0.109	-1.18	1.18	-12.5	10.7
	Level 6	-3.58	6.79	-4.85	9.07	-6.35	12.3	-17.9	32.1
# 13	Level 1	-1.22	0.792	-2.52	1.63	-5.08	3.25	-4.14	2.81
	Level 2	-3.82	4.86	-2.12	2.94	-2.43	3.19	-1.03	1.74
	Level 3	-3.60	9.45	-5.52	14.0	-9.94	25.3	-19.4	48.7
	Level 4	-5.14	2.19	-12.3	5.08	-6.85	2.91	-13.3	6.16
	Level 5	0.478	-0.364	-0.789	0.955	-2.31	2.31	-12.2	11.4
	Level 6	-2.42	4.51	-4.31	7.49	-8.30	15.0	-24.8	42.6
# 14	Level 1	0.362	0	0.229	0	-0.830	0.662	-4.70	3.16
	Level 2	-9.89	12.7	-11.4	14.4	-16.8	21.2	-36.5	46.3
	Level 3	-11.1	28.8	-12.3	31.7	-28.7	74.9	-50.8	131
	Level 4	-1.76	0.857	1.76	-1.03	-8.00	3.12	-46.6	20.4
	Level 5	-17.6	15.7	-15.5	13.8	-19.8	17.7	-51.7	456
	Level 6	-17.4	31.1	-23.0	41.0	-40.9	74.0	-68.9	124

Table 16. Mean and standard deviation of the noise parameters estimation errors. The statistics are computed for sequences 1-7 and 8-14 separately, and for each considered number of available frames (F).

PARAMETER ERROR	SEQUENCES	ERRORS ON NOISE PARAMETERS ESTIMATES							
		F = 100		F = 50		F = 25		F = 10	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
a_{err_p}	1-7	-0.0025	0.0056	-0.0042	0.0065	-0.0062	0.0105	-0.0126	0.0176
	8-14	-0.0226	0.0422	-0.0280	0.0503	-0.0467	0.0825	-0.1078	0.1643
b_{err_p}	1-7	0.0051	0.0083	0.0068	0.0094	0.0115	0.0144	0.0175	0.0235
	8-14	0.0362	0.0724	0.0458	0.0920	0.0751	0.1610	0.1686	0.2793
$b_{err_{abs}}$	1-7	0.2755	0.5082	0.3763	0.7232	0.6742	0.8471	1.5837	2.1418
	8-14	3.2800	6.7123	4.0795	7.5838	7.5403	16.832	14.885	30.1688

Considering that the actual concern of denoising are not the mere errors on noise parameters, but rather the reconstruction errors on the processed images, a qualitative and quantitative performance assessment of the noise estimation algorithm was carried out by comparing the denoising results achieved via the NVCA algorithm with the actual and the estimated noise parameters. To this aim, two synthetic noiseless test sequences were designed, which represented a dark rectangle moving from the left to the right at a speed of 1 pixel per frame over a brighter, uniform background. The

sequences were corrupted with two different levels of mixed Poisson-Gaussian noise, corresponding to level 3 and level 6 reported in Table 13. The contrast in the two noiseless sequences was set to obtain a CNR of 4 for both sequences. In practice, the same grey level was assigned to the moving rectangle in both sequences, while two different grey levels were assigned to their backgrounds. Figure 48 shows the final frame of the two noiseless sequences, their noisy versions and the results obtained via NVCA filtering by using the actual noise parameters.
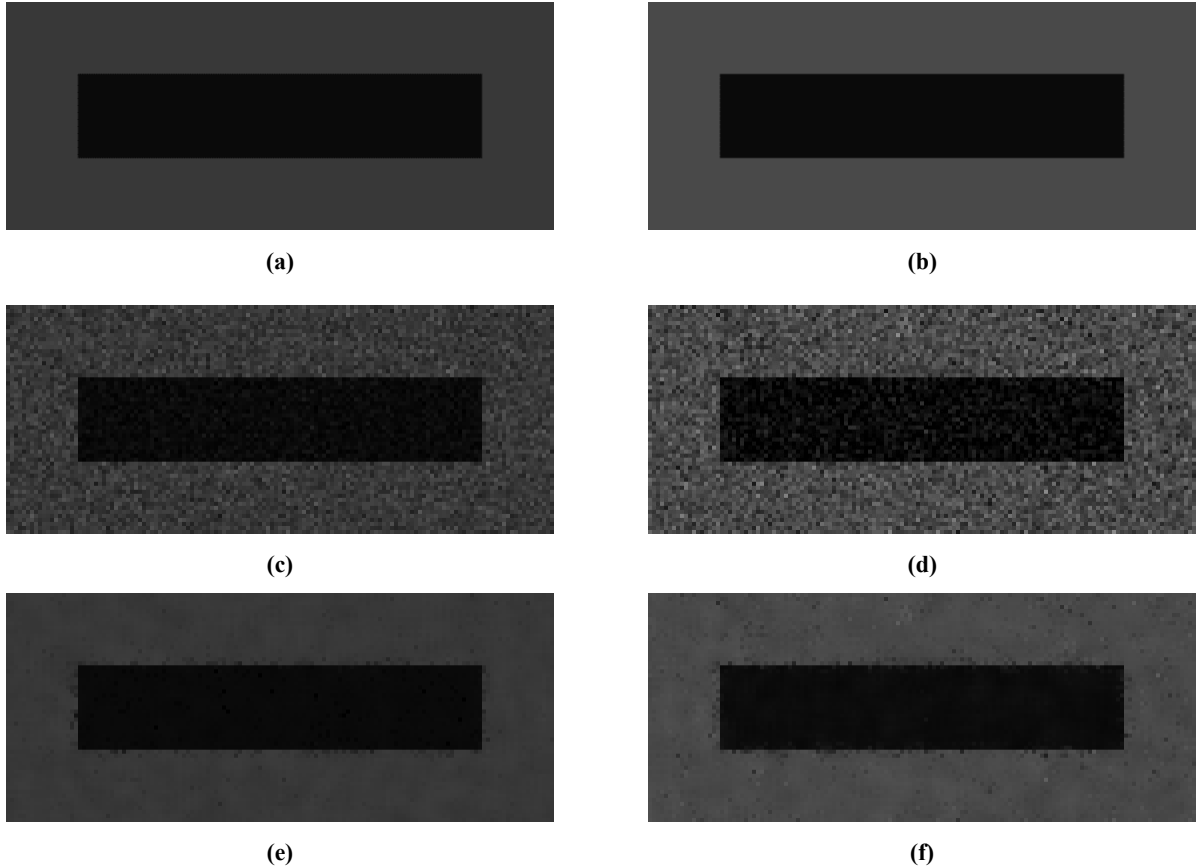
Figure 48. Synthetic sequences with a moving rectangle, adopted to test the effect of the noise estimation errors on NVCA filtering performances. In the first row, panels (a)-(b), the noiseless sequences are reported. The contrast was set in order to obtain a CNR = 4 in both related noisy sequences, which are depicted in the second row, panels (c)-(d). In the last row, panels (e)-(f), the sequences denoised via the NVCA algorithm by using the actual noise parameters are shown.

Afterward, the noisy synthetic test sequences were filtered via the NVCA algorithm, by using the most inaccurate noise parameters estimates, so as to identify the worst cases from the denoising point of view. Then, the worst results obtained for estimates extracted by using 25 and 10 frames, from sequences 1-7 and 8-14, were identified according to measures of dissimilarity between the sequences filtered with inaccurate noise parameters, referred to as the sub-optimal filtered sequences, and the sequence filtered with the actual noise parameters, referred to as the optimal filtered sequence. Two well-established image quality assessment indices were adopted, namely the Mean Squared Error (MSE), which is a global measure of dissimilarity between images, and the Full Width at Half Maximum (FWHM) of the edge spread function, which is a no-reference local measure of edge sharpness. As a first dissimilarity measure to quantify the global deviation from the optimal denoising result, the MSE between the sub-optimal and the optimal filtered sequences was computed. However, MSE is known to have high sensitivity to the overall image noise, but poor sensitivity to edge blurring

effects, especially in noisy conditions like those encountered in low-dose fluoroscopy, as already shown in Chapter 5. Since the edge-awareness is a major concern of medical image denoising, the local loss of edge sharpness due to the noise parameters estimation errors was considered as a further measure of dissimilarity, and was evaluated by estimating the Δ FWHM, that is the difference in FWHM between the sub-optimal and the optimal filtered sequences. The quantitative results of this analysis are summarized in Table 17, where it could be noticed that the worst results were always obtained in the sequences 8-14.

Table 17. Results of the denoising performance analysis. The dissimilarity scores between the sub-optimal and optimal filtered sequences are reported, along with the corresponding noise parameters estimates, the number of frames (F) used in the estimation, and the actual noise parameters values.

SEQUENCES	a	b	Frames	a_{est}	b_{est}	$MSE(S_{filt}, S_{filt_{err}})$	Δ FWHM
1 - 7	2	0	25	1.979	3.007	0.090284	0.001851
	2	0	10	1.937	7.186	0.199485	0.002836
8 - 14	2	0	25	1.425	74.866	2.276551	0.041098
	2	0	10	0.984	131.379	5.324343	0.256335

Figure 49 depicts the sub-optimal filtered sequences (end frame of each sequence), as well as the corresponding image differences with the optimal filtered sequence, where it can be observed that the pixels with the highest differences in luminance are almost all distributed in the edges neighbourhood, which is consistent with the measured increase in Δ FWHM. However, it can be assessed by visual inspection that the sub-optimal results shown in Figure 49.(a-c) and the optimal result shown in Figure 48.(e) are very similar.

The noise parameters estimation algorithm achieved very low estimation errors in the first seven sequences, while performing substantially worse on the last seven sequences. Indeed, it is worth noting that, even with only 10 frames (i.e. 10 distinct realizations of the stochastic processes representing the random distribution of the grey level of each pixel along the temporal dimension),the average percentage error in the first 7 sequences was lower than 2 %, whereas the average error in the last 7 sequences turned out to exceed these value, even for estimations performed by using the maximum number of available frames. Furthermore, it could be assessed by visual inspection that, even the worst errors achieved in the first 7 sequences, produced sub-optimal filtered sequences which were very similar to the optimal one, as opposed to the sub-optimal sequences produced by the estimates from the last 7 sequences, which clearly showed edge blurring (confirmed by the increases in Δ FWHM). These results clarify that ensuring a reasonable contrast in the test sequences to be used for noise characterization is mandatory to achieve reliable estimates. Moreover, the very small number of frames required by the noise estimation algorithm to achieve a reasonable accuracy allows for its application also to very short static scenes.
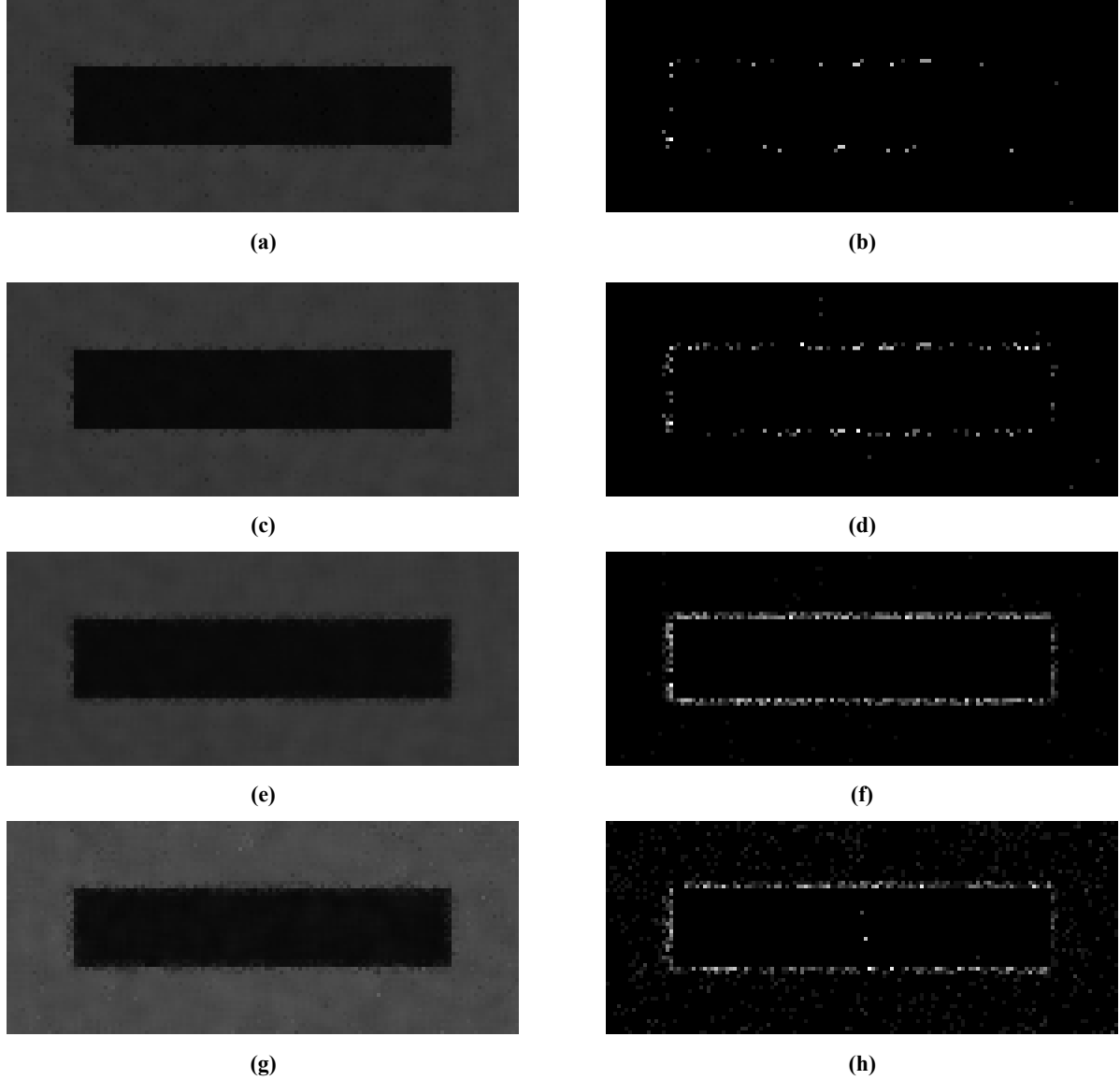
Figure 49. Synthetic sequences with a moving rectangle filtered via the NVCA algorithm by considering the noise parameters estimates reported in Table 17. The images in each row were obtained by using the noise parameters in the corresponding row of Table 17. On the first column the end frames of the filtered sequences were depicted, while the differences of the same images with the end frame of the sequence filtered with the actual noise parameters were reported on the second column.

8.3 Validation on sequences from an X-ray simulator

The fourteen synthetic sequences described in the previous paragraph are characterized by scenes that are very uncommon in medical applications, therefore four additional sequences were produced via an X-ray simulator [155 – 157] and depicted in Figure 50. This allowed testing the noise estimation algorithm on scenes with content of clinical relevance, while still having a ground truth to derive quantitative measures for performance assessment. The five sequences devised via the X-ray simulator were corrupted with the same noise levels reported in Table 13. An example of noiseless and noisy versions of a sequence is depicted in Figure 51). Noise estimation was performed by using 25 frames, which turned out to be the minimum number of frames to retrieve noise parameters with

a reasonable accuracy, according to the results obtained in the validation on the 14 sequences with a variable number of grey levels.

<p style="text-align:center">(a)</p>

<p style="text-align:center">(b)</p>

<p style="text-align:center">(c)</p>

<p style="text-align:center">(d)</p>
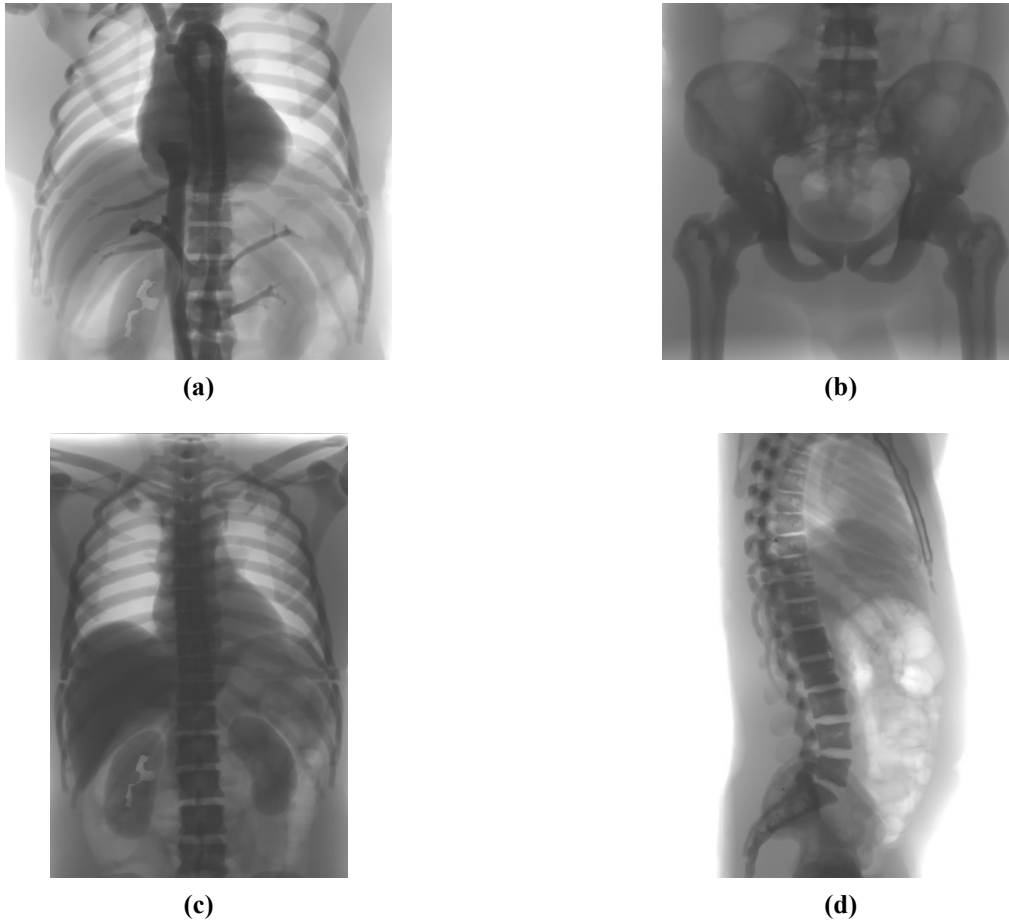
Figure 50. Static frames of the four synthetic sequences devised via the X-ray simulator.

The results of the noise parameters extraction from the four synthetic sequences designed via the X-ray simulator were reported in Table 18. The mean and SD of relative errors, outlined in Table 19 turned out to be almost comparable with those obtained in the 14 synthetic sequences, thus proving that the presence of clinically relevant structures does not alter the estimates accuracy, which, more generally, is not influenced by the particular informational content of the scene. In light of the results presented so far, the algorithm under test proved reliable in extracting noise parameters estimates with high accuracy, even from very short static scenes of only 10 frames, and can be used to assess the feasibility of the a priori noise characterization approach.

8.4   A priori noise characterization on real fluoroscopic sequences

Real fluoroscopic sequences were acquired by imaging two commercial X-ray phantoms, namely TOR-18FG [158] and TOR-CDR [159] (Leeds Test Objects, 7 Becklands Cl, Roecliffe, York YO51 9NR, UK), via a commercial fluoroscopic device (INTERMEDICAL S.r.l. IMD Group, Via E. Fermi 26, 24050 Grassobbio (BG), Italy). The fluoroscope acquired frames of 1536×1536 pixels, represented on 16 bits, with a pulsed protocol at 15 fps. Each phantom was placed over an anti-scatter grid, just above the flat panel detector, between two blocks of five plexiglass square sheets of 25 cm

x 25 cm x 1 cm (see Figure 52), which were used to produce an equivalent Compton scattering noise that would occur when imaging the human body. Five sequences were acquired for each phantom by using the X-ray tube settings reported in Table 20. The noise estimation algorithm was applied to extract the noise parameters estimates from the ten acquired fluoroscopic sequences. Then, a comparison was carried out between parameters extracted from each couple of sequences acquired with the same X-ray tube settings.

(a)
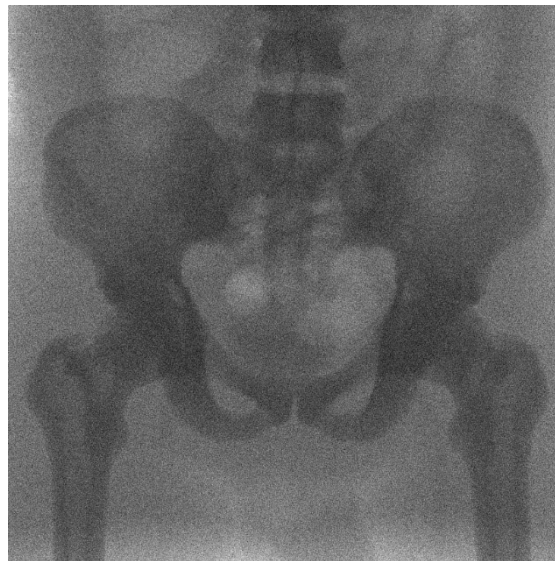
(b)

Figure 51. Comparison of noiseless and noisy versions (noise level 6) of the same static frame of sequence #2 from X-ray simulator, depicted in Figure 50.(b).

Table 18. Noise parameters estimates and related errors extracted from the synthetic sequences designed via the X-ray simulator. All errors are expressed as relative errors, except for those reported in blue, which are expressed as absolute errors, since they are referred to a null parameter (b = 0).

SEQUENCE	NOISE LEVEL	NOISE PARAMETERS ESTIMATES		ERRORS ON NOISE PARAMETERS ESTIMATES	
		F = 25			
		a	b	a	b
# 1	Level 1	0.500	0.060	-0.000727048	0.059883648
	Level 2	0.997	0.471	-0.003291033	0.470661055
	Level 3	1.975	3.426	-0.012272794	3.425821807
	Level 4	0.498	144.398	-0.004674595	0.002767017
	Level 5	0.986	145.765	-0.013538244	0.012257184
	Level 6	1.925	153.412	-0.037355982	0.065361524
# 2	Level 1	0.495	0.475	-0.010395139	0.475398655
	Level 2	0.992	0.895	-0.007838176	0.894760806
	Level 3	1.983	1.892	-0.008364067	1.892067765
	Level 4	0.486	145.463	-0.027658002	0.010161948
	Level 5	0.978	146.466	-0.022149186	0.017126109
	Level 6	1.958	148.589	-0.021242895	0.031867285
# 3	Level 1	0.497	0.321	-0.00529215	0.320866941
	Level 2	1.000	0.145	-0.000132407	0.145309017
	Level 3	1.964	4.633	-0.017766988	4.632563092
	Level 4	0.496	144.592	-0.008127859	0.004107655
	Level 5	0.984	146.179	-0.015820705	0.015130572
	Level 6	1.894	156.968	-0.053154461	0.09005548
# 4	Level 1	0.497	0.217	-0.005063311	0.216921236
	Level 2	0.996	0.480	-0.003516649	0.479907074
	Level 3	1.990	1.128	-0.004842556	1.127681369
	Level 4	0.492	144.824	-0.016299257	0.005725547
	Level 5	0.987	145.610	-0.013160495	0.0111835
	Level 6	1.968	147.608	-0.01615601	0.025058435

Table 19. Mean and standard deviation of the errors on noise parameters estimated in the sequences designed via the X-ray simulator by using 25 frames.

PARAMETER ERROR	ERRORS ON NOISE PARAMETERS ESTIMATES	
	F = 25	
	Mean	SD
a_{err_p}	-0.0135	0.0114
b_{err_p}	0.0226	0.0244
$b_{err_{abs}}$	1.0811	1.3120

X-ray tube setting	Kilovoltage peak (kV)	Current (mA)
#1	40	10
# 2	40	20
# 3	40	30
# 4	40	40
# 5	40	50

Figure 53 shows four frames of the real fluoroscopic sequences: in particular, the frames in the left column depict the TOR 18FG phantom, while the ones in the right column refer to the TOR CDR phantom. The frames in the first row were acquired with X-ray tube setting #5 (40 kVp, 50 mA), while those in the second row with setting #1 (40 kVp, 10 mA).

(a)

(b)

(c)

(d)
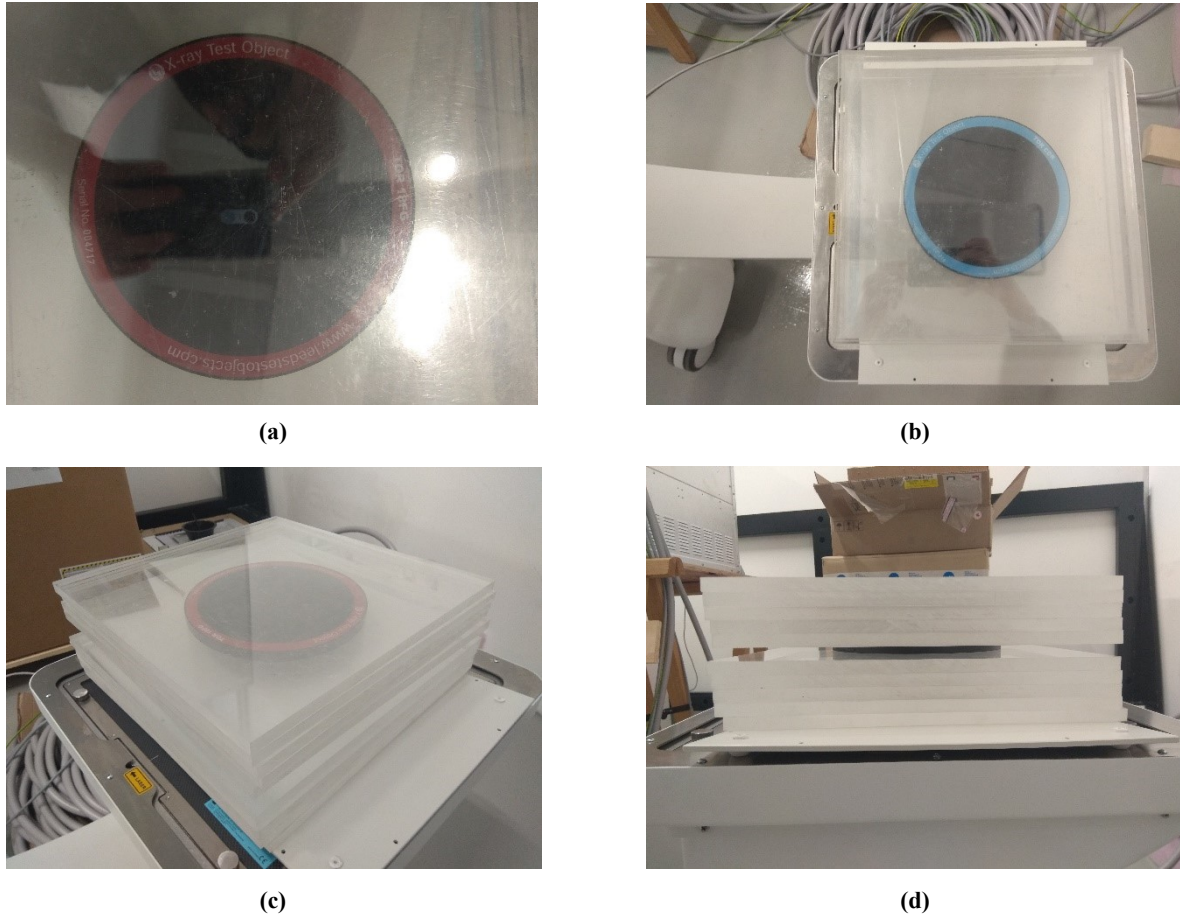
Figure 52. Pictures of the X-ray phantoms with plexiglass sheets: (a) top view of TOR-18FG ; (b) top view of TOR-CDR; (c) side view of TOR-18FG; (d) side view of TOR-CDR.

Due to the very low tube currents involved, the original images turned out to be too dark for practical visualization, as indeed the luminance values were confined within very narrow ranges in the lower part of the representation interval. For this reason, the images in Figure 53 have been processed with a full-scale histogram stretch, disregarding the grey levels of the few lightest pixels in the leftmost part of the images. This processing obviously altered the mean luminance, which was

originally much lower in the images acquired at 10 mA compared to those acquired at 50 mA, but made the noise much more visible, allowing easier comprehension of the effect of X-ray tube current reduction on the SNR of the images.

(a)　　　　　　　　　　　　　　　　　(b)

(c)　　　　　　　　　　　　　　　　　(d)

Figure 53. Static frames from the real fluoroscopic sequences. The frames shown in the first and second rows were acquired at 50 mA and 10 mA (40 kVp), respectively. The frames in the left column depict the TOR 18FG phantom, while those in the right column depict the TOR CDR.

The noise parameters extracted from the real fluoroscopic sequences are reported in Table 21, along with the relative errors of the parameters retrieved from TOR CDR sequences with respect to those extracted from the TOR 18FG sequences, for corresponding X-ray tube settings. The relative errors (mean and SD) were -0.36% $\pm$ 0.90 % and 0.30 % $\pm$ 2.8 %, for parameters a and b, respectively, and turned out to be comparable to those obtained in the analyses of the performances of the noise parameters estimation algorithm. The noise parameters extracted from the two phantoms sequences

are also plotted in Figure 54, where it can be verified that their trends with the tube current are very similar.

Table 21. Noise parameters estimates retrieved from the real fluoroscopic sequences, with relative errors on single parameters extracted from TOR CDR sequences with respect to TOR 18FG ones for each tube setting.

kVp	mA	TOR 18FG		TOR CDR		ERRORS	
		a	b	a	b	a_{err}	b_{err}
40	10	7.15091	123.033	7.20909	125.312	0.00814	0.01852
40	20	4.37212	401.807	4.32512	384.611	-0.01075	-0.04280
40	30	3.45425	522.778	3.45683	526.265	0.00075	0.00667
40	40	2.99311	615.161	2.95042	633.000	-0.01426	0.02900
40	50	2.66481	702.262	2.65917	704.725	-0.00212	0.00351

To further demonstrate the viability of the proposed approach and, at the same time, also the effectiveness of the NVCA denoising in very noisy conditions, the fluoroscopic sequence obtained by imaging the TOR18FG phantom at 40 kVp, 10 mA (i.e. the one with the most degraded quality) was filtered via NVCA, by using the same mask size and noise SD threshold multiplier (5x5x7, N_σ = 1.5) and by considering both the noise parameters extracted from the same sequence and the ones extracted from the sequence acquired by imaging the TOR CDR phantom with the same X-ray tube settings. The differences between the filtered sequences were quantitatively assessed by computing their MSE, which turned out to be equal to 1.7×10^{-4} and was mainly due to less than 500 unmatching pixels (lower than 5 ppm of the whole sequence). Undoubtedly, the MSE is extremely small and practically negligible, thus supporting the feasibility of the a priori noise characterization (recall that MSE as high as 0.2, reported in Table 17, resulted in imperceptible differences from a qualitative visual point of view, as in Figure 49.c-d). The raw and filtered sequences are shown in Figure 55.

It is the first time that the trends of Poisson – Gaussian noise parameters are investigated as a function of the X-ray tube settings. It can be noted that parameters a and b show increasing and decreasing trends with tube current, respectively. While the reduction of parameter a with tube current is understandable, as for increasing tube current (i.e. increasing luminance) it is reasonable that the detector gain is reduced to avoid clipping effects, the increasing trend of b is not clear. Hypotheses have been made on its cause but none of them seems to provide a proper explanation for the observed trend. One possible source could be the electronic noise from the detector and the readout circuitry. However, there is no clue on why the circuitry noise (referred to the circuit after the gain stage of each detector) should have a dependency on the mean luminance. It could be possible for the detector noise, which arises before the gain stage and is amplified by the detector gain.

However, the detector gain showed a decreasing trend with the tube current, so the equivalent AWGN noise contribution to the final signal should have been reduced by the detector gain for increasing tube currents. Considering the presence of the square plexiglass plates that increase Compton scattering, this phenomenon was ascribed to the increased number of Compton photons. However, Compton scattering contribute to the image formation with fluxes of scattered X-ray quanta, whose arrival on the detector is still described by a Poisson process, hence it could not have contributed to the AWGN component of the noise, i.e. the signal-independent variance contribution.
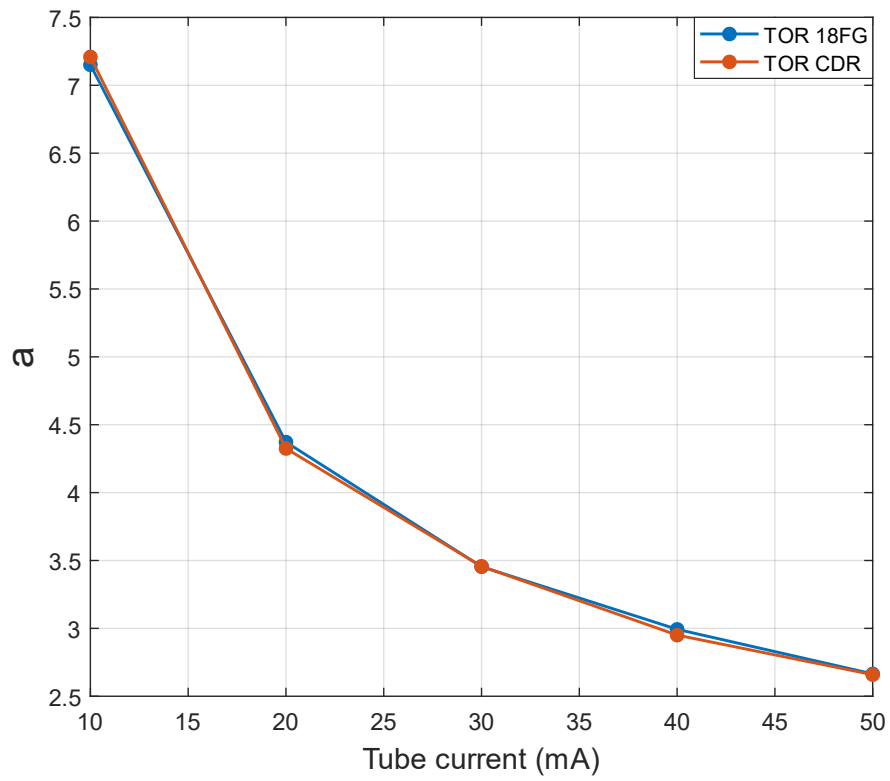
(a)
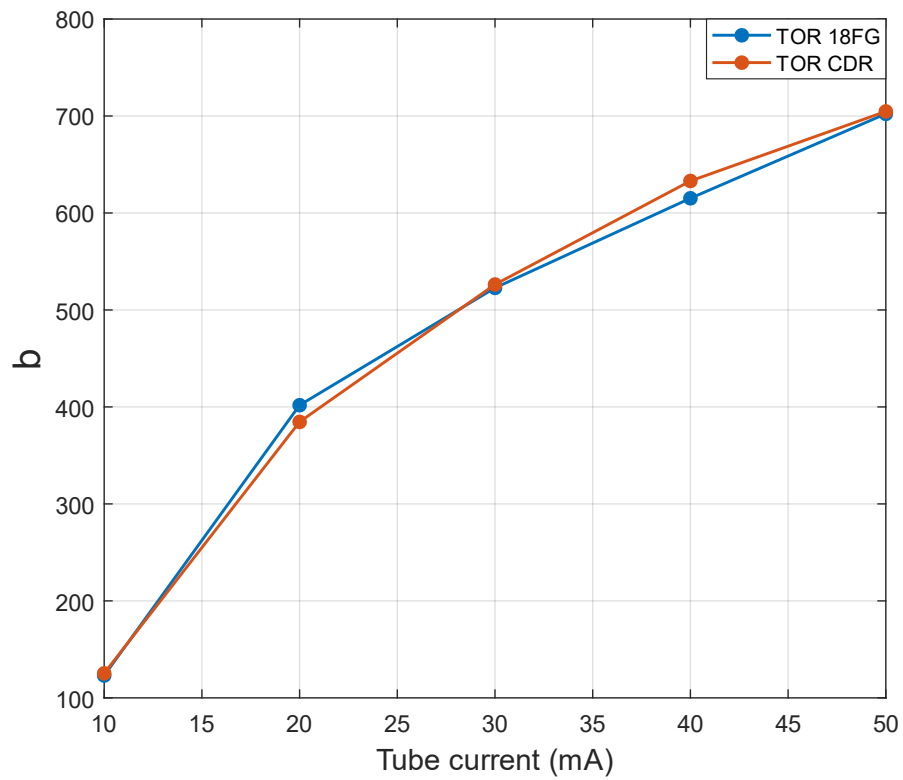
(b)

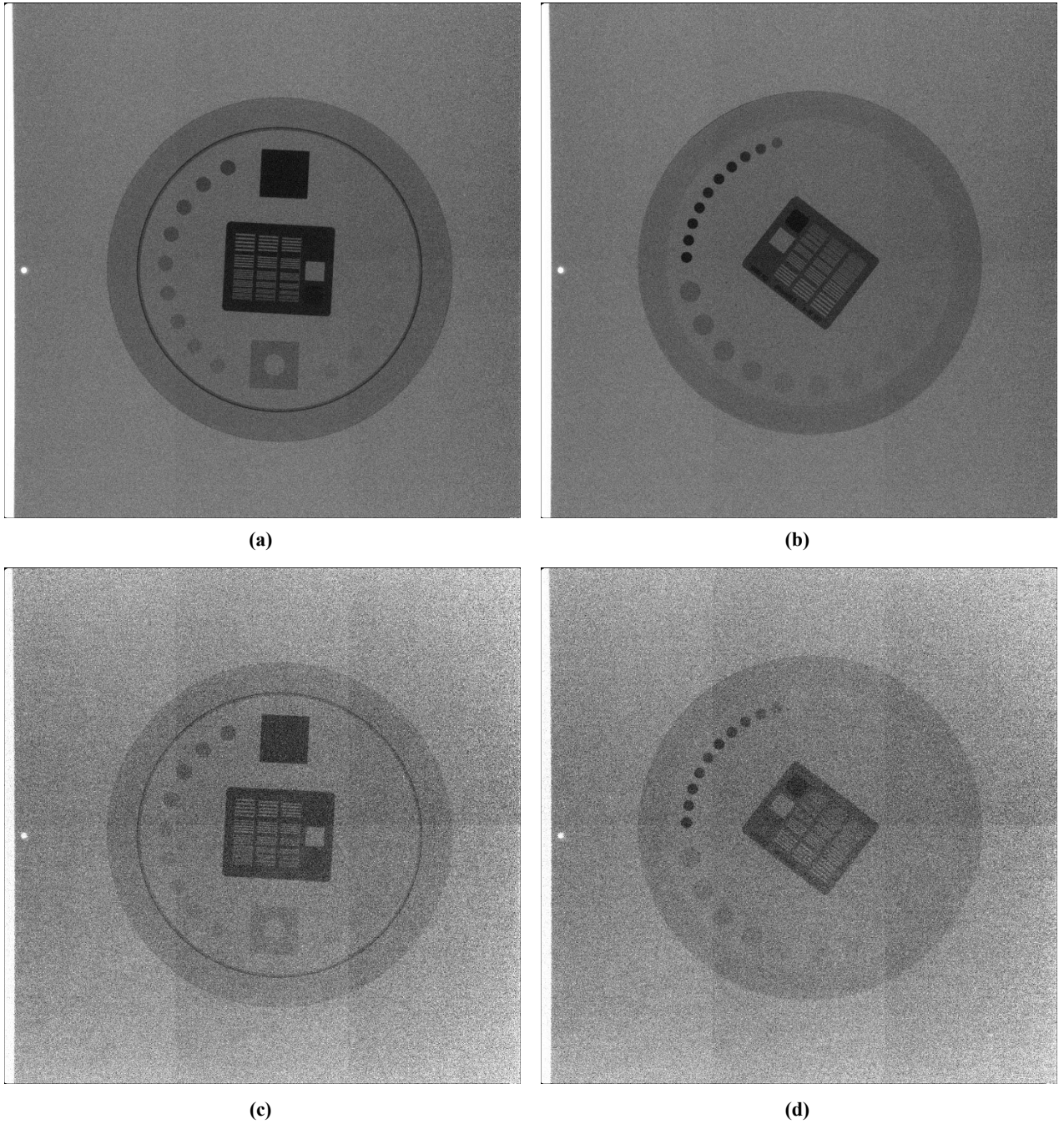*Figure 54. Noise parameters estimated from real fluoroscopic sequences. Static frames from the real fluoroscopic sequences. The frames shown in the first and second rows were acquired at 50 mA and 10 mA (40 kVp), respectively. The frames in the left column depict the TOR 18FG phantom, while those in the right column depict the TOR CDR.*
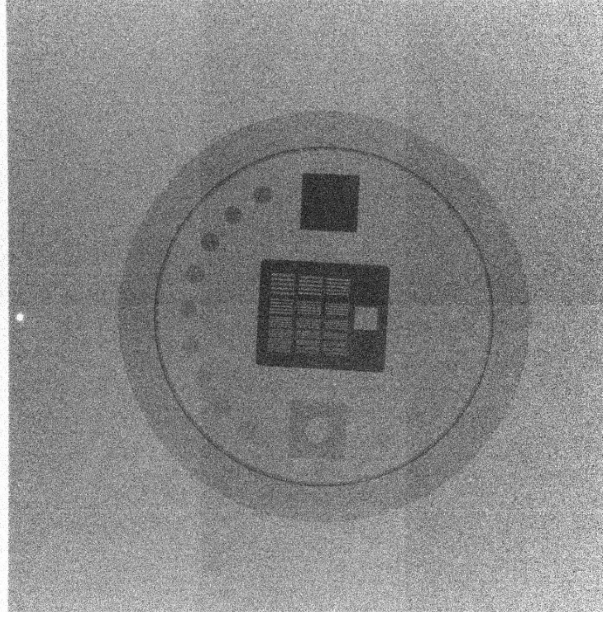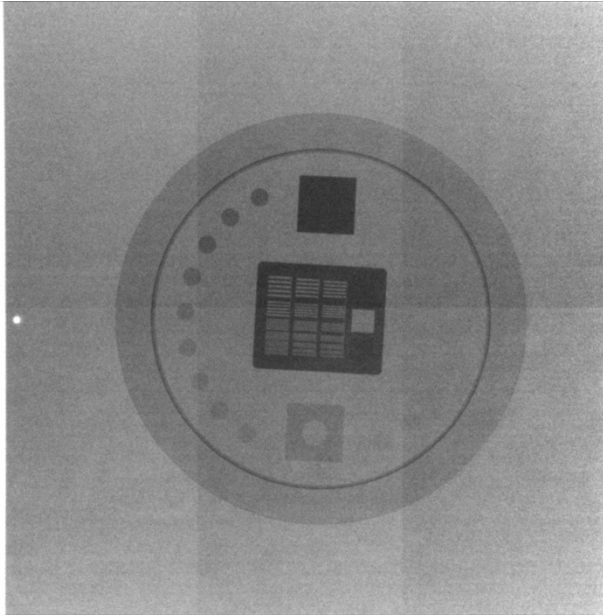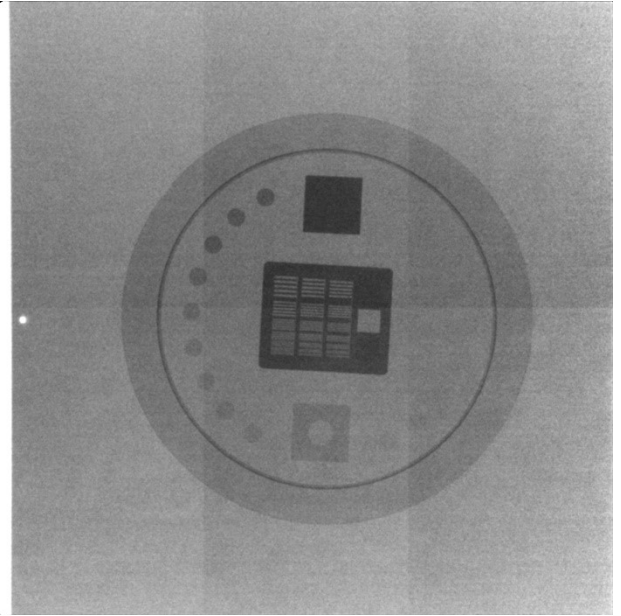
Figure 55. (a) Raw fluoroscopic sequence acquired by imaging the TOR 18FG phantom at 40 kVp, 10 mA; Filtered sequences obtained via NVCA with 5×5×7 mask and $N_\sigma = 1.5$, by using the noise parameters extracted from: (b) the raw TOR 18FG phantom sequence depicted in (a); (c) the raw TOR CDR phantom sequence acquired at 40 kVp, 10 mA. The two sequences are practically identical, as their MSE is lower than 2×10^{-4} and the number of unmatching pixels is lower than 5 parts per million.

In conclusion, the present analysis first ascertained the reliability of the estimation algorithm described in Chapter 4 in retrieving the noise parameters with high accuracy, as it was critical to detect even subtle differences that could potentially show up between noise parameters extracted from distinct sequences. Then, the algorithm was used to verify the hypothesis that noise parameters in the same fluoroscopic device mainly depend on the particular choice of X-ray tube settings, which would support the feasibility of the proposed *a priori* noise characterization approach. To this aim, the match between noise parameters extracted from pairs of real fluoroscopic sequences, acquired independently by imaging two different X-ray phantoms via the same fluoroscopic device, with the same X-ray tube settings, was assessed. The encouraging results of this analysis suggest, for the first time in literature, that an *a priori* characterization of noise for a single fluoroscopic device is feasible and could support the actual real-time implementation of edge-aware denoising strategies that take advantage of noise statistics to improve the trade-off between noise reduction and details preservation.

Conclusions

In this dissertation, different issues have been addressed, which currently turn out to limit the actual use of very low-dose X-ray fluoroscopy imaging in clinical practice. Indeed, the main challenge of very low-dose protocols is represented by the significant degradation of details visibility due to quantum noise, which could be dangerous for patient safety. Being inherent to the image formation process, quantum noise is a physical limit and cannot be addressed by improving sensors technology, rather demanding efficient real-time denoising strategies to improve image quality in image-guided interventional procedures. This dissertation focused on the Noise Variance Conditioned Average algorithm, which had already demonstrated to outperform many state-of-the-art methods in the denoising of images corrupted by quantum noise, while also keeping a reasonable computational burden that allows for its real-time hardware implementation.

In Chapter 5, NVCA was compared against VBM4D, which is widely recognized as the current state-of-the-art denoising method for AWGN. VBM4D was adapted to Poisson – Gaussian noise via the generalised Anscombe transform. While showing higher performances in static scenes, VBM4D was outperformed by NVCA in moving scenes, especially with real fluoroscopic noise. Indeed, VBM4D provided significant blurring and cartoon-like effects on the edges of low-contrast noisy areas, also degrading the visibility of small objects (e.g. catheters) as well as the quality of object contours in image derivatives. NVCA, instead, attained decidedly better edge preservation, especially for low-contrast moving objects, while providing a slightly lower noise reduction. Moreover, a lower bound for the computational complexity of VBM4D was found to be several orders of magnitude higher than the complexity of NVCA, thus confirming the practical unsuitability of VBM4D for real-time implementations. The results of these analyses also highlighted two further issues, namely the reduction of the improvements in image quality with the reduction of the quality of the input noisy images, as well as the unreliability of both classical and recent IQA indices (e.g. MSE, PSNR, CNR, FSIM) in predicting the image quality of X-ray images as perceived by the human eye.

The first issue was investigated in Chapter 3 via a theoretical statistics approach. In particular, the analysis focused on the statistical discrimination criterion at the core of the NVCA operation, which is aimed at classifying pixels as belonging or not to the same object, in order to avoid blurring effects by selectively performing an average operation only on similar pixels. Analytical expressions of true positive and false positive rates of such a classification were derived as a function of both the input CNR, which is a measure of the input image quality, and the noise SD threshold coefficient of NVCA (N_σ), which regulates the trade-off between noise reduction and edge preservation. The analysis showed that FP grows quite rapidly for CNR decreasing below 1.5, and that it can be reduced by lowering N_σ, which however reduces the TP as well (i.e. the noise reduction ratio), thus confirming the observed reduction in denoising power. An approximate relationship between N_σ and CNR was also found, which could be used to devise a more complex adaptation scheme for the local threshold, based on an automatic selection of N_σ, in order to ensure a desired FP and improve the preservation of very low-contrast edges without preventing NVCA from attaining a deeper noise reduction on

edges featuring higher contrasts. However, this would demand also a robust automated estimation of local CNR, which is not straightforward and requires further studies.

In Chapter 6, a novel edge-aware IQA index was proposed to address the unsuitability of common IQA indices for image quality prediction in medical X-ray images. The novel SED index was assessed against FSIM and CNR and proved capable of suggesting the optimal values for the tuning parameters of NVCA, which yielded the best image quality also according to visual perception. This IQA index could be used to accomplish more objective comparisons between different strategies for quantum noise suppression, as well as to carry out more in-depth investigations of the dependency of the optimal NVCA parameters on different features, such as CNR, shape, dimension and speed of moving objects within the scenes imaged via X-ray fluoroscopy.

Chapter 7 presented an improved version of the NVCA algorithm, which is based on the separation of the filtering operation in the space and time domains, along with its hardware implementation, which attained a real-time operation for sequences of 1024×1024 pixels frames at 49 fps. The cascaded temporal and spatial filtering granted an improved capability of recognition and restoration of low-contrast moving edges, as well as a substantial reduction of hardware complexity, which allowed implementing the improved NVCA algorithm on the smallest FPGA platform of the Altera StratixIV family, by using at most the 22 % of its hardware resources. These improvements support the actual implementation of the NVCA algorithm within a real fluoroscopic device, which would provide a real-time improvement of fluoroscopic image quality and is expected to foster the use of very low-dose protocols in fluoroscopy-guided interventional procedures. However, besides the selection of the optimal parameters, which is still to be solved, another issue should be addressed to enable the effective real-time denoising of fluoroscopic sequences, namely the estimation of time-varying noise parameters.

Chapter 8 addressed this issue by presenting an innovative approach based on the *a priori* characterisation of noise parameters at many different X-ray tube settings, to obviate the need for inferring noise statics after any change of tube settings, and to be able to switch among different predetermined parameters in real time, according to the particular choice of the X-ray tube settings. The feasibility of such an approach was demonstrated by first assessing the performances of a noise parameters estimation algorithm (already used in other studies focused on NVCA) and then verifying the match between noise parameters extracted from pairs of real fluoroscopic sequences, acquired independently by imaging two different X-ray phantoms via the same commercial fluoroscopic device, with the same X-ray tube settings. In addition to confirming the feasibility of the a priori noise characterization approach, the results also lay the foundations to design calibration procedures for fluoroscopic devices that embed statistics-based edge-aware strategies for the real-time denoising of fluoroscopic sequences.

Further studies are foreseen to carry out a more extensive noise characterization on a single device, e.g. over an expanded grid of X-ray tube settings (mA, kVp), to demonstrate the robustness of the proposed approach. Furthermore, looking upon the considerable number of noise estimations that must be performed to cover the actual grid of commonly used settings, it is reasonable that complete noise characterization procedures would be very long. In order to shorten the time required by these procedures, an approach could be investigated, which is based on the interpolation of noise parameters estimates corresponding to certain points of the tube settings grid (mA, kVp), to determine the noise parameters for different tube settings, without performing the explicit noise estimation.

Indeed, this approach would grant the possibility of performing the actual noise estimation for a small number of X-ray tube settings and then retrieving a conspicuous number of noise parameters estimates via interpolation, thus speeding up the overall calibration procedure. Moreover, the performances of the proposed approach must be effectively assessed in real-time denoising of low-dose fluoroscopic sequences, e.g. by embedding the *a priori* noise characterization in the hardware implementation of the NVCA algorithm presented in Chapter 7. To this aim, another interesting study would be the investigation of the potential implementation of NVCA on a System-on-Chip (SoC) platform [160], which combines the computing power of custom hardware accelerators with the flexibility of dedicated software run by embedded microprocessors. Very appealing is also the possibility to exploit the High-Level Synthesis (HLS) tools for SoC design [161], which allow hardware descriptions to be synthesized from software descriptions (e.g. in C language) and could potentially enable faster and more flexible design flows to bring improvements at algorithmic level directly into hardware. Finally, the overall improved NVCA denoising scheme should be embedded in a commercial fluoroscopic device to test its performances in its real operating environment.

References

1. Moradi, M., Mahdavi, S.S., Dehghan, E., Lobo, J.R., Deshmukh, S., Morris, W.J., et al. (2012). Seed localization in ultrasound and registration to C-Arm fluoroscopy using matched needle tracks for prostate brachytherapy. *IEEE Trans Biomed Eng*, *59*, 2558–2567 https://doi.org/10.1109/TBME.2012.2206808

2. Weese, J., Penney, G.P., Desmedt, P., Buzug, T.M., Hill, D.L.G., Hawkes, D.J. (1997). Voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for image-guided surgery. *IEEE Trans Inf Technol Biomed. 1*, 284–293. https://doi.org/10.1109/4233.681173

3. Bifulco, P., Cesarelli, M., Cerciello, T., Romano, M. (2012). A continuous description of intervertebral motion by means of spline interpolation of kinematic data extracted by video fluoroscopy. *J Biomech, 45*, 634–641 https://doi.org/10.1016/j.jbiomech.2011.12.022

4. Cerciello, T., Romano, M., Bifulco, P., Cesarelli, M., Allen, R. (2011). Advanced template matching method for estimation of intervertebral kinematics of lumbar spine. *Med. Eng. Phys.*, *33(10)*, 1293-1302. https://doi.org/10.1016/j.medengphy.2011.06.009

5. Yamazaki, T., Watanabe, T., Nakajima, Y., Sugamoto, K., Tomita, T, Yoshikawa, H., Tamura, S. (2004). Improvement of depth position in 2-D/3-D registration of knee implants using single-plane fluoroscopy. *IEEE Trans Med Imaging*, *23*, 602–612. https://doi.org/10.1109/tmi.2004.826051

6. Wang, J., Zhu, L., Xing, L. (2009). Noise reduction in low-dose X-ray fluoroscopy for image-guided radiation therapy. *Int J Radiat Oncol Biol Phys*, *74*, 637–643. https://doi.org/10.1016/j.ijrobp.2009.01.020

7. Kim, J., Yoon, C. J., Seong, N. J., Jeong, S. H., Kim, J. W. (2015). Fluoroscopy-guided radiofrequency ablation for small hepatocellular carcinoma: a retrospective comparison with ultrasound-guided ablation. *Clinical radiology*, *70(9)*, 1009–1015. https://doi.org/10.1016/j.crad.2015.05.008

8. Jakobs, R., Adamek, H. E., Maier, M., Kromer, M., Benz, C., Martin, W. R., Riemann, J. F. (1997). Fluoroscopically guided laser lithotripsy versus extracorporeal shock wave lithotripsy for retained bile duct stones: a prospective randomised study. *Gut*, *40(5)*, 678–682. https://doi.org/10.1136/gut.40.5.678

9. Bush, W. H., Jones, D., Gibbons, R. P. (1987). Radiation dose to patient and personnel during extracorporeal shock wave lithotripsy. *J Urol.*, *138(4)*, 716-719. https://doi.org/10.1016/s0022-5347(17)43350-7

10. Arenas, J. L., Baldwin, D. D. (2015). Techniques for Minimizing Radiation Exposure During Evaluation, Surgical Treatment, and Follow-up of Urinary Lithiasis. *Curr Urol Rep*, *16(7)*, 45. https://doi.org/10.1007/s11934-015-0517-7

11. Perisinakis, K., Damilakis, J., Anezinis, P., Tzagaraki, I., Varveris, H., Cranidis, A., Gourtsoyiannis, N. (2002). Assessment of patient effective radiation dose and associated radiogenic risk from extracorporeal shock-wave lithotripsy. *Health Physics*, *83(6)*, 847-853. https://doi.org/10.1097/00004032-200212000-00012

12. Wilson, D. L., Jabri, K. N., Aufrichtig, R. (1999). Perception of temporally filtered X-ray fluoroscopy images. *IEEE Transactions on Medical Imaging*, *18(1)*, 22–31. https://doi.org/10.1109/42.750250

13. Kim, K. P., Miller, D. L., de Gonzalez, A. B., Balter, S., Kleinerman, R. A., Ostroumova, E., Simon, S. L., Linet, M. S. (2012). Occupational Radiation Doses to Operators Performing Fluoroscopically-Guided Procedures. *Health Physics*, *103(1)*, 80–99. https://doi.org/10.1097/hp.0b013e31824dae76

14. Liu, B., Hirsch, J. A., Li, X., Sheridan, R. M., Rehani, M. M., Zheng, H., Rabinov, J. D. (2019). Radiation Dose Monitoring for Fluoroscopically Guided Interventional Procedures: Effect on Patient Radiation Exposure. *Radiology*, 180799. https://doi.org/10.1148/radiol.2019180799

15. Miller, D. L., Balter, S., Schueler, B. A., Wagner, L. K., Strauss, K. J., Vañó, E. (2010). Clinical Radiation Management for Fluoroscopically Guided Interventional Procedures. *Radiology*, *257(2)*, 321–332. https://doi.org/10.1148/radiol.10091269

16. Linet, M. S., Kitahara, C. M., Ntowe, E., Kleinerman, R. A., Gilbert, E. S., Naito, N., Lipner, R. S., Miller, D. L., Berrington de Gonzalez, A., Multi-Specialty Occupational Health Group (2017). Mortality in U.S. Physicians Likely to Perform Fluoroscopy-guided Interventional Procedures Compared with Psychiatrists, 1979 to 2008. *Radiology*, *284(2)*, 482–494. https://doi.org/10.1148/radiol.2017161306

17. Karellas, A. (2020) Challenges in Dosimetry and Radiation Dose Trends. *Radiology*, *294(2)*, 360-361. https://doi.org/10.1148/radiol.2019192414

18. Osei, B., Xu, L., Johnston, A., Darko, S., Darko, J., Osei, E. (2019). Retrospective study of patients radiation dose during cardiac catheterization procedures. *The British journal of radiology*, *92(1099)*, 20181021. https://doi.org/10.1259/bjr.20181021

19. Heidbuchel, H., Wittkampf, F. H. M., Vano, E., Ernst, S., Schilling, R., Picano, E, Mont, L., ESC Scientific Document Group (2014). Practical ways to reduce radiation dose for patients and staff during device implantations and electrophysiological procedures. *EP Europace*, *16(7)*, 946–964. https://doi.org/10.1093/europace/eut409

20. Li, X., Hirsch, J. A., Rehani, M. M., Yang, K., Liu, B. (2019). Effective Dose Assessment for Patients Undergoing Contemporary Fluoroscopically Guided Interventional Procedures. *American Journal of Roentgenology*, 1–13. https://doi.org/10.2214/ajr.19.21804

21. Ko, S., Chung, H. H., Cho, S. B., Jin, Y. W., Kim, K. P., Ha, M., Bang, Y. J., Ha, Y. W., Lee, W. J. (2017). Occupational radiation exposure and its health effects on interventional medical workers: study protocol for a prospective cohort study. *BMJ open*, *7(12)*, e018333. https://doi.org/10.1136/bmjopen-2017-018333

22. Borrego, D., Kitahara, C. M., Balter, S., Yoder, C. (2019). Occupational Doses to Medical Staff Performing or Assisting with Fluoroscopically Guided Interventional Procedures. *Radiology*, 190018. https://doi.org/10.1148/radiol.2019190018

23. Dörr, W. (2015). Radiobiology of tissue reactions. *Annals of the ICRP*, *44(1 Suppl)*, 58–68. https://doi.org/10.1177/0146645314560686

24. Shin, E., Lee, S., Kang, H., Kim, J., Kim, K., Youn, H., Jin, Y. W., Seo, S. and Youn, B. (2020). Organ-Specific Effects of Low Dose Radiation Exposure: A Comprehensive Review. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/fgene.2020.566244

25. Loganovsky, K. N., Marazziti, D., Fedirko, P. A. et al. (2020). Radiation-Induced Cerebro-Ophthalmic Effects in Humans. *Life*, *10(4)*, 41. https://doi.org/10.3390/life10040041

26. Jinnai, Y., Baba, T., Zhuang, X., Tanabe, H., Banno, S., Watari, T., Homma, Y., Kaneko, K. (2020). Does a fluoro-assisted direct anterior approach for total hip arthroplasty pose an excessive risk of radiation exposure to the surgeon? *SICOT-J*, *6*, 6. https://doi.org/10.1051/sicotj/2020004

27. European Society of Radiology (ESR) (2015). Summary of the European Directive 2013/59/Euratom: essentials for health professionals in radiology. Insights into imaging, *6(4)*, 411–417. https://doi.org/10.1007/s13244-015-0410-4

28. Lois, A. Killewich, Terrel,l A. (2011). Singleton, Governmental regulations and radiation exposure, *Journal of Vascular Surgery*, *53(1 Suppl)*, 44S-46S. https://doi.org/10.1016/j.jvs.2010.06.177

29. Bjarnason, T. A., Rees, R., Kainz, J. et al(2020).:COMP Report: A survey of radiation safety regulations for medical imaging x-ray equipment in Canada. *Journal of applied clinical medical physics*, *21(3)*, 10–19 https://doi.org/10.1002/acm2.12708

30. Berkhout, W. E. (2015). Het ALARA-principe. Achtergronden en toepassing in de praktijk [The ALARA-principle. Backgrounds and enforcement in dental practices]. *Nederlands tijdschrift voor tandheelkunde*, *122(5)*, 263–270. https://doi.org/10.5177/ntvt.2015.5.14227

31. Uffmann, M., Schaefer-Prokop, C. (2009) Digital radiography: the balance between image quality and required radiation dose. *Eur J Radiol*, *72(2)*, 202-208. https://doi.org/10.1016/j.ejrad.2009.05.060

32. Managing patient dose in digital radiology (2004). ICRP Publication 93. Elsevier: ICRP.

33. Vano, E. (2005). ICRP recommendations on 'Managing patient dose in digital radiology. *Radiation Protection Dosimetry*, *114 (1-3)*, 126–130. https://doi.org/10.1093/rpd/nch533

34. Recommendation of the international commission on radiological protection — users' edition. (1992) *ICRP Publication 60*. Oxford: ICRP: Pergamon Press.

35. Seibert, J. A. (2004). Tradeoffs between image quality and dose. *Pediatr Radiol*, *34(suppl. 3)*, S183–95 [discussion S234–141].

36. Busch, H. P., Busch, S., Decker, C., Schilz, C. (2003). Image quality and exposure dose in digital projection radiography. *Rofo*, *175*, 32–7.

37. Busch, H.P, Faulkner, K. (2005). Image quality and dose management in digital radiography: a new paradigm for optimisation. *Radiat Prot Dosimetry*, *117*, 143–147.

38. Heidbuchel, H., Wittkampf, F. H. M., Vano, E. et al. (2014). Practical ways to reduce radiation dose for patients and staff during device implantations and electrophysiological procedures. *EP Europace*, *16(7)*, 946–964. https://doi.org/10.1093/europace/eut409

39. Ozbir, S., Atalay, H. A., Canat, H. L., Culha, M. G., Cakır, S. S., Can, O., Otunctemur, A. (2019). Factors affecting fluoroscopy time during percutaneous nephrolithotomy: Impact of stone volume distribution in renal collecting system. *International braz j urol*, *45(6)*, 1153–1160. https://doi.org/10.1590/S1677-5538.IBJU.2019.0111

40. Cesarelli, M., Bifulco, P., Cerciello, T., Romano, M., Paura, L. (2013). X-ray fluoroscopy noise modeling for filter design. *Int J Comput Assist Radiol Surg*, *8*, 269–278. https://doi.org/10.1007/s11548-012-0772-8

41. Wang J, Blackburn TJ. (2000) The AAPM/RSNA physics tutorial for residents: X-ray image intensifiers for fluoroscopy. *Radiographics*, *20*, 1471–1477.

42. Jezierska, A., Talbot, H., Chaux, C. Pesquet, J. C., Engler (GI). (2012). Poisson – Gaussian noise parameter estimation in fluorescence microscopy imaging. *In: Proceedings of IEEE International Symposium on Biomedical Imaging* (pp.1663–1666).

43. Furnival, T., Leary, R. K. and Midgley, P. A. (2017). Denoising time-resolved microscopy image sequences with singular value thresholding. *Ultramicroscopy*, *178*, 112−124.

44. Kolaczyk, E.D., Dixon, D.D. (2000). Nonparametric estimation of intensity maps using Haar wavelets and Poisson noise characteristics. *The Astrophysical Journal*, *534*, 490–505.

45. Chatterjee, P., Joshi, N., Kang, S. B., Matsushita, Y. (2011). Noise suppression in low-light images through joint denoising and demosaicing. *In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (pp. 321–328).

46. Foi, A. (2009). Clipped noisy images: Heteroskedastic modeling and practical denoising. *Signal Processing*, *89(12)*, 2609–2629.

47. Foi, A., Alenius, S., Katkovnik, V., Egiazarian, K. (2007). Noise measurement for raw-data of digital imaging sensors by automatic segmentation of non-uniform targets. *IEEE Sensors Journal*, *7*, 1456−1461. https://doi.org/10.1109/JSEN.2007.904864

48. Foi, A., Trimeche, M., Katkovnik, V., Egiazarian, K. (2008). Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data. *IEEE Trans Image Process*, *17*, 1737−1754. https://doi.org/10.1109/tip.2008.2001399

49. Makitalo, M., Foi, A. (2014). Noise Parameter Mismatch in Variance Stabilization, With an Application to Poisson–Gaussian Noise Estimation. *IEEE Transactions on Image Processing*, *23(12)*, 5348–5359. https://doi.org/10.1109/tip.2014.2363735

50. Mäkitalo, M., Foi, A. (2013). Optimal inversion of the generalised Anscombe for Poisson-Gaussian noise. *IEEE Transactions on Image Processing*, *22(1)*, 91-103. https://doi.org/10.1109/TIP.2012.2202675

51. Mäkitalo, M., Foi, A. (2011). Optimal inversion of the Anscombe transformation in low-count Poisson image denoising. *IEEE Transactions on Image Processing*, 2099−2109.

52. Azzari, L., Foi, A. (2014). Gaussian-Cauchy mixture modeling for robust signal-dependent noise estimation. *Proc. 2014 IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP 2014)* (pp- 5357−5361).

53. Tan, X., Liu, Y., Zuo, C., Zhang, M. (2014). A real-time video denoising algorithm with FPGA implementation for Poisson–Gaussian noise. *Journal of Real-Time Image Processing*, *13(2)*, 327–343. https://doi.org/10.1007/s11554-014-0405-2

54. Ehmann, J., Chu, L.,Tsai, S. and Liang, C. (2018). Real- Time Video Denoising on Mobile Phones. *25th IEEE International Conference on Image Processing (ICIP), Athens* (pp. 505-509).

55. Ding, Y. (2018). *Visual Quality Assessment for Natural and Medical Image*. Zhejiang University Press (Springer), Beijing, China.

56. Monnin, P., Bosmans, H., Verdun, F. R., & Marshall, N. W. (2016). A comprehensive model for quantum noise characterization in digital mammography. Physics in Medicine and Biology, 61(5), 2083–2108. https://doi.org/10.1088/0031-9155/61/5/2083

57. Monnin, P., Bosmans, H., Verdun, F. R., Marshall, N. W. (2014). Comparison of the polynomial model against explicit measurements of noise components for different mammography systems. *Physics in Medicine and Biology*, *59(19)*, 5741–5761. https://doi.org/10.1088/0031-9155/59/19/5741

58. Burgess, A. (2004). On the noise variance of a digital mammography system. *Medical Physics*, *31(7)*, 1987–1995. https://doi.org/10.1118/1.1758791

59. Illers, H., Vandenbroucke, D., Buhr, E. (2004). Measurement of correlated noise in images of computed radiography systems and its influence on the detective quantum efficiency. *Medical Imaging 2004: Physics of Medical Imaging*. https://doi.org/10.1117/12.535807

60. Cerciello, T., Bifulco, P., Cesarelli, M., Fratini, A. (2012). A comparison of denoising methods for X-ray fluoroscopic images. *Biomed Signal Process Control*, *7*, 550–559. https://doi.org/10.1016/j.bspc.2012.06.004

61. Genovese, M., Bifulco, P., De Caro, D., Napoli, E., Petra, N., Romano, M., Cesarelli, M., Strollo, A.G.M. (2015). Hardware implementation of a spatio-temporal average filter for real-time denoising of fluoroscopic images. *J VLSI*, *49*, 114–124. https://doi.org/10.1016/j.vlsi.2014.10.004

62. Aufrichtig, R., Wilson, D. L. (1995). X-ray fluoroscopy spatio-temporal filtering with object detection. *IEEE Transactions on Medical Imaging*, *14*, 733–746. https://doi.org/10.1109/42.476114

63. Chan, C. L., Katsaggelos, A. K, Sahakian, A.V. (1993). Image sequence filtering in quantum-limited noise with applications to low-dose fluoroscopy. *IEEE Transactions on Medical Imaging*, *12(3)*, 610–621.

64. Bindilatti, A. A., Mascarenhas, N. D. A. (2013). A non local Poisson denoising algorithm based on stochastic distances. *IEEE Signal Process Lett*, *20(11)*, 1010–1013. http://dx.doi.org/10.1109/LSP.2013.2277111

65. Hensel, M., Pralow, T., Grigat, R. R. (2006). LAST Filter for Artifact-Free Noise Reduction of Fluoroscopic Sequences in Real-Time. *In: Handels H., Ehrhardt J., Horsch A., Meinzer HP., Tolxdorff T. (eds) Bildverarbeitung für die Medizin 2006. Informatik aktuell. Springer, Berlin, Heidelberg.* https://doi.org/10.1007/3-540-32137-3_4

66. Hensel, M., Pralow, T., Grigat, R. R. (2006). Real-Time Denoising of Medical X-Ray Image Sequences: Three Entirely Different Approaches. *In: Campilho A., Kamel M. (eds) Image Analysis and Recognition. ICIAR 2006. Lecture Notes in Computer Science, vol 4142. Springer, Berlin, Heidelberg.* https://doi.org/10.1007/11867661_43

67. Hensel, M., Wiesner, G., Kuhrmann, B., Pralow, T., Grigat, R. R. (2005). Motion Detection for Adaptive Spatio-temporal Filtering of Medical X-Ray Image Sequences. *In: Meinzer HP., Handels H., Horsch A., Tolxdorff T. (eds) Bildverarbeitung für die Medizin 2005. Informatik aktuell. Springer, Berlin, Heidelberg.* https://doi.org/10.1007/3-540-26431-0_15

68. Wagner, M., Yang, P., Schafer, S., Strother, C., & Mistretta, C. (2015). Noise reduction for curve-linear structures in real time fluoroscopy applications using directional binary masks. Medical Physics, 42(8), 4645–4653. https://doi.org/10.1118/1.4924266

69. Amiot, C., Girard, C., Chanussot, J., Pescatore, J., Desvignes, M. (2015). Curvelet Based Contrast Enhancement in Fluoroscopic Sequences. *IEEE Transactions on Medical Imaging*, *34(1)*, 137–147. https://doi.org/10.1109/tmi.2014.2349034

70. Amiot, C., Girard, C., Chanussot, J., Pescatore, J., Desvignes, M. (2016). Spatio-Temporal Multiscale Denoising of Fluoroscopic Sequence. *IEEE Transactions on Medical Imaging*, *35(6)*, 1565–1574.

71. Irrera, P., Bloch, I., Delplanque, M. (2016). A flexible patch based approach for combined denoising and contrast enhancement of digital X-ray images. *Medical Image Analysis*, *28*, 33–45. https://doi.org/10.1016/j.media.2015.11.002

72. Vieira, M. A. C., Bakic, P. R., Maidment, A. D. A., Schiabel, H., Mascarenhas, N. D. A. (2012). Filtering of Poisson Noise in Digital Mammography Using Local Statistics and Adaptive Wiener Filter. *In: Maidment A.D.A., Bakic P.R., Gavenonis S. (eds) Breast Imaging. IWDM 2012. Lecture Notes in Computer Science, vol 7361. Springer, Berlin, Heidelberg.* https://doi.org/10.1007/978-3-642-31271-7_35

73. Lee, M. S, Park, C. H., Kang, M. G. (2017). Edge enhancement algorithm for low-dose X-ray fluoroscopic imaging. *Comput Meth Prog Bio*, *152*, 45−52.

74. Leem, M. S., Park, S.W., Lee, S. Y., Kang, M.G. (2017). Motion-adaptive 3D nonlocal means filter based on stochastic distance for low-dose X-ray fluoroscopy. *Biomed Signal Process*, *38*, 74−85.

75. Hariharan, S. G., Strobel, N., Kaethner, C., Kowarschik, M., Demirci, S., Albarqouni, S. et al. (2018). A photon recycling approach to the denoising of ultra-low dose X-ray sequences. *Int J Comput Ass Rad*, *13,* 847−854.

76. Shi, L., Hu, Y., Chen, Y., Yin, X., Shu, H., Luo, L., Coatrieux, J. L. (2016). Improving Low-dose Cardiac CT Images based on 3D Sparse Representation. *Sci Rep*, *6*, 22804.

77. Zhao, T., Hoffman, J., McNitt-Gray, M., Ruan, D. (2019). Ultra-low-dose CT image denoising using modified BM3D scheme tailored to data statistics. *Med Phys.*, *46*, 190–198.

78. Bovik, A.C. (2009). The Essential Guide to Image Processing. *Academic Press. (Elsevier)*, Burlington, MA.

79. Chow, L.S., Paramesran, R. (2016): "Review of medical image quality assessment", *Biomed Signal Process Control 27*, 145–154.

80. Thanki, R., Borra, S., Dey, N., Ashour, A.S. (2018). Medical imaging and its objective quality assessment: an introduction. *In: Classification in BioApps* (pp. 3-32). Springer, Cham.

81. Sadda, P., Qarni, T. (2018). Real-Time Medical Video Denoising with Deep Learning: Application to Angiography. *International journal of applied information systems*, *12(13)*, 22–28. https://doi.org/10.5120/ijais2018451755

82. Lefkimmiatis, S., Maragos, P., Papandreou, G. (2009). Bayesian inference on multiscale models for Poisson intensity estimation: applications to photon-limited image denoising. *IEEE Transactions of Image Processing*, *18*, 1724–1741. https://doi.org/10.1109/TIP.2009.2022008

83. Tapiovaara, M. J. (1993). SNR and noise measurements for medical imaging: II. Application to fluoroscopic X-ray equipment. *Phys Med Biol*, *38*, 1761–1788. https://doi.org/10.1088/0031-9155/38/12/006

84. Lee, S., Lee, M. S., Kang, M. G. (2018). Poisson-Gaussian Noise Analysis and Estimation for Low-Dose X-ray Images in the NSCT Domain. *Sensors*, *18(4)*, 1019. https://doi.org/10.3390/s18041019

85. Hensel, M., Pralow, T., Grigat, R. R. (2007). Modeling and real-time estimation of signal-dependent noise in quantum-limited imaging. *In Proceedings of the 6th WSEAS International Conference on Signal Processing, Robotics and Automation (ISPRA'07). World Scientific and Engineering Academy and Society (WSEAS), Stevens Point, Wisconsin, USA* (pp. 183–191).

86. Hubbard, V. M. (1970). The approximation of a Poisson distribution by a Gaussian distribution. *Proc. of the IEEE*, *58(9)*, 1374-1375.

87. Niklason, L. T., Sorenson, J. A., Nelson, J. A. (1981). Scattered radiation in chest radiography. *Medical Physics*, *8(5)*, 677–681. https://doi.org/10.1118/1.595027

88. Swank, R. K. (1973). Absorption and noise in x-ray phosphors. *J Appl Phys*, 444199–203

89. Bürgisser, P., Clausen, M., Shokrollahi, A. (1997). *Algebraic Complexity Theory*. Springer.

90. Atallah, M. J. (Ed.) (1999). *Algorithms and Theory of Computation Handbook*. CRC Press.

91. Bogduk, N. and Mercer, S. (2000). Biomechanics of the cervical spine. I: Normal kinematics. *Clin. Biomech*, *15(9)*, 663-648, , https://doi.org/10.1016/S0268-0033(00)00034-6

92. Vogt, L., Pfeifer, K., Portscher, M., Banzer, W. (2001). Influences of Nonspecific Low Back Pain on Three-Dimensional Lumbar Spine Kinematics in Locomotion. *Spine*, *17*, 1910-1919.

93. Yeo, C.G., Jeon, I., Kim, S.W. (2015). Delayed or Missed Diagnosis of Cervical Instability after Traumatic Injury: Usefulness of Dynamic Flexion and Extension Radiographs. *Korean J. Spine*, *12(3)*, 146-149.

94. Bundschuh, C., Modic, M., Kearney, F., Morris, R., Deal, C. (1988). Rheumatoid arthritis of the cervical spine: surface-coil MR imaging. *Am J Roentgenol*, *151(1)*, 181-187. https://doi.org/10.2214/ajr.151.1.181

95. Ali, F.E., Al-Bustan, M.A., Al-Busairi, W.A., Al-Mulla, F.A., Esbaita, E.Y. (2006). Cervical spine abnormalities associated with Down syndrome. *Int Orthop*, *30(4)*, 284-289. https://doi.org/10.1007/s00264-005-0070-y

96. Panjabi, M.M. (2003). Clinical spinal instability and low back pain. *J Electromyogr Kines*, *13*, 371-379.

97. Leone, A., Guglielmi, G., Cassar-Pullicino, V.N., Bonomo, L. (2007). Lumar inter-vertebral instability: a review, *Radiology*, *245(1)*, 62-77.

98. Panjabi, M.M. (1979). Centers and angles of rotation of body joints: a study of errors and optimization", *J Biomech*, *12*, pp. 991-920.

99. Panjabi, M.M., Chang, D., Dvorak, J. (1992). Analysis of errors in kinematics parameters associated with in vivo functional radiographs", *Spine*, *2*, 200-205.

100. Van Mameren, H., Sanches, H., Beurgens, J., Drukker, J. (1992) Cervical spine motion in the sagittal plane (II) position of segmental averaged instantaneous centers of rotation: a cineradiographic study. *Spine*, *17*, 467-474.

101. Lee, S.W., Draper, E.R., Hughes, S.P. (1997). Instantaneous center of rotation and instability of the cervical spine. A clinical study. *Spine*, *22(6)*, 641-648.

102. Patwardhan, A.G., Havey, R.M., Carandang, G., Simonds, J., Voronov, L.I., Ghanayem, A.J., Meade, K.P., Gavin, T.M. and Paxinos, O. (2003). Effect of compressive follower preload on the flexion–extension response of the human lumbar spine. *J. Orthop. Res.*, *21(3)*, 549-546. https://doi.org/10.1016/s0736-0266(02)00202-4

103. Gonzalez, R. C., Woods, R. E. (1992). Digital image processing. 3rd ed. Boston: Addison-Wesley, Reading.

104. Bhardwaj, S., Mittal, A. A Survey on Various Edge Detector Techniques. *Proc. Technology*, *4*, 220-226, https://doi.org/10.1016/j.protcy.2012.05.033

105. Trichili, H., Bouhlel, M. S., Derbel, N., Kamoun, L. (2002). A survey and evaluation of edge detection operators application to medical images", *IEEE Sys Man Cybern*. https://doi.org/10.1109/icsmc.2002.1173373

106. Jain, A., Gupta, M., Tazi, S.N., Deepika. (2014). Comparison of edge detectors, MedCom, https://doi.org/10.1109/medcom.2014.7006020

107. Sarkar, S., Mandal, A. (2015) Comparison of Some Classical Edge Detection Techniques with their Suitability Analysis for Medical Images Processing, *Int J Comput Sci. Eng.*, *3(1)*, 81-87,

108. M. Kondracki (2001). "Clinical applications of digitized videofluoroscopy in the lumbar spine", PhD Thesis, University of Southampton,

109. Breen, A.C., Muggleton, J.M. Mellor, F.E. (2006). An objective spinal motion imaging assessment (OSMIA): reliability, accuracy and exposure data", *BMC Musculoskelet Disord*, *7(1)*, pp. 1-10,.

110. Gonzalez, R. C., Woods, R. E. (1992). Digital image processing (3rd ed., Boston): Addison-Wesley, Reading.

111. Lo, C. M., Sawchuk, A. A. (1973). Nonlinear restoration of filtered images with Poisson noise. *SPIE Proc Applications of Digital Image Processing III, San Diego* (pp. 84–95).

112. Gilboa, G., Sochen, N., Zeevi Y. Y. (2006). Variational denoising of partly textured images by spatially varying constraints. *IEEE Transactions on Image Processing*, *15(8)*, 2281–2289.

113. You, Y. L., Kaveh, M. (2000). Fourth-order partial differential equations for noise removal. *IEEE Transactions on Image Processing*, *9(10)*, 1723–1730.

114. Hirakawa, K., Parks, T. W. (2006). Image denoising using total least squares. *IEEE Transactions on Image Processing*, *15(9)*, 2730–2742.

115. Elad, M., Aharon, M. (2006). Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, *15(12)*, 3736–3745.

116. Crouse, M. S., Nowak, R. D., Baraniuk, R. G. (1998). Wavelet-based statistical signal processing using hidden Markov models. *IEEE Transactions on Signal Processing*, *46(4)*, 886–902.

117. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K. (2007). Image denoising by sparse 3D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, *16(8)*, 2080-2095 https://doi.org/10.1109/TIP.2007.901238

118. Burger, H. C., Schuler, C.J. and Harmeling. S. (2012). Image denoising: Can plain neural networks compete with BM3D?. *2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI* (pp. 2392−2399).

119. Nishio, M., Nagashima, C., Hirabayashi, S., Ohnishi, A., Sasaki, K., Sagawa, T., Hamada, M., Yamashita, T. (2017) Convolutional auto-encoder for image denoising of ultra-low-dose CT. *Heliyon*, *3*, e00393.

120. Anaya, J. and Barbu, A. (2018). RENOIR - A Dataset for Real Low-Light Image Noise Reduction. *J Vis Commun Image R*, *51*, 144−154.

121. Hasan, M. and El-Sakka, M. J. (2018). Improved BM3D image denoising using SSIM-optimized Wiener filter. *EURASIP J Image Vide*, 25.

122. Anscombe, F. J. (1948) The transformation of Poisson binomial and negative-binomial data. *Biometrika*, *35*, no. 3/4, 246-254.

123. Maggioni, M., Boracchi, G., Foi, A., Egiazarian, K. (2012). Video Denoising, Deblocking and Enhancement Through Separable 4-D Nonlocal Spatiotemporal Transforms. *IEEE Transactions on Image Processing*, *21(9)*, 3952-3966 https://doi.org/10.1109/TIP.2012.2199324

124. Mildenhall, B., Barron, J. T., Chen, J., Sharlet, D., Ng, R., Carroll R. (2018). Burst Denoising with Kernel Prediction Networks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT*, 2502−2510.

125. Arias, P., Morel, J. M. (2018). Video Denoising via Empirical Bayesian Estimation of Space-Time Patches. *J Math Imaging Vis*; *60*, 70−93.

126. Brox, T., Kleinschmidt, O., Cremers, D. (2008). Efficient non-local means for denoising of textural patterns. *IEEE Trans Image Proc*, *17*, 1083−1092.

127. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K. (2006). Image denoising with block-matching and 3D filtering. *Electronic Imaging'06, Proc. SPIE 6064, no. 6064A-30, San Jose, California USA*.

128. Foi, A., Katvonik, V., Egiazarian, K. (2007). Pointwise Shape-Adaptive DCT for High-Quality Denoising and Deblocking of Grayscale and Color Images. *IEEE Transactions on Image Processing*, *16*, 1395−1411.

129. Zhang, L., Zhang, L., Mou, X., Zhang, D. (2011). FSIM: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*, *20*, 2378−2386.

130. Nilsson, J., Akenine-Möller, T. (2020). Understanding SSIM. arXiv/2006.13846.

131. Henriksson, L., Hyvärinen, A., Vanni, S. (2009). Representation of cross-frequency spatial phase relationships in human visual cortex. *J Neurosci*, *29(45)*, 14342–14351.

132. http://www3.gehealthcare.ca/en/products/categories/surgical_imaging/oec_c-arms/oec_9900_elite (accessed online on 07/15/2019).

133. Angelino, C.V., Debreuve, E., Barlaud, M. (2011). Confidence-based Denoising Relying on a Transformation-invariant, Robust Patch Similarity - Exploring Ways to Improve Patch Synchronous Summation. *International Conference on Imaging Theory and Applications. Proceedings of the International Conference on Imaging Theory and Applications and International Conference on Information Visualization Theory and Applications, Vilamoura, Algarve, Portugal*.

134. Wang, Z., Bovik, A.C, Sheikh, H.R., Simoncelli, E.P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13(4)*, 600-612.

135. Marr, D. (1980). *Vision*. New York: Freeman.

136. Marr, D. and Hildreth, E. (1980). Theory of edge detection. *Proc. R. Soc.Lond. B*, *207(1167)*, 187–217.

137. Morrone, M. C. and Burr, D. C. (1988). Feature detection in human vision: A phase-dependent energy model. *Proc. R. Soc. Lond. B*, *235(1280)*, 221–245.

138. Li, C. and Bovik, A. C. (2009). Three-component weighted structural similarity index. In *Proc. SPIE: Vol. 7242, Image Quality and System Performance VI, 72420Q*. https://doi.org/10.1117/12.811821

139. Larson, E. C. and Chandler, D. M. (2008). Unveiling relationships between regions of interest and image fidelity metrics. *Proc. SPIE Vis. Comm. Image Process.*, *6822*, 6822A1–6822A16.

140. Larson, E. C., Vu, C., Chandler, D. M. (2008). Can visual fixation patterns improve image fidelity assessment? *In Proc IEEE Int Conf Image Process* (pp. 2572–2575).

141. Attar, A., Rad, R.M., Shahbahrami, A. (2011). EBIQA: An edge based image quality assessment. *In: 7th Iranian Conference on Machine Vision and Image Processing* (pp. 1–4).

142. Zhang, M., Mou, X., Zhang, L. (2011) Non-shift edge based ratio (NSER): An image quality assessment metric based on early vision features. *IEEE Signal Process. Lett.*, *18(5)*, 315–318.

143. Canny, J. A computational approach to edge detection. (1986).*IEEE Trans Pattern Anal Mach Intell*, *8(6)*, 679–698.

144. Paranjape, R.B. *Fundamental Enhanced Techniques*, in *Handbook of Medical imaging*, Academic Press, London, ISBN: 0-12-077790-8.

145. Jiang, A., Kwan, H.K. (2010). Minimax Design of IIR Digital Filters Using Iterative SOCP. *IEEE Trans. on Circuits and Systems—I: Regular Papers*, *57(6)*, 1326 1337.

146. Lang, M. C. (2000). Least-squares design of IIR filters with prescribed magnitude and phase responses and a pole radius constraint. *IEEE Transactions on Signal Processing*, *48(11)*, 3109-3121.

147. Parks, T.W., Burrus, C.S. (1987). *Digital filter design*. John Wiley & Sons.

148. Steiglitz, K., McBride, L.E. (1965). A technique for the identification of linear system. *IEEE Trans. on Automatic Control*, *AC-10*, 461-464,.

149. Castellano, G., De Caro, D., Esposito, D., Bifulco, P., Napoli, E., Petra, N., Andreozzi, E., Cesarelli, M., Strollo, A.G.M. (2019). An FPGA-oriented Algorithm for Real-Time Filtering of Poisson Noise in Video Streams, with Application to X-Ray Fluoroscopy. *Circuits Syst Signal Process*, *38*, 3269–3294. https://doi.org/10.1007/s00034-018-01020-x

150. Awate, S. P., Whitaker, R. T. (2005). Higher-order image statistics for unsupervised, information-theoretic, adaptive, image filtering. *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR'05)*: *Vol. 2* (pp.44-51).

151. Buades, A., Coll, B., Morel, J.M. (2005). A Non-Local Algorithm for Image Denoising. *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR'05): Vol. 2* (pp. 60-65).

152. Dabov, K., Foi, A., Egiazarian, K. (2007). Video denoising by sparse 3D trasform-domain collaborative filtering. *In Proc. of European Signal Processing Conference (EUSIPCO), Poznań, Poland*.

153. Varghese, G., Wang, Z. (2010). Video Denoising Based on a Spatiotemporal Gaussian Scale Mixture Model. *IEEE Trans. on Circuits and Systems for Video Technology*, *20(7)*, 1032 – 1040.

154. Video Sequence Database [Online]. Available: https://media.xiph.org/video/derf/.

155. Vidal, F.P., Villard, P.F. (2016). Development and validation of real-time simulation of X-ray imaging with respiratory motion. *Comput Med Imaging Graph, 49*, 1–15. https://doi.org/10.1016/j.compmedimag.2015.12.002

156. Súgar, A., Meuleman, A., Villard, P.-F., Garcia, M., Vidal, F. (2017). gVirtualXRay: Virtual X-Ray Imaging Library on GPU. *CGVC*, 61-68. https://doi.org/10.2312/cgvc.20171279

157. Súgar, A., Kelly, G., García, M., Vidal, F. (2019). Projectional Radiography Simulator: an Interactive Teaching Tool. *CGVC*, 125 – 128. https://doi.org/10.2312/cgvc.20191267

158. Leeds Test Objects: TOR 18FG product specifications (2015). https://www.leedstestobjects.com/wp-content/uploads/TOR-18FG-product-specifications-1.pdf?x78567 (Accessed 14/10/2020)

159. Leeds Test Objects: TOR CDR product specifications (2017). https://www.leedstestobjects.com/wp-content/uploads/TOR-CDR-product-specifications-1.pdf?x78567 (Accessed 14/10/2020).

160. Badawy, W.; Graham, A. J. (Eds.) (2003). *System-on-Chip for Real-Time Applications*. Springer US. https://doi.org/10.1007/978-1-4615-0351-4

161. Oruklu, E., Hanley, R., Aslan, S., Desmouliers, C., Vallina F., Saniie J. (2012). System-on-Chip Design Using High-Level Synthesis Tools. *Circuits and Systems*, *3(1)*, pp. 1 – 9. https://doi.org/10.4236/cs.2012.31001