Rapid and flexible high-resolution scanning enabled by cycloidal computed tomography and convolutional neural network (CNN) based data recovery

Daniël M. Pelt, Charlotte Maughan-Jones, Oriol Roche i Morgó, Alessandro Olivo, Charlotte Hagen

Abstract—**We have combined a recently developed imaging concept ("cycloidal computed tomography") with convolutional neural network (CNN) based data recovery. The imaging concept is enabled by exploiting, in synergy, the benefits of probing the sample with a structured x-ray beam and applying a cycloidal acquisition scheme by which the sample is simultaneously rotated and laterally translated. The beam structuring provides a means of increasing the in-slice spatial resolution in tomographic images irrespective of the blur imposed by the x-ray source and detector, while the "roto-translation" sampling allows for rapid scanning. Data recovery based on the recently proposed Mixed-Scale Dense (MSD) CNN architecture enables an efficient reconstruction of high-quality, high-resolution images despite the fact that cycloidal computed tomography data are highly incomplete. In the following, we review the basic principles underpinning cycloidal computed tomography, introduce the CNN based data recovery method and discuss the benefit of combining both.**

Index Terms—**computed tomography, micro-CT, convolutional neural networks, machine learning**

I. INTRODUCTION

X RAY computed tomography (CT) plays a crucial role in various areas spanning from medicine and biomedical research to security, non-destructive testing of materials and study of archeological artefacts. A strength of this imaging modality is the broad range of spatial resolution levels that can be accessed – hospital scanners can typically resolve hundreds of μm, compact micro-CT machines can access resolutions of a few to tens of μm, and more specialized implementations like nano-CT enable resolutions on the sub-μm scale. A scanner's spatial resolution is, on the most basic level, defined by the x-ray source and detector, which introduce a combined blur to the images that defines (besides the signal-to-noise ratio) what can and cannot be detected[1]. While some x-ray sources (e.g. those with a variable focal spot) and detectors (e.g. CCD cameras

with variable optics) allow controlling the blur to some extent[2], allowing to increase resolution within a limited range, generally this is a restrictive and inflexible constraint.

In an effort to overcome this, we have developed a (micro-CT) scanner that features a mask with alternating absorbing and transmitting septa in the x-ray beam path. Provided that the mask fulfills certain design criteria (see Section II), its presence introduces spatial frequencies higher than the cut-off normally imposed by the combined source and detector blur into the image formation process[3,4]. Therefore, the mask provides the ability of high-resolution imaging even with x-ray sources and detectors that would not normally allow this due their large focal spot and/or pixel size, respectively. Although the mask creates an under-sampling problem as some areas of the sample are shielded from radiation, this does not compromise the ability of reconstructing high-resolution images, provided that a dedicated cycloidal ("roto-translation") acquisition scheme is applied in combination with a suitable method for data recovery. The synergy between the beam structuring, "roto-translation" sampling and data recovery has led to a unique imaging concept ("cycloidal computed tomography") that enables rapid and highly flexible high-resolution CT scanning[5]. Here, we report on utilizing a machine learning technique (the Mixed-Scale Dense CNN architecture[6]) for data recovery, which outperforms previously applied "naive" approaches like bivariate interpolation.

II. COMBINING A STRUCTURED X-RAY BEAM SETUP WITH CYCLOIDAL SAMPLING

In our scanner, the x-ray beam is structured into an array of long, thin beamlets (a few μm to tens of μm laterally, and extending uniformly in the y-direction) by means of a mask positioned immediately upstream of the sample (Fig. 1a). The mask period (p) matches the effective pixel size (s), i.e. p = s/m, where m is the magnification between the mask and detector (for simplicity, the distance between mask and sample, which

D. M. Pelt is with the Centrum Wiskunde & Informatica (CWI), 1090 GB Amsterdam, The Netherlands.

C. Maughan-Jones, O. Roche i Morgò, M. Endrizzi, A. Olivo and C. K. Hagen are with the Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, United Kingdom. C. K. Hagen is the corresponding author (email: charlotte.hagen.10@ucl.ac.uk).
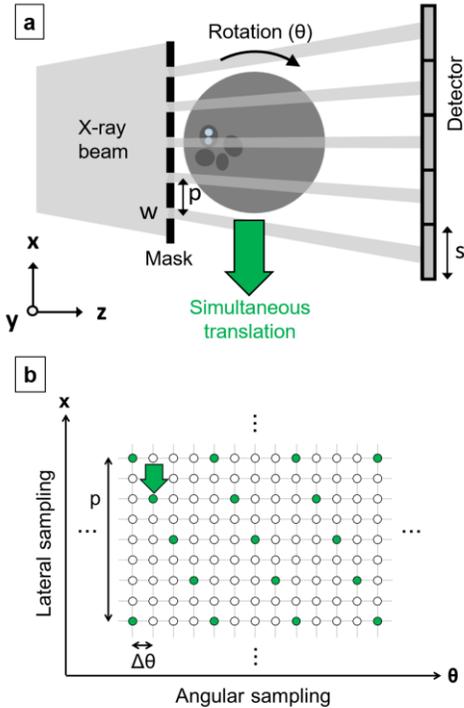
Figure 1. **a)** Schematic of the micro-CT scanner layout (seen from top), based on structuring the incoming x-ray beam into an array of narrow beamlets. Note the schematic is not to scale; in a real system the mask apertures are on the order of a few to tens of μm, the detector pixel size is around 50-100 μm, and the distance between mask and detector is between 15 cm and 1 m. **b)** Sinogram sampling grid realized through cycloidal ("roto-translation") sampling (see text). The grid is shown for a single mask period and a subset of angular increments. The filled circles indicate the available data points, whereas the empty circles indicate the entries which are to be recovered through an appropriate data recovery method.

is typically < 5% of the mask-detector distance, is considered negligible). It is important that the mask fulfills two design criteria; (1) its apertures (w) must be smaller than both the effective (i.e. scaled to the mask plane) projected focal spot of the x-ray source and the effective detector pixel size; (2) the mask period (p) must be sufficiently large to keep the beamlets physically separated. While the first criterion leads to the presence of additional spatial frequencies in the image formation process, the second criterion ensures that these frequencies remain accessible (by not being blurred). To a first approximation, spatial frequencies up to the inverse of the aperture width (1/w) are now present and can contribute to the in-slice spatial resolution of an image, provided that diffraction effects and detector cross-talk are negligible. It should be noted that the second criterion does not impose a constraint on the detector pixel size, as one can divert from the p = s/m relationship and employ a "line skipping" mask, in which p is an integer multiple (>1) of the effective pixel size such that only every second (or third etc.) detector pixel column receives one beamlet. This is also an effective strategy when working with detectors with non-negligible cross-talk between pixels; by skipping one or more pixel columns, any signal diffusion due to cross-talk can be minimized.

Generally, the use of beamlets leads to an under-sampled sinogram, i.e. the lateral sampling step defined by the mask

period (p) is too large to adequately capture the newly gained spatial frequencies. To manage this problem efficiently, we have applied a cycloidal scanning scheme by which the sample is translated laterally (along x) simultaneously with being rotated around the vertical (y-) axis. This results in a "roto-translation" motion, where, if the translation is uni-directional, each feature in the sample follows a cycloidal trajectory. Although "roto-translation" sampling does not solve the under-sampling problem as such (as no additional data are acquired), it transforms a highly unbalanced sampling grid, which would correspond to rotation-only sampling, into a balanced grid where the available data are more evenly distributed (Fig. 1b). A balanced sinogram sampling grid is advantageous, as it provides a much better basis for restoring the missing data via a mathematical recovery method, to an extent that high-resolution images can be reconstructed.

We would briefly like to point out that there is another solution to the under-sampling problem that is, however, much less efficient than the cycloidal approach. In particular, the sample can be scanned laterally through the setup in multiple steps at each rotation angle; this leads to the actual acquisition of the missing entries in the sinogram. This procedure, which is often termed dithering, is incompatible with rapid scanning, as – although the overall exposure time can be re-distributed among the scanning steps – dead times are required for the necessary repeated stop-starting of the sample rotation and translation motors. This is not required in a cycloidal acquisition; in fact, the latter enables continuous scans (sometimes termed "fly-scans"), which do not suffer from any dead times, and in which the total scan time is solely determined by the exposure time. In the following, images obtained via dithering will be considered the "gold-standard" to which cycloidally sampled data are compared.

III. CONVOLUTIONAL NEURAL NETWORK (CNN) RECONSTRUCTION

To improve the quality of cycloidal computed tomography images, we have replaced "naive" data recovery based on bivariate interpolation by a machine learning approach. Specifically, we have trained and applied the recently proposed Mixed-Scale Dense (MSD) CNN architecture[6] to recover those sinogram entries that are not acquired during cycloidal acquisitions. In brief, the MSD CNN architecture differs from popular existing CNNs by using dilated convolutions instead of scaling operations to capture image features at different scales, and by connecting all network layers with each other. As a result, MSD networks typically require fewer intermediate images and learned parameters to achieve accurate results compared with existing CNNs, and are therefore well-suited for efficiently processing large datasets and accurately learning from relatively few training images. It has recently been demonstrated that the MSD CNN can improve the quality of images reconstructed from incomplete "conventional" CT measurements[7] (acquired without any beam structuring or cycloidal scanning). While previously the CNN was applied as a post-processing tool in the image domain, we have adapted it to and applied it in the sinogram domain. This has led to the recovery of complete sinograms (with as many entries as
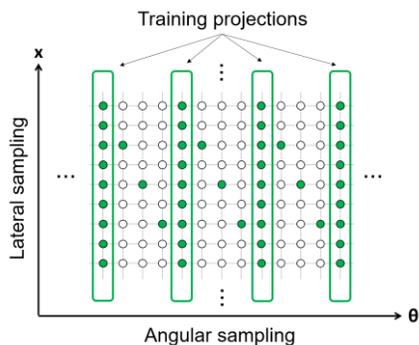
Figure 2. Sinogram sampling grid (shown for one mask period) corresponding to a cycloidal ("roto-translation") scan with an interleaved acquisition of training data. The angular intervals at which training data are acquired are variable.

present in dithered sinograms), containing all the data required for the reconstruction of high-resolution tomographic images.

A particular challenge lay in the task of generating suitable training data for the CNN. We had aimed for a general approach that does not rely on scanning many samples of the same type and using these data for the training. Instead, we have developed a strategy by which training data are acquired as part of the scan of each individual sample. Specifically, the acquisition of a number of dithered projections, which would become the training data, was integrated into a cycloidal acquisition. The angles at which the training projections are acquired should be distributed evenly across the full angular range. The sinogram sampling grid that corresponds to this scanning sequence is shown in Fig. 2. For the most part, the grid is the same as the one shown on Fig. 1b, but for some angles all sinogram entries are now available.

IV. RESULTS

To demonstrate the performance of CNN-based cycloidal computed tomography, we have acquired experimental data of a chicken bone (fixed in formalin) that was placed in a plastic container of approximately 9 mm diameter. To stabilize the chicken bone in the container, it was surrounded by agarose. The experimental setup comprised a MicroMax-007 HF x-ray tube (Rigaku, Japan) with a rotating molybdenum anode, operated at 40 kV and 25 mA, resulting in a horizontal focal spot of approximately 70 μm (full width half maximum). The detector was the Pixirad-2 photon counter with a pixel size of 62 μm. The mask (Creatv Microtec, USA) had a 79 μm period and 10 μm apertures. The mask and detector were positioned at 1.6 m and 2.53 m from the source, respectively. With these distances, the mask period covered two detector pixel columns when projected to the detector ("line-skipping"), which is equivalent to using a detector with twice as large pixels horizontally. A single, dithered scan of the sample was carried out, and the data were subsequently sub-sampled to mimic cycloidal acquisitions. During the scan, the sample was translated laterally in eight steps of 10 μm at each rotation angle, as this ensured that all sinogram entries were captured. Data were acquired as the sample was rotated in steps of 0.2 degrees over 180 degrees, corresponding to the acquisition of 900 x 8 = 7200 frames in total. The exposure time per frame

was 2s.

A dithered image was reconstructed from all acquired data. Since dithering provides complete sinograms, no data recovery was required, and reconstruction was performed with filtered back projection (FBP). To generate cycloidally sampled data, only a single dithering step was considered, but this was a different step (corresponding to a different lateral sample offset) at each rotation angle, chosen in such a way that an interlaced sampling pattern like the one shown in Fig. 1b was created. However, at every 10th angle, the cycloidal sub-sampling was interrupted and all dithering steps were included; the data from these angles were used as training projections (90 projections in total). Specifically, a 100-layer MSD network was trained with bicubic interpolated cycloidal sinograms as network input, and sinograms including only the dithered projections as training target. Out of 300 available sinograms (obtained for 300 cross-sections of the scanned sample), 270 were used for training the network, while the remaining 30 were used as a validation set to monitor performance during training. After training the CNN on the dithered projections, the CNN was applied to the remaining, cycloidally sampled sinogram areas, so as to restore the missing entries in these areas. Once the sinogram had been completed, image reconstruction was performed with FBP. For comparison, we have also applied bicubic interpolation to the same sinogram, and reconstructed an image via FBP. In both cases (CNN and bicubic interpolation), the 90 dithered training projections had been used in the tomographic reconstruction.

The results are shown in Fig. 3. Panel 3a shows the dithered image, panel 3b shows the cycloidal image reconstructed with bicubic interpolation, and panel 3c shows the cycloidal image reconstructed with CNN-based data recovery. Both of the cycloidal images were reconstructed from 22% of the data that contributed to the dithered image (this can calculated be from the fraction of training projections and the fact that cycloidal sampling produces an eighth of the frames produced by dithering). It can be seen that the spatial resolution in both of the cycloidal images is comparable to that in the dithered image, but that the CNN based recovery (panel c) leads to a better overall image quality than bicubic interpolation (panel b). To quantitatively analyse the performance of bicubic interpolation and CNN based data recovery, we have calculated the peak signal-to-noise ratio (PSNR) between the respective cycloidal
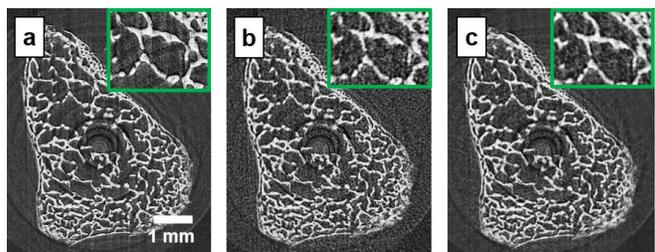
Figure 3. Tomographic images of a chicken bone: **a)** dithered image, **b)** cycloidally sampled image processed with bicubic interpolation, **c)** cycloidally sampled image processed using CNN based data recovery. The images shown in b) and c) were reconstructed from 22% of the data that contributed to the dithered image shown in a).

image and the dithered image; the results are:

PSNR (bicubic interpolation) = 28.9

PSNR (CNN based recovery) = 30.9,

confirming the visual observation on image quality.

V. CONCLUSION

Cycloidal computed tomography offers an opportunity for flexible and efficient high-resolution scanning. The flexibility stems from the beam structuring, as it provides access to higher resolutions than those "dictated" by the source-detector blur. This implies that high-resolution imaging becomes possible also with "low-resolution equipment", such as x-ray sources with relatively broad focal spots and detectors with relatively large pixels, and that fast switching between resolution levels can be achieved simply by changing the mask inside the scanner. The flexibility of the concept also lies in another aspect: while the images shown in this document show attenuation contrast, cycloidal computed tomography can also be implemented in phase contrast mode. This can be achieved by including an array of beam stops, positioned immediately upstream of the detector, into the scanner. The beam stop array transforms the scanner into an edge illumination x-ray phase contrast imaging device[8], which has been demonstrated to produce high-quality phase contrast images for a broad range of samples and applications[9-11]. Phase contrast imaging has proven to be especially beneficial when a sample exhibits weak x-ray attenuation, as phase effects can be relatively stronger than attenuation ones for such samples[12].

The efficiency of cycloidal computed tomography stems from the dedicated sampling scheme ("roto-translation"); this removes the need for excessive, sub-pixel lateral scanning of the sample per rotation angle (i.e. dithering, an extremely cumbersome way to acquire data). A key advantage is that "roto-translation" sampling is compatible with continuous acquisitions ("fly-scans"). In this work, we have explored the benefit of applying the Mixed-Scale Dense (MSD) CNN architecture to the cycloidally sampled sinogram. We have shown that high-quality, high-resolution images can be obtained despite the fact that the sinogram is highly incomplete. Specifically, we have shown that the CNN outperforms bicubic interpolation. In future studies, the CNN should also be compared against an iterative reconstruction approach to fully appreciate its performance.

A strategy for generating training data has been proposed; this method is very general, and does not rely on previously acquired images of similar samples. In practical terms, when scanning in step-and-shoot mode, training projections can be acquired by applying dithering at a number of rotation angles. When scanning continuously ("fly-scan"), the training projections would have to be acquired before or after the scan. Although this leads to an increase in scan time, this can be relatively small (e.g. training projections were acquired at 10% of all angles in order to generate Fig. 3c).

In summary, we believe that, in combination, cycloidal computed tomography and CNN based data recovery bear great potential to increase the flexibility of CT machines in terms of spatial resolution and overall image quality, while also allowing to perform scans in a rapid and efficient manner.

REFERENCES

[1] T. Buzug, "Computed Tomography; From Photon Statistics to Modern Cone-Beam CT," Berlin, Heidelberg: Springer, 2008

[2] J. Grimes, X. Duan, L. Yu, A. Halaweish, N. Haag, S. Leng, and C. McCollough, "The influence of focal spot blooming on high-contrast spatial resolution in CT imaging", Med. Phys. 42, 6011-6020, 2015

[3] P. Diemoz, F. A. Vittoria, and A. Olivo, "Spatial resolution of edge illumination x-ray phase contrast imaging", Opt. Express 22, 15514-15529

[4] C. K. Hagen, F. A. Vittoria, M. Endrizzi, and A. Olivo, "Theoretical framework for spatial resolution in edge illumination x-ray tomography", Phys. Rev. Appl. 10, 054050, 2018

[5] C. K. Hagen, F. A. Vittoria, M. Endrizzi, and A. Olivo, "Cycloidal Computed Tomography", UK Patent Application 1820362.0 (Filed Dec 13, 2018)

[6] D. M. Pelt and J. A. Sethian, "A mixed-scale dense convolutional neural network for image analysis", Proc. Natl. Acad. Sci. U.S.A. 115, 254-259, 2018

[7] D. M. Pelt, K. J. Batenburg, and J. A. Sethian, "Improving tomographic reconstruction from limited data using mixed-scale dense convolutional neural networks", J. Imaging 4, 128, 2018

[8] Zamir A, C. K. Hagen, P. C. Diemoz, M. Endrizzi, F. A. Vittoria, Y. Chen, M. A. Anastasio, and A. Olivo, "Recent advances in edge illumination x-ray phase contrast computed tomography", J. Med. Imaging 4, 040901, 2017

[9] C. K. Hagen, P. Maghsoudlou, G. Totonelli, P. C. Diemoz, M. Endrizzi, L. Rigon, R. H. Menk, F. Arfelli, D. Dreossi, E. Brun, P. Coan, A. Bravin, P. De Coppi, and A. Olivo, "High contrast microstructural visualization of natural acellular matrices by means of phase-based x-ray tomography", Sci. Rep. 5, 18156, 2015

[10] L. Massimi, C. K. Hagen, M. Endrizzi, P. R. T. Munro, G. Havariyoun, P. M. S. Hawker, B. Smit, A. Astolfo, O. J. Larkin, R. M. Waltham, Z. Shah, S. W. Duffy, R. L. Nelan, A. Peel, T. Suaris, J. L. Jones, I. G. Haig, and A. Olivo, "Laboratory-based x-ray phase contrast CT technology for clinical intra-operative specimen imaging", Proc. SPIE Vol. 10948, 2019

[11] D. Shoukroun, L. Massimi, F. Iacoviello, M. Endrizzi, D. Bate, A. Olivo, and P. Fromme, "Enhanced composite plate impact damage detection and chaacterisation using x-ray refraction and scattering contrast combined with ultrasonic imaging", Composites Part B 181, 107579, 2020

[12] A. Momose, "Development toward high-resolution x-ray phase imaging", Microcopy 66, 155-166, 2017