Case Studies in Nondestructive Testing and Evaluation 6 (2016) 14–20

Contents lists available at ScienceDirect

Case Studies in Nondestructive Testing and Evaluation

www.elsevier.com/locate/csndt

# Fast inline inspection by Neural Network Based Filtered Backprojection: Application to apple inspection

Eline Janssens [a,*], Luis F. Alves Pereira [b], Jan De Beenhouwer [a], Ing Ren Tsang [b], Mattias Van Dael [c], Pieter Verboven [c], Bart Nicolaï [c], Jan Sijbers [a]

[a] iMinds-VisionLab, University of Antwerp, Universiteitsplein 1, Antwerp, Belgium
[b] Centro de Informática, Universidade Federal de Pernambuco, Av. Jornalista Anibal Fernandes s/n, Recife, Brazil
[c] BIOSYST-MeBios, KU Leuven, Willem De Croylaan 42, Heverlee, Belgium

ARTICLE INFO

*Article history:*
Available online 31 March 2016

ABSTRACT

Speed is an important parameter of an inspection system. Inline computed tomography systems exist but are generally expensive. Moreover, their throughput is limited by the speed of the reconstruction algorithm. In this work, we propose a Neural Network-based Hilbert transform Filtered Backprojection (NN-hFBP) method to reconstruct objects in an inline scanning environment in a fast and accurate way. Experiments based on apple X-ray scans show that the NN-hFBP method allows to reconstruct images with a substantially better tradeoff between image quality and reconstruction time.



## 1. Introduction

The fruit industry employs various inspection techniques to assure the quality of their products. Common defects in, for example, apples are browning and the appearance of holes. It is desirable to detect these defects to evaluate the quality of the apple. Several inspection techniques exist to evaluate this quality. The most common are Magnetic Resonance Imaging (MRI) [1], X-ray Imaging [2], Color cameras [3] and Near Infrared (NIR) reflectance and fluoroscopy [4,5]. Color cameras can only inspect the outside of fruit. While MRI generates excellent images, it is a very expensive and slow technique and therefore not suitable for fast inspection of large amounts of apples. NIR reflectance and fluoroscopy on the other hand are fast, but it is difficult to define both the type of defect and the location at the same time. Therefore, a suitable inspection technique for fast inspection of apples is X-ray imaging.

The simplest method for inspection using X-rays is radiography. Here, a two-dimensional radiographic projection of the apple is acquired, after which analysis can be performed directly on the acquired projection. Advantages of this technique are that the method is very fast, the infrastructure relatively inexpensive and it allows detecting certain defects in samples like foreign objects, pits in cherries [6], etc.

A single X-ray projection, however, does not provide depth information and features of interest may be hidden before or behind other structures such as the core of the apple. Hence, for optimal inspection, 3D information is needed. To provide this information, Computed Tomography (CT) is used. CT obtains multiple projections from a large number of angles and combines them with a reconstruction algorithm to create a detailed reconstruction image. The most common reconstruction algorithms are divided into two groups: analytical and iterative algorithms. Filtered Back Projection (FBP) [7]

* Corresponding author.
*E-mail addresses:* eline.janssens@uantwerpen.be (E. Janssens), lfap@cin.ufpe.br (L.F. Alves Pereira), tir@cin.ufpe.br (I.R. Tsang), bart.nicolai@biw.kuleuven.be (B. Nicolaï).

http://dx.doi.org/10.1016/j.csndt.2016.03.003

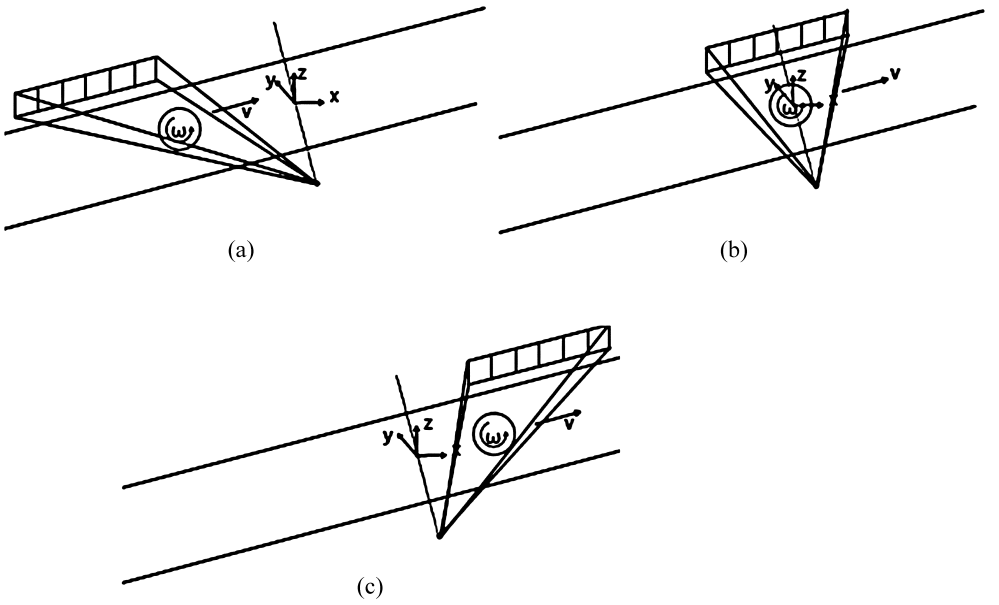


**Fig. 1.** Experimental setup for the inline inspection of apples on a conveyor belt. The apple translates with a speed v and rotates around its central axis with rotation speed $\omega$: (a) first projection, (b) central position, (c) last projection.

is the most common analytical reconstruction algorithm. FBP reconstructions are obtained very fast, but projections from a large number of angles are required to obtain a high quality reconstruction. Iterative reconstruction algorithms like SIRT, SART [7] and DART [8], discretise the reconstruction problem into a set of linear equations and, by solving these equations, a final reconstruction image is obtained. Iterative methods often generate better reconstructions for limited data problems, but at the expense of a longer reconstruction time.

In this paper, an inline inspection system is proposed where an apple moves with a constant speed through a source/detector system where the source is fixed and the detector moves over a fixed distance along with the apple as shown in Fig. 1. To avoid the limited wedge problem, the apple rotates around its central axis while it moves on the belt. This way, projections from the whole angular range can be taken very fast without tempering the throughput of the inspection system. The range of projections can then be chosen freely, dependent on the rotation speed $\omega$.

However, to maintain a high throughput, the reconstruction algorithm should be very fast. Recently, an FBP-based reconstruction algorithm, the neural network FBP (NN-FBP) [9,10] was introduced. The method computes different FBP reconstructions, each with a specific filter, and combines them into the final reconstruction image. The different filters are first trained by a neural network. Since the reconstruction algorithm is FBP-based, the method is very fast while adequate reconstructions are provided even for limited data problems. The method is, however, only applicable for parallel-beam data in a circular scanning environment. Therefore, a first version of the NN-hFBP was introduced in [11] for a circular scanning setup. In this paper, we extend the Neural Network Hilbert-transform based FBP (NN-hFBP) for an inline inspection geometry.

## 2. Methods

In this section, artificial neural networks (Section 2.1) and the Hilbert transform based FBP (Section 2.2) are presented, which are needed to understand the new NN-hFBP described in Section 2.3.

### 2.1. Artificial Neural Networks

An Artificial Neural Network (ANN) is a mathematical model that approximates an unknown function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ by learning from a training dataset. One type of ANN is a multilayer perceptron, which consists of three layers as shown in Fig. 2: an input layer with $n$ input nodes, a hidden layer with $N$ hidden nodes and an output layer with $m$ output nodes. Every input node is connected to every hidden node according to a weight matrix $\boldsymbol{W} \in \mathbb{R}^{n \times N}$ and every hidden node is connected to every output node according to a weight matrix $\boldsymbol{Q} \in \mathbb{R}^{N \times m}$. Furthermore, on every hidden node and output node an activation function is applied, which allows the network to model non-linear functions as well. In this paper, we use the sigmoid function as the activation function. To alter the decision boundary of the sigmoid function, a bias $b$ is subtracted from every hidden and output node before the activation function is applied. This allows the network to approximate more functions.

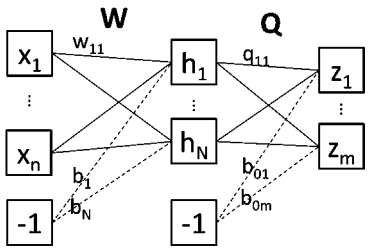


**Fig. 2.** A multilayer perceptron with n input nodes, N hidden nodes and m output nodes.

In this paper, the network is used to reconstruct a single pixel. This means that the network has only one output node and the value of this node corresponds to the gray value of the reconstructed pixel. Mathematically, the function $f$ that describes the ANN from the input $\boldsymbol{z}$ to the output $o$ is defined as follows:

$$f(\boldsymbol{z}) = o_{\boldsymbol{W},\boldsymbol{q},\boldsymbol{b},b_0}(\boldsymbol{z}) = \sigma\left(\sum_{i=0}^{N-1} q_i \sigma(\boldsymbol{w_i} \cdot \boldsymbol{z} - b_i) - b_0\right) \tag{1}$$

To approximate a certain function, the ANN should be trained. Therefore, a training dataset should be available with input vectors $\boldsymbol{z}_k$ and their corresponding output gray values $\boldsymbol{y}_k$. Then, the weight matrices $\boldsymbol{W}$ and $\boldsymbol{Q}$ are updated by the Levenberg–Marquardt algorithm, after calculating the sum of squared errors between the correct outputs $\boldsymbol{y}_k$ and the outputs obtained by the ANN.

### 2.2. Hilbert-transform based FBP

FBP is a very fast analytical reconstruction technique that is commonly used for reconstructing X-ray images when many projections are available. In this work, we use a variant of the conventional FBP algorithm, namely the Hilbert transform-based FBP (hFBP) [12] for fan-beam projection data. We have chosen this method, because, after reordering the algorithm, it can be written as a multiplication of certain input data with a filter. This will be useful in the next section. Another possibility is to rebin fan-beam data first into parallel-beam data. However, this would slow down the reconstruction and could cause interpolation artefacts.

The hFBP algorithm consists of three steps. First, the Hilbert transform of the projection data is calculated. The discrete Hilbert transform for a scanning geometry with a source and a flat panel detector is given by [13]

$$p_H^{fan}(u_i, \beta_j) = \frac{\Delta\sqrt{u^2 + D^2}}{\pi} \sum_{j,i\neq j} \frac{p^{fan}(v_j, \beta)}{(u_i - v_j)\sqrt{v_j^2 + D^2}} \tag{2}$$

In Eq. (2), $u$ and $v$ are detector pixels, $\beta$ is the projection angle, $D$ the distance from the source to the detector and $\Delta$ the sampling grid interval with respect to the variable $u$. In the second step, the differential with respect to $u$ is taken by the central difference.

$$p_{F_0}^{fan}(u_i, \beta_j) = \frac{\sqrt{(u_i^2 + D^2)^3}}{D^3}\left[\frac{p_H^{fan}(u_{i+1}, \beta_j) - p_H^{fan}(u_{i-1}, \beta_j)}{2\Delta}\right] \tag{3}$$

The final reconstruction is then obtained in the third step by back projecting the differential onto the reconstruction plane.

### 2.3. Neural Network hFBP for inline fan-beam data

The neural network hFBP (NN-hFBP) is based on the NN-FBP introduced by D. Pelt et al. in 2013 [9,10]. The NN-FBP algorithm uses a neural network to train several filters. A number of FBP reconstructions are then made, each with one of the trained filters. Afterwards, the FBP reconstructions are combined into a final image. To replace the FBP in the NN-FBP with the hFBP for fan-beam data, the algorithm should be rewritten as the product of a position and angle independent filter and a certain input as shown in Eq. (4). The input is created based on the projection data by performing several angle and position dependent multiplications from the algorithm. The input consists of 2 vectors of $M$ input data, with $M$ the number of detector pixels. It is obtained as seen in Eq. (4). The filter can afterwards be trained by the neural network. After reorganizing the projection data of Eq. (3) and replacing $\boldsymbol{v_j}$ by $\boldsymbol{u_i} - \boldsymbol{\tau}$ with $\tau$ the number of pixels between detector pixels $\boldsymbol{v_j}$ and pixels $\boldsymbol{u_i}$, the final reconstruction image can be obtained as follows:

$$f(r, \varphi) = \sum_{\tau} h_1(\tau) I_1(u_j - \tau) + h_2(\tau) I_2(u_j - \tau)$$

$$= \sum_{\beta_j \in B} \left( \sum_{\tau} \left( h_1(\tau) \sqrt{u_{j+1}^2 + D^2} + h_2(\tau) \sqrt{u_{j-1}^2 + D^2} \right) \cdot \frac{p^{fan}(u_j - \tau, \beta_j) \sqrt{(u_j^2 + D^2)^3}}{D^3 \sqrt{(u_j - \tau\Delta)^2 + D^2}} \right) \tag{4}$$

with $u_j$ the detector position where the ray that passes through $(r, \varphi)$ hits the detector at projection angle $\beta_j$. The two filters $h_1(\tau)$ and $h_2(\tau)$ are defined by Eq. (5) and Eq. (6), respectively:

$$h_1(\tau) = \frac{1}{16\pi^2 \Delta(\tau + 1)} \tag{5}$$

$$h_2(\tau) = -\frac{1}{16\pi^2 \Delta(\tau - 1)} \tag{6}$$

In an inline geometry, the projection data can be expressed in function of the detector pixel $u$ and the distance $d$ over which the object is translated compared to the central position. From every detector pixel $u$ in Eq. (4), the distance over which the apple is translated at that projection should be subtracted to calculate the correct hFBP. The filters of Eqs. (5) and (6) are independent of the detector pixel and the translation distance and are the filters that will be trained by the neural network. The inputs of Eq. (4), adapted to the inline geometry, can then be used as input data for the neural network so that the hFBP reconstruction consists of the multiplication of half the input data with the first filter and half the input data with the second filter. Summing the results yields the final reconstruction value. Hence, the input data is twice as large as the number of detector pixels. By substituting the input values into the neural network, the output value $o$ is calculated as follows:

$$o_{w,b} = \sigma \left( \sum_{l=0}^{N_h - 1} q_l \sigma \left( \begin{array}{l} \sum_{k=1}^{N} h_{l1k} \sum_{d_j \in Y} \frac{p^{fan}(u_j - \tau_k, d_j)}{\sqrt{(u_j - \tau_k \Delta - d_j)^2 + D^2}} \frac{\sqrt{(u_{j+1} - d_j)^2 + D^2} \sqrt{((u_j - d_j)^2 + D^2)^3}}{D^3} + \\ \sum_{k=1}^{N} h_{l2k} \sum_{d_j \in Y} \frac{p^{fan}(u_j - \tau_k, d_j)}{\sqrt{(u_j - \tau_k \Delta - d_j)^2 + D^2}} \frac{\sqrt{(u_{j-1} - d_j)^2 + D^2} \sqrt{((u_j - d_j)^2 + D^2)^3}}{D^3} - b_k \end{array} \right) - b_o \right) \tag{7}$$

In Eq. (7), $Y$ is the set of displacement values, $N$ the number of detector pixels and $h_{l1k}$ and $h_{l2k}$ the elements of the weight matrix $W$. After training the network with a specific training and validation dataset, it can be used to reconstruct new images by performing several hFBP reconstructions with the filters trained by the neural network and combining them according to the trained weights $q$ so that a good reconstruction of the interior of the apple can be obtained in a very fast way.

## 3. Experiments and results

The proposed methodology was validated by comparing the performance of NN-hFBP compared to FBP and SIRT, both in terms of reconstruction quality and reconstruction time. Due to the absence of real data, the method was validated on semi-real data. This data was obtained from 27 Jonagold apples that were scanned in a circular scanning geometry with 470 projections over a range of 188 degrees. The semi-real data was then obtained by converting the circular scans of the Jonagold apples into linear data scans to mimic the acquisition geometry of Fig. 1. As ground truth images for the training, validation and evaluation, we used a SIRT reconstruction with 200 iterations from the circular scans. Specifications of the imaging geometry were a detector of 1148 pixels with 127 μm pixelpitch and a fixed X-ray source. The projection data was then downsampled 4 times to 287 detector pixels to reduce the reconstruction time. Since the browning areas and holes that should be detectable in this stage were relatively large, downsampling the data did not result in worse detection but instead sped up the algorithm 4 times. The image was reconstructed on a $256 \times 256$ pixel grid and the first projection (Fig. 1(a)) was taken when the apple was located 25 cm before the central position (Fig. 1(b)) on the conveyor belt between the source and the detector. The last image (Fig. 1(c)) was taken 25 cm after this position, so that images were acquired over a distance of 50 cm. The network was trained on 100000 pixels from 100 training images of 10 apples and 10000 pixels from 10 validation images of 10 apples and the network was tested on 100 images from 8 apples. The neural network contained 4 hidden nodes, the source object distance is 84.527 mm and the object detector distance is 1 m. The values were chosen in accordance with an experimental setup that is still under construction to approximate the real data experiments as close as possible. The ASTRA toolbox [14,17] was used to generate the reconstructions.

The reconstruction quality was evaluated with three generic quality measures: the root mean squared error (RMSE), the mean apparent distortion (MAD) [15] and the Feature similarity index (FSIM) [16]. The RMSE was both evaluated on the whole image and only on the apple region while MAD and FSIM were evaluated on the whole image. These measures take into account the perception of the human visual system (HVS) to compare the reconstructions with the ground truth. They therefore better reflect visual perception of the image quality. The MAD combines two measures into one, based on the detection based strategy of the HVS and the appearance based strategy of the HVS. The detection based strategy states that, when the distortions are near-threshold, the HVS tends to look past the image to the distortions. These distortions are calculated based on the local luminance and contrast masking. The appearance based strategy on the other hand states that, when the distortions are dominant, the image tends to look past the distortions to detect the image. These distortions

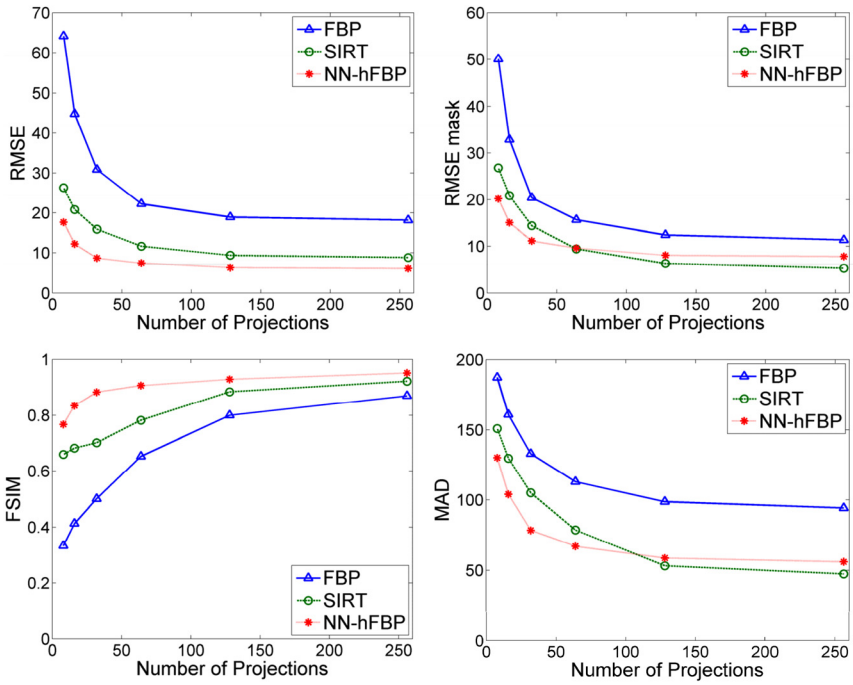


**Fig. 3.** Evaluation of FBP, SIRT and NN-hFBP reconstructions for different number of equidistant projections with different quality measures: (a) RMSE over the whole image, (b) RMSE on the apple, (c) FSIM and (d) MAD.

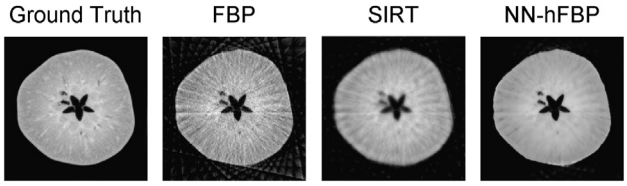


**Fig. 4.** Ground Truth image and reconstructions from 32 projections with FBP, SIRT and NN-hFBP with equidistant projections.

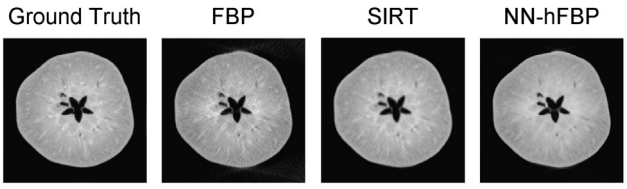


**Fig. 5.** Ground Truth image and reconstructions from 32 projections with FBP, SIRT and NN-hFBP without equidistant projections.

are calculated by looking at the changes in the local statistics of the spatial-frequency components. The FSIM uses the idea that the HVS understands an image based on its low-level features. The two features that are exploited here are the phase congruency and the image gradient magnitude. Combining these in a final measure gives a value between 0 and 1.

We performed an experiment for a limited number of equidistantly spaced angles in an inline environment. In such a geometry, the projection angles are not equiangularly distributed. While the apple is moving on the conveyor belt, it will counterclockwise rotate around its central axis (Fig. 1). The apple rotates at a constant speed $\omega = 2\pi/50$ cm so that it will have rotated $2\pi$ around its central axis between the first and the last projection. The projections were taken at equidistant positions, so that the projection angles were not equiangularly distributed. They covered a range larger than $2\pi$. In Fig. 3, the four quality measures are evaluated for the three different reconstruction algorithms in function of the number of projections. Reconstructions with 32 projections and 128 projections are shown in Figs. 4 and 5. Table 1 gives the values of the different quality measures for the different reconstruction techniques for 32 and 128 projections and Fig. 6 shows the reconstruction time of the three reconstruction algorithms in function of the number of projections.

From Fig. 3, one can observe that, for a low number of projections, the NN-hFBP obtains better reconstructions than SIRT and FBP in terms of the RMSE, FSIM and MAD when looking at the whole image. Note that 32 projections are already enough to obtain high quality reconstructions to detect holes in the apple. For a high number of projections, the NN-hFBP

**Table 1**
Average RMSE over the whole image, FSIM and MAD for the different reconstruction methods with 32 and 128 projections.

| Reconstruction method | RMSE | | FSIM | | MAD | |
|---|---|---|---|---|---|---|
| | 32 proj | 128 proj | 32 proj | 128 proj | 32 proj | 128 proj |
| FBP | 30.892 | 18.942 | 0.501 | 0.800 | 132.743 | 98.735 |
| SIRT | 15.920 | 9.390 | 0.702 | 0.885 | 105.039 | **53.227** |
| NN-hFBP | **8.677** | **6.289** | **0.883** | **0.929** | **78.331** | 58.709 |

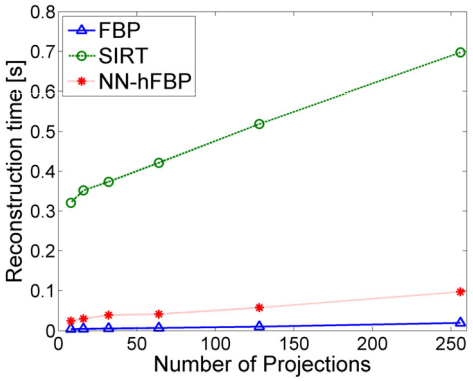


**Fig. 6.** Reconstruction time of FBP, SIRT and NN-hFBP in function of the number of projections.

reconstruction is better than FBP and SIRT in terms of the reconstruction quality evaluated by the RMSE and the FSIM. From the difference between the RMSE and the masked RMSE, it can be concluded that the NN-hFBP is able to filter out the noise in the background of the image as the RMSE of the NN-hFBP is still lower than the one of FBP or SIRT, but the masked RMSE is not. Furthermore, for a high number of projections the SIRT reconstruction of the apple has a better reconstruction quality than the one of NN-hFBP in terms of the MAD and the masked RMSE. This can be explained by the characteristics of the FBP. An FBP reconstruction is best when the angles are equiangularly distributed. Since the NN-hFBP is based on FBP reconstructions, it holds the same property. Although the quality of SIRT is better than the one of NN-hFBP, undesired artefacts are visible in both reconstructions. With respect to the reconstruction time (see Fig. 6), we see that NN-hFBP is much faster than SIRT. This allows a higher throughput of the inspection system. We can conclude that for a number of projections up to 64 projections, the NN-hFBP outperforms the FBP and SIRT reconstructions both in reconstruction quality and reconstruction time. For a number of projections larger than 100, the quality of the SIRT reconstruction is better than the one of NN-hFBP according to the masked RMSE and the MAD, but this is compensated by the short reconstruction time of NN-hFBP.

The quality measures that were used to evaluate the NN-hFBP reconstruction technique are general measures to evaluate image quality. For inline inspection, dependent on the specific application and the details that are required in the reconstruction, it is useful to use more specific quality measures to evaluate the technique. However, often in inline inspection, after the reconstruction step, a postprocessing step is needed to evaluate whether the quality of the apple is sufficient or the apple should be removed from the conveyor belt. Generally it can be stated that the time of this postprocessing step depends on the quality of the reconstruction. Hence, it is reasonable to assume that if the reconstructions are better according to the RMSE, FSIM and MAD, also the post-processing will benefit from the improved input image quality. Furthermore, the advantage of the NN-hFBP is that due to the training phase, the post processing step can be incorporated in the reconstruction, which allows for a very fast decision. It is, for example, possible to train the network on segmented images, so that the output of the network will also be segmented. We could also train the network to classify pixels and by optimizing the training phase, the output of the NN-hFBP can be a map of those classes. The post processing will then become trivial which allows a very fast inspection.

## 4. Conclusion

To conclude, we can state that the NN-hFBP has a high potential for fast inline inspection of apples. For a small number of projections taken, the NN-hFBP method outperforms the SIRT and FBP reconstruction methods in terms of reconstruction quality and reconstruction time. For a higher number of projections, the SIRT reconstruction is slightly better, but the very short reconstruction time of the NN-hFBP compensates the slightly lower reconstruction quality. Furthermore, due to its capacity to combine reconstruction and postprocessing, the NN-hFBP is very promising for inline inspection of apples.

## Acknowledgements

This work was supported by the SBO TomFood project of the Agency for Innovation by Science and Technology in Flanders (IWT SBO 120033).

## References

[1] Herremans E, Melado-Herreros A, Defraeye T, Verlinden B, Hertog M, Verboven P, et al. Comparison of X-ray CT and MRI of watercore disorder of different apple cultivars. Postharv Biol Technol 2014;87:42–50.
[2] Herremans E, Verboven P, Bongaers E, Estrade P, Verlinden BE, Wevers M, et al. Characterisation of 'Braeburn' browning disorder by means of X-ray micro-CT. Postharv Biol Technol 2013;75:114–24.
[3] Xiao-bo Z, Jie-wen Z, Yanxiao L, Holmes M. In-line detection of apple defects using three color cameras system. Comput Electron Agric 2010;70:129–34.
[4] Clark CJ, Mcglone VA, Jordan RB. Detection of brownheart in braeburn apple by transmission NIR spectroscopy. Postharv Biol Technol 2003;28:87–96.
[5] Kim MS, Chen YR, Cho BK, Chao K, Yang CC, Lefcourt AM, et al. In-line detection of apple defects using three color cameras system. Comput Electron Agric 2010;7:129–34.
[6] Haff RP, Toyofuku N. X-ray detection of defects and contaminants in the food industry. Sens Instrumen Food Qual 2008;2:262–73.
[7] Kak AC, Slaney M. Principles of computerized tomographic imaging. Philadelphia, USA: SIAM; 2001.
[8] Batenburg KJ, Sijbers J. DART: a practical reconstruction algorithm for discrete tomography. IEEE Trans Image Process 2011;20:2542–53.
[9] Pelt D, Sijbers J, Batenburg KJ. Fast tomographic reconstruction from highly limited data using artificial neural networks. In: 1st international conference on tomography of materials and structures (ICTMS). IEEE; 2013. p. 109–12.
[10] Pelt D, Batenburg KJ. Fast tomographic reconstruction from limited data using artificial neural networks. IEEE Trans Image Process 2013;22:5238–51.
[11] Janssens E, Pelt D, De Beenhouwer J, Van Dael M, Verboven P, Nicolai B, et al. Fast neural network based X-ray tomography of fruit on a conveyor belt. Chem Eng Trans 2015;44:181–6.
[12] You J, Zeng GL. Hilbert transform based FBP algorithm for fan-beam CT full and partial scans. IEEE Trans Med Imaging 2006;26:190–9.
[13] You J, Zeng GL, Liang Z. FBP algorithms for attenuated fan-beam projections. Inverse Probl 2005;21:1801.
[14] Palenstijn WJ, Batenburg KJ, Sijbers J. Performance improvements for iterative electron tomography reconstruction using graphics processing units (GPUs). J Struct Biol 2011;176(2):250–3.
[15] Larson EC, Chandler DM. Most apparent distortion: full-reference image quality assessment and the role of strategy. J Electron Imaging 2010;19(1):011006.
[16] Zhang L, Zhang L, Mou X, Zhang D. FSIM: a feature similarity index for image quality assessment. IEEE Trans Image Process 2011;20:2378–86.
[17] Van Aarle W, Palenstijn WJ, De Beenhouwer J, Altantzis T, Bals S, Batenburg KJ, et al. The ASTRA Toolbox: a platform for advanced algorithm development in electron tomography. Ultramicroscopy 2015;157:35–47.