

A Tensor Factorization Method for 3D Super-Resolution with Application to Dental CT

Janka Hatvani, Adrian Basarab, Jean-Yves Tournet, Miklós Gyöngy and Denis Kouamé

Abstract—Available super-resolution techniques for 3D images are either computationally inefficient prior-knowledge-based iterative techniques or deep learning methods which require a large database of known low- and high-resolution image pairs. A recently introduced tensor-factorization-based approach offers a fast solution without the use of known image pairs or strict prior assumptions. In this article this factorization framework is investigated for single image resolution enhancement with an off-line estimate of the system point spread function. The technique is applied to 3D cone beam computed tomography for dental image resolution enhancement. To demonstrate the efficiency of our method, it is compared to a recent state-of-the-art iterative technique using low-rank and total variation regularizations. In contrast to this comparative technique, the proposed reconstruction technique gives a 2-order-of-magnitude improvement in running time – 2 minutes compared to 2 hours for a dental volume of $282 \times 266 \times 392$ voxels. Furthermore, it also offers slightly improved quantitative results (peak signal-to-noise ratio, segmentation quality). Another advantage of the presented technique is the low number of hyperparameters. As demonstrated in this paper, the framework is not sensitive to small changes of its parameters, proposing an ease of use.

Index Terms—3D super-resolution, single image super-resolution, tensor factorization, cone beam computed tomography, dental application

I. INTRODUCTION

ROOT canal treatment is carried out on a regular basis in dental centers in order to save decayed and infected teeth. In spite of their popularity, the success rate of the treatment is only 60-85% [1], [2]. For an improvement of the process the dentists need a better visualization of the canal, as its length, diameter and curvature are all important factors for planning the therapy [3]. Therefore, further research on the visualization possibilities of the pulp cavity is necessary, as stated by the European Commission on Radiation Protection in 2012 [4, pp. 61-65].

This work has been submitted to the IEEE for possible publication. Copyright may be transferred without notice, after which this version may no longer be accessible.

This research has been partially supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.3-VEKOP- 16-2017-00002 and by Pázmány University KAP17-19.

This work was supported by the thematic trimester on image processing of the CIMI Labex, Toulouse, France, with the Program ANR-11-IDEX-0002-02 under Grant ANR-11-LABX-0040-CIMI.

J. Hatvani, A. Basarab, and D. Kouamé are with IRIT, CNRS UMR 5505, Universit Paul Sabatier Toulouse 3, University of Toulouse, France.

J. Hatvani, and M. Gyöngy are with the Faculty of Information Technology and Bionics, Pazmany Peter Catholic University, Budapest Hungary

J. Y. Tournet is with IRIT/INP-ENSEEIH/Tésa, University of Toulouse, France.

Dental offices use cone beam computed tomography (CBCT) for determining the 3D structure of the teeth. Apart from the detector size, the spatial resolution of such imaging devices is also affected by partial volume effect, noise, and beam hardening, resulting in a typical value of 500 μm . This resolution is not sufficient in endodontology since the diameter of the canal is usually in the range of 0.16-1.60 mm and the apical, narrower segment is more important for planning the treatment [5]. On the other hand, the resolution of micro-CT (μCT) is sufficient for precise measurements on the cavity, but the physical dimensions of the system only permit the imaging of extracted teeth. The long acquisition time and high radiation dose also prohibit *in vivo* measurements.

The post-processing application of image super-resolution (SR) algorithms is an intensely investigated field in the image processing community. Classical SR techniques can combine information from a sequence of measurements [6], from different modalities [7], or in the simplest case they try to improve the resolution of a single image [8]. Many SR methods assume that the low resolution (LR) image of interest is obtained from the high resolution (HR) image by blurring and decimation with a residual additive noise. The SR problem can then be formulated as an inverse problem, which is ill-posed and thus requires an appropriate regularization to provide suitable solutions. A standard regularization often employed is total variation (TV) leading to piecewise smooth solutions [9]. Low-rank [10], or wavelet representations [11] have also proved to be efficient tools for SR. A method based on a sparse representation was applied to 3D MRI images with a patch-based structural similarity constraint in [12]. Convolutional neural networks have also shown interesting properties for SR, where the network is trained to map an LR image to its HR counterpart [13]–[16]. However, this technique requires a large training dataset, which is not always available. Furthermore, only a few of the above-mentioned techniques is available for 3D volumes (*e.g.*, [10], [12], [13]), and they all suffer from heavy computational costs, preventing them from application in practice.

A new hyperspectral-multispectral image fusion technique using tensor factorization (TF) was introduced in [17]. This technique combines a multispectral image (with high spatial and low spectral resolutions) and a hyperspectral image (with low spatial and high spectral resolution) to obtain an SR image (with high spatial and high spectral resolutions). One advantage of the tensor-based method of [17] is that it does not need to unfold the image of interest into a 2D matrix as

in many existing SR methods [10], [16]. As a consequence, this method avoids any loss of information about the locality of the image pixels and does not require to introduce spatial regularization (such as the TV of the image).

This paper investigates a 3D single image SR (SISR) method based on TF for CBCT images of teeth as an attempt to approximate their HR μ CT pairs. The idea is to decompose the image of interest using its canonical polyadic decomposition (CPD). The CPD of a tensor is a representation based on a sum of an appropriate number of rank-1 tensors, which number depends on the structure of the image. It will be shown in this paper that this representation leads to a notably fast and efficient reconstruction method. The described method is compared to a state-of-the-art iterative deconvolution technique with low-rank and TV regularization (LRTV). For validation the peak signal-to-noise ratio (PSNR) is calculated and the canal is segmented, permitting volumetric and diametric comparison.

The rest of this paper is organized as follows. First, the tensor operations used in this paper are defined and a connection between image complexity and tensor decomposition is drawn. In Section III the proposed TF-SISR method is first defined for the CBCT resolution enhancement problem, followed by the data acquisition and the estimation of the blurring point spread function (PSF), ending with the introduction of the evaluation metrics. Section IV compares the images obtained by the two different SISR methods and discusses the possibilities and limits of tensor factorization. Finally a conclusion is drawn about the applicability of the introduced SR technique to dental imaging with some possible future work.

II. TENSORS AND IMAGE COMPLEXITY

A. Notations

For easier distinction, 2D matrices are denoted using uppercase letters (e.g., A) and 3D tensors by bold uppercase letters (e.g., \mathbf{A}). The uppercase letter with an overline (eg., \overline{A}) denotes a set of 2D matrices.

B. Factorization, mode product, matricization

In this section, operations from tensor algebra necessary for the proposed method are summarized. Readers may refer to [17] and [18] for further details.

A tensor is a generalization of vectors and matrices, where the order of the tensor indicates the dimensionality. A 3D CT image volume is a third-order tensor $\mathbf{X} \in \mathbb{R}^{I \times J \times K}$ from which one dimensional fibers can be extracted. Depending on which indices are fixed, there are mode-1 fibers denoted as $\mathbf{X}(:, j, k)$ vectors (columns), mode-2 fibers denoted as $\mathbf{X}(i, :, k)$ vectors (rows) and mode-3 fibers denoted as $\mathbf{X}(i, j, :)$ vectors. The outer product (denoted by \circ) between one mode-1, one mode-2 and one mode-3 array forms a rank-1 third order tensor, written as

$$\begin{aligned} \mathbf{X} &= u \circ v \circ w, \\ u &\in \mathbb{R}^I, v \in \mathbb{R}^J, w \in \mathbb{R}^K, \mathbf{X} \in \mathbb{R}^{I \times J \times K} \\ \text{where} \\ \mathbf{X}(i, j, k) &= u(i)v(j)w(k). \end{aligned} \quad (1)$$

The smallest number of rank-1 tensors that can sum up to form the tensor \mathbf{X} is called the tensor rank of \mathbf{X} , denoted by F . The resulting factorization of \mathbf{X} is called the CPD of \mathbf{X} expressed as

$$\mathbf{X} = \sum_{f=1}^F U^1(:, f) \circ U^2(:, f) \circ U^3(:, f) \quad (2)$$

where

$$\mathbf{X}(i, j, k) = \sum_{f=1}^F U^1(i, f)U^2(j, f)U^3(k, f).$$

$\overline{U} = \{U^1, U^2, U^3\}$ is a set of three 2D matrices, $\{U^1 \in \mathbb{R}^{I \times F}, U^2 \in \mathbb{R}^{J \times F}, U^3 \in \mathbb{R}^{K \times F}\}$, known as the decomposition of the tensor \mathbf{X} . For illustration, the reader may refer to Fig. 1. In the following, the operation in (2) will be denoted as

$$[[U^1, U^2, U^3]] = \sum_{f=1}^F U^1(:, f) \circ U^2(:, f) \circ U^3(:, f). \quad (3)$$

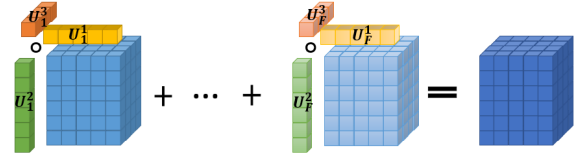


Fig. 1. Illustration of tensor factorization. F is the number of outer products formed by mode-1 ($U_i^1 := U^1(:, i)$), mode-2 ($U_i^2 := U^2(:, i)$) and mode-3 ($U_i^3 := U^3(:, i)$) fibers summing up to a rank- F tensor.

An important property of the CPD is that this decomposition is essentially unique (allowing permutations within U^1, U^2, U^3). Thus \overline{U} can be identified almost surely if its tensor rank F is smaller than an upper bound. Chiantini *et al.* [19] proved that if $I \geq J \geq K$, with $F \leq 2^{\lfloor \log_2 J \rfloor + \lfloor \log_2 K \rfloor - 2}$, the CPD of the rank- F tensor $\mathbf{X} \in \mathbb{R}^{I \times J \times K}$ is essentially unique. This condition allows identifiability of the CPD even for tensors with high rank. For example, in the application addressed herein, a typical CBCT volume with $260 \times 260 \times 300$ pixels can be decomposed uniquely even if the tensor rank of the image is as high as $2^{14} = 16384$.

Next, the multiplication between a 2D matrix and a 3D tensor referred to as the mode product is defined. This multiplication can be performed along all three dimensions, and in each case the mode- n fibers of the tensor are extracted and are pre-multiplied by the matrix one-by-one. The mode- n products ($n \in \{1, 2, 3\}$) of $\mathbf{X} \in \mathbb{R}^{I \times J \times K}$ with $P_1 \in \mathbb{R}^{I^* \times I}$, $P_2 \in \mathbb{R}^{J^* \times J}$, $P_3 \in \mathbb{R}^{K^* \times K}$ are denoted as \mathbf{x}_n , and are defined as

$$\begin{aligned} \mathbf{X} \times_1 P_1 &= \mathbf{X}_1 \in \mathbb{R}^{I^* \times J \times K} \\ \text{where } \mathbf{X}_1(:, j, k) &= P_1 \mathbf{X}(:, j, k) \\ \mathbf{X} \times_2 P_2 &= \mathbf{X}_2 \in \mathbb{R}^{I \times J^* \times K} \\ \text{where } \mathbf{X}_2(i, :, k) &= P_2 \mathbf{X}(i, :, k) \\ \mathbf{X} \times_3 P_3 &= \mathbf{X}_3 \in \mathbb{R}^{I \times J \times K^*} \\ \text{where } \mathbf{X}_3(i, j, :) &= P_3 \mathbf{X}(i, j, :) \end{aligned} \quad (4)$$

where I^*, J^*, K^* are arbitrary integer numbers. In Fig. 2, the principle of the mode-1 product, $\mathbf{X} \times_1 P_1 = \mathbf{X}_1$ is illustrated, where the columns of the tensor are premultiplied by P_1 , leading to a shrinkage along the first dimension.

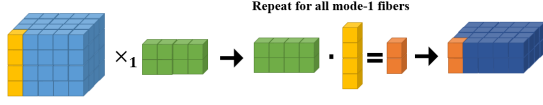


Fig. 2. Illustration of the mode-1 product. The mode-1 fibers of the 3D tensor are extracted and pre-multiplied by the 2D matrix. This example can illustrate a downsample operation with rate 2 in the first dimension.

Using the factorization of \mathbf{X} in (2) and (3) the mode- n products can also be written as

$$\mathbf{X} \times_1 P_1 \times_2 P_2 \times_3 P_3 = \llbracket P_1 U^1, P_2 U^2, P_3 U^3 \rrbracket. \quad (5)$$

Finally, the matricization or unfolding of the tensor $\mathbf{X} \in \mathbb{R}^{I \times J \times K}$ from 3D to 2D is defined. Note that this matricization can be realized again along the three dimensions. For a mode- n matricization the mode- n fibers are extracted and form the columns of $\mathbf{X}^{(n)}$ in lexicographical order expressed as

$$\begin{aligned} \mathbf{X}^{(1)} &= [\mathbf{X}(:, 1, 1), \mathbf{X}(:, 2, 1), \dots \\ &\quad \mathbf{X}(:, J, 1), \mathbf{X}(:, 1, 2), \dots, \mathbf{X}(:, J, K)] \\ \mathbf{X}^{(1)} &\in \mathbb{R}^{I \times JK} \\ \mathbf{X}^{(2)} &= [\mathbf{X}(1, :, 1), \mathbf{X}(2, :, 1), \dots \\ &\quad \mathbf{X}(I, :, 1), \mathbf{X}(1, :, 2), \dots, \mathbf{X}(I, :, K)] \\ \mathbf{X}^{(2)} &\in \mathbb{R}^{J \times IK} \\ \mathbf{X}^{(3)} &= [\mathbf{X}(1, 1, :), \mathbf{X}(2, 1, :), \dots \\ &\quad \mathbf{X}(I, 1, :), \mathbf{X}(1, 2, :), \dots, \mathbf{X}(I, J, :)] \\ \mathbf{X}^{(3)} &\in \mathbb{R}^{K \times IJ}. \end{aligned} \quad (6)$$

The same operation can be realized using the decomposition \bar{U} of \mathbf{X} . For this the Khatri-Rao product – denoted as \odot – is necessary. It operates on matrices having the same number of columns, and calculates their column-wise Kronecker-product ($A \odot B = C$, where $A \in \mathbb{R}^{I \times F}$, $B \in \mathbb{R}^{J \times F}$, $C \in \mathbb{R}^{IJ \times F}$). Using this notation the matricization can be written as

$$\begin{aligned} \mathbf{X}^{(1)} &= U^1 (U^3 \odot U^2)^T \\ \mathbf{X}^{(2)} &= U^2 (U^3 \odot U^1)^T \\ \mathbf{X}^{(3)} &= U^3 (U^2 \odot U^1)^T. \end{aligned} \quad (7)$$

C. Connection between the tensor rank and the image complexity

The tensor rank is related to the complexity of the image. To illustrate this claim, an illustrative set of examples is provided in Fig. 3. In these examples the notations and dimensionality of (1) are used with $I = J = K = 2$. Fig. 3 a) shows that the image with a single dark pixel (representing 1) in the white volume (representing 0) has a tensor rank of $F = 1$. More precisely

$$\begin{aligned} u_a \circ v_a \circ w_a &= [1, 0] \circ [0, 1] \circ [1, 0] = \mathbf{X}_a \\ \mathbf{X}_a^{(3)} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned} \quad (8)$$

In Fig. 3 b), two neighboring pixels are dark. This does not change the complexity of the image, since one outer product can still describe this volume. Indeed, we have

$$\begin{aligned} u_b \circ v_b \circ w_b &= [1, 0] \circ [0, 1] \circ [1, 1] = \mathbf{X}_b \\ \mathbf{X}_b^{(3)} &= \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned} \quad (9)$$

In Fig. 3 c), two fibers are linearly dependent with $2 \times [21, 35] = [42, 70]$, which also makes one outer product sufficient for decomposing the tensor since

$$\begin{aligned} u_c \circ v_c \circ w_c &= [5, 3] \circ [1, 2] \circ [7, 0] = \mathbf{X}_c \\ \mathbf{X}_c^{(3)} &= \begin{bmatrix} 35 & 70 & 0 & 0 \\ 21 & 42 & 0 & 0 \end{bmatrix}. \end{aligned} \quad (10)$$

This set of illustrative examples shows that for images with piecewise constant regions (like the neighboring cells in Fig. 3 b)) or with low matrix rank (as the linearly dependent fibers in Fig. 3 c)) a smaller tensor rank can be expected. More generally, the tensor decomposition (2) tends to promote solutions with small tensor ranks. This property is useful in the case of denoising, when independent outlier pixels have to be eliminated. A degraded image may contain larger constant areas, with higher dependency between neighboring rows and columns. It means that describing these images will also be more efficient with a tensor of small rank. Thus these simple examples allow us to understand why CBCT images can be represented by a reduced number of rank-1 tensors.

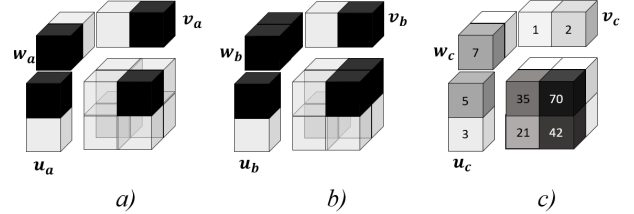


Fig. 3. Tensor rank and image complexity. In example a) a single dark pixel (representing 1) in the white (representing 0) volume can be expressed by one outer product. In b) two neighboring pixels are dark, making one outer product sufficient for their description. In c) the pixel value is printed on the cell, equals 0 if not present. Two fibers are linearly dependent ($2 \times [21, 35] = [42, 70]$), so the volume can be decomposed using a tensor rank of $F = 1$.

III. METHODS

A. Problem formulation

The image degradation model considered herein is the one classically used in SISR methods. It relates the LR image (CBCT in the case of the current dental application) to an HR image (considered to be close to the μ CT). The HR image $\mathbf{X} \in \mathbb{R}^{I \times J \times K}$ is corrupted by a decimation operator D with rate r , a blurring kernel H , and some added noise N , resulting in the LR image $\mathbf{Y} \in \mathbb{R}^{I/r \times J/r \times K/r}$ such that

$$\text{vec}(\mathbf{Y}) = DH \text{vec}(\mathbf{X}) + \text{vec}(N) \quad (11)$$

where $\text{vec}(\cdot)$ vectorizes the elements of the 3D tensor in lexicographical order. We assume that $H \in \mathbb{R}^{IJK \times IJK}$ is the block-circulant version of the 3D Gaussian kernel

h to avoid circular convolution. A 3D Gaussian kernel h is separable along the three dimensions to 1D kernels as $h = h_1 \circ h_2 \circ h_3$ and is usually assumed for a blurring PSF [20]. For Gaussian kernels h_1, h_2, h_3 with standard deviations $\sigma_1, \sigma_2, \sigma_3$ the corresponding block-circulant matrices are $H_1 \in \mathbb{R}^{I \times I}, H_2 \in \mathbb{R}^{J \times J}, H_3 \in \mathbb{R}^{K \times K}$. The decimation operator downsamples the image by an integer number, by averaging blocks of r neighboring pixels in each direction. In matrix form the downsampling operators for the three dimensions are $D_1 \in \mathbb{R}^{I/r \times I}, D_2 \in \mathbb{R}^{J/r \times J}, D_3 \in \mathbb{R}^{K/r \times K}$. This formulation of D corresponds for instance to the physical process of a large CBCT detector element collecting rays over a larger area, than μ CT does. This matrix also has better inversion properties compared to the regular decimation operator which discards pixels at a rate r .

Let $\bar{U} = \{U^1 \in \mathbb{R}^{I \times F}, U^2 \in \mathbb{R}^{J \times F}, U^3 \in \mathbb{R}^{K \times F}\}$ be the CPD of \mathbf{X} . The image degradation problem can be rewritten following (5) using the separated kernels

$$\begin{aligned} \mathbf{Y} &= \mathbf{X} \times_1 D_1 H_1 \times_2 D_2 H_2 \times_3 D_3 H_3 + \mathbf{N} \\ &= \llbracket D_1 H_1 U^1, D_2 H_2 U^2, D_3 H_3 U^3 \rrbracket + \mathbf{N}. \end{aligned} \quad (12)$$

The SR task can be defined as finding the set of matrices \bar{U} which is the solution of the following minimization problem

$$\min_{\bar{U}} \|\mathbf{Y} - \llbracket D_1 H_1 U^1, D_2 H_2 U^2, D_3 H_3 U^3 \rrbracket\|_F^2 \quad (13)$$

where $\|\cdot\|_F$ denotes the Frobenius norm of a tensor defined as the square root of the sum of its squared elements. This cost-function is different from the minimization problem of [17] in the following aspects. First, only one measured datavolume is used here in contrast to the two measurements in the fusion problem of [17]. Second, here the degradation happens in all three dimensions between the HR and LR image, while in [17] the hyperspectral measurement is degraded in the first two dimensions, the multispectral volume in the third dimension.

As problem (13) is NP-hard an alternating optimization method is investigated, minimizing the cost function in (13) sequentially for U^1, U^2, U^3 . Building a tensor from its decomposition (2) consists of a summation of F outer products. Taking the square of this sum of F terms in (13) would result in the sum of $\binom{F+2-1}{2}$ terms, where $\binom{(\cdot)}{(\cdot)}$ denotes the binomial coefficient. Instead when minimizing over U_n , the tensors are mode- n -matricized using (6) and (7) leading to

$$\begin{aligned} \min_{U^1} \frac{1}{2} \left\| \mathbf{Y}^{(1)} - D_1 H_1 U^1 (D_3 H_3 U^3 \odot D_2 H_2 U^2)^T \right\|_F^2 \\ \min_{U^2} \frac{1}{2} \left\| \mathbf{Y}^{(2)} - D_2 H_2 U^2 (D_3 H_3 U^3 \odot D_1 H_1 U^1)^T \right\|_F^2 \\ \min_{U^3} \frac{1}{2} \left\| \mathbf{Y}^{(3)} - D_3 H_3 U^3 (D_2 H_2 U^2 \odot D_1 H_1 U^1)^T \right\|_F^2. \end{aligned} \quad (14)$$

Note that the unfolding is performed in each direction sequentially, conserving the 3D local information. These minimizations can now be carried out easily, leading to

$$\begin{aligned} U^1 &= (D_1 H_1)^+ \mathbf{Y}^{(1)} (D_3 H_3 U^3 \odot D_2 H_2 U^2)^{+T} \\ U^2 &= (D_2 H_2)^+ \mathbf{Y}^{(2)} (D_3 H_3 U^3 \odot D_1 H_1 U^1)^{+T} \\ U^3 &= (D_3 H_3)^+ \mathbf{Y}^{(3)} (D_2 H_2 U^2 \odot D_1 H_1 U^1)^{+T} \end{aligned} \quad (15)$$

where $^+$ is the regularized Moore-Penrose pseudo-inverse defined as

$$A^+ = A^T A + \epsilon^2 I)^{-1} A^T \quad (16)$$

where ϵ is a hyperparameter used to provide a stable inverse (this procedure is classically referred to as diagonal loading).

To implement the proposed TF-SISR method the TensorLab toolbox [21] was used in Matlab 2017b. In the algorithm \bar{U} was initialized with elements from the standard normal distribution and U^1, U^2, U^3 were updated iteratively several times as described in Algo. 1.

Algorithm 1 TF-SISR algorithm

Input: $\mathbf{Y} \in \mathbb{R}^{I/r \times J/r \times K/r}, F, [\sigma_1, \sigma_2, \sigma_3], r$
1: **Initialize** $\bar{U} = \{U^1 \in \mathbb{R}^{I \times F}, U^2 \in \mathbb{R}^{J \times F}, U^3 \in \mathbb{R}^{K \times F}\}$ with normally distributed values
2: $D_1, D_2, D_3 \leftarrow$ decimation operator with a factor r
3: $H_1, H_2, H_3 \leftarrow$ Gaussian kernels with standard deviations $[\sigma_1, \sigma_2, \sigma_3]$
4: **while** stopping criteria is not met **do**
5: $U^1 \leftarrow Y^{(1)}, U^2, U^3$
6: $U^2 \leftarrow Y^{(2)}, U^1, U^3$ \triangleright update using (15)
7: $U^3 \leftarrow Y^{(3)}, U^1, U^2$
8: **end while**
9: $\mathbf{X} \leftarrow \bar{U}$ \triangleright build using (2)
Output: \mathbf{X} , the estimated high resolution image

B. Data acquisition

The dataset used for testing contains images of 13 teeth which were extracted for health reasons and donated anonymously for research. This set consists of all different tooth types, including incisors, canines, premolars and molars. A Carestream 81003D system was used for CBCT imaging. The linewidth resolution of the CBCT machine was 500 μm and the volumes had a voxel size of $80 \times 80 \times 80 \mu\text{m}^3$.

For evaluation purposes the reconstructed HR images were compared to μ CT images acquired from the same samples. The μ CT acquisitions were obtained with a Quantum FX system from Perkin Elmer, with a voxel size of $40 \times 40 \times 40 \mu\text{m}^3$ and linewidth resolution of 50 μm .

C. PSF estimation

The proposed method in Algo. 1 assumes that the PSF is known. In practice the blurring kernel has to be measured or estimated, which is usually carried out empirically in many existing works [9]. Here the blurring kernel was assumed to be Gaussian and its standard deviation was estimated from the observed data. Employing direct inverse filtering on each sample image, the μ CT volume was divided by the CBCT volume in the frequency domain to obtain the Fourier transform of the PSF. The high-frequency band was suppressed by a Hanning-window before computing the inverse Fourier-transform and averaging the 13 estimated PSFs. For further details see [16]. Finally a 3D Gaussian function was fitted to the averaged PSF to estimate the standard deviations σ_1, σ_2 and σ_3 .

D. Metrics

The comparison of the image volumes was carried out through two metrics. The first one measures the PSNR between the estimated SR CBCT and μ CT volumes. It is calculated by dividing the square of the dynamic range with the mean square error between the enhanced image and the μ CT image, expressed in dB.

The second, more application oriented metric consists of comparing volumes segmented from μ CT and SR CBCT. The canal root was segmented with a dedicated adaptive local thresholding method (see [22] for details). The canal area and Feret's diameter (the longest distance between two parallel straight lines that are tangent to the shape) were calculated for each radicular axial slice. The measured values are compared through the mean of absolute differences. The differences of the canal volumes were also measured using the Dice coefficient [23]. Finally MeVisLab [24] was used for visualizing the segmentation results.

IV. RESULTS

A. Comparison to an existing 3D SR method

The state-of-the-art LRTV introduced in 2013 in [10], [25] was used as a benchmark to compare the performance of the proposed method¹. Among the relatively small collection of 3D SISR techniques, LRTV provided competitive results compared to other popular methods (cubic interpolation, non-local means, TV-based up-sampling) [10]. It uses low-rank and total-variation regularizers, minimizing the following cost-function

$$\hat{\mathbf{X}} = \underset{\mathbf{X}}{\operatorname{argmin}} \|\mathbf{D}\mathbf{H}\mathbf{X} - \mathbf{Y}\|^2 + \lambda_R \operatorname{Rank}(\mathbf{X}) + \lambda_{TV} \operatorname{TV}(\mathbf{X}), \quad (17)$$

where λ_R and λ_{TV} are hyperparameters. The minimization problem 17 is solved by the alternating direction method of multipliers (ADMM), which requires to adjust two additional hyperparameters, the penalty term ρ and an iteration number n_{ADMM} . One of the subproblems within the ADMM scheme is solved by gradient descent with an additional iteration number n_{grad} and an update rate denoted as dt .

The parameters used for testing can be seen in Table I. They were tuned manually to get the highest possible improvement of the PSNR. The tests were run on a standard PC with an Intel(R) Core(TM) i7 2x2.5GHz processor and 16 GB of RAM.

TABLE I
PARAMETERS

| LRTV | TF-SISR |
|----------------------------|----------------------------|
| $n_{\text{ADMM}} = 5$ | $n_{\text{TF}} = 10$ |
| $\sigma = [5.8, 5.3, 0.9]$ | $\sigma = [5.8, 5.3, 0.9]$ |
| $\lambda_{TV} = 0.02$ | $F = 500$ |
| $\lambda_R = 0.05$ | $\epsilon = 1$ |
| $\rho = 0.05$ | |
| $n_{\text{grad}} = 100$ | |
| $dt = 0.05$ | |

¹The Matlab code associated with LRTV is available at <https://bitbucket.org/fengshi421/superresolutiontoolkit>

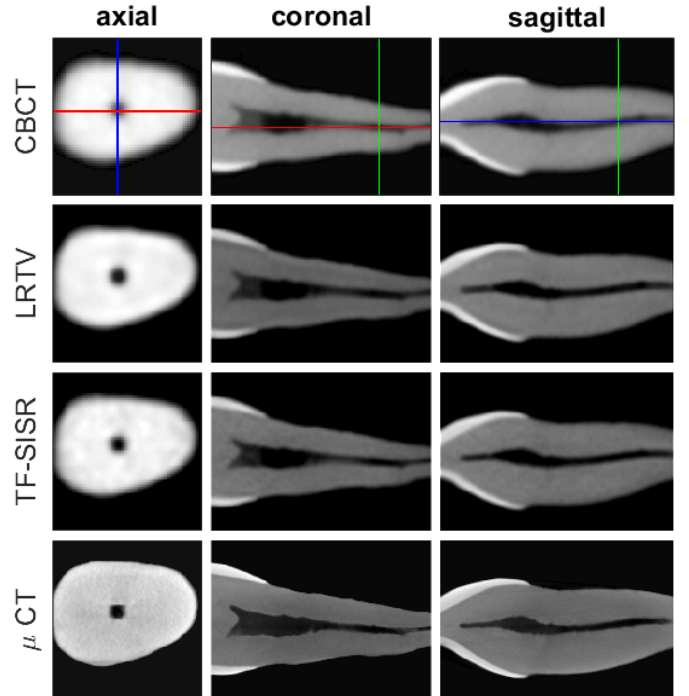


Fig. 4. Results on Sample #1. In the rows the CBCT, LRTV output, TF-SISR output and μ CT images can be seen, whereas the columns correspond to one axial, a coronal and a sagittal slice. The CBCT image is shown on the same scale as the HR images, for better comparison. The location of the slices within the volume is illustrated on the CBCT images in colored lines.

The two methods were tested for three samples from the dataset, including an incisor, a premolar and a molar. The sizes of the sample volumes, the PSNR calculated against the μ CT images and the execution times are provided in Table II. Compared to the CBCT images the PSNR improves similarly for the LRTV (average of 1.2 dB) and the TF-SISR (average of 1.5 dB) methods with the chosen parameters. However, this enhancement is achieved at a much lower computational cost: 10 iterations of TF-SISR run 100 times faster than 5 iterations of LRTV. This faster execution time is important since it permits a wider range of applications, including those requiring a rapid diagnosis during a medical examination. In Fig. 4, the quality of the enhanced volumes is visualized, showing that the canal is better defined and contrasted compared to the CBCT image, suggesting better segmentation properties.

TABLE II
TEST RESULTS

| | Sample #1 | Sample #2 | Sample #3 |
|---------------------|-----------------------------|-----------------------------|-----------------------------|
| tooth type | upper incisor | lower premolar | lower molar |
| μ CT image size | $282 \times 266 \times 392$ | $280 \times 268 \times 492$ | $324 \times 306 \times 402$ |
| CBCT PSNR | 23.17 dB | 22.67 dB | 24.14 dB |
| LRTV PSNR | 24.32 dB | 24.65 dB | 24.61 dB |
| TF-SISR PSNR | 24.32 dB | 24.48 dB | 25.71 dB |
| LRTV time | 6988 s | 9059 s | 10301 s |
| TF-SISR time | 71 s | 92 s | 104 s |

For further analysis the root canal was segmented from each

TABLE III
CANAL SEGMENTATION METRICS

| | method | Sample #1 | Sample #2 | Sample #3 | mean |
|---|---------|-----------|-----------|-----------|--------|
| Mean of Diff. - Feret (μm) | CBCT | 96 | 89 | 341 | 176 |
| | LRTV | 74 | 71 | 196 | 113 |
| | TF-SISR | 50 | 57 | 177 | 95 |
| Mean of Diff. - Area (mm^2) | CBCT | 0.0463 | 0.0461 | 0.2492 | 0.1139 |
| | LRTV | 0.0914 | 0.0920 | 0.2350 | 0.1395 |
| | TF-SISR | 0.0447 | 0.0271 | 0.2243 | 0.0987 |
| Dice coefficient | CBCT | 0.88 | 0.88 | 0.90 | 0.88 |
| | LRTV | 0.87 | 0.88 | 0.90 | 0.89 |
| | TF-SISR | 0.90 | 0.91 | 0.91 | 0.90 |

volume, using the segmentation method described in Section III-D. Qualitative and quantitative results are provided in Fig. 5 and in Table III. In particular, Table III shows differences between the estimates and the values obtained using the μCT image for three parameters (Feret diameter, area of the canal and Dice coefficient). The estimated Feret diameter improves similarly with both SR techniques compared to the CBCT images with an averaged improvement of $63 \mu\text{m}$ for LRTV and $81 \mu\text{m}$ for TF-SISR. The second line of the table shows how the area of the canal on the axial slices is changing from one method to another. Note that the LRTV method shows a higher difference compared to the original CBCT (by 0.0256 mm^2), suggesting that the TV regularization overestimates the canal. This observation is also confirmed in Fig. 5, as the LRTV volumes have a more blueish color corresponding to positive differences. The TF-SISR provides the best overall performance with an improvement of 0.0152 mm^2 on average. The last metric in the table is the Dice-coefficient, also showing some improvement in the overlap of the canals, by 1% using the LRTV and 2% with the TF-SISR. Fig. 5 displays zoomed-in sections of the apical part of the canal, as this part is the most important during the treatment. Considering these results, the TF-SISR method shows slightly better segmentation properties than the LRTV technique, while offering a great reduction in running time.

B. Adjusting the parameters of the TF-SISR method

The impact of the tensor rank and the iteration number was investigated using Sample #1. Fig. 6 a) shows that the runtime increases linearly with the number of iterations, as expected. Fig. 6 a) also shows that the PSNR converges rapidly to its maximum value (close to 24.5), which is an interesting property of the proposed method. Fig. 7 a) shows how the solutions qualitatively evolve with the iteration number. For improved visibility the difference from the $n_{\text{TF}} = 10$ case is shown in the figure. In the case of $n_{\text{TF}} = 10$ a second test was run, and the difference was calculated compared to this result, as the random initialization of \bar{U} results in slightly different outputs. It can be seen that as the iteration number increases, the difference becomes less structured: in the third column the shape of the tooth is almost invisible and it is lost in random noise.

According to the upper limit of F for a unique CPD, in the case of Sample #1, $F \leq 2^{14}$ should be efficient. However, numbers higher than 2000 caused memory problems,

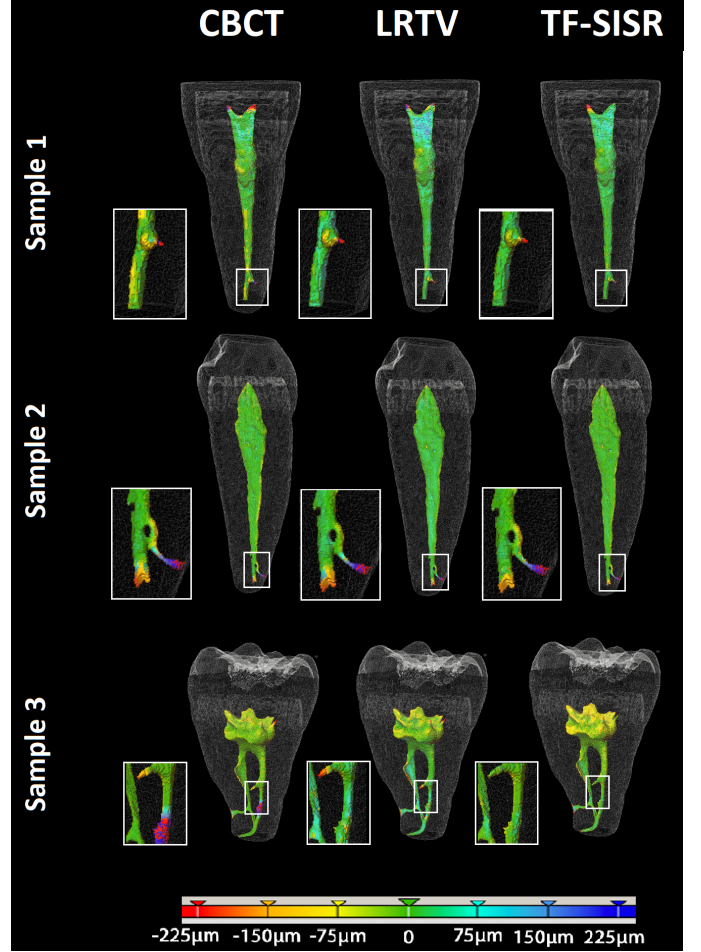


Fig. 5. Segmentation results for CBCT, LRTV and TF-SISR for the 3 samples. The color-bar visualizes the distance between the estimated surface of the canal and the one obtained with μCT segmentation.

and were therefore not tested. Fig. 6 b) shows that the computational time increases exponentially with the rank F since the algorithm requires the inversion of larger matrices in \bar{U} . It can also be seen that the PSNR stabilizes for ranks larger than $F = 500$. Some sample images can be seen in Fig. 7 b) showing that low values of the rank F lead to large constant

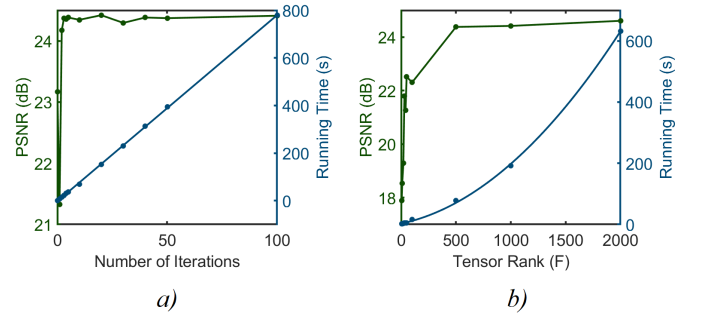


Fig. 6. Effect of the iteration number and the tensor rank on the PSNR and runtime. The rest of the parameters are as in Table I. In a) the PSNR saturates after a small number of iterations, while the runtime increases linearly. In b) the runtime has an exponential growth versus the tensor rank and the PSNR saturates around $F = 500$.

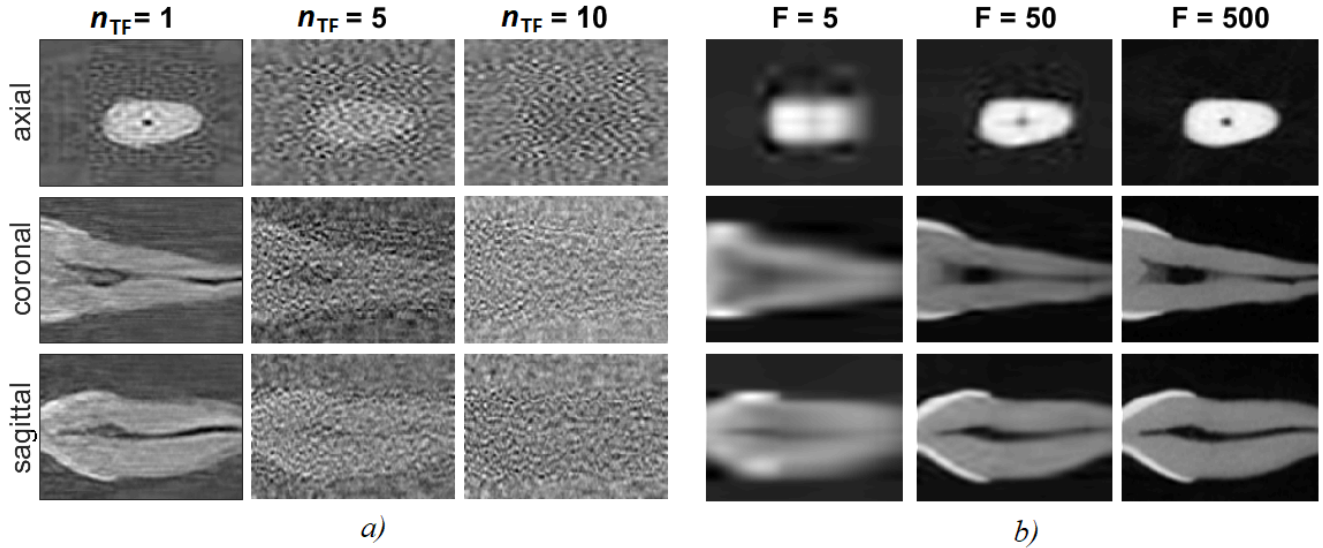


Fig. 7. Effect of the iteration number it and the tensor rank F in the reconstructed images. Sample #1 is visualized through 3 slices from the axial, coronal and sagittal directions. In panel *a*) the difference compared to a result obtained after 10 iterations is shown. In case of $n_{TF}=10$, a second test run was used for calculating the difference (note that the algorithm has random initialization, therefore different runs result in slightly different outputs). With more iterations the difference becomes less structured, more random. In panel *b*) the change with F can be seen: low numbers cause large blocks in the images, and the higher the tensor rank, the more detailed the output is.

blocks in the image, which is characteristic of a low-rank or TV regularization. For higher numbers finer details become visible.

Note that our results indicate that neither of the above parameters have to be estimated precisely. After a small number of iterations the result converges. F can be considered as a prior information on the complexity of the image. Using higher values a more natural result can be obtained, but above a threshold the method will not give more precise outputs.

V. CONCLUSION

In this work, a tensor-factorization-based method was proposed for the 3D single image super-resolution problem. This method showed interesting computational advantages compared to currently available regularization-based methods, with slightly improved image quality compared to the investigated LRTV technique. The runtime of this method was about 100 times faster than with LRTV, allowing a wider field of applications. The method also uses significantly less parameters (tensor rank and iteration number) that can be easily adjusted by visual inspection of the reconstruction results. Dental CBCT volumes used as experimental data showed improved PSNR and canal-segmentation properties, with moderately better results than the LRTV method. Considering these results, the method was found to be promising for 3D single image super-resolution.

The prior information in regularization-based techniques is often empirical and guides the solution. Future work can investigate if such classical priors could be included in this framework, and whether they would improve the result. As [17] proposed a solution also for embedded PSF estimation, its application to TF-SISR could be a potential direction of further research.

ACKNOWLEDGMENT

The authors would like to thank Jérôme Michetti for the data acquisition and the work on the visualization of the teeth.

REFERENCES

- [1] Y.-L. Ng, V. Mann, S. Rahbaran, J. Lewsey, and K. Gulabivala, "Outcome of primary root canal treatment: systematic review of the literature—part 1. effects of study characteristics on probability of success," *International endodontic journal*, vol. 40, no. 12, pp. 921–939, 2007.
- [2] H. M. Eriksen, L.-L. Kirkevang, and K. Petersson, "Endodontic epidemiology and treatment outcome: general considerations," *Endodontic Topics*, vol. 2, no. 1, pp. 1–9, 2002.
- [3] O. A. Peters, "Current challenges and concepts in the preparation of root canal systems: a review," *Journal of endodontics*, vol. 30, no. 8, pp. 559–567, 2004.
- [4] K. Horner and S. Panel, *Cone Beam CT for Dental and Maxillofacial Radiology (Evidence Based Guidelines)*, ser. Radiation Protection series. European Commission: Directorate-General for Energy, 2012, the Seventh Framework Programme of the European Atomic Energy Community (Euratom) for nuclear research and training activities (2007 to 2011) <http://cordis.europa.eu/fp7/euratom/>.
- [5] J. Martos, G. H. Tatsch, A. C. Tatsch, L. F. M. Silveira, and C. M. Ferrer-Luque, "Anatomical evaluation of the root canal diameter and root thickness on the apical third of mesial roots of molars," *Anat. Sci. Int.*, vol. 86, no. 3, pp. 146–150, Sep. 2011.
- [6] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Trans. Image Process.*
- [7] H. Yin, S. Li, and L. Fang, "Simultaneous image fusion and super-resolution using sparse representation," *Information Fusion*, vol. 14, no. 3, pp. 229–240, 2013.
- [8] K. I. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1127–1133, 2010.
- [9] A. Toma, L. Denis, B. Sixou, J.-B. Pialat, and F. Peyrin, "Total variation super-resolution for 3D trabecular bone micro-structure segmentation," in *Proc. 22nd European Signal Processing Conference (EUSIPCO)*, Lisbon, Portugal, Sep. 2014, pp. 2220–2224.
- [10] F. Shi, J. Cheng, L. Wang, P.-T. Yap, and D. Shen, "LRTV: MR image super-resolution with low-rank and total variation regularizations," *IEEE Trans. Med. Imag.*, vol. 34, no. 12, pp. 2459–2466, 2015.

- [11] W. Zhang, H. Zhang, L. Wang, A. Cai, L. Li, and B. Yan, "Limited angle CT reconstruction by simultaneous spatial and radon domain regularization based on TV and data-driven tight frame," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 880, pp. 107–117, 2018.
- [12] J. V. Manjón, P. Coupé, A. Buades, V. Fonov, D. L. Collins, and M. Robles, "Non-local MRI upsampling," *Medical image analysis*, vol. 14, no. 6, pp. 784–792, 2010.
- [13] O. Oktay, W. Bai, M. Lee, R. Guerrero, K. Kamnitsas, J. Caballero, A. de Marvao, S. Cook, D. O'Regan, and D. Rueckert, "Multi-input cardiac image super-resolution using convolutional neural networks," in *Proc. 19th MICCAI Int. Conf.* Athens, Greece: Springer, Oct.17–21 2016, pp. 246–254.
- [14] S. Cengiz, M. d. C. Valdes-Hernandez, and E. Ozturk-Isik, "Super resolution convolutional neural networks for increasing spatial resolution of ^1h magnetic resonance spectroscopic imaging," in *Proc. 21st MIUA Annu. Conf.* Edinburgh, UK: Springer, Jul.11-13 2017, pp. 641–650.
- [15] Y. Zhang and M. An, "Deep learning- and transfer learning-based super resolution reconstruction from single medical image," *J. Healthc. Eng.*, vol. 2017, 2017.
- [16] J. Hatvani, A. Horváth, J. Michetti, A. Basarab, D. Kouamé, and M. Gyöngy, "Deep learning-based super-resolution applied to dental computed tomography," *Special Issue of IEEE Trans. on Radiation and Plasma Med. Sci.*, 2018, accepted.
- [17] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W. Ma, "Hyperspectral super-resolution: A coupled tensor factorization approach," *arXiv preprint arXiv:1804.05307*, 2018.
- [18] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.
- [19] L. Chiantini and G. Ottaviani, "On generic identifiability of 3-tensors of small rank," *SIAM Journal on Matrix Analysis and Applications*, vol. 33, no. 3, pp. 1018–1037, 2012.
- [20] J. H. Elder and S. W. Zucker, "Local scale control for edge detection and blur estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 7, pp. 699–716, 1998.
- [21] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer. (2016, Mar.) Tensorlab 3.0. Available online. [Online]. Available: <https://www.tensorlab.net>
- [22] J. Michetti, A. Basarab, F. Diemer, and D. Kouame, "Comparison of an adaptive local thresholding method on CBCT and μCT endodontic images," *Physics in Medicine & Biology*, vol. 63, no. 1, p. 015020, 2017.
- [23] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [24] MeVis Medical Solutions AG. (2017) MeVisLab. Available online. [Online]. Available: <https://www.mevislab.de>
- [25] F. Shi, J. Cheng, L. Wang, P.-T. Yap, and D. Shen, "Low-rank total variation for image super-resolution," in *Proc. 16th Int. Conf. Med. Image Comput. Assist. Interv. (MICCAI)*. Nagoya, Japan: Springer, Sep. 2013, pp. 155–162.