Improving Segmentation Quality in Rotational Angiography Using Epipolar Consistency

Mathias Unberath[1,2,*], André Aichert[1,*], Stephan Achenbach[3], and Andreas Maier[1,2]

[1] Pattern Recognition Lab., Computer Science Department,
Friedrich-Alexander University Erlangen-Nuremberg, Germany
[2] Graduate School in Advanced Optical Technologies (SAOT), Erlangen, Germany
[3] University Hospital, Friedrich-Alexander University Erlangen-Nuremberg, Germany

Abstract. This paper suggests vessel segmentation and refinement in rotational cone-beam CT angiography based on virtual digital subtraction imaging and epipolar consistency conditions (ECC). Independent 2D segmentations allow for background estimation and generation of non-truncated but imperfect angiograms. A new image domain consistency metric then introduces 3D information to visualize regions of over- and under-segmentation with respect to multiple views. We devise an algorithm to successively optimize consistency by removing structures not present in all views. The algorithm is evaluated in three phantom studies of varying trajectory. We achieved an improvement in Dice score to 0.89 ± 0.01 up from 0.87 ± 0.04 without the proposed consistency requirement. We conclude, that epipolar consistency has great potential in correcting false segmentations in individual views that result from overlapping structures, such as bones or surgical instruments.

1 Introduction

Due to the high prevalence of cardiovascular diseases (CVD) coronary angiography is of high importance in clinical routine [1]. Consequently, vessel segmentation in diagnostic 3D CT angiography has been studied extensively as it may ease quantification and assessment [1]. However, treatment of CVD is impeded by the low temporal resolution of cone-beam CT. A straight-forward volumetric reconstruction of rotational angiograms is impossible, hence, aforementioned methods cannot be applied. Recent developments enable symbolic reconstruction of the artery tree from few views but require segmentations of the arteries in all projection images [2, 3]. Hessian- or gradient-based filters work well in 2D [4, 5]. Yet, segmentations are usually not consistent throughout the acquisition due to overlap with high contrast structures such as vertebrae.

As the geometry of the acquisition is well known, it seems natural to incorporate 3D information, for instance by enforcing consistency among all segmentations. However, it is not clear how information from one view should be propagated to

---

[*] Both authors contributed equally to this paper.

the others. We propose the use of epipolar consistency conditions (ECC) [6] to identify inconsistent regions in the images using novel image domain inconsistency (IDI) maps. These maps enable a refinement of the segmentation, which is the major contribution of this work.

As ECC assume non-truncated projections, we rely on a virtual single-frame subtraction imaging technique suitable for spatially limited structures such as coronary arteries [7]. We report quantitative results for subtraction imaging and segmentation refinement of a phantom study. The phantom is designed similarly to the widely known XCAT [8, 9]. We performed experiments using multiple saddle trajectories [10] that yield different orientations of the epipolar lines.

2   Material and Methods

We describe a pipeline targeted at vessel segmentation in multiple X-ray projection images of the same scene with known geometry. We suggest an adaptation of the ECC which projects inconsistencies to the image domain and allows for the localization of inconsistent structures. First, vessels are segmented separately in each view. Second, the original projections are masked using the segmentations and the virtual background is estimated [7]. This estimate is then subtracted from the original projections yielding images that only contain information within the segmentation mask. Finally, IDI maps are used to propagate 3D information and enforce consistency of the individual segmentation masks.

2.1   Preprocessing and segmentation

Coronary arteries appear as bright tubular structures on a darker background. We detect them using high pass filters, while suppressing noise with the bilateral filter [11]. To determine the filter response at position $\mathbf{u}$ in a noise-filtered image $\mathcal{I}$ the Hessian is calculated for different scales σ

$$\mathbf{H}(\mathbf{u}, \sigma) = \begin{pmatrix} \frac{\partial^2 \mathcal{I}_\sigma(\mathbf{u})}{\partial u_1^2} & \frac{\partial^2 \mathcal{I}_\sigma(\mathbf{u})}{\partial u_1 \, \partial u_2} \\ \frac{\partial^2 \mathcal{I}_\sigma(\mathbf{u})}{\partial u_2 \, \partial u_1} & \frac{\partial^2 \mathcal{I}_\sigma(\mathbf{u})}{\partial u_2^2} \end{pmatrix}, \tag{1}$$

where $\mathcal{I}_\sigma = \mathcal{I} * G_\sigma$, and G_σ is a 2D Gaussian kernel with standard deviation σ. Given $\mathbf{H}(\mathbf{u}, \sigma)$ one can calculate its Eigenvalues $\lambda_{1/2}$. They allow for the definition of meaningful measures such as the blobness $R_b = \lambda_2/\lambda_1$, and the structureness $\mathcal{S} = \sqrt{\lambda_1^2 + \lambda_2^2}$. The vesselness is then calculated as

$$\mathcal{V}(\mathbf{u}) = \max_\sigma \; \exp(-c_1 R_b) \cdot (1 - \exp(-c_2 \mathcal{S})), \tag{2}$$

where $c_{1/2}$ are constants [5]. Moreover, using Eq. 1 the orientation α of the underlying structure is given by $\tan(2\alpha) = 2H_{12}(H_{11} - H_{22})$. Let $\mathbf{e}_\alpha$ be a unit vector orthogonal to the structure's orientation. Vessels of scale σ exhibit large negative gradients at positions $\mathbf{u} \pm \sigma \mathbf{e}_\alpha$, yielding a response [4]

$$\mathcal{K}(\mathbf{u}) = \max_\sigma \left\{ \min \left\{ -\left(\nabla \mathcal{I}_\sigma(\mathbf{u} + \sigma \mathbf{e}_\alpha) \right)^\top \mathbf{e}_\alpha, \left(\nabla \mathcal{I}_\sigma(\mathbf{u} - \sigma \mathbf{e}_\alpha) \right)^\top \mathbf{e}_\alpha \right\} \right\}. \tag{3}$$

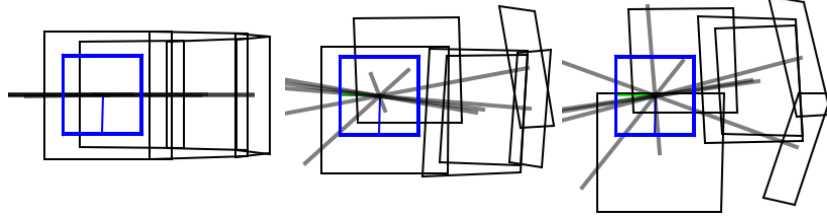**Fig. 1.** Geometries of the projections remaining after ECG-gating. Left to right: the geometry for the parameter sets $(0,0), (10°, 8)$, and $(20°, 4)$.

Eq. 2 gives strong responses for vessels, but also for other structures with high curvature. Hysteresis thresholding is used to suppress low, isolated responses yielding a binary mask $\mathcal{V}_b$. Eq. 3 has a lower false positive rate than the vesselness but only exhibits high values in the center of vessels. Let $Q_K = \{\mathbf{u} \,|\, \mathcal{K}(\mathbf{u}) > t_K\}$ be the set of points with responses larger than a heuristic threshold t_k and let $g(\mathbf{u}) = \min\{\|\mathbf{u} - \mathbf{v}\|_2\}, \; \mathbf{v} \in Q_K$ be the distance of position $\mathbf{u}$ to one such point. Finally, the segmentation mask $\mathcal{W}$ is obtained by combining both filters such that $W(\mathbf{u}) = 1$ if $\mathcal{V}_b(\mathbf{u}) \neq 0 \wedge g(\mathbf{u}) < c_3$ and 0 otherwise. c_3 is a constant related to the maximal vessel radius. The segmentation mask is subdivided into connected components yielding a set of components $\mathcal{S}^{(j)} \subset \mathcal{W}$.

2.2   Virtual subtraction imaging

The initial segmentation according to Sec. 2.1 can be used to remove all pixels of the target structures, i.e. contrasted vessels. For digital subtraction, we seek to recover a background image $\mathcal{B}$. This task is similar to defect pixel interpolation if the corruption process is modeled as multiplication $\mathcal{G} = \mathcal{I} \cdot \overline{\mathcal{W}}$, where $\overline{\mathcal{W}}(\mathbf{u}) = 1 - \mathcal{W}(\mathbf{u})$. The original projection $\mathcal{I}$ and the background image $\mathcal{B}$ are similar in the way that their masked versions cannot be distinguished. As the artificially corrupted observations $\mathcal{G}$ do not contain information about bright, spatially sparse objects an estimate of the unobserved background $\mathcal{B}$ solely relies on the remaining information thus estimating a virtual non-contrast image [7]. In order to avoid a patchy, unnatural appearance of the estimation, a frequency domain technique, namely spectral deconvolution, is used [12]. According to the convolution theorem, the corruption in frequency domain reads $g = \frac{1}{N}(b * \overline{w})$, where N is the number of pixels. As $\mathcal{G}, \mathcal{B}$ are real valued, their Fourier transforms g, b are symmetric, exemplary $b(k) = b^*(N-k)$. Assuming that b consisted only of two spectral lines at s and $N - s$ the observation g after convolution of b with the mask reads

$$g(s) = \frac{1}{N}\left(\hat{b}(s) \cdot \overline{w}(0) + \hat{b}^*(s) \cdot \overline{w}(2s)\right), \quad \text{and} \quad g(N - s) = g^*(s). \quad (4)$$

Eq. 4 can be solved for the coefficients $\hat{b}$. However, b consists of more than two spectral lines such that in practice the dominant line pairs are estimated sequentially [12].

The inpainting is applied on image patches to preserve the locality of image appearance. Moreover, the extracted patches are weighted with a Blackman window. Finally, the estimated background can be subtracted from the original projections yielding virtual subtraction images $\mathcal{D}(\mathbf{u}) = \mathcal{I}(\mathbf{u}) - \mathcal{B}(\mathbf{u})$ [7].

2.3   Epipolar consistency and segmentation refinement

The series of difference images $\mathcal{D}_i$ with $i = 1, ..., M$ are not truncated and show the contrasted vessels and structures resulting from erroneous segmentation. Recently, we showed that ECC can be used to estimate rigid components of intra-scan motion [7]. Here, we consider a stationary scene and use ECC to identify inconsistencies arising from over-complete segmentation.

When the projection matrices $\mathbf{P}_i \in \mathbb{R}^{3 \times 4}$ are known, it is possible to establish a relation between two views $i_1 \neq i_2$, where $i_{1/2} \in \{1, ..., M\}$. The join of the two source positions $\mathbf{o}_{i_1}$ and $\mathbf{o}_{i_2}$ is called the baseline. Around this line there exists a pencil of epipolar planes $\mathbf{E}_{i_1}^{i_2}(\kappa)$, where $\mathbf{E}$ is a homogeneous 3-vector, and κ denotes the angle of the epipolar plane [6]. The meet of an epipolar plane with the image planes $\mathcal{D}_{i_1}$ and D_{i_2} is called the epipolar line $\mathbf{l}_{i_1}^{i_2}(\kappa)$ and $\mathbf{l}_{i_2}^{i_1}(\kappa)$, respectively. It has been shown that in transmission imaging the two lines carry redundant information

$$\frac{\partial}{\partial t} \rho_{i_1}\left(\mathbf{l}_{i_1}^{i_2}(\kappa)\right) - \frac{\partial}{\partial t} \rho_{i_2}\left(\mathbf{l}_{i_2}^{i_1}(\kappa)\right) \approx 0, \tag{5}$$

where $\rho_i(\mathbf{l}_i)$ denotes integration over the line $\mathbf{l}_i$ in image $\mathcal{D}_i$. In the original formulation (Eq. 5), a derivative in orthogonal direction t to the lines $\mathbf{l}_{i_1}^{i_2}$ and $\mathbf{l}_{i_2}^{i_1}$ accounts for a weighting by the distance to the source position in cone-beam geometries [6]. However, the derivative acts as a high-pass filter that favors correct alignment of edges. Despite being theoretically sound, said derivative is not beneficial for image segmentation, where the total amount of absorption within structures is more important than the alignment of their boundaries. Therefore, we omit the derivative from Eq. 5 yielding $\rho_{i_1}\left(\mathbf{l}_{i_1}^{i_2}(\kappa)\right) - \rho_{i_2}\left(\mathbf{l}_{i_2}^{i_1}(\kappa)\right) \approx 0$. It states that the integral over the two epipolar lines $\mathbf{l}_{i_1}^{i_2}$ and $\mathbf{l}_{i_2}^{i_1}$ should be equal. Consequently, the sign indicates the lack or surplus of structure.

We accumulate the signed inconsistencies of a particular view and devise an image domain inconsistency (IDI) map

$$\mathcal{E}_{i_1}^{\pm} = \sum_{i_2 \neq i_1} \int \mathcal{E}_{i_1}(\mathbf{l}_{i_1}^{i_2}(\kappa)) \cdot H(\pm\mathcal{E}_{i_1}(\mathbf{l}_{i_1}^{i_2}))\mathrm{d}\kappa \ , \tag{6}$$

where $H(x)$ is the Heaviside step function, and

$$\mathcal{E}_{i_1}(\mathbf{l}_{i_1}^{i_2}(\kappa)) = \rho_{i_1}\left(\mathbf{l}_{i_1}^{i_2}(\kappa)\right) - \rho_{i_2}\left(\mathbf{l}_{i_2}^{i_1}(\kappa)\right) \ . \tag{7}$$

The total consistency of the current scene $\mathcal{D}_i(\{\mathcal{S}_i^{(j)} \,|\, j = 1, ..., S_i\}), i = 1, ..., M$ is given by the total energy $E(\mathbb{S}) = \sum_i \sum_{\Omega_i} (|\mathcal{E}_i^+| + |\mathcal{E}_i^-|)$, where Ω_i is the domain of image i, and $\mathbb{S} = \dot{\cup}\mathcal{S}_i$. $\mathcal{E}_i$ implicitly depends on $\mathbb{S}$ as removing or adding
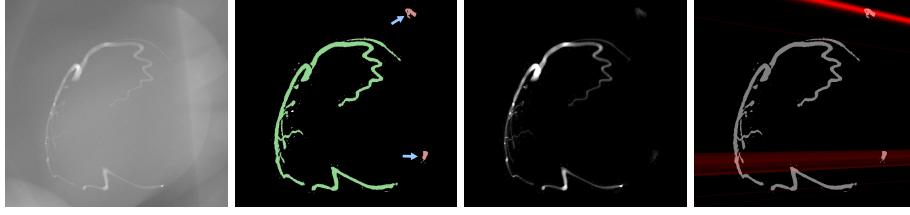
Fig. 2. Example view. Left to right: Original projection; initial segmentation mask (components are colored according to their score shown in Fig. 4, left), subtraction image and overlay of segmentation and corresponding inconsistency map $\mathcal{E}_i^+$ (red).

components influences the scene and, therefore, the overall consistency $E(\mathbb{S})$. In order to refine the segmentation we loop over all views i and perform the following steps for each view:

1. Remove each component $\mathcal{S}_i^{(j)}$ exactly once from $\mathbb{S}$ yielding $\mathbb{S}_i^{(j)} = \mathbb{S} \setminus \mathcal{S}_i^{(j)}$
2. Calculate the overall consistencies $E(\mathbb{S}_i^{(j)})$
3. Determine $\mathbb{S}_i^{(j)}$ yielding the largest improvement $\Delta E = E(\mathbb{S}) - E(\mathbb{S}_i^{(j)})$
4. If $\Delta E / E(\mathbb{S}) < 1\%$, set $\mathbb{S} = \mathbb{S}_i^{(j)}$, $E(\mathbb{S}) = E(\mathbb{S}_i^{(j)})$, recalculate all $\mathcal{E}_i$, and repeat

The process continues until no component satisfying the condition in 4 can be found. Requiring changes $< 1\%$ instead of $< 0\%$ in 4 accounts for a limited accuracy of the global consistency score. The components left after refinement are then assembled into segmentation masks $\mathcal{W}_i'$ such that $\mathcal{W}_i'(\{\dot{\cup}\mathcal{S}_i^{(j)}\}) = 1$ for all remaining components j, and 0 everywhere else.

2.4   Experiments

We simulated coronary angiography acquisitions using a phantom similar to the XCAT [8, 9]. Two additional scenarios were simulated to provide a ground-truth (the phantom without contrast agent, and contrast agent only). We added zero-mean Gaussian noise with an average signal-to-noise ratio (SNR) of 33.30 ± 0.17. The phantom was projected using trajectories with a different secondary angle modulation that is described by a sinusoid $\phi_2(i) = a_2 \sin(\frac{f_2 \phi_1}{2\pi^2})$, where ϕ_1 is the primary angle and $(a_2, f_2) \in \{(0°, 0), (10°, 8), (20°, 4)\}$. Out-of-plane acquisition leads to a broader range of epipolar line directions, which is potentially useful to detect inconsistent structures. We then applied hard ECG-gating yielding $M = 6$ stationary projections, the geometries of which are shown in Fig. 1. Using the methods described earlier, the contrasted vessels were segmented and masked. Virtual subtraction images were obtained by subtraction of the estimated background. Finally, the segmentation was refined by optimization of epipolar consistency according to Section 2.3. All intermediate and final results are evaluated quantitatively. We state the structural similarity (SSIM) and the mean-absolute-difference (MAD) between the virtual background images $\mathcal{B}_i$ and

Table 1. Average SSIM and MAD of the virtual background images and the ground-truth for the different trajectories compared to the noisy ground-truth.

|	$(0°, 0)$	$(10°, 8)$	$(20°, 4)$
SSIM	0.73 ± 0.06	0.73 ± 0.05	0.74 ± 0.05
MAD $(\cdot 10^{-2})$	2.16 ± 0.23	2.12 ± 0.21	1.84 ± 0.60

Table 2. Average Dice scores for the initial, independent segmentations and the refined segmentations $\mathcal{W}_i$ and $\mathcal{W}_i'$, respectively, with the ground-truth.

	$(0°, 0)$	$(10°, 8)$	$(20°, 4)$
Independent	0.86 ± 0.04	0.88 ± 0.04	0.86 ± 0.03
Refined	0.89 ± 0.01	0.89 ± 0.01	0.89 ± 0.01

the ground-truth. Both values were computed in the vicinity of the segmentations only. It was defined as the morphological dilation of $\mathcal{W}_i$ with a circular structural element of 10 pixels, which is approximately twice the average vessel diameter. Moreover, the average Dice score was calculated for the initial and the refined segmentations $\mathcal{W}_i$ and $\mathcal{W}_i'$, respectively.

3   Results and discussion

The results of the virtual subtraction imaging pipeline are stated in Tab. 1 and in the first row of Tab. 2. The segmentation algorithm described in Sec. 2.1 achieved an average Dice score of 0.87 ± 0.04. This result indicates that several non-vessel structures are included in the resulting binary mask, as the parameters of the algorithm were tuned for over-complete segmentation. The background estimation yielded average SSIM values of 0.73 ± 0.05 and MAD values of $(2.04 \pm 0.39) \cdot 10^{-2}$ when a ground-truth with different noise floor was used which would correspond to a real digital subtraction angiography. When comparing the virtual background images to a noise-free ground-truth, however, much higher values for the SSIM (0.91 ± 0.00) and much lower values for the MAD ($(1.03 \pm 0.16) \cdot 10^{-2}$)
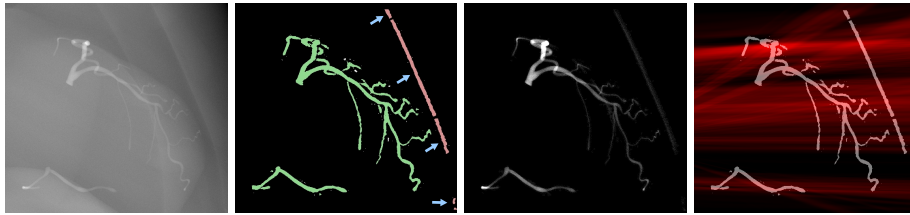
Fig. 3. Another view. Left to right: Original projection; initial segmentation mask (colors of the distinct components corresponds to the bars in Fig. 4, right), subtraction image and overlay of segmentation and corresponding inconsistency map $\mathcal{E}_i^+$ (red).
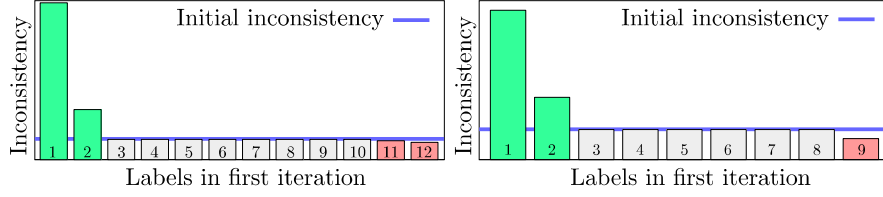
Fig. 4. Global consistency $E(\mathbb{S}_i^{(j)})$ of each label for the views shown in Fig. 2 and 3.

were achieved. Qualitative and quantitative results indicate that spectral deconvolution performs well for this task and is capable of robustly estimating the corrupted image pixels.

Fig. 2 and 3 show representative views of the scan with $(a_2, f_2) = (0°, 0)$. The initial segmentation mask shows the correctly segmented vessels in green and erroneous segmentations, also highlighted by arrows, in red. Fig. 4 shows bar plots of our image-space metric of global consistency for both views in corresponding colors. Each bar reports the global inconsistency, if the corresponding label were removed. Inconsistency spikes, if any of the correctly segmented vessels (green) are removed. Inconsistency is reduced, when over-segmented structures are removed (red). The removal of these inconsistent components yielded an improved average Dice score of 0.89 ± 0.01, compared to 0.87 ± 0.04 without the proposed consistency requirement. Qualitatively, the IDI maps provide a very intuitive visualization of inconsistent regions. Epipolar lines passing through the two over segmented components highlighted in Fig. 2, right, exhibit an excessive total absorption and, therefore, yield bright, well localized streaks in the corresponding IDI map. However, from Fig. 3, right, it becomes apparent that such maxima are not necessarily well confined if the erroneous components are large. This complicates a straight-forward analysis in the image domain but requires the use of a global inconsistency measure. We did not find any substantial difference for the three trajectories investigated here. Yet, the range of epipolar line orientations heavily depends on the underlying geometry [6]. Including few projections acquired at large secondary angles introduces steeper epipolar lines that may improve IDI map rasterization and, therefore, facilitate their interpretation.

4   Conclusion

This work presented a new approach to segmentation refinement in cone-beam CT angiography. We used virtual subtraction angiography based on an initial 2D segmentation to produce contrast-only, non-truncated views of the vessels. This enabled the investigation of consistency between individual views, thus incorporating 3D knowledge. We devised an image domain formulation of epipolar consistency, which allowed us to identify image regions of over- and under-segmentation. A proof-of-concept implementation successively removed inconsistent structures in phantom experiments. Several issues remain with the current

method. First, we only remove over-segmented structures. We currently have no way to add structures to under-segmented regions or sub-divide large components into parts, which is important in case of overlap. Faint, small structures are naturally more difficult to classify. Future research should address the effect of inconsistencies from other sources, notably residual motion from inaccurate gating and variations in contrast due to virtual subtraction angiography. Our ad-hoc approach also ignores a derivative in the original formulation of ECC, which possibly leads to inconsistencies due to non-parallel rays in cone-beam geometries. We conclude, that our proposed application of consistency to the problem of segmentation is general, in that it only relies on an initial segmentation mask and the projection data, it is straight-forward, as it uses 3D information from multiple views, without ever establishing direct correspondences, and it intuitively visualizes regions of over- and under-segmentation. We plan to investigate how the results presented here translate to clinically acquired data.

References

1. Lesage, D., Angelini, E.D., Bloch, I., Funka-Lea, G.: A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes. Med. Image Anal. **13**(6) (2009) 819–845
2. Blondel, C., Malandain, G., Vaillant, R., Ayache, N.: Reconstruction of Coronary Arteries from a Single rotational X-ray projection sequence. IEEE Trans. Med. Imaging **25**(5) (2006) 653–663
3. Unberath, M., Achenbach, S., Fahrig, R., Maier, A.: Exhaustive Graph Cut-based Vasculature Reconstruction. In: Biomedical Imaging, International Symposium on, IEEE (2016) 1143–1146
4. Koller, T.M., Gerig, G., Szekely, G., Dettwiler, D.: Multiscale detection of curvilinear structures in 2-D and 3-D image data. In: Computer Vision, International Conference on, IEEE (1995) 864–869
5. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. In: MICCAI, Proceedings of the, Springer (1998) 130–137
6. Aichert, A., Berger, M., Wang, J., Maass, N., Doerfler, A., Hornegger, J., Maier, A.: Epipolar Consistency in Transmission Imaging. IEEE Trans. Med. Imaging **34**(10) (2015) 2205–2219
7. Unberath, M., Aichert, A., Achenbach, S., Maier, A.: Single-frame Subtraction Imaging. In: Image Formation in X-ray Computed Tomography, International Conference on. (2016) 89–92
8. Segars, W.P., Mahesh, M., Beck, T.J., Frey, E.C., Tsui, B.M.W.: Realistic CT simulation using the 4D XCAT phantom. Med. Phys. **35**(8) (2008) 3800–3808
9. Maier, A., Hofmann, H.G., Schwemmer, C., Hornegger, J., Keil, A., Fahrig, R.: Fast simulation of x-ray projections of spline-based surfaces using an append buffer. Phys. Med. Biol. **57**(19) (2012) 6193–6210
10. Pack, J.D., Noo, F., Kudo, H.: Investigation of saddle trajectories for cardiac CT imaging in cone-beam geometry. Phys. Med. Biol. **49**(11) (2004) 2317
11. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Computer Vision, International Conference on, IEEE (1998) 839–846
12. Aach, A., Metzler, V.: Defect interpolation in digital radiography - how object-oriented transform coding helps. In: SPIE Medical Imaging. (2001) 824–835