Sharp as a Tack

Measuring and Comparing Edge Sharpness in Motion-Compensated Medical Image Reconstruction

Oliver Taubmann[1,2], Jens Wetzl[1,2], Günter Lauritsch[3], Andreas Maier[1,2], Joachim Hornegger[1,2]

[1]Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg
[2]Graduate School in Advanced Optical Technologies (SAOT), Erlangen, Germany
[3]Siemens AG, Healthcare, Forchheim, Germany
oliver.taubmann@fau.de

Abstract. Organ motion occuring during acquisition of medical images can cause motion blur artifacts, thus posing a major problem for many commonly employed modalities. Therefore, compensating for that motion during image reconstruction has been a focus of research for several years. However, objectively comparing the quality of different motion compensated reconstructions is no easy task. Often, intensity profiles across image edges are utilized to compare their sharpness. Manually positioning such a profile line is highly subjective and prone to bias. Expanding on this notion, we propose a robust, semi-automatic scheme for comparing edge sharpness using an ensemble of profiles. We study the behavior of our approach, which was implemented as an open-source tool, for synthetic data in the presence of noise and artifacts and demonstrate its practical use in respiratory motion-compensated MRI as well as cardiac motion-compensated C-arm CT.

1 Introduction

Motion artifacts such as image blurring are caused by movements of the imaged objects during acquisition, leading to inconsistent raw data. To obtain high-quality images using all of the acquired data—as opposed to a retrospective gating—this inconsistency must be addressed during image reconstruction. A typical approach is to estimate the motion from an initial reconstruction and subsequently or jointly perform another motion-compensated reconstruction from all data [1, 2]. Comparing the quality of motion-compensated reconstructions in terms of common measures such as the signal-to-noise ratio of the reconstructed image is of limited value: A noise-free image may be obtained even if the motion was estimated incorrectly. In fact, many pixel-wise distance measures used in phantom studies are dominated by homogeneous regions as well. Therefore, it makes sense to attempt to directly measure the entity that is supposed to be improved, i. e. the sharpness of an edge. For this purpose, analyses of modulation transfer functions (MTF) of slanted edges [3] used to be common, but they are not applicable to non-linear reconstruction methods which are dependent on the

H. Handels et al. (Hrsg.), *Bildverarbeitung für die Medizin 2015,* Informatik aktuell,
DOI 10.1007/978-3-662-46224-9_73,

imaged object. Here, the measurements need to be performed on the representative object itself. Typically, in order to perform such measurements, intensity profile lines across edges are considered [4]. Proper selection of these lines is crucial, but typically amounts to visual inspection based on personal experience.

We propose a more robust, objective approach, comparable to e. g. [5], that aims to measure edge sharpness reliably from a large number of semi-automatically placed profile lines and is explained in section 2. In section 3, its merits are demonstrated using a synthetic phantom and shown exemplarily for respiratory motion-compensated 3-D whole-heart coronary magnetic resonance imaging (MRI) [2] and cardiac motion-compensated C-arm computed tomography (CT) of the heart [1]. We discuss and conclude the paper in section 4.

2 Materials and methods

Let us first identify the main problems of manual placement of a profile line:

P1 *Susceptibility to noise.* Even noise that is incoherent with the signal and randomly distributed will influence the measurement if it cannot be separated reliably in the 1-D profile, which is often the case. As this influence is generally not consistent for different reconstructions, comparing a profile line across images may be problematic.

P2 *Susceptibility to artifacts.* For similar reasons, artifacts pose a problem to this approach. In contrast to (moderate) noise, however, they may completely render a profile invalid in one image but not the other. For instance, consider a streak artifact coinciding with the profile line in one image, which happens to appear at a different position along the same edge in the other.

P3 *Placement bias.* With no automated process for positioning the profiles, there is no way to avoid a subjective placement, potentially unconsciously biased toward a certain outcome.

P4 *Mismatch of desired and measured entity.* This is a more abstract, conceptual issue than (P1) through (P3). The entity we are interested in is the sharpness of an originally motion blurred edge. Reducing this problem to measuring the sharpness of a profile is a simplification which may or may not be valid

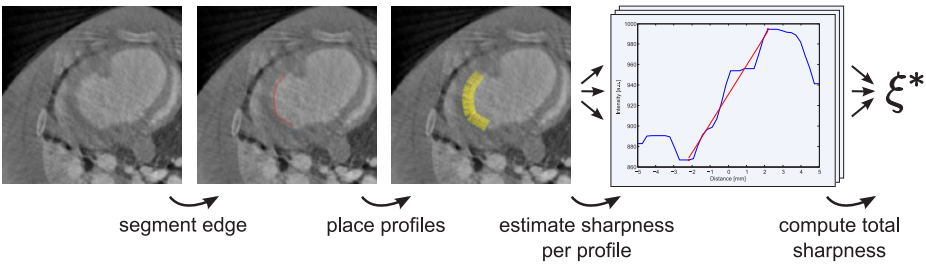

Fig. 1. A schematic overview of our approach to measuring edge sharpness from a large number of semi-automatically placed intensity profile lines.

in any given case. The closer we can get to measuring the relevant entity, i. e. the less we have to simplify, the more trustworthy our results.

We address each of these problems with our approach, which is summarized below and illustrated in Fig. 1. An obvious and common choice for alleviating the susceptibility to noise and artifacts is using multiple profile lines and some form of averaging of the measurements for increased robustness. Manually choosing a large number of lines, though, is time-consuming and may even increase the effect of placement bias. If they are to be averaged, it also requires the chosen profiles to be somewhat comparable. Therefore, starting from the core idea behind (P4), the first and only manual step in our method is to select the desired edge by roughly tracing it with a connected sequence of line segments. Along this segmentation, profile lines of equal length are generated at short equidistant steps, oriented orthogonally to the segmented line. They cover the edge densely and completely, eliminating placement bias. Naturally, there still remains a possibility for bias while choosing the edge, although we believe that it is considerably less dangerous than manually placing profiles. For each profile, an estimate of its sharpness is obtained. For this purpose, the image is sampled densely along the profile line, which is subsequently reoriented, if necessary, such that the intensities are rising. A region of interest is defined starting at the location of the minimum intensity in the first half of the profile and ending at that of the maximum in the second. In this region, the slope of the least squares regression line fit is computed as

$$\xi_\ell = \frac{\operatorname{cov}[\boldsymbol{s}_\ell, \boldsymbol{I}_\ell]}{\operatorname{var}[\boldsymbol{s}_\ell]} = \frac{\sum_i (s_{\ell,i} - \frac{1}{N_\ell}\sum_j s_{\ell,j})(I_{\ell,i} - \frac{1}{N_\ell}\sum_j I_{\ell,j})}{\sum_i (s_{\ell,i} - \frac{1}{N_\ell}\sum_j s_{\ell,j})^2} \tag{1}$$

where $\boldsymbol{s}_\ell$ contains the distances in physical units of the N_ℓ sample locations within the region of interest and $\boldsymbol{I}_\ell$ the corresponding intensities along profile line ℓ. The slope ξ_ℓ serves as a rough measure of the edge sharpness locally, but as explained above, we do not expect it to be very robust. Averaging all the ξ_ℓ would now reliably eliminate susceptibility to noise, yet artifacts can be the cause of outliers strongly affecting the mean value. Hence, we select the

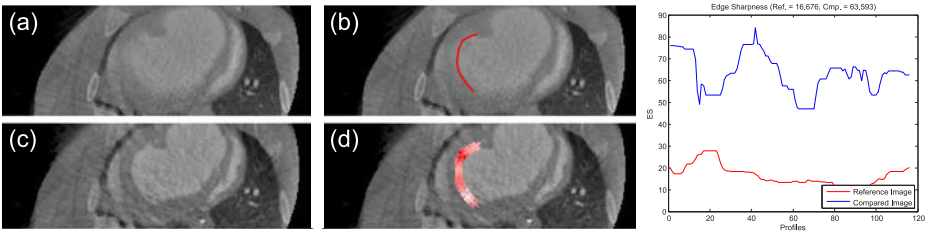

Fig. 2. The upper row shows a motion-blurred image (a) and the edge selected for comparison (b). Below, a motion-compensated reconstruction (c) and the edge sharpness increase (d) can be seen, with stronger hues of red indicating a larger relative change. The plot shows the corresponding sharpness estimates along the edge for both images.

total edge sharpness as $\xi^* = \text{median}[\boldsymbol{\xi}]$, where $\boldsymbol{\xi} = [\xi_0, \xi_1, \ldots, \xi_{M-1}]^\top$ with M the number of profiles. For two comparable images showing the same edge, the sharpness can now be compared quantitatively using the ratio of their respective ξ^*. Our approach can also support a qualitative, visual comparison. Instead of computing a single sharpness value for the whole edge, we can alternatively restrict the median computation to a smaller window to retain the spatial information, i. e. apply a median filter to $\boldsymbol{\xi}$. Plotting the color-coded ratios of two vectors filtered in that manner on top of the corresponding edge shows in which regions the sharpness has increased the most (Fig. 2). It facilitates an intuitive understanding and "sanity check" of the performed measurements.

2.1 Experiments

For a first validation of our method, we use a Gaussian-blurred Shepp-Logan phantom as well as a slightly sharper version (Fig. 3, left). The latter is additionally corrupted with noise or artifacts to obtain test images for which we expect roughly the same relative increase in edge sharpness compared to the original blurred image. Next, we assess the performance in real-world scenarios by comparing measured edge sharpness values to scores assigned by people working in the field of medical image reconstruction (hereafter referred to as experts). The first data set we use for this purpose is a C-arm CT acquisition of a porcine model [6]. Over 14.5 s, 381 projection images are acquired in a single sweep. Right atrial pacing and systemic injection of contrast agent ensure a sufficient quality of the initial electrocardiography-gated reconstructions. Each phase is deformably registered to the reference phase using a uniform B-Spline motion model [1, 7]. Subsequently, a motion-compensated reconstruction is performed for the reference phase (Fig. 4, left). The second data set stems from an in-vivo volunteer experiment using a 1.5 T clinical MR scanner. The measurement was taken with electrocardiography triggering to avoid cardiac motion artifacts, but in free breathing. Respiratory motion is taken into account with a motion-compensated reconstruction [2] using displacement fields computed with Demons at half resolution and Bilateral Demons [8] (Fig. 5, left). The edge sharpness measurement tool is openly available on http://www5.cs.fau.de/research/software/.

3 Results

The table in Fig. 3 shows the results of our phantom study. The proposed approach achieved the most consistent measurements (smallest standard deviation) compared to manually placed smaller sets of profile lines. Note that the sets in (ii) and (iv) do not touch the simulated artifact; hence, they measure the same values in the corresponding case (d). Nevertheless, they show higher standard deviations due to the noise case (c) alone. In the tables of Figs. 4 and 5, the sharpness scores given to each image by experts are compared to the computed edge sharpness estimates. Expert annotation was performed without any

prior knowledge of the employed reconstruction algorithms or the computed estimates. In both experiments, the sharpest (least sharp) image according to our measurements also received the highest (lowest) score. The difference between the uncompensated and the motion compensated images is larger than the difference between the two motion compensated versions consistently in both scores and measurements. For the C-arm CT images, even the relative distances are preserved well. As there still is a slight slope present in the uncompensated reconstructions despite the edge appearing strongly blurred visually, our approach yields comparatively larger values here.

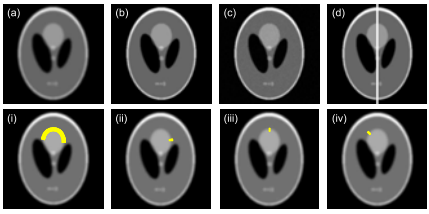

Image	Profiles			
	(i)	(ii)	(iii)	(iv)
(b) vs. (a)	1.477	1.504	1.412	1.345
(c) vs. (a)	1.459	1.936	1.334	1.487
(d) vs. (a)	1.463	1.504	0.133	1.345
Std. dev.	**0.0095**	0.25	0.72	0.082

Fig. 3. The top row shows a blurred phantom image (a) and a sharper version (b) corrupted with noise (c) and a large artifact (d). In the bottom row, profile lines covering the whole edge (i) or parts of it (ii, iii, iv) are plotted on top of the blurred image. The table on the right lists the edge sharpness increase as described in the text for different combinations of images and profile lines used to measure the sharpness.

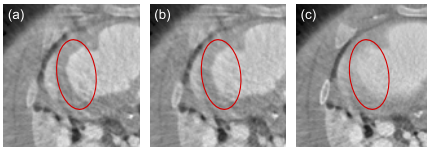

Image	Edge sharpness	Expert score
(a)	0.0402	3.43 ± 0.53
(b)	0.0296	2.14 ± 0.69
(c)	0.0114	0.00 ± 0.00

Fig. 4. Two different motion-compensated C-arm CT images (a,b) of a porcine model are shown together with an uncompensated FDK reconstruction (c) for comparison. B-Spline based deformable motion estimation [7] was performed with a control point spacing of 8 mm (a) and 16 mm (b). The table lists the computed sharpness as well as the average scores (4 = very sharp, 0 = blurred) of 7 experts for the highlighted edge.

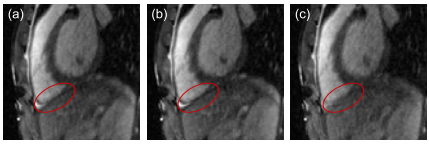

Image	Edge sharpness	Expert score
(a)	0.0604	2.00 ± 0.58
(b)	0.0724	3.57 ± 0.53
(c)	0.0367	0.29 ± 0.49

Fig. 5. Two different motion-compensated MR images (a,b) of a volunteer are shown together with an uncompensated reconstruction (c) for comparison. Demons at half resolution (a) and Bilateral Demons [8] (b) motion estimation was performed. The table lists the computed sharpness as well as the average scores (4 = very sharp, 0 = blurred) of 7 experts for the highlighted edge.

4 Discussion

In summary, the edge sharpness values computed by our method are in agreement with the visual impression and exhibit increased robustness compared to the manual approach. Artifacts and noise still have an influence on the measurements as they locally affect the observed sharpness, but it is reduced considerably. A limitation of our method lies in reconstruction methods that enforce sharp edges. In this case, the sharpness of an edge is no longer a clear indication that the motion of this edge was indeed estimated correctly, eliminating the validity of our main criterion. However, note that this also applies to the manual placement of profile lines we compare our method with. Further improvements of our approach in the future could include an integrated intensity normalization scheme, the use of higher-order splines instead of line segments for representing the edge, a pre-segmentation step to automatically generate a set of edges to select from, as well as a variety of different per-profile estimators of sharpness.

Acknowledgement. The authors gratefully acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German Research Foundation (DFG) in the framework of the German excellence initiative, and would like to thank S. De Buck, J.-Y. Wielandts and H. Heidbüchel from the University of Leuven for providing the C-arm image data. The concepts and information presented in this paper are based on research and are not commercially available.

References

1. Müller K, Maier A, Schwemmer C, et al. Image artefact propagation in motion estimation and reconstruction in interventional cardiac c-arm CT. Phys Med Biol. 2014;59(12):3121–38.
2. Forman C, Grimm R, Hutter J, et al. Free-breathing whole-heart coronary MRA: motion compensation integrated into 3D cartesian compressed sensing reconstruction. Proc MICCAI. 2013; p. 575–82.
3. Reichenbach SE, Park SK, Narayanswamy R. Characterizing digital image acquisition devices. Opt Eng. 1991;30(2):170–7.
4. Chung YC, Wang JM, Bailey RR, et al. A non-parametric blur measure based on edge analysis for image processing applications. IEEE Int Syst Cybern. 2004;1:356–60.
5. Schwemmer C, Forman C, Wetzl J, et al. CoroEval: a multi-platform, multi-modality tool for the evaluation of 3D coronary vessel reconstructions. Phys Med Biol. 2014;59(17):5163.
6. De Buck S, Dauwe D, Wielandts JY, et al. A new approach for prospectively gated cardiac rotational angiography. Proc SPIE. 2013;8668.
7. Klein S, Staring M, Murphy K, et al. Elastix: a toolbox for intensity based medical image registration. IEEE Trans Med Image. 2010;29(1):196–205.
8. Papież BW, Heinrich MP, Risser L, et al. Complex lung motion estimation via adaptive bilateral filtering of the deformation field. Lect Notes Computer Sci. 2013;8151:25–32.