Assessment of perceptual quality measures for multi-exposure radiography and tomography

Joaquim G. Sanctorum*[a], Sam Van Der Jeught[a], Sam Van Wassenbergh[b], Joris J.J. Dirckx[a]
[a]BIMEF, Dept. of Physics, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerpen, Antwerp, Belgium [b]FunMorph, Dept. of Biology, University of Antwerp, Universiteitsplein 1, 2610 Wilrijk, Antwerp, Belgium

ABSTRACT

Fusion of X-ray projection images obtained with different exposure levels is a promising technique for studying objects with features beyond the dynamic range of the X-ray detector. Various multi-exposure fusion techniques are described in the literature, yet a direct comparison between these methods is not available. This was mainly due to the absence of objective quality measures dedicated to multi-exposure X-ray images and tomographic reconstructions, a problem remaining unsolved to this day. Therefore, in this work, we compare several fusion algorithms in terms of perceptual quality using recently reported quality measures based on structural similarity. Moreover, we investigate whether these quality measures apply to tomographic slices as well. Our results indicate that the reliability of the quality measures is more convincing for fused projection images as opposed to reconstructed slices. Additionally, it is shown that fusion algorithms developed for optical photography are also suitable for multi-exposure X-ray image fusion to increase perceptual quality.

Keywords: Dynamic range, multi-exposure fusion, structural similarity, X-ray, quality measures

1. INTRODUCTION

Due to the limited dynamic range of X-ray detectors, attenuation information may be incomplete in a single exposure image or projection series. In radiography, this problem can occur when the object is a heterogeneous mixture of materials with highly differing attenuation properties[1] or when the object has a wide range of thickness components[2]. In computed tomography (CT), since the object is scanned under many different angles, objects may have an aspect ratio that is too large for the dynamic range of the detector[3]. These issues cause under- and overexposure in the recorded projections, inhibiting correct image evaluation or subsequent tomographic reconstruction.

Apart from hardware modifications to increase the dynamic range of detectors, various methods have been proposed to increase the information contained in X-ray projections through multi-exposure fusion (MEF). In essence, projections are recorded with different exposure parameters to capture information in multiple attenuation ranges, which are then fused to gather the information in different attenuation ranges in the same (fused) projection image. Different exposure situations and fusion methods are created by altering the integration time[1], the tube current[3,4], or the voltage[2,5].

Although various acquisition and fusion schemes were reported, to the best of our knowledge, no direct comparison has been made between those methods. This is partly because the methods were developed in different fields and for different purposes (although the problem is general), but more importantly, there was no reported objective quality measure for MEF X-ray images. For the fusion methods, it is vital to investigate which method represents best the fused information from the input images in the fused image. This work compares several fusion methods and evaluates their perceptual quality using MEF quality measures reported recently in the literature. Additionally, we investigate the possibility of using these measures to assess the perceptual quality of CT slices obtained with MEF projection data. The great strength of such a measure is that a subjective property, like perceptual quality, can be described with an objective measure, independent of the observer.

*joaquim.sanctorum@uantwerpen.be

7th International Conference on Image Formation in X-Ray Computed Tomography,
edited by Joseph Webster Stayman, Proc. of SPIE Vol. 12304, 123042K
 · 0277-786X · doi: 10.1117/12.2646445

2. MATERIALS AND METHODS

2.1 Image acquisition

Radiographs of a preserved piglet specimen were acquired using a single source-detector pair of our stereoscopic X-ray set-up, the 3D²YMOX system[6,7]. The piglet specimen was chosen for its high aspect ratio, resulting in a considerable difference in absorption between longitudinal and lateral transmission. During the acquisition, the voltage was set to 60 kVp. A low-exposure dataset was recorded using a tube current of 45 mA, and the tube current was put to 90 mA for a high-exposure dataset. For each dataset, 450 projections (2048 pixels × 2048 pixels, pixel size of 0.143 mm) were recorded during continuous rotation of the rotation stage (rotation period of 3 s, stationary source and detector) with a shutter speed of 0.5 ms. The distances from the X-ray source (SRD) and the detector (DRD) to the axis of rotation were 1025 mm and 278 mm, respectively. As the set-up is highly modular, the geometry of the set-up was calibrated using a phantom-based method[8,9]. Prior to geometry calibration and reconstruction, the geometric distortion in the projections induced by the X-ray image intensifier was corrected using a method based on digital image correlation[10,11].

2.2 Reconstruction

Tomographic reconstruction was carried out using the ASTRA-toolbox[12] with a Matlab interface (version 2021b). The reconstruction algorithm of choice was the simultaneous iterative reconstruction technique (SIRT), of which 150 iterations were performed with a non-negativity constraint. The voxel size was isotropic and measured 0.246 mm. As the imaging set-up was not calibrated for the attenuation coefficient of water, the gray values of the reconstructed volume were not scaled to Hounsfield units. Instead, to use the objective quality measures, the gray values were scaled between 0 and 4095, which is the gray value range of the X-ray detector.

2.3 Quality measures and assessment

Ma et al.[13] proposed an objective quality measure for MEF images in optical photography based on the structural similarity index measure (SSIM)[14], denoted as multi-exposure-fusion SSIM, or MEF-SSIM. In their work, the MEF-SSIM is calculated from the local contrast and structural components of corresponding image patches of the input sequence and the fused image. The calculation is done on multiple scales to incorporate both small-range and large-range luminance patterns and structures.

A perceptual quality measure dedicated to X-ray projections was proposed very recently by Qi et al.[15], which is also based on the SSIM. Their method calculates the SSIM from the fused maximum gradient amplitude map from the input sequence and the gradient map from the fused image, weighted with the contrast sensitivity function (CSF) to obtain a final measure compatible with the human visual system (HVS). The authors did not explicitly name their proposed method, but since it is based on the gradient amplitude map weighted with the CSF, we refer to their method as weighted-gradient-amplitude SSIM or WGA-SSIM. The mathematical framework of both measures is not included, as it is well documented in the original publications[13,15].

Both of these measures will be used to assess the quality of exposure-fused X-ray projections and CT reconstructions obtained with different fusion methods. In the present work, we consider two fusion methods originating from optical photography proposed by Mertens *et* al.[16] and by Paul *et al.*[17], and two fusion methods dedicated to X-ray images recorded with different tube currents proposed by Krämer *et al*.[3] and by Sisniega *et al*.[4]. The fusion methods will be referred to by the corresponding name of their first author.

3. RESULTS

3.1 Radiographs

High and low exposure sets of projections of the piglet specimen were recorded and fused with the fusion methods mentioned before. Examples of projections are shown in Fig. 1. Visually, from Fig. 1(a), there is little contrast between the cervical vertebrae and the surrounding soft tissue when a low exposure is used. The contrast is increased when a high exposure is used, at the expense of loss of soft tissue attenuation (for example, at the snout and tail). By fusion of these projections, the contrast in certain regions can be increased while retaining the soft tissue attenuation, as is shown in Figs. 1(c)-(e). The method of Krämer was included intentionally, as it produces nearly the same projection as the low-exposure one, but with a higher signal-to-noise ratio (SNR), as was reported by the authors and observed in our results.

The fusion methods all produce different results, and the main question is which one has the best perceptual quality. To answer this question, the MEF-SSIM and WGA-SSIM scores of the 450 fused images in the projection series were calculated, and the average values with corresponding standard deviation are shown in Table I. The MEF-SSIM measure favors the method of Mertens, whereas the WGA-SSIM measure favors the method of Paul, as a higher score indicates a higher quality (with a maximum of 1).

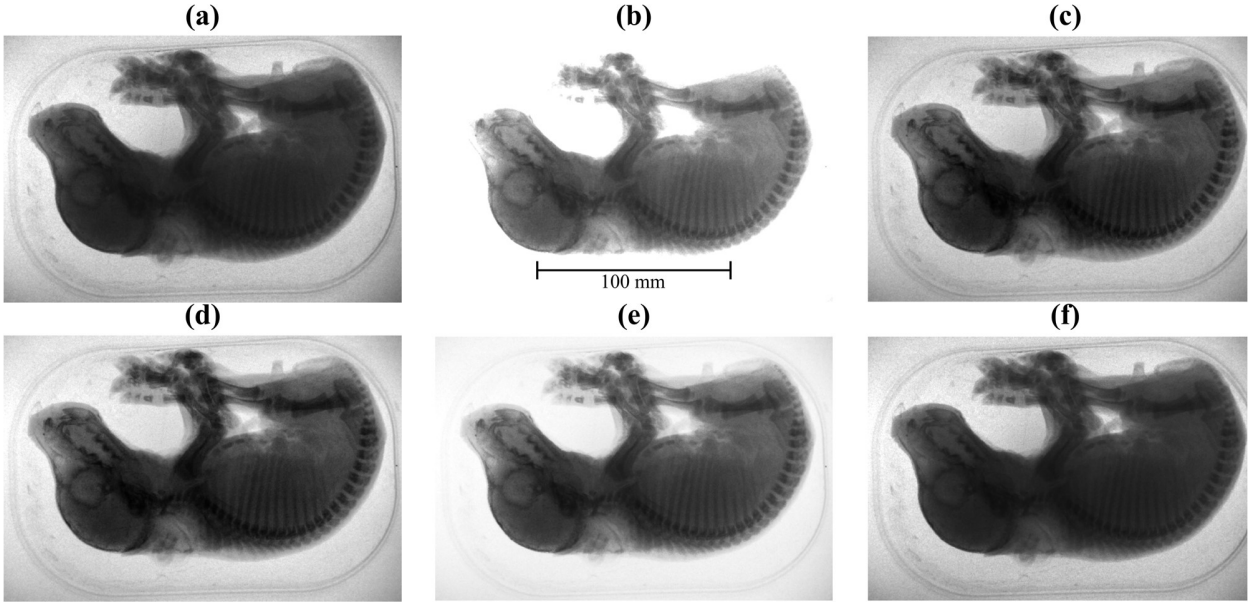

Fig. 1. The original and fused radiographs of the piglet specimen (lateral transmission). **(a)** Low exposure. **(b)** High exposure. **(c)** Mertens. **(d)** Paul. **(e)** Sisniega. **(f)** Krämer. In these projections, the axis of rotation is positioned in the center, from top to bottom. In each panel, the gray values range from 0 to 4095. The scale bar in the top center panel applies to all panels.

Table 1. Quality measures for fused projections and slices. In the names of the SSIM quality measures, subscripts P and S denote projection and slice, respectively. The subscripts CV and ST represent the CNR values for the cervical vertebrae and the sternum. The highest reported values are indicated in bold.

	Mertens	Paul	Sisniega	Krämer
MEF-SSIM$_P$	**0.988±0.004**	0.986±0.002	0.897±0.003	0.96±0.02
WGA-SSIM$_P$	0.986±0.002	**0.9903±0.0007**	0.984±0.003	0.981±0.003
MEF-SSIM$_S$	**0.93±0.06**	0.91±0.06	0.90±0.06	0.92±0.06
WGA-SSIM$_S$	**0.986±0.008**	0.981±0.006	0.979±0.005	0.98±0.01
CNR$_{CV}$	1.65	1.94	**2.08**	1.73
CNR$_{ST}$	3.03	3.51	**3.53**	3.39

3.2 Reconstructed slices

In addition to the fused projection series, a fused flat-field image was obtained with each fusion method from low-exposure and high-exposure flat-field images, subsequently used for flat-field and log correction prior to tomographic reconstruction. Examples of reconstructed slices are shown in Fig 2. In contrast to the radiographs, the difference between the slices is less apparent. Visually, the high-exposure reconstructed slice yields a better contrast between bones and soft tissue (for example, at the cervical vertebrae), at the expense of loss of signal in the soft tissue (missing tail and parts of the trotters, top center) and even some bones in the tail are not reconstructed. The MEF-SSIM and WGA-SSIM scores of the reconstructed slices are shown in Table I. From these measures, the method of Mertens appears to yield the best results.

As the HVS is sensitive to contrast, we also present contrast-to-noise (CNR) values between soft tissue and bone in one of the cervical vertebrae and the sternum, as found in Table I. It is shown that the method of Sisniega yields the greatest

CNR in both locations. The CNR values were calculated as [4]:

$$CNR = \frac{|\mu_{bone} - \mu_{tissue}|}{\sqrt{\sigma_{bone}^2 + \sigma_{tissue}^2}}, \quad (1)$$

in which and σ are the mean pixel gray value and corresponding standard deviation in an 8 pixels × 8 pixels image region.

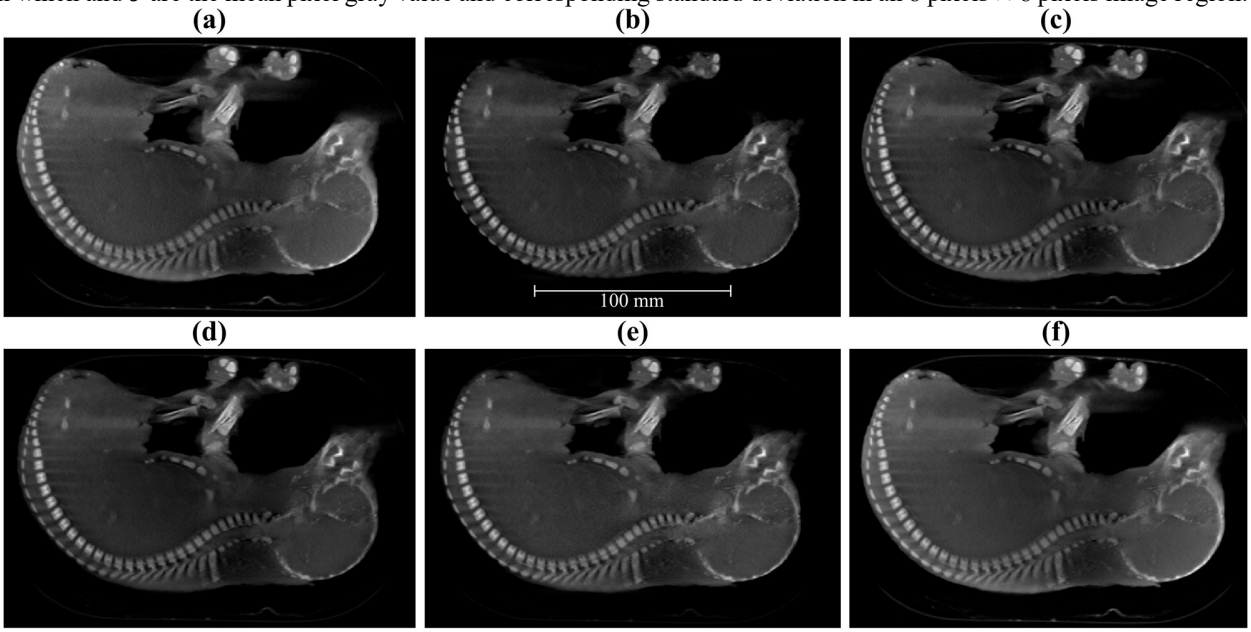

Fig. 2. Reconstructed slices of the piglet specimen. **(a)** Low exposure. **(b)** High exposure. **(c)** Mertens. **(d)** Paul. **(e)** Sisniega. **(f)** Krämer.

4. DISCUSSION

Table I shows that the MEF-SSIM measure indicates that the projections fused with the method of Mertens yield the best perceptual result. However, the error bar overlaps with that of the method of Paul. Visually, both projections are indeed very similar (Fig. 1). The method of Paul is favored by the WGA-SSIM measure, which is partly in agreement with the MEF-SSIM measure.

It is quite interesting that both quality measures indicate that the fusion methods originating from optical photography outperform those dedicated to X-ray imaging, as these optical methods were not considered for X-ray imaging before, to the best of our knowledge. Visually, the result obtained using the method of Sisniega is most dissimilar compared to the others, which is well represented by its lower MEF-SSIM score, possibly caused by the low contrast between the background and the soft tissue. On the other hand, all WGA-SSIM scores are very similar (all error bars overlap, except for the method of Paul), while the fused projections are visually not. The relative differences between the MEF-SSIM scores are more representative of the perceptual differences of the projections than the WGA-SSIM scores.

In the case of the slices, the quality scores indicate that the method of Mertens yields the best perceptual results. However, it is important to notice that all error bars overlap for both quality measures. Visually, the method of Sisniega provides good contrast between soft tissue and bone (for example, in the cervical vertebrae), which is also represented by the CNR values. Moreover, the tail is reconstructed, and most soft tissue is visible. The highest quality scores for the method of Mertens are not supported by the CNR values and actual visual perception. The soft tissue near the cervical vertebrae is quite dark, as is the case when the method of Paul is used, which is probably an amplified cupping artifact due to beam-hardening. In comparison, the gray values of the soft tissue are more uniform in the result obtained with the method of Sisniega.

These results suggest that the proposed quality measures are unreliable for reconstructed slices, yet the perceptual quality

scores of the fused projections are very reasonable and in agreement with the visual quality. Moreover, the results imply that high perceptual quality in the fused projections does not guarantee high perceptual quality in the reconstructed slices. This indicates that quality measures dedicated to tomographic reconstructions are needed to validate the quality of multi-exposure methods for tomographic purposes.

5. CONCLUSIONS

The results demonstrate that the quality measures are reliable for selecting the fusion method that yields the highest perceptual quality in fused projections and that optical fusion methods are suitable for MEF X-ray imaging as well. Yet, the same is not valid for reconstructed slices. As reconstructed slices have other demands than projection data, dedicated quality measures are needed. In future research, we plan to investigate the possibilities for a new quality measure suitable for reconstructed slices obtained from fused projection data.

REFERENCES

[1] Haidekker, M. A., Morrison, L. D., Sharma, A. and Burke, E., "Enhanced dynamic range X-ray imaging," Computers in Biology and Medicine **82**, 40–48 (2017).

[2] Bin, L., Yan, H., Jinxiao, P. and Ping, C., "Multi-energy image sequence fusion based on variable energy X-ray imaging," Journal of X-ray Science and Technology(2), 241–251 (2014).

[3] Krämer, P. and Weckenmann, A., "Multi-energy image stack fusion in computed tomography," Measurement Science and Technology **21**(4), 045105 (2010).

[4] Sisniega, A., Abella, M., Desco, M. and Vaquero, J. J., "Dual-exposure technique for extending the dynamic range of X-ray flat panel detectors," Physics in Medicine and Biology **59**(2), 421–439 (2014).

[5] Chen, P., Yang, S., Han, Y., Pan, J. and Li, Y., "High-dynamic-range X-ray CT imaging method based on energy self-adaptation between scanning angles," OSA Continuum **3**(2), 253 (2020).

[6] Sanctorum, J. G., Adriaens, D., Dirckx, J. J. J., Sijbers, J., Van Ginneken, C., Aerts, P. and Van Wassenbergh, S., "Methods for characterization and optimisation of measuring performance of stereoscopic X-ray systems with image intensifiers," Meas. Sci. Technol. **30**(10), 105701 (2019).

[7] Van Wassenbergh, S., "3D2YMOX | 3-Dimensional DYnamic MOrphology using X-rays (3D2YMOX) | University of Antwerp," <https://www.uantwerpen.be/en/research-groups/3d2ymox/> (3 January 2022).

[8] Nguyen, V., De Beenhouwer, J., Sanctorum, J. G., Van Wassenbergh, S., Aerts, P., Van Ginneken, C., Dirckx, J. J. J. and Sijbers, J., "A low-cost and easy-to-use phantom for cone-beam geometry calibration of a tomographic X-ray system," NDT.net Issue: 2019-03, 9, Padova, Italy (2019).

[9] Nguyen, V., Sanctorum, J. G., Van Wassenbergh, S., Dirckx, J. J. J., Sijbers, J. and De Beenhouwer, J., "Geometry Calibration of a Modular Stereo Cone-Beam X-ray CT System," 3, Journal of Imaging **7**(3), 54 (2021).

[10] Sanctorum, J. G., Van Wassenbergh, S., Aerts, P. and Dirckx, J. J. J., "Technical Note: Correction of geometric X-ray image intensifier distortion based on digital image correlation," Medical Physics **47**(2), 597–603 (2020).

[11] Sanctorum, J. G., Van Wassenbergh, S., Nguyen, V. T. H., De Beenhouwer, J., Sijbers, J. and Dirckx, J. J. J., "Projection-angle-dependent distortion correction in high-speed image-intensifier-based X-ray computed tomography," Measurement Science and Technology (2020).

[12] Aarle, W. van, Palenstijn, W. J., Cant, J., Janssens, E., Bleichrodt, F., Dabravolski, A., Beenhouwer, J. D., Batenburg, K. J. and Sijbers, J., "Fast and flexible X-ray tomography using the ASTRA toolbox," Opt. Express, OE **24**(22), 25129–25147 (2016).

[13] Ma, K., Zeng, K. and Wang, Z., "Perceptual Quality Assessment for Multi-Exposure Image Fusion," IEEE Transactions on Image Processing **24**(11), 3345–3356 (2015).

[14] Wang, Z., Bovik, A. C., Sheikh, H. R. and Simoncelli, E. P., "Image Quality Assessment: From Error Visibility to Structural Similarity," IEEE Trans. on Image Process. **13**(4), 600–612 (2004).

[15] Qi, Y., Yang, Z. and Kang, L., "Multi-exposure X-ray image fusion quality evaluation based on CSF and gradient amplitude similarity," XST **29**(4), 697–709 (2021).

[16] Mertens, T., Kautz, J. and Reeth, F. V., "Exposure Fusion," 15th Pacific Conference on Computer Graphics and Applications (PG'07), 382–390 (2007).

[17] Paul, S., Sevcenco, I. S. and Agathoklis, P., "Multi-Exposure and Multi-Focus Image Fusion in Gradient Domain," J CIRCUIT SYST COMP **25**(10), 1650123 (2016).