Non-local Super Resolution in Ultrasound Imaging

Parviz Khavari
Department of Electrical and Computer Engineering
Concordia University
Montreal, Canada
Email: p_khavar@encs.concordia.ca

Amir Asif
Department of Electrical and Computer Engineering
Concordia University
Montreal, Canada
Email: amir.asif@concordia.ca

Hassan Rivaz
Department of Electrical and Computer Engineering
Concordia University
Montreal, Canada
Email: hrivaz@ece.concordia.ca

***Abstract*—The resolution of ultrasound (US) images is limited by physical constraints and hardware restrictions, such as the frequency, width and focal zone of the US beam. Different interpolation methods are often used to increase the sampling rate of ultrasound images. However, interpolation methods generally introduce blur in images. Herein, we present a super resolution (SR) algorithm for reconstruction of the B-mode images using the information from the envelope of radio frequency (RF) data. Our method is based on utilizing repetitive data in the non-local neighborhood of samples. The performance of the proposed approach is determined both qualitatively and quantitatively using phantom and *in-vivo* data.**

I. Introduction

Ultrasound is a commonly used medical imaging modality since it is non-invasive, real-time, portable, and inexpensive. However, US images are intrinsically noisy and, therefore, numerous methods are used to increase the quality of the data. These procedures are applied to improve the quality of ultrasound (US) images, or, as a preprocessing step to perform high-level image analysis tasks [1], [2], [3]. Extensive previous work exists on the enhancement of US images and on the impact it has on image analysis techniques, such as image segmentation [4], [5] and registration [6], [7]. Various interpolation approaches have been proposed to increase the number of samples in US images [8], [9].

Several factors such as physical hardware, limitation on the processing time, and poor signal to noise ratio limit the quality of the images in medical imaging modalities such as US. Interpolation techniques are therefore applied in the aforementioned scenarios to improve the quality of the images [10], [11].

Interpolation methods assume that high resolution (HR) patches can be represented by a polynomial function under some smoothing assumptions. However, in US imaging, the pattern of B-mode images is very variable and may invalidate these assumptions. An alternative approach to improve the resolution in the B-mode images of US is based on the super resolution (SR) techniques. In the SR process, one or more HR patches are reconstructed based on corresponding one or more low resolution (LR) observations [12], [13]. Multi image SR techniques try to restore the HR images based on samples of LR patches in the temporal domain [14], or using machine learning techniques [15]. Nevertheless, due to scarcity of ground truth and training data sets in the US images, we choose the single image SR path. In the single image SR, which can be classified as interpolation- and recursion-based, the process often attempts to recreate the HR images from information in LR observations of a single image.

In this paper, we propose a recursive image reconstruction on the envelope of US images that preserves delicate structure while maintaining the margin between different issues. The rest of the paper is organized as follows. In Section II, we explore the proposed method in detail and provide a mathematical framework that validates the proposed approach as a Bayesian estimation. In Section III, we evaluate the performance of the proposed method using different measures on both *in-vivo* and phantom data, and we conclude the paper in Section IV.

II. Super Resolution

In US imaging, interpolations for upsampling the US images are based on an assumption that HR pixels can be expressed in terms of LR pixels. To simplify our explanation, we focus on 1D signals but our results are generalizable to 2D images. Mathematically speaking, let y_n be a LR noisy pixel in the image. Assume pixels in the HR image are denoted as x_i, then y_n can be expressed as:

$$y_n = \frac{1}{N}\sum_{i=1}^{N} x_i + n, \tag{1}$$

where in RF data, n can be consider as Gaussian noise and N is the size of the averaging filter. Within *neighborhoods* of LR pixels, the interpolation algorithm can be used to reconstruct the HR pixels by assigning the weights to them as follows:

$$x_i = \begin{cases} \sum_{j\in\Omega} w_{ij} y_j & \text{if } i \neq j \\ y_j & \text{if } i = j \end{cases} \tag{2}$$

where symbol Ω is a neighborhood of pixels in the HR image. The weights, ω_{ij}, are calculated as a function based on the distance between the location of HR pixel and LR ones. If the location of the pixels are the same, the value of LR pixel y_j is kept for HR pixel x_i.

The goal of SR or interpolation algorithms is to find values for x_i using the eq. (1). This problem is an ill-posed problem because there are infinite values for x_i that satisfy eq. (1). In the next section, a new algorithm is proposed for upsampling

the US images using information in a *non-local neighborhood* of pixels.

A. Non-local Super Resolution Approach

The proposed method (illustrated by the block diagram in fig. 1) is based on denoising the noisy data first and then applying the SR method on the denoised data. The input RF data is denoised using the non-local means (NLM) based approach on RF data. The output of the denoising filter will be y that is the denoised version of y_n.

$$y = \frac{1}{N}\sum_{i=1}^{N} x_i. \tag{3}$$

The eq. (3) can be rewritten as

$$y - \frac{1}{N}\sum_{i=1}^{N} x_i = 0. \tag{4}$$

This equation enforces the upscaled image to be consistent and is referred to as the *subsampling consistency* [16]. We used this equation as the stopping criterion for the SR algorithm.

In the next step, an iterative process is exploited to extract the information of LR patches. It consists of two separate blocks, used to: (i) smooth the image without losing edges (non-local means), and (ii) update the HR part from the LR denoised pixels. The input to the algorithm is the preprocessed LR image y and the stopping criterion sc for the algorithm. sc is set to 0.01% for err function output value and algorithm halts in cases when the err function returns a value less than sc. The err function gives the average of absolute value difference between its arguments.

The upsampling and downsampling operations are used in the SR algorithm for making the number of computations and comparisons equal. Firstly, y is upsampled to the desire aspect ratio and dimensions. In our case, the upscaling factor is set to 2 and has been done using bicubic interpolation. Assuming that x' is the approximation output of x and considering the interpolation operation as U, the output will be

$$x' = U(y). \tag{5}$$

The other blocks of this figure (fig. 1) are elaborated in the next subsections. We investigate the NLM algorithm for assuring smoothing in the neighborhood of pixel and then by using the subsampling consistency of eq. (4). The error of difference between y and downsampled version of x' is used to refine the algorithm output x'.

1) Patch-wise Smoothing: Let $x'(i)$ be the observed value of the discretized image for pixel i and $x(i)$ be its true value. Due to the presence of noise $n(i)$, we have

$$x'(i) = x(i) + n(i). \tag{6}$$

To remove the distortion present in image for each pixel i, NLM searches a reference area of the image within a rectangular *search window* Ω_i, which is centered around pixel i. A neighborhood N_i of known dimensions is selected around pixel i and compared to neighborhood N_j around pixel j for

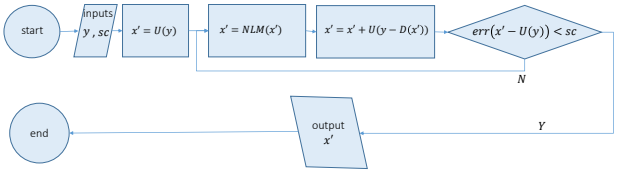

Fig. 1. The flowchart for the proposed method. The inputs are the LR patch and stopping criterion (y and sc), and the output is x' that is the HR patch. The operator U stands for upsampling using bicubic interpolation. The err function returns the mean of absolute value difference between its arguments.

all $j \in \Omega_i$. For pixel i, weight $w(i,j)$ is assigned to each pixel j. The value of pixel i is then replaced by

$$\mathrm{NLM}[x'](i) = \sum_{j\in\Omega_i} w(i,j) * x'(j). \tag{7}$$

The distance metric is proportional to the square of Euclidian distance between the two patches. The weight is then calculated as

$$w(i,j) = \frac{1}{Z_i}\exp\left\{-\frac{||x'(N_i) - x'(N_j)||^2_{2,a}}{h^2}\right\}. \tag{8}$$

Based on (8), it is clear that the weight is the convolution of a Gaussian with standard deviation $a > 0$ and the squared Euclidean distance between two neighborhoods $||x'(N_i) - x'(N_j)||^2_2$, for N_i and N_j. The smoothing parameter h controls the contribution of the Gaussian-Euclidean distance exponent in the weights. The normalization factor Z_i for pixel i is given by

$$Z(i) = \sum_{j\in\Delta_i}\exp\left\{-\frac{||x'(N_i) - x'(N_j)||^2_{2,a}}{h^2}\right\}, \tag{9}$$

where the weight is normalized to ensure that the dynamic range of the $\mathrm{NLM}[x'](i)$ is the same as that of its counterpart $x'(i)$. The NLM is applied to the envelope of RF data. More details of the NLM implementation are presented in Section II-B.

2) Retrieving Low Resolution Pixels: The LR pixels, y, are observations that are assumed to be the ground truth for reconstruction of the HR image. For retrieving the pixels in the LR patch, the algorithm upsamples the difference between LR patch and corresponding pixels in the upsampled image, x' using a downsampling operation. This difference is upsampled using interpolation and added to the existing x' to create the updated x':

$$x' = x' + U(y - D(x')), \tag{10}$$

where D and U are downsampling and upsampling operations, respectively. Finally, the error is evaluated based on the upsampled image x' and upsampled version of the LR patch y. If the stopping criterion is met, the algorithm halts. Fig. 1 shows the flowchart of the proposed algorithm. Similar to Coupe et al. [2] and Kevrann et al. [17], next we suggest a Bayesian framework that provides the mathematical foundation of the proposed algorithm.

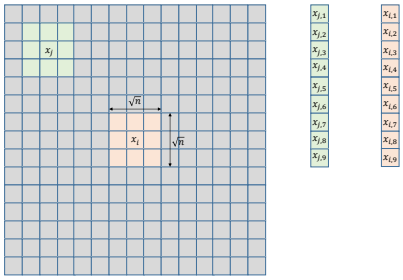

Fig. 2. An example of illustration patches of an observed radio frequency image. Patches and corresponding vectorized indices are presented in this figure for $n = 9$.

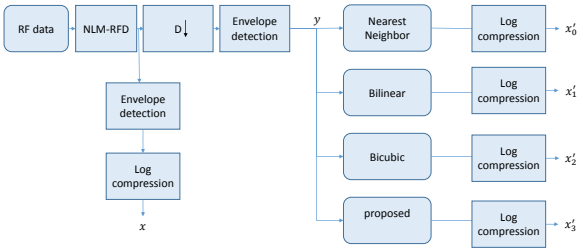

Fig. 3. The flowchart for comparison of the proposed method.

B. Bayesian Framework

Suppose **g(x)** and **o(x)** are vectorized ground truth and upscale of observed version patches of size n centered at x of the radio frequency data, defined as **g(x)** $:= g(x_k)$ and $x_k \in N_g(x)$ and **o(x)** $:= o(x_k)$, where $x_k \in N_o(x)$ (N_o and N_g are the neighborhoods (patch) of size $\sqrt{n}$ by $\sqrt{n}$ around the central pixel x in the ground truth and observed images.) The goal is to get the Bayesian estimator, $\widehat{\mathbf{g}}(x)$, of patch **g(x)** based on the observed patch **o(x)**. Let's define optimal estimator by minimizing the posterior expected loss as follows

$$E[L(\mathbf{g(x)}, \widehat{\mathbf{g}}(x))] = \sum_{\mathbf{g(x)} \in \Gamma} [L(\mathbf{g(x)}, \widehat{\mathbf{g}}(x))] p(\mathbf{g(x)}|\mathbf{o(x)}), \quad (11)$$

where Γ is all possible configurations of **g(x)**. The loss function is defined as

$$L(\mathbf{g(x)}, \widehat{\mathbf{g}}(x)) = \|\mathbf{g(x)} - \widehat{\mathbf{g}}(x)\|^2. \quad (12)$$

If eq. (12) is substituted in eq. (11), the optimal Bayesian estimator, $\widehat{\mathbf{g}}(x)_{opt}$, is given by

$$\begin{aligned} \widehat{\mathbf{g}}(x)_{opt} &= arg\min_{\widehat{\mathbf{g}}(x)} \sum_{\mathbf{g}(x)} \|\mathbf{g(x)} - \widehat{\mathbf{g}}(x)\|^2 p(\mathbf{g(x)}|\mathbf{o(x)}) \\ &= \sum_{\mathbf{g(x)}} \mathbf{g(x)} p(\mathbf{g(x)}|\mathbf{o(x)}). \end{aligned} \quad (13)$$

Then eq. (13) can be expressed as

$$\widehat{\mathbf{g}}(x)_{opt} = \sum_{\mathbf{g(x)}} \mathbf{g(x)} \frac{p(\mathbf{g(x)}, \mathbf{o(x)})}{p(\mathbf{o(x)})} = \frac{\sum_{\mathbf{g(x)}} \mathbf{g(x)} p(\mathbf{o(x)}|\mathbf{g(x)}) p(\mathbf{g(x)})}{\sum_{\mathbf{g(x)}} p(\mathbf{o(x)}|\mathbf{g(x)}) p(\mathbf{g(x)})}. \quad (14)$$

Note that only a subset of Γ is accessible in the search region of central pixel x_i. We refer to this subset as the search region, $SR(\mathbf{x})) = \{\mathbf{g}_1(x), \mathbf{g}_2(x), \mathbf{g}_3(x), ..., \mathbf{g}_K(x)\}$. Assuming that K is number of samples in the SR (the number of members in, we get the mentioned subset) and $p(\mathbf{g(x)})$ is uniformly distributed

$$p(\mathbf{g}_i(x)) = \frac{1}{K}, \quad \text{for all } 0 \le i \le K. \quad (15)$$

Based on these assumptions, eq. (14) leads to

$$\begin{aligned} \widehat{\mathbf{g}}(x_i) &= \frac{\frac{1}{K}\sum_{j=1}^{K} \mathbf{g}(x_j) p(\mathbf{o}(x_i)|\mathbf{g}(x_j))}{\frac{1}{K}\sum_{j=1}^{K} p(\mathbf{o}(x_i)|\mathbf{g}(x_j))} \\ &= \frac{\sum_{j=1}^{K} \mathbf{g}(x_j) p(\mathbf{o}(x_i)|\mathbf{g}(x_j))}{\sum_{j=1}^{K} p(\mathbf{o}(x_i)|\mathbf{g}(x_j))}, \end{aligned} \quad (16)$$

where $\widehat{\mathbf{g}}(x_i)$ is the optimal estimator based on assumption (15). However, there is only the upscaled of observed values and the ground truth is not accessible. Consequently, we substitute the observed values of the neighborhood patches to get.

$$\widehat{\mathbf{g}}(x_i) = \frac{\sum_{j=1}^{K} \mathbf{o}(x_j) p(\mathbf{o}(x_i)|\mathbf{o}(x_j))}{\sum_{j=1}^{K} p(\mathbf{o}(x_i)|\mathbf{o}(x_j))}. \quad (17)$$

By assuming that the likelihood can be considered as (see Fig. 2)

$$p(\mathbf{o}(x_i)|\mathbf{o}(x_j)) \propto \exp^{-\|\mathbf{o}(x_i) - \mathbf{o}(x_j)\|^2}, \quad (18)$$

the estimator for $\widehat{\mathbf{g}}(x_i)$ can be rewritten as

$$\begin{aligned} \widehat{\mathbf{g}}(x_i) &= \frac{1}{C(x_i)} \sum_{j=1}^{K} \exp^{-\|\mathbf{o}(x_i) - \mathbf{o}(x_j)\|^2} \mathbf{o}(x_j); \\ C(x_i) &= \sum_{j=1}^{K} \exp^{-\|\mathbf{o}(x_i) - \mathbf{o}(x_j)\|^2}. \end{aligned} \quad (19)$$

For smoothing the similarities, the parameter h is introduced and final estimator for $\widehat{\mathbf{g}}(x_i)$ is

$$\begin{aligned} \widehat{\mathbf{g}}(x_i) &= \frac{1}{C(x_i)} \sum_{j=1}^{K} \exp^{-\frac{\|\mathbf{o}(x_i) - \mathbf{o}(x_j)\|^2}{h^2}} \mathbf{o}(x_j); \\ C(x_i) &= \sum_{j=1}^{K} \exp^{-\frac{\|\mathbf{o}(x_i) - \mathbf{o}(x_j)\|^2}{h^2}}. \end{aligned} \quad (20)$$

The pixels of upsampled image can be grouped into two kinds of pixels, the synthesized pixels created by upsampling operation and the input pixels to upsampling algorithm. The Baye sian estimator can be used for the former group and the value of pixels in the latter group are retrieved from the output of SR algorithm as discussed in the algorithm illustrated in Fig. 1.

Method	Patient 1	Patient 2	Patient 3	Phantom
NN	0.0538	0.1358	0.1232	0.1078
Bilinear	0.0472	0.1224	0.1085	0.0985
Bicubic	0.0461	0.1188	0.1057	0.0940
Proposed	**0.0225**	**0.0597**	**0.0544**	**0.0481**
decreased	**48.80%**	**50.25%**	**51.46%**	**51.70%**

TABLE I
SUM OF ABSOLUTE DIFFERENCES (SAD) RESULTS OF PROPOSED METHOD VERSUS CONVENTIONAL INTERPOLATION METHODS

Method	Patient 1	Patient 2	Patient 3	Phantom
NN	13.0904	11.5248	14.0850	11.6921
Bilinear	13.9813	11.8812	14.4330	12.2291
Bicubic	14.9000	13.0660	15.4788	13.4417
Proposed	**17.2063**	**15.5290**	**18.0046**	**15.3727**
improved	**15.43%**	**18.78%**	**16.88%**	**14.06%**

TABLE II
PEAK SIGNAL-TO-NOISE RATIO (PSNR) RESULTS OF PROPOSED METHOD VERSUS CONVENTIONAL INTERPOLATION METHODS IN DB.

III. EXPERIMENTS AND RESULTS

In our experiments, phantom and *in-vivo* data have been studied. Phantom data in Fig 5 is obtained from a CIRS breast phantom (Norfolk, VA). Patients' data were acquired from patients undergoing open surgical radio frequency thermal ablation for primary or secondary liver cancer. The Institutional Ethical Review Board at Johns Hopkins University approved all experimental procedures involving human subjects. All data is collected using a Siemens Antares ultrasound machine (Issaquah, WA) with a VF10-5 linear probe.

For the sake of comparison, three upsampling interpolations are used: the nearest neighbor interpolation, bilinear interpolation, and bicubic interpolation. First, we apply NLM for denoising the RF data and afterwards, by envelope detection and log compression, the HR image for comparison is constructed (as depicted in Fig. 3), called x of size R^{n*m}. The denoised RF data is then fed to a downsampler followed by an envelope detector to construct the LR image, y. We used the downsampling with a factor of 2. The LR image is upsampled using different interpolations and proposed method and after logarithmic compression the outputs are used for comparison, or equivalently x'_0 to x'_3 of size R^{n*m}. The parameters values for proposed algorithm are $\Omega_i = 7$, $N_i = 3$, $h = 1$ and $sc = 0.01\%$. Fig. 3 highlights the steps used to compare the outputs.

The output of the proposed algorithm is compared using two different measures. We used Sum of Absolute Difference (SAD) between each $x'_i, \quad 0 \leq i \leq 3$, and x using

$$SAD(x'_i, x) = \frac{1}{m * n} \sum_{k=1}^{n} \sum_{t=1}^{m} |x'_i(k,t) - x(k,t)|. \qquad (21)$$

The results for this comparison are shown in Table I. The proposed algorithm outperforms the conventional interpolation methods and produce the least SAD error. Another measure for comparing the results is the Peak Signal to Noise Ratio or PSNR computed as

$$PSNR(x'_i, x) = 10 \log_{10}\left(\frac{Max(x'_i) - Min(x'_i)}{\sqrt{\frac{1}{m*n}\sum_{k=1}^{n}\sum_{t=1}^{m}|x'_i(k,t) - x(k,t)|^2}}\right). \qquad (22)$$

The higher PSNR implies the more accurate reconstruction of the LR image. Table II shows the PSNRs for different interpolations. The proposed algorithm surpasses other interpolation methods there were considered in our experiments. To present the qualitative improvement, different US images and the LR and HR counterparts obtained from different interpolation approaches are shown in Fig. 4 for patient 1 and Fig. 5 for phantom data. The residual pattern for each interpolation and proposed method is displayed by calculating the SAD between the upsampled images and HR image. The residual patterns in each upsampled patient data using existing interpolation techniques contain substantial information from the HR image, . In contrast, the SAD of the proposed algorithm is uncorrelated and does not show any pattern, which confirms the superiority of the proposed method in reconstructing the pattern of HR image.

IV. CONCLUSIONS

In this paper, we proposed an iterative SR approach for US imaging. The proposed method is applied to the envelope of RF data and has the ability to enhance the resolution of the image while retaining fine structures of tissue. This is demonstrated using phantom data and experiments based on *in-vivo* data. The results of the proposed algorithm are compared with interpolation approaches, which verifies the effectiveness of the proposed method. More studies are needed to fully verify the benefits of the algorithm method in US medical applications, such as US image registration and segmentation.

V. ACKNOWLEDGEMENT

This research has been provided by NSERC Discovery grants RGPIN-2015-04136 and RGPIN-2017-06629.

REFERENCES

[1] G. Ramos-Llordén, G. Vegas-Sánchez-Ferrero, M. Martin-Fernandez, C. Alberola-López, and S. Aja-Fernández, "Anisotropic diffusion filter with memory based on speckle statistics for ultrasound images," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 345–358, 2015.

[2] P. Coupé, P. Hellier, C. Kervrann, and C. Barillot, "Nonlocal means-based speckle filtering for ultrasound images," *IEEE transactions on image processing*, vol. 18, no. 10, pp. 2221–2229, 2009.

[3] G. Montaldo, M. Tanter, J. Bercoff, N. Benech, and M. Fink, "Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 56, no. 3, pp. 489–506, 2009.

[4] R. J. Cunningham, P. J. Harding, and I. D. Loram, "Real-time ultrasound segmentation, analysis and visualisation of deep cervical muscle structure," *IEEE transactions on medical imaging*, vol. 36, no. 2, pp. 653–665, 2017.

[5] G. N. Stevenson, S. L. Collins, J. Ding, L. Impey, and J. A. Noble, "3-d ultrasound segmentation of the placenta using the random walker algorithm: Reliability and agreement," *Ultrasound in Medicine and Biology*, vol. 41, no. 12, pp. 3182–3193, 2015.

[6] R. Shams, Y. Xiao, F. Hébert, M. Abramowitz, R. Brooks, and H. Rivaz, "Assessment of rigid registration quality measures in ultrasound-guided radiotherapy," *IEEE transactions on medical imaging*, vol. 37, no. 2, pp. 428–437, 2018.

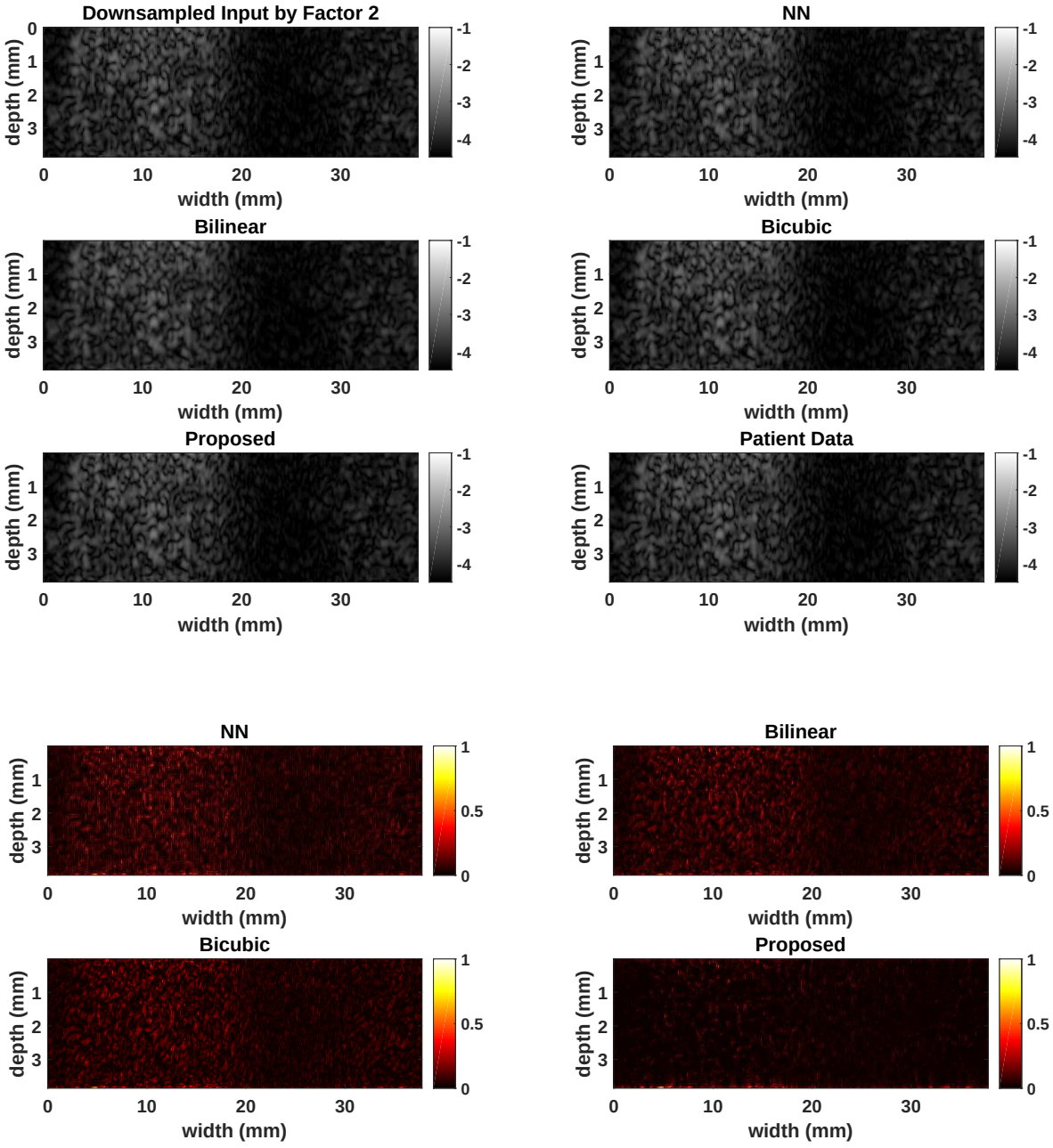

Fig. 4. B-mode of different interpolation and corresponding residual patterns of patient1.

[7] H. Rivaz, S. J.-S. Chen, and D. L. Collins, "Automatic deformable mr-ultrasound registration for image-guided neurosurgery," *IEEE transactions on medical imaging*, vol. 34, no. 2, pp. 366–380, 2015.

[8] R. J. Housden, A. H. Gee, G. M. Treece, and R. W. Prager, "Subsample interpolation strategies for sensorless freehand 3d ultrasound," *Ultrasound in Medicine and Biology*, vol. 32, no. 12, pp. 1897–1904, 2006.

[9] J. Kortbek, H. Andresen, S. Nikolov, and J. A. Jensen, "Comparing interpolation schemes in dynamic receive ultrasound beamforming," in *2005 IEEE Ultrasonics Symposium*. IEEE, 2005, pp. 1972–1975.

[10] T. M. Lehmann, C. Gonner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE transactions on medical imaging*, vol. 18, no. 11, pp. 1049–1075, 1999.

[11] P. Thévenaz, T. Blu, and M. Unser, "Interpolation revisited [medical images application]," *IEEE Transactions on medical imaging*, vol. 19, no. 7, pp. 739–758, 2000.

[12] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," *arXiv preprint*, 2016.

[13] K. Nasrollahi and T. B. Moeslund, "Super-resolution: a comprehensive survey," *Machine vision and applications*, vol. 25, no. 6, pp. 1423–1468, 2014.

[14] M. Protter, M. Elad, H. Takeda, and P. Milanfar, "Generalizing the nonlocal-means to super-resolution reconstruction," *IEEE Transactions on image processing*, vol. 18, no. 1, pp. 36–51, 2009.

[15] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using

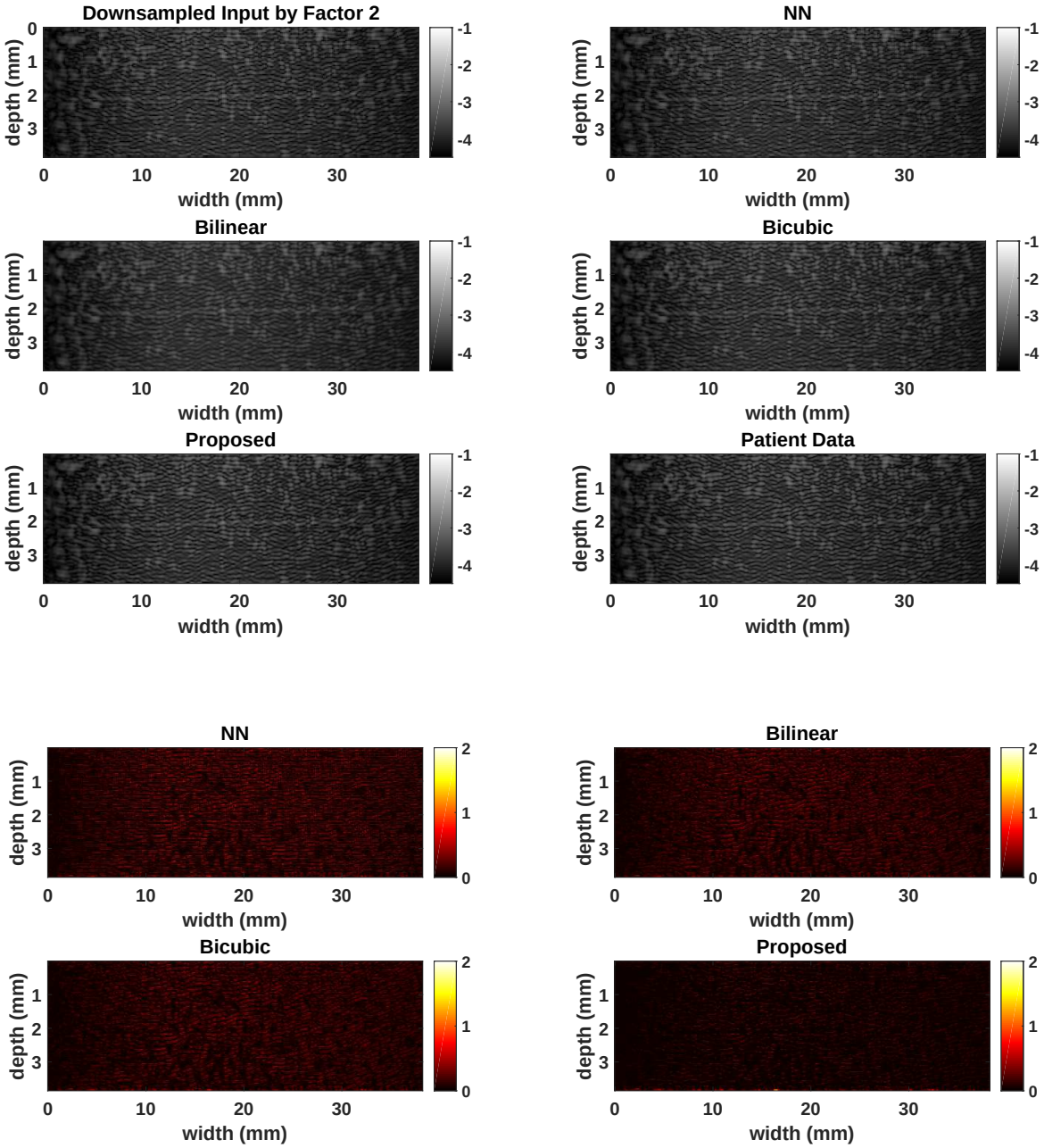

Fig. 5. B-mode of different interpolation and corresponding residual patterns of the phantom.

deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 2, pp. 295–307, 2016.

[16] J. Banerjee and C. Jawahar, "Super-resolution of text images using edge-directed tangent field," in *Document Analysis Systems, 2008. DAS'08. The Eighth IAPR International Workshop on*. IEEE, 2008, pp. 76–83.

[17] C. Kervrann, J. Boulanger, and P. Coupé, "Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal," *Scale Space and Variational Methods in Computer Vision*, pp. 520–532, 2007.