Performance evaluation of image reconstruction algorithms for photoacoustic tomography

I. Pi-Martín, J.J. García-Garrigós, A. Cebrecos, N. Jiménez, F. Camarena

Instituto de Instrumentación para Imagen Molecular (i3M), Consejo Superior de Investigaciones Científicas (CSIC),
Universitat Politècnica de València (UPV), Camino de Vera s/n, 46022, Valencia, Spain
alcebrui@upv.es

Abstract

Photoacoustic (PA) imaging uses the acoustic energy produced by some materials when heated with a laser due to the photoacoustic effect. Photoacoustic tomography (PAT) can be a powerful tool in medical imaging, since it combines the optical contrast with the high penetration of ultrasound to generate anatomical, molecular and functional images. To create the images, beamforming algorithms are used as commonly in ultrasonic imaging. The most used algorithm is Delay-and-Sum (DAS) beamforming because is computationally efficient and very easy to implement, however, it has some limitations as it offers low image resolution and contrast. To overcome these image quality problems, improvements have emerged over the typical DAS to obtain better image quality and narrower main lobes. They are based on combining the signals of the different elements, windowing the signal, or even calculating the optimal apodization for each pixel based on the signals received by the ultrasonic elements. This work is a numerical and experimental study that uses the spatial resolution and the signal to noise ratio to evaluate the performance of the main reconstruction algorithms used in PAT.

Keywords: Ultrasound, photoacoustic tomography, beamforming.

1   Introduction

Photoacoustic tomography (PAT) is an emerging modality of high-resolution biomedical imaging combining the use of optical and acoustic techniques for imaging biological tissue [1]. The technique consists of illuminating a tissue by means of a laser beam causing a local heating of the tissue. Due to the photoacoustic effect, in which optical energy is absorbed and converted into acoustic energy by thermoelastic expansion of the tissue, tissue rarefaction produces a pulsed ultrasonic signal that can be detected and converted into images by using beamforming algorithms [2]. The received acoustic signals come from the light absorbers or chromophores that have emitted the ultrasonic signal after laser excitation. The chromophores are excited at a specific laser wavelengths, and they can be classified as endogenous, such as hemoglobin or oxyhemoglobin, which allow to obtain anatomical and functional images monitoring parameters like the level of oxygen saturation, blood flow velocity, or the basal metabolic rate [3]; or exogenous, such as photosensitive inks or nanoparticles, which can be used as a contrast for molecular imaging [1].

Photoacoustic tomography presents better penetration into the tissue than purely optical tomographic methods, and substantially improving spatial resolution and contrast compared to classical ultrasonic methods. In addition, it maintains many of the advantages of ultrasonic imaging, being a non-invasive, non-ionizing radiation, fast and low-cost [4]. Classical ultrasonic linear phased arrays, often employed in both commercial and research ultrasound imaging systems, have been adapted for their use in photoacoustic imaging systems. However, the use of these receptors reduces the quality of the obtained PAT images, since the laser-generated

acoustic signals are travelling in all directions in the 3D-space. Thus, improved PAT images can be achieved using concave or even semi-spherical sensors which almost surround the sample [5]. Ideally, if the sample is completely enclosed by the sensor, an error-free tomographic image could be reconstructed by delaying and adding the registered signals from each element of the array [2, 6]. However, since it is usually unfeasible to completely wrap the object, basic reconstruction algorithms result in images having artifacts, aside from other limitations inherent to the experimental setup that complicate the reconstructions. Hence, it becomes necessary to use more complex reconstruction algorithms that take advantage of the properties of the recorded ultrasonic signals, obtaining images with improved resolution and contrast. For example, the time reversal algorithm projects the signals in the reverse path instead of delaying them, which allows reconstructions in inhomogeneous medium [7]; there are also other algorithms based on FFT [8] and even techniques that use artificial intelligence for image reconstruction and even for quantitative photoacoustic imaging [9]. Nevertheless, since soft tissues have similar sound propagation speeds, it is common to use reconstruction algorithms that assume homogenous sound speed and are just based on delaying and combining the received signals [6], those algorithms can be mainly classified into two groups: adaptive and non-adaptive techniques.

This work aims to evaluate and compare adaptive and non-adaptive algorithms, seeking to be able to characterize them and proving what advantages and disadvantages each one of them has and in what circumstances they should be used to optimize the reconstructed image properties. For this, the Delay and Sum (DAS), Delay Multiply and Sum (DMAS), Minimum Variance (MV) and MV-DMAS [10] beamforming algorithms have been implemented and have been numerically and experimentally evaluated using qualitative and quantitative criteria. The coherence factor has also been implemented in order to reduce lateral lobes.

2   Materials and methods

2.1   Beamforming algorithms

The basic beamforming of acoustic signals is based on aligning the signals received in each element of the sensor and adding them, Delay and Sum (DAS), as can be seen in Eq. (1), where t is the time index, P is the pixel index, $x_i(t - \Delta_{iP})$ is the delayed signal from the element i for the pixel P, and M is the number of elements of the array [6]. The intensity at a pixel P is calculated as

$$I_{DAS_P}(t) = \sum_{i=1}^{M} x_i(t - \Delta_{iP}).$$
(1)

From this base, different variations have emerged introducing pre-addition operations to enhance the signal and improve the resolution of the image. To improve the quality of the tomographic images, in Delay Multiply and Sum (DMAS), a combinatorial multiplication of the signals received in the different elements of the transducer (i and j in Eq. (2)) is introduced into the beamforming leading to a non-linear beamforming based on the spatial coherence of the photoacoustic signals. The objective of this modification is to improve the signal to noise ratio (SNR), enhancing the signals of the photoacoustic effect and reducing the noise [11] as

$$I_{DMAS_P}(t) = \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} x_i(t - \Delta_{iP}) \, x_j(t - \Delta_{jP}).$$
(2)

Due to the products in the time domain, the spectrum of the signal after beamforming has components at $2f_c$, $-2f_c$ and 0 Hz, so a band-pass filter that allows us to keep only the information in $2f_c$ is necessary [11]. However, if we calculate the analytic signal of the RF recorded by the transducers, we are left with only the

information at f_c. Furthermore, the multiplication of complex signals does not cause information to appear at 0 Hz, as it happens with real signals, since these components are canceled in the multiplication process. Therefore, the resulting signal after beamforming only has the component in $2f_c$, hence, it does not need band-pass filtering.

Although this technique improves the quality of the reconstructions, it does not make the lateral lobes disappear. To minimize them, adaptive algorithms, such as MV can be used [12]. The beamformed signals are given by

$$I_{MV_P}(t) = \sum_{i=1}^{M} w_{iP} \; x_i(t - \Delta_{iP}), \tag{3}$$

where w_{iP} are weighting factors. The success of adaptive algorithms lies in the calculation of the optimal weights based on the received signal. The vector of weights that maximizes the SINR (signal-to-interference-plus-noise-ratio) is defined as

$$W_{opt} = \frac{R^{-1} \; a}{a^H \; R^{-1} \; a} \;, \tag{4}$$

where R is the covariance matrix (interference-plus-noise) and a the steering vector, which will be a vector of ones, since a response without distortion is assumed and the signals from the different transducers have been aligned [12]. However, in practice R is not known, therefore, it is estimated using the covariance matrix ($\hat{R}$) of the received signals using the N samples received in the M elements of the transducer as

$$\hat{R} = \; X_d \; X_d^H, \tag{5}$$

where X_d is the matrix of delayed signals. To obtain better results, a spatial and temporal averaging can be performed, by windowing the M elements of the transducer in windows of L elements, using K samples before and after the calculated sample. In this case, it has been applied an spatial averaging considering a window of $L = M/2$.

The DMAS technique, as its name suggests, has a simple sum in each term. This sum is the common part with the classic DAS technique. If, instead of using DAS for each term, the adaptive weighted sum MV is used, the MV-DMAS technique is obtained [10]. In this case the signal intensity is given by

$$I_{MV-DMAS_P}(t) = \sum_{i=1}^{M} x_i(t - \Delta_{iP}) \left[\sum_{j=1, j \neq i}^{M} w_{jP} \; x_j\left(t - \Delta_{jP}\right) \right]. \tag{6}$$

In addition to these beamforming algorithms, a coherence factor (CF), which has already been used for ultrasonic imaging [13], can also be employed in photoacoustic. It is based on applying a weight function to each pixel, which is neither linear nor adaptive, to drastically reduce the side lobes [14]. It involves multiplying the beamforming of each pixel by its CF, which is defined as

$$CF = \frac{(\sum_{i=1}^{M} x_i(\Delta_i))^2}{M \; (\sum_{i=1}^{M} (x_i(\Delta_i))^2)}. \tag{7}$$

2.2 Evaluation of the beamformers

2.2.1 Numerical simulations

The reconstruction algorithms described above have been numerically evaluated using the k-Wave tool. For the simulation, an area of 100 mm x 60 mm has been simulated using an orthotropic mesh of 10 µm, resulting in 10000 rows x 6000 columns. The size of the Perfectly Matched Layer (PML) is 4 mm, surrounding all four sides of the simulated area. The medium was water, 1445 m/s is set as the speed of sound and 1000 kg/m^3 as the density. The attenuation of ultrasound in water has been neglected. The design of the initial pressure distribution produced by the simulated photoacoustic effect should allow an easy comparison and evaluation of the different image shaping algorithms that are implemented. To do this, an initial pressure consisting of a series of discs of diameter 0.1 mm placed on the vertical axis (90 ° with respect to the sensor) spaced 10 mm, from 10 mm to 80 mm of depth has been configured. A 128-element linear phased array with a pitch of 3 mm and a central frequency of 7.6 MHz has been simulated. A sampling frequency of 30.4 MHz has been selected. With these characteristics, the maximum frequency supported in the simulation is 72.25 MHz, which is large enough considering the central frequency of the transducer, which has been set at 7.6 MHz. White noise has been added to the simulated sensor data, setting the SNR to 50 dB.

2.2.2 Experimental setup

A gelatin phantom with a sound propagation speed of 1525 m/s, whose properties are similar to those of soft tissues [15], has been fabricated. Eight black nylon threads of 0.1 mm in diameter separated by 5 mm, who will act as chromophores, have been evenly located through the gelatin, crossing transversally the sample. For the laser excitation, an OPO (Optical Parametric Oscillator) laser tuned at 532nm wavelength and with 10 Hz pulse repetition rate (EKSPLA-NT350) has been used, complemented with a beam expander that magnifies the collimated beam to achieve a 20 mm diameter illumination area, representing the maximum image field of view, and with average pulse energy of 60 mJ. A Verasonics Vantage 256^{TM} ultrasound phased-array system has been used for the registration of the ultrasonic signals emitted by the chromophores of the phantom. To synchronize the laser and the acquisition system a trigger signal was sent from the phased-array to a function generator, which generated a rectangular pulse whose duration determines the energy emission of the laser pulse, as shown in Figure 1.
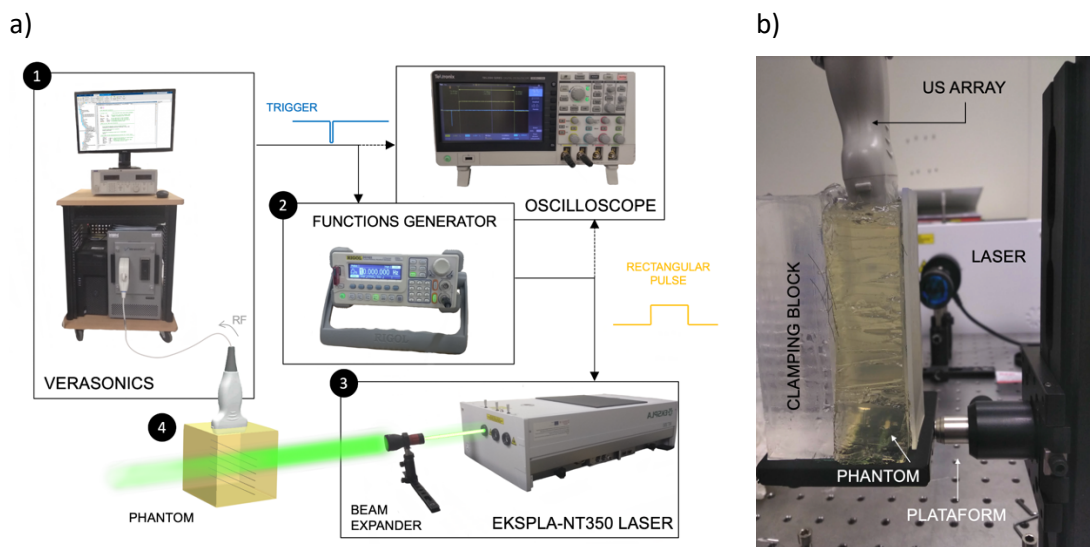

Figure 1. Experimental set-up (a) diagram and (b) photograph.

2.2.3 Parameters

Full width at half maximum (FWHM) was measured to quantify the performance of the reconstruction algorithms. FWHM indicates the minimum distance that two-point sources can be separated so that their image is interpreted as two independent sources. To calculate it, the lateral distance (width) of the main lobe at half the maximum power of the reconstructed image has been considered.

To evaluate the contrast between signal and noise, SNR has been calculated as

$$SNR = 10 \log_{10} \frac{P_{signal}}{P_{noise}}, \tag{8}$$

where P_{signal} and P_{noise} are calculated considering the intensity of the image at the zones marked in Figure 2 as the power signal and noise regions.
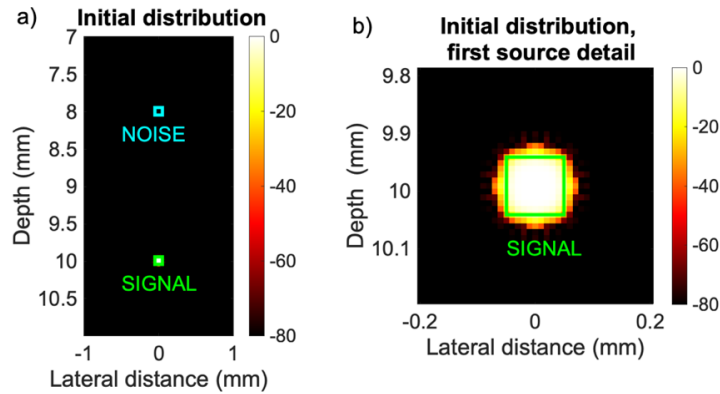

Figure 2. Initial pressure distribution for the simulation. Color bar represents power in a 80dB logarithmic scale. A small region around the first absorbent is represented, in which (a) the signal and noise areas to be considered and (b) the detail of the first absorbent (b) have been marked; these proportions are maintained for all sources, from 10 mm to 80 mm.

3   Results

3.1   Numerical results

As the sensor simulated is a linear array, the Point Spread Function (PSF) is not symmetric and the lateral and axial resolution do not depend on the same factors. The axial resolution is limited by the bandwidth and frequency of the signal, which are linked in photoacoustic imaging by the attenuation of the medium and the bandwidth of the receptor. Since the simulated medium is water, the attenuation is almost negligible for the range of frequencies used in this study. The main limitation for the axial resolution is the frequency response of the array. As shown in Figure 3, the axial resolution does not vary with depth neither in DAS nor in DMAS. On the contrary, the lateral resolution is highly influenced, besides the above factors, by the effective aperture of the array, i.e., the f-number, and, therefore, it does vary with depth.

The main effect of DMAS (Figure 3(c, d)) is to boost the signal with respect to background noise, however, the main lobe is too wide, and this means that reconstructions with this technique will not have good resolution. To improve this aspect, the use of adaptive methods becomes necessary. The combination of MV (Figure 3 (g, h)), which improves resolution, and DMAS, which enhances the sources with respect to the background, maintains the good properties of both techniques (Figure 3(i, j)). Finally, the CF has been applied on DAS and

MV-DMAS (Figure 3(e, f, k, l)) showing that lateral lobes are improved using the CF in both techniques as compared to Figure 3(a, b, i, j).
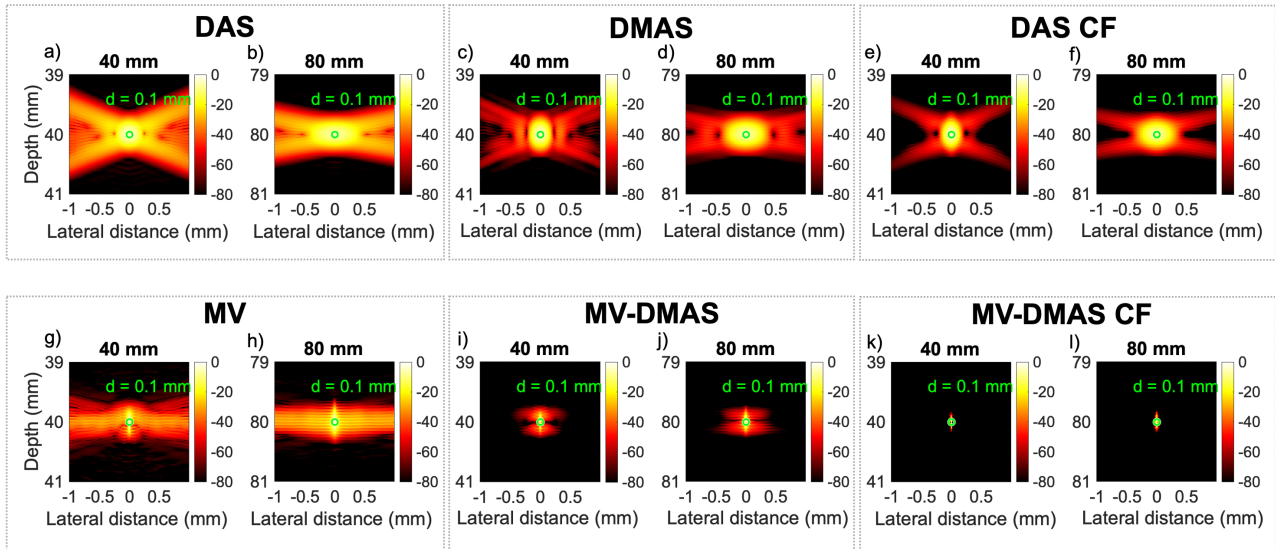

Figure 3. Numerical reconstructions at 40 and 80 mm depth using different beamformers.

To understand and evaluate how the source reconstruction varies depending on depth and the beamforming algorithm used, the lateral and axial profiles of the sources located at 40 mm and 80 mm depths have also been studied. Figure 4 illustrates the lateral and axial profiles obtained using different reconstructions methods, being expressed in power logarithmic scale. Note that no limitation in the dynamic range is set. Since it is a simulation, the initial pressure distribution is known so, to use it as a size reference, the profile of the initial pressure distribution has also been included and labelled as "Simulated".
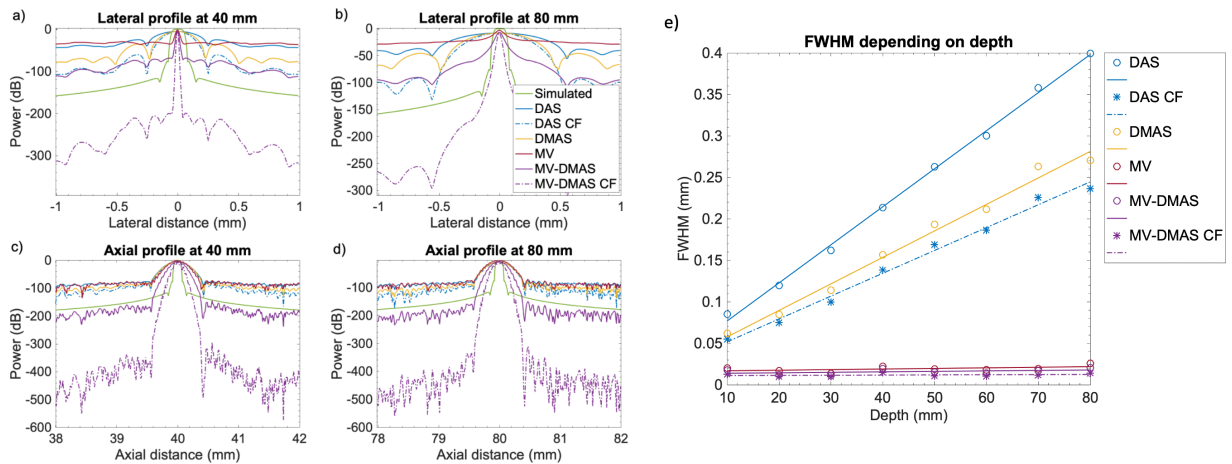

Figure 4. (a, b) Lateral and (c, d) axial profiles of the numerical reconstructions and (e) FWHM depending on depth and the beamforming algorithm used for a source of 0.1 mm.

As shown in Figure 4(a, b), increasing depth increases the width of the main lobe for all the algorithms. The DMAS, compared to the DAS, increases the difference between the signal and noise levels and vaguely reduces the width of the main lobe, but does not make the secondary lobes disappear. The MV algorithm manage to almost eliminate the side lobes, however the MV-DMAS combination slightly reintroduces the lateral lobes of

6

the DMAS, although reducing both the width of the main lobe and the baseline level with respect to the standard MV. Axial resolution (Figure 4(c, d)) is less variable in depth and between algorithms than lateral resolution. The main difference in the axial profile is observed between the MV-DMAS algorithm and the others, since it slightly reduces the width of the main lobe and considerably reduces the background noise level of the image. Regarding lateral resolution, as shown in Figure 4(e), adaptive algorithms practically maintain the FWHM with depth, while non-adaptive DAS and DMAS increase the width of their main lobe linearly as the depth increases. This result is consistent with the discussion in the previous sections.

Finally, regarding the signal to noise ratio (SNR), the one with the best ratio is the MV-DMAS CF method, which is also capable of keeping it practically constant around 44, as shown in Table 1.

Table 1. SNR as a function of depth and the beamforming algorithm used.

Depth (mm)	DAS (dB)	DAS CF (dB)	DMAS (dB)	MV (dB)	MV-DMAS (dB)	MV-DMAS CF (dB)
10	36.0	39.8	38.7	41.1	43.5	44.2
20	36.9	39.7	38.8	41.6	43.5	44.2
30	37.3	39.9	39.1	41.9	43.9	44.2
40	37.3	39.8	39.1	41.7	43.7	44.2
50	37.1	39.6	38.8	41.1	43.5	43.9
60	37.6	40.1	39.3	42.6	44.1	44.3
70	37.3	39.8	38.9	42.2	43.9	44.3
80	37.6	40.0	39.2	42.4	43.6	44.2

3.2 Experimental results

To obtain the experimental data at different depths, several measurements were made with the laser illuminating the different absorbers threads. The reconstructions of one of the excitations are shown below. The upper part of the phantom, where there was no signal, has been omitted in the images because the laser was not exciting those absorbers. The pixel size is 50 μm x 50 μm.
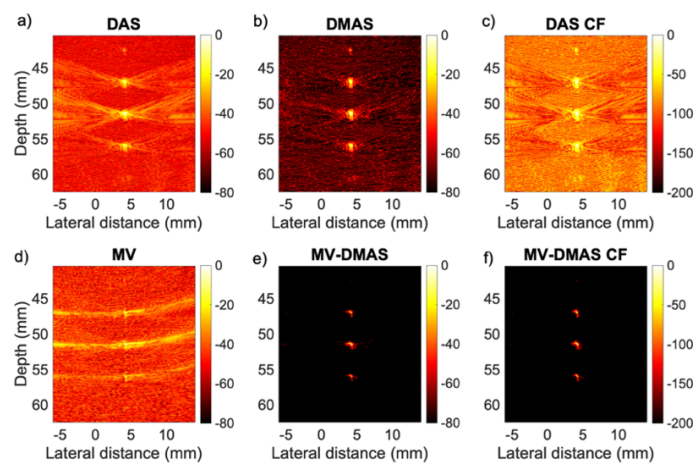

Figure 5. Experimental reconstructions.

The laser beam was centered at 50 mm depth and it can be seen that the absorber at 51.4 mm receives the maximum energy. The others have not received as much energy as the central one and, therefore, the emitted signals are weaker. This is noticeable for all algorithms, but it is especially significant in the adaptive ones since they enhance more the areas in which there are abrupt pulses. The main difference between threads at different depths is the maximum power level of each lobe, which is maximum for the central thread , whose profiles are represented in Figure 6(a, b). In this figure, it can also be seen that the dynamic range of 80 dB is not the optimum for the MV-DMAS CF, which would be best represented with a dynamic range of 200 dB, which is why results in Figure 5(c, f) have been represented with more dynamic range than the rest.
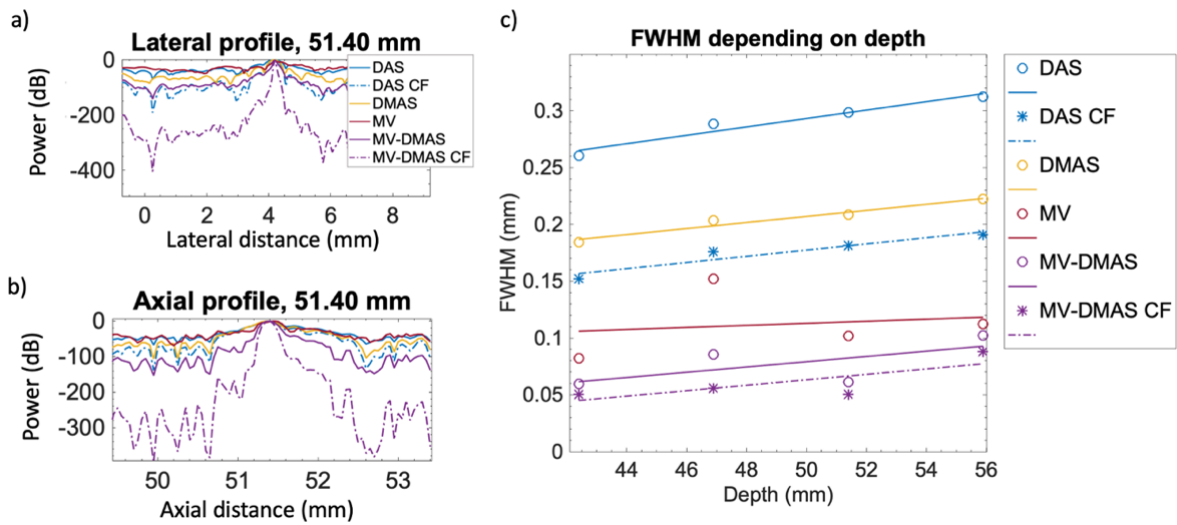

Figure 6. (a) Lateral and (b) axial profiles of the experimental reconstructions for the central thread, located at 51.40 mm depth. (c) Dependence of FWHM with depth.

MV-DMAS CF considerably improves resolution and SNR for both the central thread and its neighbors. To evaluate this quantitatively, the same parameter as for the numerical case has been used and, as shown in Figure 6(c), depth plays a fundamental role in the lateral resolution for the non-adaptive algorithms, since with just 15 mm extra in depth there is a difference of about 0.05 mm in the FWHM. However, for the adaptive algorithms, the position of the center of the laser spot becomes more important since they are just based on enhancing where there is already more signal. Therefore, when performing the linear adjustment of the data, it is found that the FWHM grows linearly with depth for non-adaptive techniques (blue and yellow lines in Fig. 6(c)), while for adaptive algorithms (red and purple lines) depth does not govern the FWHM. The location of the center of the laser spot area for MV and MV-DMAS has a greater effect than for non-adaptive algorithms. Therefore, the dependence of FWHM with depth does not grow linearly. Data does not fit to an increasing line with depth as in the simulation, where all sources had the same power, but the best resolutions are obtained for the thread located at 51.4 mm, which is the closest to the center of the laser spot.

Finally, the SNR has been evaluated and the obtained results are presented in Table 2. The greatest contrast between the signal and the background, as advanced by the results of the previous sections, is obtained for the central thread (51.400 mm) in all the reconstruction algorithms used. The thread at the top (46.852 mm), in most cases, has a better ratio than the bottom one (55.957 mm) because the center of the laser was slightly offset upwards. Furthermore, it is quantitatively confirmed that the MV-DMAS CF is the algorithm with the best SNR.

Table 2. Experimental SNR for the three ranges of depth.

	DAS (dB)	DAS CF (dB)	DMAS (dB)	MV (dB)	MV-DMAS (dB)	MV-DMAS CF (dB)
Upper	29.7	32.3	31.8	30.2	34.5	36.4
Central	34.7	37.7	37.0	30.0	43.2	44.6
Lower	27.9	30.8	30.2	28.6	33.7	36.2

4   Conclusions

In this study we quantify the performance of several reconstruction algorithms for photoacoustic tomography. A PAT system was developed using a linear array, and several reconstruction algorithms were implemented. In particular, Delay and Sum (DAS), Delay Multiply and Sum (DMAS), Minimum Variance (MV) and MV-DMAS have been numerically and experimentally evaluated, along with the coherence factor. It is observed that, under realistic conditions, the DAS algorithm involves low-resolution reconstructions and prominent lateral lobes, as observed in both numerical and experimental results. This makes it necessary to use more complex algorithms to compensate the limited-aperture detector. After the implementation and the numerical and experimental evaluation of both the adaptive and non-adaptive algorithms, it is concluded that the non-adaptive algorithms (DAS and DMAS) are more appropriate for real-time image reconstruction of large scenarios, they allow obtaining a general image of the reconstructed scenario in real time. In this way, non-adaptive algorithms are especially useful for applications in which the goal is to quantify the concentration of a particular photoacoustic contrast or to visualize large objects. However, adaptive algorithms (MV and MV-DMAS) provide high resolution and high contrast images. Therefore, they are best suited when looking to image small and close objects and identify them as two different entities. It should be noticed that they can differentiate very close objects, but they could also give a wrong perception of the real dimension of the object. Finally, it is concluded that the CF improves the properties of the reconstructed images for all cases.

Acknowledgements

References

[1]   P. Beard, "Biomedical photoacoustic imaging," *Interface Focus*, vol. 1, no. 4, pp. 602–631, 2011, doi: 10.1098/rsfs.2011.0028.

[2]   C. Tian, M. Pei, K. Shen, S. Liu, Z. Hu, and T. Feng, "Impact of System Factors on the Performance of Photoacoustic Tomography Scanners," *Phys. Rev. Appl.*, vol. 13, no. 1, p. 1, 2020, doi: 10.1103/PhysRevApplied.13.014001.

[3]   J. Joly, J. Vanlauwe, D. Van Assche, I. Reischl, and F. P. Luyten, "High-speed three-dimensional photoacoustic computed tomography for preclinical research and clinical translation," *Tissue Eng. Second Ed.*, no. 2021, pp. 783–807, 2014, doi: 10.1016/B978-0-12-420145-3.00022-5.

[4]   H. Jiang, *Photoacoustic Tomography*. 2014.

[5]   R. A. Kruger, R. B. Lam, D. R. Reinecke, S. P. Del Rio, and R. P. Doyle, "Photoacoustic angiography

of the breast," *Med. Phys.*, vol. 37, no. 11, pp. 6096–6100, 2010, doi: 10.1118/1.3497677.

[6]   V. Perrot, M. Polichetti, F. Varray, and D. Garcia, "So you think you can DAS? A viewpoint on delay-and-sum beamforming," *Ultrasonics*, vol. 111, no. May 2020, p. 106309, 2021, doi: 10.1016/j.ultras.2020.106309.

[7]   B. E. Treeby, E. Z. Zhang, and B. T. Cox, "Photoacoustic tomography in absorbing acoustic media using time reversal," *Inverse Probl.*, vol. 26, no. 11, 2010, doi: 10.1088/0266-5611/26/11/115003.

[8]   R. Schulze *et al.*, "On the use of frequency-domain reconstruction algorithms for photoacoustic imaging," *J. Biomed. Opt.*, vol. 16, no. 8, p. 086002, 2011, doi: 10.1117/1.3605696.

[9]   H. Deng, H. Qiao, Q. Dai, and C. Ma, "Deep learning in photoacoustic imaging: a review," *J. Biomed. Opt.*, vol. 26, no. 04, pp. 1–32, 2021, doi: 10.1117/1.jbo.26.4.040901.

[10] M. Mozaffarzadeh, A. Mahloojifar, M. Orooji, K. Kratkiewicz, S. Adabi, and M. Nasiriavanaki, "Linear-array photoacoustic imaging using mínimum variance-based delay multiply and sum adaptative beamforming algorithm," *J. Biomed. Opt.*, vol. 23, 2018, doi: 10.1117/1.JBO.23.2.026002.

[11] G. Matrone, A. S. Savoia, G. Caliano, and G. Magenes, "The delay multiply and sum beamforming algorithm in ultrasound B-mode medical imaging," *IEEE Trans. Med. Imaging*, vol. 34, no. 4, pp. 940–949, 2015, doi: 10.1109/TMI.2014.2371235.

[12] N. Q. Nguyen and R. W. Prager, "Minimum Variance Approaches to Ultrasound Pixel-Based Beamforming," *IEEE Trans. Med. Imaging*, vol. 36, no. 2, pp. 374–384, 2017, doi: 10.1109/TMI.2016.2609889.

[13] J. Camacho, M. Parrilla, and C. Fritsch, "Phase coherence imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 56, no. 5, pp. 958–974, 2009, doi: 10.1109/TUFFC.2009.1128.

[14] M. Mozaffarzadeh, M. Mehrmohammadi, and B. Makkiabadi, "Image improvement in linear-array photoacoustic imaging using high resolution coherence factor weighting technique," *BMC Biomed. Eng.*, vol. 9, pp. 1–11, 2019, doi: 10.1186/s42490-019-0009-9.

[15] J. R. Cook, R. R. Bouchard, and S. Y. Emelianov, "Tissue-mimicking phantoms for photoacoustic and ultrasonic imaging," *Biomed. Opt. Express*, vol. 2, no. 11, p. 3193, 2011, doi: 10.1364/boe.2.003193.