Image reconstruction from photoacoustic projections

Chao Tian,[a,b,c,d] Kang Shen,[b] Wende Dong,[e] Fei Gao,[f,g] Kun Wang,[h] Jiao Li,[i] Songde Liu,[b,c] Ting Feng,[j] Chengbo Liu,[k] Changhui Li,[l,m] Meng Yang,[n,*] Sheng Wang,[c,*] and Jie Tian[h,o,p,*]

[a]Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, China
[b]School of Engineering Science, University of Science and Technology of China, Hefei, China
[c]Department of Anesthesiology, the First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China
[d]Anhui Province Key Laboratory of Biomedical Imaging and Intelligent Processing, Hefei, China
[e]College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China
[f]School of Information Science and Technology, ShanghaiTech University, Shanghai, China
[g]Shanghai Clinical Research and Trial Center, Shanghai, China
[h]CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, China
[i]School of Precision Instruments and Optoelectronics Engineering, Tianjin University, Tianjin, China
[j]Academy for Engineering and Technology, Fudan University, Shanghai, China
[k]Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China
[l]Department of Biomedical Engineering, College of Future Technology, Peking University, Beijing, China
[m]National Biomedical Imaging Center, Peking University, Beijing, China
[n]Departments of Ultrasound, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
[o]School of Engineering Medicine, Beihang University, Beijing, China
[p]Key Laboratory of Big Data-Based Precision Medicine, Beihang University, Ministry of Industry and Information Technology, Beijing, China

Abstract. Photoacoustic computed tomography (PACT) is a rapidly developing biomedical imaging modality and has attracted substantial attention in recent years. Image reconstruction from photoacoustic projections plays a critical role in image formation in PACT. Here we review six major classes of image reconstruction approaches developed in the past three decades, including delay and sum, filtered back projection, series expansion, time reversal, iterative reconstruction, and deep-learning-based reconstruction. The principal ideas and implementations of the algorithms are summarized, and their reconstruction performances under different imaging scenarios are compared. Major challenges, future directions, and perspectives for the development of image reconstruction algorithms in PACT are also discussed. This review provides a self-contained reference guide for beginners and specialists in the photoacoustic community, to facilitate the development and application of novel photoacoustic image reconstruction algorithms.

Keywords: optical imaging; photoacoustic tomography; image reconstruction; deep learning.

1 Introduction

Photoacoustic tomography (PAT), also referred to as opto-acoustic tomography, is a cross-sectional or three-dimensional (3D) biomedical imaging technique. The physical foundation for PAT is the photoacoustic effect discovered in 1880 by Alexander Graham Bell[1], who found that modulated light can excite sound waves in materials. Modern PAT typically uses a short-pulsed laser to illuminate biological tissues. The laser energy is absorbed by tissue chromophores and converted into heat. Under the conditions of stress confinement and thermal confinement, acoustic signals are generated due to the thermo-elastic effect and are recorded by ultrasound detectors to recover

*Address all correspondence to Meng Yang, yangmeng_pumch@126.com; Sheng Wang, iamsheng2020@ustc.edu.cn; and Jie Tian, tian@ieee.org

the optical absorption property of tissues[2–5]. By combining optical excitation and acoustic detection, PAT has the advantages of rich optical contrast and high ultrasonic resolution in deep tissues and has found unique applications in a range of biomedical fields[6–9].

PAT has three major embodiments: photoacoustic computed tomography (PACT), photoacoustic microscopy (PAM), and photoacoustic endoscopy (PAE). Among them, PACT uses diffused light to illuminate biological tissues and can achieve hundred-micron-level imaging resolution at centimeter-level tissue depth[10–13]. Analogous to X-ray computed tomography, high-quality image formation in PACT relies on sophisticated image reconstruction procedures, without which the details in an image may not be able to be properly resolved. In 1981, Bowen studied thermoacoustic imaging of soft tissues using ionizing radiation (high-energy electrons, X-ray photons, neutrons, and other charged particles) and non-ionizing radiation (radio waves, microwaves, and ultrasonic waves)[14–16] and presented one-dimensional (1D) depth-resolved thermoacoustic signals from a human upper arm using radio-wave excitation. However, the results are only 1D depth-resolved and have no lateral resolution. In the mid-1990s, Oraevsky[17–19] and Kruger[20,21] independently studied laser-induced photoacoustic imaging of biological tissues and presented experimental 1D signals and two-dimensional (2D) scan images (see Fig. 1). The results are depth-resolved but poorly laterally resolved due to the lack of image reconstruction procedures. To obtain cross-sectional or 3D images resolved both axially and laterally, Kruger *et al*. in 1995 developed an approximate filtered back projection (FBP) image reconstruction algorithm for laser-induced PACT analogous to the FBP algorithm used in X-ray computed tomography[22]. The results suggest that the approximate FBP algorithm helps resolve imaging targets both in the axial and lateral directions (see Fig. 1).

High-performance image reconstruction is critical to image formation in modern PACT imaging systems. Ever since the publication of the approximate FBP algorithm by Kruger and colleagues, several novel image reconstruction approaches have been proposed to reconstruct high-quality PACT images. According to whether the emerging deep learning technique is involved, image reconstruction approaches in PACT can be divided into two main categories: conventional image reconstruction methods and deep learning (DL)-based image reconstruction methods. Conventional image reconstruction methods do not involve deep learning and mainly include five classes of algorithms: FBP, delay and sum (DAS), series expansion (SE), time reversal (TR), and iterative reconstruction (IR).

FBP is a class of analytical image reconstruction algorithms originating from X-ray computed tomography that project filtered photoacoustic signals back to target regions according to the time of flight (TOF) of photoacoustic signals. Kruger adopted this idea and proposed an approximate version of FBP for image reconstruction in PACT in 1995, as mentioned previously[22]. However, this algorithm can only achieve approximate image reconstruction. To reduce reconstruction error, Finch *et al*. developed an exact FBP algorithm for image reconstruction in a spherical detection geometry in 2004[23], and Xu and Wang developed a universal FBP algorithm for exact image reconstruction in infinite planar, cylindrical, and spherical detection geometries in 2005[24]. FBP algorithms are simple yet efficient and have been widely used for image reconstruction in PACT.

DAS-based reconstruction is essentially a simplified version of the FBP algorithm, which directly projects measured photoacoustic signals back to target regions without filtering. It is one of the most widely used beamforming methods in ultrasound imaging and can potentially be adapted for image reconstruction in PACT. In 1998, Hoelen *et al*. first used a simple weighted DAS algorithm for image reconstruction in PACT and achieved good reconstruction results[25,26]. Several modified algorithms were subsequently developed to improve the reconstruction performance of DAS[27–30]. DAS-based reconstruction algorithms are typically fast and easy to implement. However, they are intuitive and empirical image reconstruction techniques and can only achieve approximate image reconstruction.

SE is a class of image reconstruction algorithms that use mathematical series to represent the image to be reconstructed. In 2001, Köstli *et al*. derived an exact SE-based inverse formula in the Fourier domain for image reconstruction in a planar detection geometry[31]. Kunyansky took a step further and derived SE reconstruction formulas for image reconstruction in circular, spherical, cylindrical, and cubic geometries[32,33]. Compared with other image reconstruction algorithms in PACT, SE algorithms may be computationally efficient for certain detection geometries (e.g., planar geometry) due to the application of the fast Fourier transform (FFT) algorithm during computation[31].
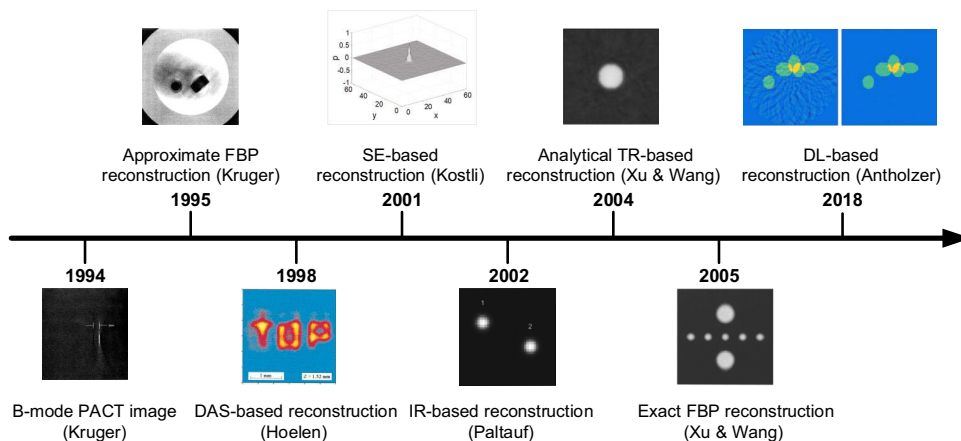

Fig. 1 Key events in the development of PACT image reconstruction algorithms.

TR algorithms recover photoacoustic images by running a forward numerical acoustic propagation model backward and re-transmitting the photoacoustic signals measured by each detector in a temporally reversed order. In 2004, Xu and Wang presented an analytical TR model for image reconstruction in arbitrary closed detection geometries[34]. Burgholzer *et al.* developed a numerical TR algorithm for image reconstruction in arbitrary detection geometries based on the finite difference method[35]. To improve computational efficiency and image quality, Treeby *et al.* further developed a numerical k-space pseudospectral TR algorithm for image reconstruction in heterogeneous media[34]. The TR algorithms can couple the acoustic properties of media (e.g., SOS, density, dispersion, and absorption) and can be used for image reconstruction in arbitrary closed detection geometries. Therefore, they are regarded as the "least restrictive" image reconstruction algorithms in PACT.

The last class of conventional image reconstruction methods is IR. It solves the image reconstruction problem by iteratively seeking the optimal solution that minimizes the error between measured projection data and the estimate from constructed mathematical models. In 2002, Paltauf *et al.* proposed the first IR method to improve image reconstruction quality under nonideal imaging conditions and achieved 3D image reconstruction with reduced artifacts[36]. After that, much research was conducted by different groups to improve the performance of IR, such as improving the computational accuracy of the system matrix in the model[37,38], compensating for the response of detectors[39,40], coupling the acoustic property of the media[41], and accelerating the reconstruction process[42,43]. Compared with other algorithms, IR algorithms are typically slow but can yield high-quality imaging results when only limited projection data are available. Therefore, they are more suitable for nonideal imaging scenarios, such as limited-view imaging, sparse-view imaging, and imaging in acoustically heterogeneous media.

The second major category of image reconstruction approaches in PACT is fast-developing DL-based methods, which are inspired by the successful use of DL in a range of fields. DL-based methods typically train neural networks and use them to automatically transform input data into output photoacoustic images. Compared with conventional image reconstruction approaches, DL methods are more efficient and can handle more complicated scenarios. In 2018, Antholzer *et al.* first proposed a deep convolutional neural network (CNN)-based method for PACT image reconstruction under sparse-view sampling and opened new opportunities for intelligent image reconstruction in PACT[44]. After that, a series of DL-based methods was developed for image reconstruction using both simulated and experimental datasets[45–47]. State-of-the-art DL methods are powerful enough to achieve preprocessing in the data domain, postprocessing in the image domain, hybrid processing in both the data and image domains, learned IR, and direct image reconstruction from the data domain to the image domain[48–50]. They have been used to address a range of image reconstruction problems in PACT, such as detector bandwidth expansion[51,52], resolution enhancement[53,54], artifact removal[55], ultralow-laser-energy imaging[56], reconstruction acceleration[57], and reconstruction enhancement in sparse-view and limited-view imaging[44,46,47,58,59].

Over the past three decades, great achievements have been made in PACT image reconstruction. To date, there have been a few excellent reviews devoted to summarizing the research progress in this field[60–63]. However, most review papers in the literature are from more than 10 years ago and cannot reflect the latest research advances[60–62]. Moreover, some review papers cover only particular subjects of the field, such as IR reconstruction[63,64] or DL-based reconstruction[48–50,65–68]. There is a pressing need to prepare new work to systematically review the recent achievements in this field. For the above reasons, we prepared this review, which differs from other works in the following aspects. First, the review was prepared from a historical perspective. We surveyed the development of each type of reconstruction algorithm and put them into the context of the entire history of PACT. Second, the review is comprehensive. It not only contains the five classes of conventional reconstruction algorithms (FBP, DAS, SE, TR, and IR) but also covers the state-of-the-art DL-based reconstruction algorithms. Third, the review contains comparative studies. A dedicated part (Sec. 5) was prepared to compare the performance of each type of image reconstruction algorithm in terms of image reconstruction quality, reconstruction speed, and memory footprint. Finally, the review is expected to be beginner-friendly. The entire review contains 51 figures, 14 tables, 126 mathematical equations, and comparative studies of each algorithm and is easy for novices to understand.

The remainder of this review is organized as follows. Section 2 describes the forward problem of PACT, including photoacoustic signal generation, propagation, detection, and the mathematical foundation (i.e., the Radon transform) for image reconstruction. Section 3 reviews the basic principles of the five classes of conventional image reconstruction algorithms, namely, FBP, DAS, SE, TR, and IR. Section 4 surveys the DL-based image reconstruction methods from the perspective of preprocessing, postprocessing, hybrid processing, learned IR, and direct image reconstruction. Section 5 provides comparative studies of major image reconstruction algorithms in PACT. Section 6 highlights the major challenges and future directions for PACT image reconstruction. Finally, Sec. 7 offers concluding remarks. Figure 2, Table 1, and Table 2 list major topics, abbreviations, and symbols used in this review, respectively.

It is worth pointing out that image reconstruction in PACT involves two aspects: acoustic inversion for determining the distribution of initial acoustic pressure and optical inversion for determining the map of optical absorption[61,62,69]. This review focuses on the acoustic inversion problem, which is important for image formation in PACT. Moreover, the mathematical foundation for image reconstruction in PACT is equivalent to that in thermoacoustic tomography[70–73]. Therefore, the image reconstruction algorithms discussed in this review can in principle be used for thermoacoustic tomography.

2 The Forward Problem

2.1 Photoacoustic Signal Generation

To achieve efficient photoacoustic signal excitation and generation, two conditions, i.e., thermal confinement and stress confinement, should be satisfied. The two conditions are related to two timescales, i.e., the thermal relaxation time τ_{th} and the stress relaxation time τ_s. The thermal relaxation time τ_{th} characterizes the diffusion of heat from the region heated by laser pulses, while the stress relaxation time τ_s describes the propagation of acoustic waves from the heated region. τ_{th} is given by[74]
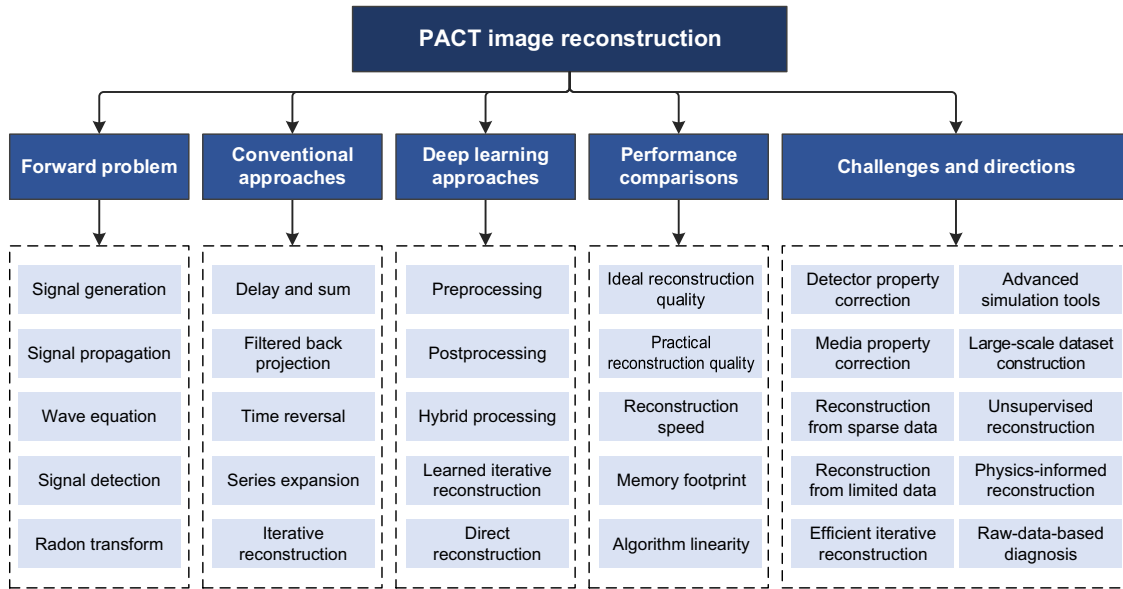
PACT image reconstruction

Forward problem	Conventional approaches	Deep learning approaches	Performance comparisons	Challenges and directions	
Signal generation	Delay and sum	Preprocessing	Ideal reconstruction quality	Detector property correction	Advanced simulation tools
Signal propagation	Filtered back projection	Postprocessing	Practical reconstruction quality	Media property correction	Large-scale dataset construction
Wave equation	Time reversal	Hybrid processing	Reconstruction speed	Reconstruction from sparse data	Unsupervised reconstruction
Signal detection	Series expansion	Learned iterative reconstruction	Memory footprint	Reconstruction from limited data	Physics-informed reconstruction
Radon transform	Iterative reconstruction	Direct reconstruction	Algorithm linearity	Efficient iterative reconstruction	Raw-data-based diagnosis

Fig. 2 Major topics discussed in this review.

Table 1 Abbreviations Used in this Review

Abbreviation	Meaning	Abbreviation	Meaning
1D	One-dimensional	2D	Two-dimensional
3D	Three-dimensional	CF	Coherence factor
CNN	Convolutional neural network	CNR	Contrast-to-noise ratio
CT	Computed tomography	DAS	Delay and sum
DMAS	Delay multiply and sum	DL	Deep learning
EIR	Electrical impulse response	FBP	Filtered back projection
FFT	Fast Fourier transform	GPU	Graphics processing unit
IR	Iterative reconstruction	MV	Minimum variance
PACT	Photoacoustic computed tomography	PAT	Photoacoustic tomography
ROI	Region of interest	SE	Series expansion
SIR	Spatial impulse response	SLSC	Short-lag spatial coherence
SNR	Signal-to-noise ratio	SOS	Speed of sound
TOF	Time of flight	TR	Time reversal

$$\tau_{\text{th}} = \frac{d_{\text{c}}^2}{\alpha_{\text{th}}}, \tag{1}$$

where d_{c} is the characteristic dimension of the heated region and α_{th} is the thermal diffusivity (m^2/s). τ_{s} is given by[74]

$$\tau_{\text{s}} = \frac{d_{\text{c}}}{v_0}, \tag{2}$$

where v_0 is the speed of sound (SOS). If the laser pulse duration is less than the thermal relaxation time, i.e., $\tau < \tau_{\text{th}}$, the thermal diffusion is negligible during laser heating, in which case the thermal confinement condition can be considered to be satisfied. Similarly, if the laser pulse duration is less than the stress relaxation time, i.e., $\tau < \tau_{\text{s}}$, the pressure propagation can be ignored during pulse heating, in which case the stress confinement condition can be considered to be satisfied. In PACT, the thermal and stress confinement conditions should be satisfied simultaneously. For instance, assuming that in soft tissues, the SOS is 1500 m/s, the thermal diffusivity is 0.14 mm^2/s, and the dimension of the heated region is 50 μm, the thermal relaxation time τ_{th} is calculated to be 18 ms, and the stress relaxation time τ_{s} is 33 ns. This indicates that the laser used for imaging should have a pulse duration of less than 33 ns to achieve efficient photoacoustic signal excitation and generation.

In PACT, the local fractional volume expansion $\Delta V/V$ induced by the photoacoustic effect can be described as[74]

Table 2 Symbols Used in this Review

Symbol	Meaning	Symbol	Meaning
	Lowercase English letters		
$b(\mathbf{r}_d, t)$	Back-projection term	$b(r)$	The Kaiser–Bessel function
d_c	Characteristic size of the heated region	$d\sigma$	Element of a detection surface S
$d\Omega$	Solid angle subtended by $d\sigma$	f	Frequency
$g(s, \theta)$	Radon transform	h_{EIR}	EIR of a detector
$\tilde{h}_{\mathrm{EIR}}$	Fourier transform of the EIR of a detector	h_{SIR}	SIR of a detector
$\tilde{h}_{\mathrm{SIR}}$	Fourier transform of the SIR of a detector	i	Imaginary unit
j_n	The spherical Bessel function of the first kind of order n	k_x, k_y, k_z	Spatial wavenumbers in the x, y, and z directions
n	A general variable	$p_0(x, y, z)$	Initial photoacoustic pressure (image to be reconstructed)
$p(\mathbf{r}, t)$	Photoacoustic signal at position $\mathbf{r}$ and time t	$p(\mathbf{r}_d, t)$	Real photoacoustic signal measured by a detector
$p_{\mathrm{ideal}}(\mathbf{r}_d, t)$	Ideal photoacoustic signal measured by a detector	$s_i(t)$	Photoacoustic signal measured by the ith detector at time t
t	Time	$u_k(\mathbf{r})$	Normalized eigenfunctions of the Dirichlet Laplacian
v_0	Sound speed		
	Uppercase English letters		
C_p	Specific heat capacity at constant pressure	C_v	Specific heat capacity at constant volume
F	Optical fluence	$G(\mathbf{r}_d, t; \mathbf{r}_s, t_s)$	The Green's function
H	Heat deposited per unit volume	$H_{|k|}^{(1)}$	The Hankel function of the first kind of order k
I_n	The modified Bessel function of the first kind of order n	$J_{|k|}$	The Bessel function of the first kind of order $|k|$
K	Sampling length	M	Total number of detectors
N	Total number of image grid points	N_x, N_y, N_z	Numbers of image grids along the x, y, and z axes
$P_0(k_x, k_y, k_z)$	Spatial Fourier transform of $p_0(x, y, z)$	$P(\mathbf{r}_d, \omega)$	Temporal Fourier transform of $p(\mathbf{r}_d, t)$
P_Ω	Poisson operator of harmonic extension	$R(\mathbf{x})$	Regularization
S	A detection surface	S_{DAS}	Image reconstructed by DAS
T	Temperature	V	Volume
$W(\omega)$	Window function in the frequency domain		
	Lowercase Greek letters		
α_0	Acoustic absorption coefficient	α_{th}	Thermal diffusivity
β	Thermal coefficient of volume expansion	δ	Dirac delta function
$\phi(\mathbf{r}_d, t)$	Velocity potential	η_{th}	Photothermal conversion efficiency
$\eta(\mathbf{r})$	Dispersion proportionality coefficient	κ	Isothermal compressibility
λ_m^2	Eigenvalues of the Dirichlet Laplacian	μ_a	Optical absorption coefficient
$\rho(\mathbf{r})$	Distribution of mass density	$\rho(\mathbf{r}, t)$	Acoustic density
$\rho_0(\mathbf{r})$	Ambient density	τ	Laser pulse duration
τ_{th}	Thermal relaxation time	τ_s	Stress relaxation time
$\tau(\mathbf{r})$	Absorption proportionality coefficient	ω	Angular frequency
$\psi(\mathbf{r})$	Expansion function		

(Table continued)

Uppercase Greek letters

Φ_{λ_k}	Free-space rotationally invariant Green's function	Γ	Grüneisen parameter
Ω	Solid angle of a detection surface or domain defined by a detection surface	Δf	Frequency sampling interval
Δp	Change in pressure	Δt	Temporal sampling interval
ΔT	Change in temperature	ΔV	Change in volume

Vectors or matrices

$\mathbf{n}_d$	Unit normal vector of a detector surface pointing to a photoacoustic source	$\mathbf{p}$	Photoacoustic signal in matrix form
$\mathbf{r} = (x, y, z)$	Rectangular coordinates in space	$\mathbf{r} = (R, \varphi, \theta)$	Spherical coordinates in space
$\mathbf{r}_d$	Detector position	$\mathbf{r}_s$	Photoacoustic source position
$\mathbf{u}(\mathbf{r}, t)$	Particle velocity	$\mathbf{x}$	Photoacoustic image in matrix form
$\mathbf{A}$	System matrix	$\mathbf{A}^\dagger$	Pseudo-inverse matrix of $\mathbf{A}$
$\mathbf{A}^*$	Adjoint matrix of $\mathbf{A}$	$\tilde{\mathbf{A}}$	The Fourier transform of $\mathbf{A}$
$\mathbf{A}^H$	Conjugate transpose of $\mathbf{A}$	$\mathbf{A}^T$	Transpose of $\mathbf{A}$
$\mathbf{D}$	Differential matrix	$\tilde{\mathbf{E}}$	Fourier transform of the EIR of a detector
$\mathbf{G}$	Spherical Radon transformation matrix	$\mathbf{I}$	Identity matrix
$\mathbf{P}$	The Fourier transform of $\mathbf{p}$	$\mathbf{R}$	Covariance matrix

Others symbols

$\mathcal{A}$	Forward acoustic propagation operator	$\mathcal{A}_{\text{modify}}^{\text{TR}}$	Modified TR operator
$\mathcal{F}$	The Fourier transform	$\mathcal{F}^{-1}$	The inverse Fourier transform
$\mathcal{H}(\mathbf{r}, t)$	Heating function	∇	Nabla operator

$$\frac{\Delta V}{V} = \beta \Delta T - \kappa \Delta p, \tag{3}$$

where ΔT and Δp represent the changes in temperature and pressure, respectively, β is the thermal coefficient of volume expansion, and κ is the isothermal compressibility, which can be written as

$$\kappa = \frac{C_p}{\rho v_0^2 C_v}, \tag{4}$$

where ρ is the mass density and C_p and C_v are the specific heat capacities at constant pressure and volume, respectively.

Under the conditions of thermal and stress confinement, the fractional volume expansion in Eq. (3) is negligible (i.e., $\Delta V/V = 0$). The initial photoacoustic pressure $p_0 = \Delta p$ and can be written as

$$p_0 = \frac{\beta \Delta T}{\kappa}, \tag{5}$$

where the temperature rise can be calculated by[75]

$$\Delta T = \frac{\eta_{\text{th}} H}{\rho C_v}. \tag{6}$$

Here η_{th} is the percentage of the laser energy converted into heat, and H is the heat energy deposited per unit volume, which is defined as the product of the absorption coefficient μ_a and the optical fluence F (i.e., $H = \mu_a F$). The initial acoustic pressure p_0 can thus be rewritten as

$$p_0 = \left(\frac{\beta v_0^2}{C_p}\right) \eta_{\text{th}} H = \Gamma \eta_{\text{th}} H, \tag{7}$$

where Γ is the Grüneisen parameter, a dimensionless constant representing the efficiency of the conversion of heat to pressure. For water at body temperature (37℃), $\Gamma \approx 0.2$, which indicates that 20% of the thermal energy deposited by a laser in water couples into acoustic energy.

The generation of photoacoustic signals can be illustrated using a numerical example. Assume that the laser employed for imaging has a fluence of 10 mJ/cm^2 ($F = 10$ mJ/cm^2), which is within the American National Standards Institute (ANSI) safety limit[76], and the biological tissue being imaged has the following physical parameters: $\Gamma = 0.2$, $\mu_a = 0.1$ cm^{-1}, $\rho = 1.0$ g/cm^3, and $C_v = 4.0$ Jg^{-1} K^{-1}. The factor η_{th} can be set to 1 because the fluorescence and nonthermal absorption of biological tissues are typically weak. In this way, the temperature rise ΔT is calculated to be 0.25 mK and the initial acoustic pressure p_0 is 200 Pa, which indicates that a 1 mK temperature rise can produce an acoustic pressure rise of 800 Pa.

2.2 Photoacoustic Signal Propagation

2.2.1 The photoacoustic wave equations

The generation and propagation of photoacoustic waves can be mathematically modeled by the photoacoustic wave equation[77]. Generally, three first-order equations, including the linearized equation of motion, the linearized equation of continuity, and the thermal elastic equation, can be used to model the properties

of acoustic wave generation and propagation[78]. If the medium is lossless and the thermal conductivity can be ignored, the three equations can be written as[79]

$$\frac{\partial}{\partial t}\mathbf{u}(\mathbf{r}, t) = -\frac{1}{\rho(\mathbf{r})}\nabla p(\mathbf{r}, t), \tag{8}$$

$$\nabla \cdot \mathbf{u}(\mathbf{r}, t) = -\frac{1}{\rho(\mathbf{r})v_0^2(\mathbf{r})}\frac{\partial}{\partial t}p(\mathbf{r}, t) + \beta\frac{\partial}{\partial t}\Delta T(\mathbf{r}, t), \tag{9}$$

$$\rho(\mathbf{r})C_\mathrm{p}\frac{\partial}{\partial t}T(\mathbf{r}, t) = \mathcal{H}(\mathbf{r}, t), \tag{10}$$

where $\mathbf{u}(\mathbf{r}, t)$ is the particle velocity, $v_0(\mathbf{r})$ is the SOS, $\rho(\mathbf{r})$ is the mass density, $p(\mathbf{r}, t)$ is the acoustic pressure at position $\mathbf{r}$ and time t, $T(\mathbf{r}, t)$ and $\Delta T(\mathbf{r}, t)$ are the temperature and temperature rise, respectively, and $\mathcal{H}(\mathbf{r}, t)$ is the heating function that represents the thermal energy deposited per unit volume and per unit time. The second-order photoacoustic wave equation can be obtained by eliminating the variable $\mathbf{u}(\mathbf{r}, t)$ from the three equations as[78,79]

$$\left\{\rho(\mathbf{r})\nabla\cdot\left[\frac{1}{\rho(\mathbf{r})}\nabla\right] - \frac{1}{v_0^2(\mathbf{r})}\frac{\partial^2}{\partial t^2}\right\}p(\mathbf{r}, t) = -\frac{\beta}{C_\mathrm{p}}\frac{\partial\mathcal{H}(\mathbf{r}, t)}{\partial t}. \tag{11}$$

Equation (11) describes the relationship between the acoustic pressure $p(\mathbf{r}, t)$ and the photoacoustic source associated with the heating function $\mathcal{H}(\mathbf{r}, t)$. The source term [right side of Eq. (11)] is proportional to the first derivative of the heating function $\mathcal{H}(\mathbf{r}, t)$, which indicates that the heating function $\mathcal{H}(\mathbf{r}, t)$ should be time-varying to achieve efficient photoacoustic signal generation.

When the medium is acoustically homogeneous, the photoacoustic wave equation in Eq. (11) reduces to[78,79]

$$\left(\nabla^2 - \frac{1}{v_0^2}\frac{\partial^2}{\partial t^2}\right)p(\mathbf{r}, t) = -\frac{\beta}{C_\mathrm{p}}\frac{\partial\mathcal{H}(\mathbf{r}, t)}{\partial t}. \tag{12}$$

2.2.2 Forward Green's function solution
The forward solution for the wave equation in a homogeneous medium in Eq. (12) can be obtained using Green's function[75,80] as

$$p(\mathbf{r}_\mathrm{d}, t) = \frac{\beta}{C_\mathrm{p}}\int_{-\infty}^{+\infty}\int_V G(\mathbf{r}_\mathrm{d}, t; \mathbf{r}_\mathrm{s}, t_\mathrm{s})\frac{\partial\mathcal{H}(\mathbf{r}_\mathrm{s}, t_\mathrm{s})}{\partial t_\mathrm{s}}\mathrm{d}\mathbf{r}_\mathrm{s}\mathrm{d}t_\mathrm{s}, \tag{13}$$

where $p(\mathbf{r}_\mathrm{d}, t)$ is the acoustic pressure detected at position $\mathbf{r}_\mathrm{d}$ and time t and $\mathbf{r}_\mathrm{s}$ and t_s represent the location and time of the photoacoustic source, respectively. The Green's function $G(\mathbf{r}_\mathrm{d}, t; \mathbf{r}_\mathrm{s}, t_\mathrm{s})$ in an infinite 3D space has the following form:

$$G(\mathbf{r}_\mathrm{d}, t; \mathbf{r}_\mathrm{s}, t_\mathrm{s}) = \frac{\delta(t - t_\mathrm{s} - |\mathbf{r}_\mathrm{d} - \mathbf{r}_\mathrm{s}|/v_0)}{4\pi|\mathbf{r}_\mathrm{d} - \mathbf{r}_\mathrm{s}|}, \tag{14}$$

where δ is the Dirac delta function.

Under the condition of stress confinement, the heating function can be decomposed as $\mathcal{H}(\mathbf{r}_\mathrm{s}, t_\mathrm{s}) = H(\mathbf{r}_\mathrm{s})\delta(t_\mathrm{s})$, where $H(\mathbf{r}_\mathrm{s})$ is the heat deposited per unit volume. In this way, Eq. (13) becomes

$$p(\mathbf{r}_\mathrm{d}, t) = \frac{\beta}{C_\mathrm{p}}\int_{-\infty}^{+\infty}\int_V G(\mathbf{r}_\mathrm{d}, t; \mathbf{r}_\mathrm{s}, t_\mathrm{s})H(\mathbf{r}_\mathrm{s})\delta'(t_\mathrm{s})\mathrm{d}\mathbf{r}_\mathrm{s}\mathrm{d}t_\mathrm{s}, \tag{15}$$

where δ' is the derivative of the Dirac delta function. Using the property $\int \delta'(t - t_0)f(t)\mathrm{d}t = -f'(t_0)$, Eq. (15) becomes

$$p(\mathbf{r}_\mathrm{d}, t) = \frac{\beta}{C_\mathrm{p}}\int_V H(\mathbf{r}_\mathrm{s})\frac{\partial G(\mathbf{r}_\mathrm{d}, t; \mathbf{r}_\mathrm{s}, t_\mathrm{s})}{\partial t}\mathrm{d}\mathbf{r}_\mathrm{s}\bigg|_{t_\mathrm{s}=0}. \tag{16}$$

Applying the relation in Eq. (7) (η_th is set to 1),

$$p_0(\mathbf{r}_\mathrm{s}) = \Gamma H(\mathbf{r}_\mathrm{s}) = \left(\frac{\beta v_0^2}{C_\mathrm{p}}\right)H(\mathbf{r}_\mathrm{s}), \tag{17}$$

and substituting the Green's function [Eq. (14)] into Eq. (16), we obtain the general forward solution of the photoacoustic wave equation, i.e.,

$$p(\mathbf{r}_\mathrm{d}, t) = \frac{1}{4\pi v_0^2}\frac{\partial}{\partial t}\int_V\frac{p_0(\mathbf{r}_\mathrm{s})}{|\mathbf{r}_\mathrm{d} - \mathbf{r}_\mathrm{s}|}\delta\left(t - \frac{|\mathbf{r}_\mathrm{d} - \mathbf{r}_\mathrm{s}|}{v_0}\right)\mathrm{d}\mathbf{r}_\mathrm{s}. \tag{18}$$

The measured pressure $p(\mathbf{r}_\mathrm{d}, t)$ is associated with the velocity potential $\phi(\mathbf{r}_\mathrm{d}, t)$ via[81]

$$p(\mathbf{r}_\mathrm{d}, t) = -\rho\frac{\partial\phi(\mathbf{r}_\mathrm{d}, t)}{\partial t}. \tag{19}$$

The velocity potential can be written as

$$\phi(\mathbf{r}_\mathrm{d}, t) = -\frac{1}{4\pi v_0^2\rho}\int_V\frac{p_0(\mathbf{r}_\mathrm{s})}{|\mathbf{r}_\mathrm{d} - \mathbf{r}_\mathrm{s}|}\delta\left(t - \frac{|\mathbf{r}_\mathrm{d} - \mathbf{r}_\mathrm{s}|}{v_0}\right)\mathrm{d}\mathbf{r}_\mathrm{s}. \tag{20}$$

Equations (18) and (20) indicate that the measured pressure $p(\mathbf{r}_\mathrm{d}, t)$ and velocity potential $\phi(\mathbf{r}_\mathrm{d}, t)$ are linearly related to the initial acoustic pressure $p_0(\mathbf{r}_\mathrm{s})$ and are inversely proportional to the distance between the detector and the source, i.e., $|\mathbf{r}_\mathrm{d}\mathbf{r}_\mathrm{s}|$.

2.2.3 Photoacoustic signal of a spherical absorber
The concept of acoustic pressure and velocity potential can be illustrated through a simple example[81]. Suppose that the photoacoustic source is a uniform sphere and that the detector is a point detector. We examine the velocity potential and acoustic pressure received by the point detector in two situations. In the first situation, the point detector is located at the center of the photoacoustic source, as shown in Fig. 3(a). The velocity potential can be obtained as [Eq. (20)]

$$\phi(\mathbf{r}_\mathrm{d}, t) = \begin{cases} -\dfrac{p_0(\mathbf{r}_\mathrm{s})}{v_0\rho}t, & \text{if } v_0 t \leq r_\mathrm{a}, \\ 0, & \text{if } v_0 t > r_\mathrm{a}, \end{cases} \tag{21}$$

where r_a is the radius of the spherical source. In the second situation, the point detector is located outside the photoacoustic source, as shown in Fig. 3(d). The velocity potential can be calculated as

$$\phi(\mathbf{r}_d, t) = \begin{cases} -\dfrac{p_0(\mathbf{r}_s)}{4v_0\rho d}[(d - v_0 t)^2 - r_a^2], & \text{if } d - r_a \leq v_0 t \leq d + r_a, \\ 0, & \text{otherwise,} \end{cases}$$

(22)

where d is the distance between the point detector and the center of the photoacoustic source. The acoustic pressure $p(\mathbf{r}_d, t)$ in these two cases can be accordingly calculated using the relationship in Eq. (19).

To quantify the velocity potential and acoustic pressure measured by the point detector, we make the following assumptions: the radius of the spherical source is 4 mm ($r_a = 4$ mm), the mass density of the medium is 1000 kg/m³ ($\rho = 1000$ kg/m³), the SOS is 1480 m/s ($v_0 = 1480$ m/s), the Grüneisen parameter at room temperature is 0.12 ($\Gamma = 0.12$), the optical absorption coefficient μ_a is 100 cm⁻¹ ($\mu_a = 100$ cm⁻¹), and the optical fluence is 10 mJ/cm² ($F = 10$ mJ/cm²). Therefore, the heat energy H deposited at time zero is 10^6 J/m³, and the initial acoustic pressure $p_0(\mathbf{r}_s)$ [Eq. (7)] is 1.2×10^5 Pa.

Figures 3(b) and 3(c) show the velocity potential and acoustic pressure measured by the point detector in the first case (the point detector is located at the center of the source). The results show that the negative velocity potential $-\phi(\mathbf{r}_d, t)$ measured by the point detector first linearly increases as the shell of the integration [Eq. (20)] increases with time and then suddenly drops to zero due to no energy deposition outside the source. Consequently, the acoustic pressure $p(\mathbf{r}_d, t)$ (i.e., the first derivative of the velocity potential) is a non-zero constant at the beginning and becomes zero with a negative impulse caused by the sudden drop of velocity potential to zero. Figures 3(e) and 3(f) show the velocity potential and acoustic pressure measured

by the point detector in the second case (the point detector is located 10 mm away from the center of the source). The results show that the negative velocity potential $-\phi(\mathbf{r}_d, t)$ measured by the point detector first increases as the incremental hemispherical shell of the integration [Eq. (20)] advances to the center of the source, and then decreases with time as the shell advances to the rear of the source. Consequently, the acoustic pressure $p(\mathbf{r}_d, t)$ is initially positive, passes through zero, then becomes negative, and has an N-shaped waveform.

Based on the example in Fig. 3, we can also deduce that the size of a photoacoustic source impacts the characteristics of the photoacoustic signals received by a detector in the time domain. To illustrate this, Fig. 4(a) shows three time-domain photoacoustic signals measured by a point detector 3 mm away from the centers of three uniform spherical sources with diameters of 1 mm, 200 µm, and 50 µm. The time-domain photoacoustic signals all have an N-shaped waveform as expected. However, the amplitude and duration of the photoacoustic signals are distinct. The photoacoustic signals produced by sources with a smaller size have smaller amplitudes and shorter durations, which correspond to higher frequencies and broader bandwidths in the frequency domain, as shown in Fig. 4(b). The center frequency of the photoacoustic signals f_c generated by a spherical source can be estimated by $f_c = v_0/3r_a$, where r_a is the radius of the spherical source[82]. This indicates that detectors with a high center frequency should be employed for high-frequency imaging.

2.2.4 Photoacoustic field visualization: the k-Wave toolbox

Numerical simulation of the forward propagation of photoacoustic fields helps visualize the sound fields in complex media
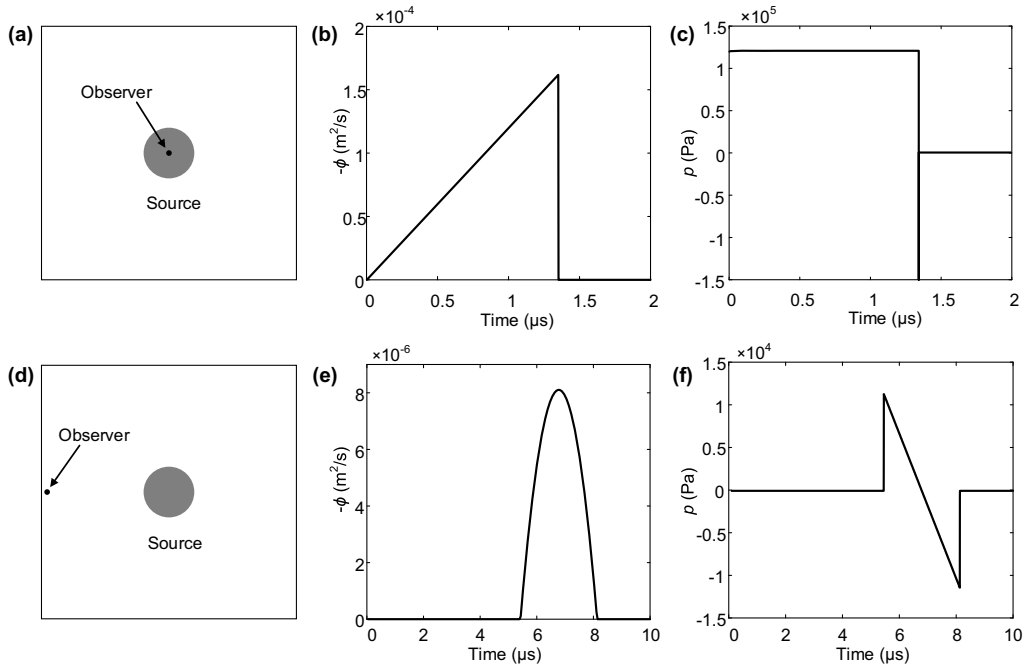

Fig. 3 Velocity potential and acoustic pressure generated from a 4-mm-diameter spherical source. The first and second rows show the results when the detector is located at the center of the source and is 10 mm away from the center of the source, respectively. (a), (d) Schematic diagrams showing the point detector and the spherical source. (b), (e) Negative velocity potentials at the point detector. (c), (f) Corresponding acoustic pressures.
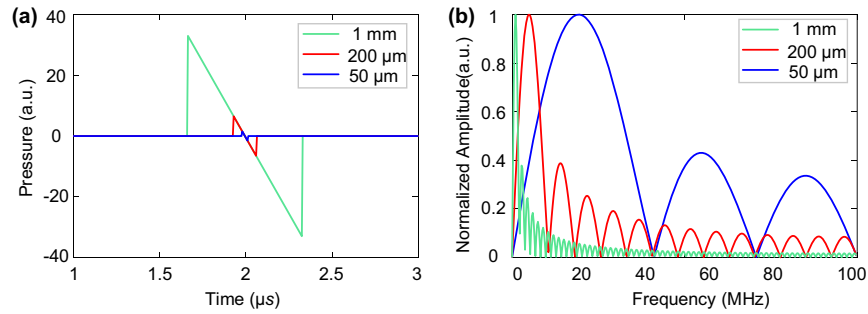
Fig. 4 Signals of spherical photoacoustic sources with different sizes and their Fourier spectra. (a) Time-domain N-shaped photoacoustic signals generated from three spherical sources with diameters of 1 mm, 200 µm, and 50 µm. (b) Normalized Fourier spectra of the corresponding photoacoustic signals. Reprinted from Ref. [82] with permission.

and solve the acoustic inversion problem in PACT. Numerical simulation can be implemented using the powerful open-source k-Wave toolbox[83], which was developed by the Photoacoustic Imaging Group at University College London (UCL). In the k-Wave toolbox, the propagation of photoacoustic fields is modeled by three coupled first-order partial differential equations[84], which are fundamentally the same as the three equations in Eqs. (8)–(10). The propagation model considers the acoustic properties of media, such as acoustic speed, dispersion, and absorption, and can characterize the acoustic propagation problem in heterogeneous media. The k-Wave toolbox solves the propagation model via a k-space pseudospectral method[83,85,86], which can perform fast and accurate computations with reduced memory. Figure 5 is an example showing the propagation of the photoacoustic fields of a 2D disk and a 3D sphere using the k-Wave toolbox. The results visualize the instant characteristics of the photoacoustic fields during propagation.

2.3 Photoacoustic Signal Detection

The photoacoustic signals propagating outward from a photoacoustic source need to be captured by ultrasound detectors

for image formation. Ideally, for photoacoustic signal detection, detector arrays that can perfectly record original signals in time and space should be used. However, perfect detector arrays never exist in reality. The characteristics of a practical detector array, such as detector aperture, detector bandwidth, detector number, and view angle, impact detected photoacoustic signals and final images.

Aperture and bandwidth are two fundamental characteristics of an ultrasound detector and have important impacts on measured photoacoustic signals. An ideal ultrasound detector should have a point-like aperture, in which way it has an omnidirectional response. Moreover, an ideal ultrasound detector should also have an infinite bandwidth so that it can respond to all frequency contents of a signal. Neither of the two conditions, however, is attainable in practice. A practical ultrasound detector always has a finite aperture size and a finite bandwidth. The finite aperture averages photoacoustic signals in space, resulting in a smoothed spatial impulse response (SIR), while the finite bandwidth affects the conversion of photoacoustic signals to electrical signals, leading to a degraded electrical impulse response (EIR). Taking the SIR and EIR of an ultrasound detector into account, the photoacoustic signal measured by a
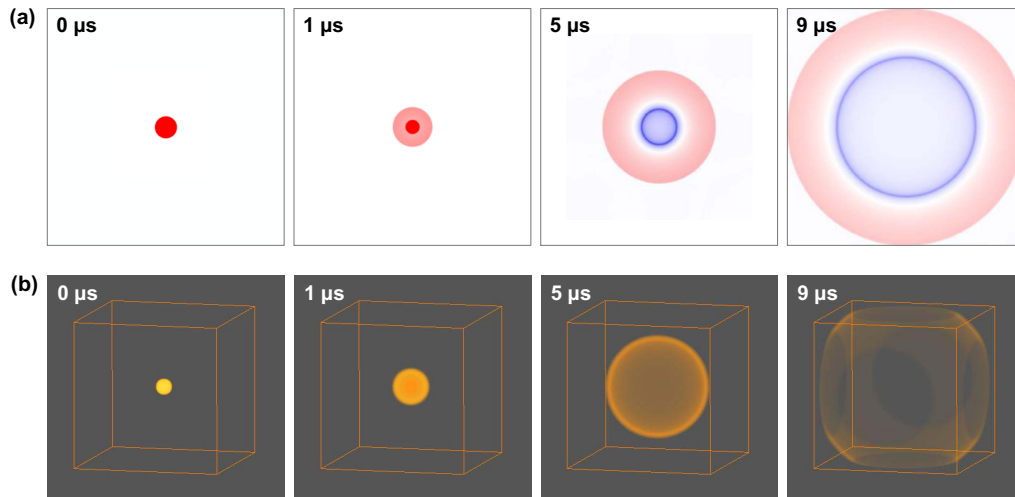

Fig. 5 Photoacoustic field visualization using the k-Wave toolbox. (a) Propagation of photoacoustic fields generated from a 2D disk (diameter: 2 mm). Red: positive pressure; blue: negative pressure. (b) Propagation of photoacoustic fields generated from a 3D sphere (diameter: 2 mm). For observation, negative pressure is not displayed in this case.
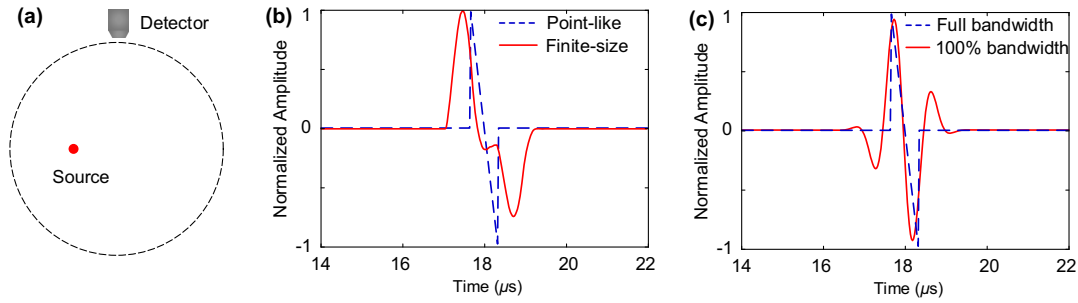
Fig. 6 Effects of the SIR and EIR of a detector on photoacoustic signals. (a) Schematic diagram showing a photoacoustic source and a circular detector array. The photoacoustic source is a 1-mm-diameter sphere and is 10 mm away from the center of the detector array, which has a radius of 25 mm. (b) Effect of the finite aperture size (SIR) on the photoacoustic signals. In this case, the height and width of each detector are set to 10 and 5 mm, respectively. (c) Effect of finite bandwidth on the photoacoustic signals. In this case, the center frequency and fractional bandwidth of the detector are set to 1 MHz and 100%, respectively.

practical ultrasound detector can be mathematically formulated as[4]

$$p(\mathbf{r}_d, t) = p_{ideal}(\mathbf{r}_d, t) * h_{SIR}(\mathbf{r}_d, \mathbf{r}_s, t) * h_{EIR}(t), \qquad (23)$$

where $p_{ideal}(\mathbf{r}_d, t)$ is the ideal photoacoustic signal, $*$ represents temporal convolution, and $h_{SIR}(\mathbf{r}_d, \mathbf{r}_s, t)$ and $h_{EIR}(t)$ represent the SIR and EIR of a detector, respectively. The non-ideal SIR and EIR of an ultrasound detector distort photoacoustic signals measured in the time domain, as illustrated in Fig. 6, which eventually degrades the image quality of photoacoustic images.

In addition to the aperture and bandwidth of a detector, the detector number and view angle of a detector array also play important roles in photoacoustic signal detection. Ideally, the number of detectors in a detector array used for photoacoustic signal measurement should meet the spatial Nyquist sampling theorem for complete signal acquisition[4]. The view angle of a detector array should be 4π steradian (full 3D view) to record complete photoacoustic signals in 3D space. However, these two conditions are actually unattainable in practice due to

the high fabrication cost of a large number of detectors and the configuration constraints of an imaging system (e.g., a separate space is required for laser illumination). Violating these conditions leads to the problems of image reconstruction from sparse-view and limited-view projections, which are mathematically challenging and will be discussed later in this review. Figure 7 showcases commonly used 2D and 3D detection geometries in PACT[87,88], which have limited view angles except for the closed spherical array.

2.4 Radon Transform

The forward signal propagation and detection processes in PACT can be described by the well-known Radon transform[89], which is the mathematical foundation of computed tomography (CT).

Before discussing the Radon transform in PACT, we first introduce the linear Radon transform in X-ray CT, which is defined as the integral of a function along a straight line. Specifically, the linear Radon transform in X-ray CT is written as[90]
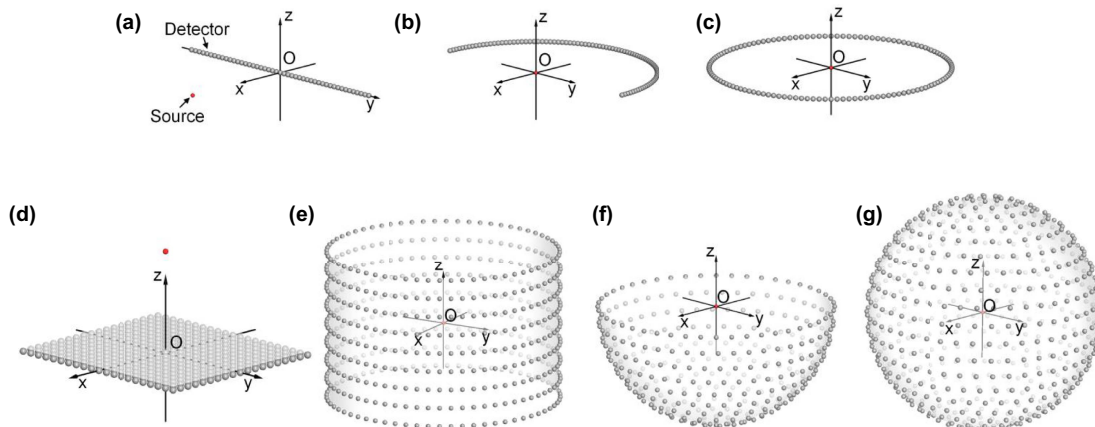

Fig. 7 Commonly used detection geometries in PACT. (a) Linear array. (b) Curved array. (c) Circular array. (d) Planar array. (e) Cylindrical array. (f) Hemispherical array. (g) Closed spherical array. Reprinted from Refs. [87,88] with permission.

$$g(s, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - s) \mathrm{d}x \mathrm{d}y, \tag{24}$$

where $f(x, y)$ is the original function, $g(s, \theta)$ is the sinogram or projection data of the function $f(x, y)$ along the straight line defined in the Dirac delta function δ, and (s, θ) are two parameters of the normal equation of the straight line in the delta function. The inverse Radon transform reverses the forward process and recovers the original function $f(x, y)$ from measured sinograms $g(s, \theta)$. A schematic representation of the linear Radon transform is shown in Fig. 8(a).

$$g(s, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta \left[s - \sqrt{(x - r \cos \theta)^2 + (y - r \sin \theta)^2} \right] \mathrm{d}x \mathrm{d}y, \tag{25}$$

where (r, θ) represents the position of the detector in a polar coordinate system and s is the radius of the integral circle. Similarly, the 3D spherical Radon transform $g(s, \theta)$ can be mathematically written as[92]

$$g(s, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z) \delta \left[s - \sqrt{(x - r \sin \varphi \cos \theta)^2 + (y - r \sin \varphi \sin \theta)^2 + (z - r \cos \varphi)^2} \right] \mathrm{d}x \mathrm{d}y \mathrm{d}z, \tag{26}$$

In PACT and thermoacoustic tomography, the Radon transform integrates a function along a circle (2D) or a sphere (3D) instead of a straight line. The 2D circular Radon transform can be written as[91]

where (r, φ, θ) represents the position of the detector in a spherical coordinate system and s is the radius of the spherical shell to be integrated. Schematic representations of the circular and spherical Radon transforms are shown in Figs. 8(b) and 8(c).
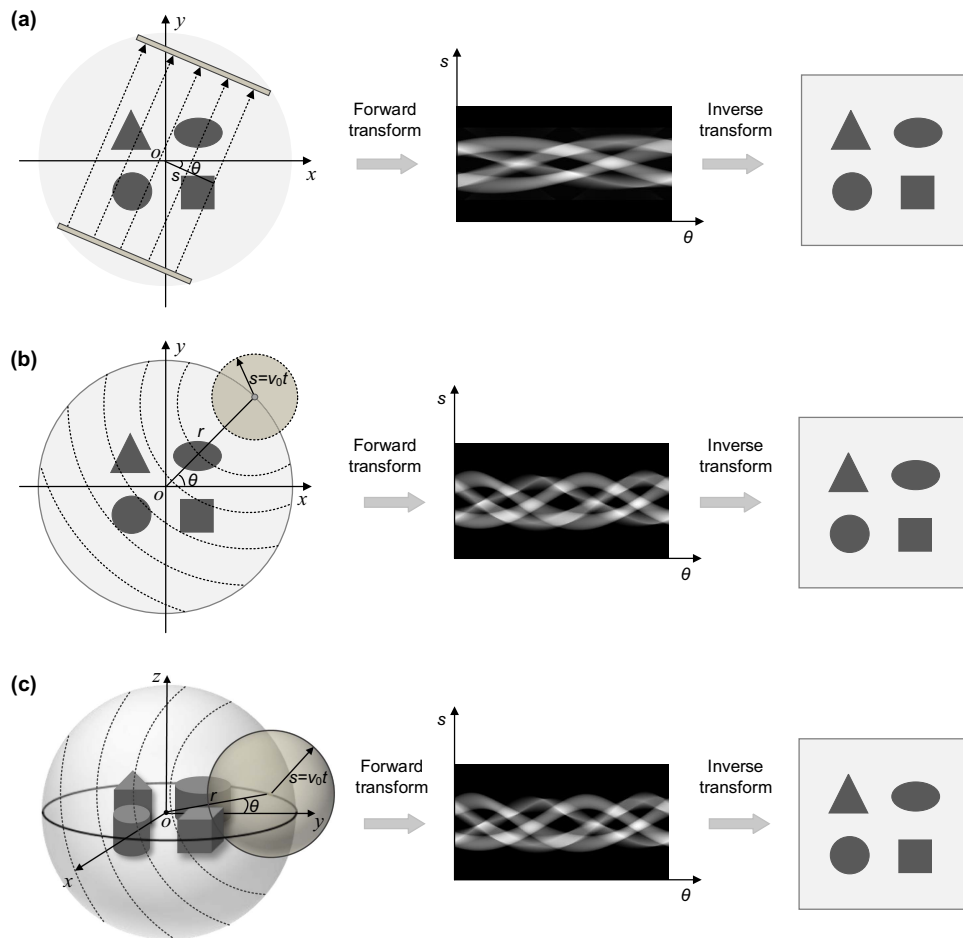
Fig. 8 Forward and inverse Radon transforms in X-ray CT and PACT. (a) Linear Radon transform and its inverse in X-ray CT. (b), (c) Circular and spherical Radon transforms and their inverses in PACT.

Using the definition in Eq. (26), the spherical Radon transform in PACT can be obtained from Eq. (18) and has the following form:

$$g(\mathbf{r}_d, t) = \int p_0(\mathbf{r}_s)\delta\left(t - \frac{|\mathbf{r}_d - \mathbf{r}_s|}{v_0}\right)d\mathbf{r}_s. \qquad (27)$$

Equation (27) is in the form of the spherical Radon transform [Eq. (26)], representing the integral over a spherical shell with a radius of $v_0 t$ centered at the detector position $\mathbf{r}_d$. The projection data $g(\mathbf{r}_d, t)$ are related to the measured photoacoustic signal $p(\mathbf{r}_d, t)$ via

$$g(\mathbf{r}_d, t) = 4\pi v_0^2 |\mathbf{r}_d - \mathbf{r}_s| \int_0^t p(\mathbf{r}_d, \tau)d\tau. \qquad (28)$$

The projection data $g(\mathbf{r}_d, t)$ have a similar definition as the velocity potential in Eq. (19) and can be calculated from the measured photoacoustic signal $p(\mathbf{r}_d, t)$. The task of image reconstruction in PACT is to find the inverse spherical Radon transform and recover the original function $f(x, y)$ or $p_0(\mathbf{r}_s)$ from the measured photoacoustic signal $p(\mathbf{r}_d, t)$, which is the focus of the following sections.

3 Conventional Approaches

3.1 DAS-Type Algorithms

Delay and sum (DAS)-based beamforming is a commonly used image reconstruction technique in ultrasound imaging[93]. Due to similar image formation processes, DAS-based beamforming is also widely used in PACT, where it reconstructs an image by summing the delayed raw photoacoustic signals of each detector. The time delay between each channel is calculated according to the acoustic TOF of the photoacoustic signals from the point of interest to each ultrasound detector. To yield high-quality images, preprocessing of raw photoacoustic signals and/or postprocessing of reconstructed images are usually needed. Here we review the five most commonly used DAS-type image reconstruction algorithms: DAS, delay multiply and sum (DMAS), short-lag spatial coherence (SLSC),

minimum variance (MV), and coherence factor (CF), which adopt different preprocessing and/or postprocessing strategies. The basic workflows of these methods are first illustrated in Fig. 9, and detailed descriptions are presented below.

3.1.1 Delay and sum

DAS is the most basic beamforming algorithm in ultrasound imaging due to its simplicity, speed, and robustness[94–96]. Hoelen *et al.* introduced a DAS method for 3D PACT imaging of blood vessels based on a planar detection geometry in 1998[26,97]. Feng *et al.* applied a DAS method in linear-scanning thermoacoustic tomography in 2001[98]. The DAS method can be mathematically formulated as

$$S_{DAS}(x, z) = \sum_{i=1}^{M} s_i(t), \qquad (29)$$

where S_{DAS} is the reconstructed image, $s_i(t)$ is the photoacoustic signal measured by the ith detector at time t, M is the total number of detectors, and (x, z) represents the position in a coordinate system. The variable t in $s_i(t)$ denotes the TOF of the photoacoustic signals from position (x, z) to the ith detector. The workflow and principle of the DAS algorithm are illustrated in Figs. 9(a) and 10, respectively.

The performance of the DAS algorithm was evaluated using a designed numerical phantom with multiple point sources equally distributed along the longitudinal centerline [Fig. 11(a)]. In the evaluation, a linear detector array with a width of 200 mm and 128 elements was placed at the top of the phantom to receive the photoacoustic signals generated from the sources. Figure 11(b) shows the image reconstructed by the DAS method. The results show that DAS can successfully reconstruct the structural information of photoacoustic sources but fails to reproduce the correct amplitude. The reconstructed image has both positive and negative components and is bipolar in magnitude. A unipolar image can be obtained by finding the envelope of the bipolar image, as shown in Fig. 11(c). Log transformation can be further used to improve the contrast of the reconstructed image, as shown in Fig. 11(d). Since DAS treats the delayed photoacoustic signals $s_i(t)$ from different detectors
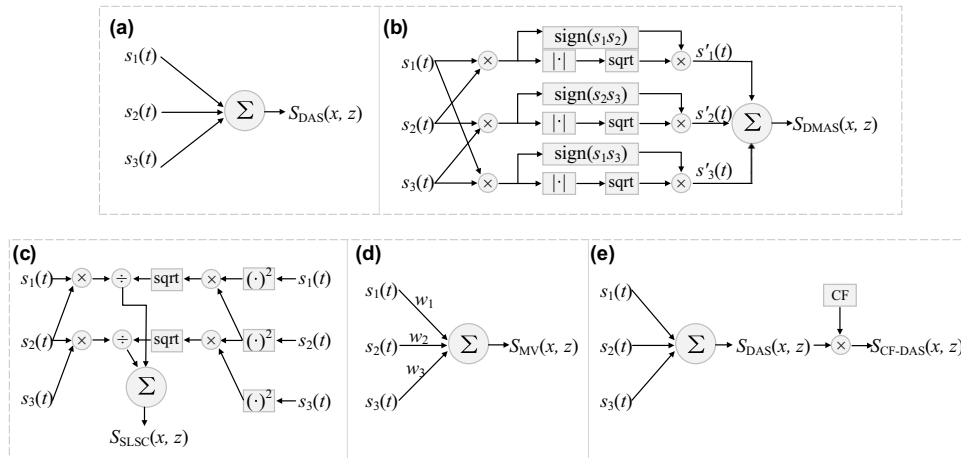

Fig. 9 Workflows of different DAS-based image reconstruction algorithms in PACT. (a) DAS. (b) DMAS. (c) SLSC ($M = 3$, $L = 1$, $l = 1$, $t_1 = t_2 = t$). (d) MV. (e) CF-DAS; sign denotes the signum function and sqrt denotes the square root.
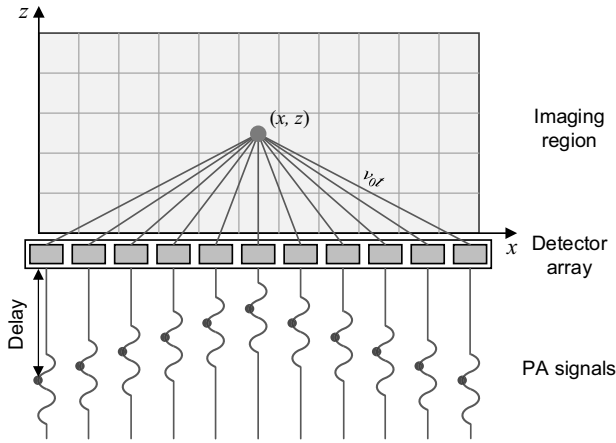
Fig. 10 Principle of DAS-based image reconstruction.

equally and is non-adaptive, significant side lobes are present in the reconstructed image in Fig. 11(d), which degrades the spatial resolution of the image.

3.1.2 Delay multiply and sum

Delay multiply and sum (DMAS) is a variant of the DAS algorithm that can provide improved contrast, signal-to-noise ratio (SNR), and resolution. Similar to DAS, DMAS also sums photoacoustic signals measured by different detectors according to calculated time delays. However, the measured photoacoustic signals in DMAS need to be combinatorically coupled and multiplied before summation. Therefore, DMAS is essentially a nonlinear algorithm.

The DMAS algorithm was first proposed by Lim *et al.* for confocal microwave detection of breast cancer in 2008 and showed improved identification of embedded malignant tumors in a variety of numerical breast phantoms compared with DAS[99]. In 2015, Matrone *et al.* modified this method for B-mode ultrasound imaging and demonstrated that DMAS can provide higher contrast resolution than DAS[100]. Inspired by the success of DMAS in confocal microwave imaging and ultrasound imaging, several research groups worldwide have conducted in-depth studies of DMAS-based photoacoustic image reconstruction. For example, Alshaya *et al.* introduced DMAS to the field of PACT in 2016 and proposed a filter DMAS to

improve the SNR of reconstructed images[101]. To further improve the performance of DMAS, in 2018, Mozaffarzadeh *et al.* developed a double-stage DMAS algorithm that can produce images with improved quality compared with DAS and DMAS but at the expense of greater computational cost[27]. In the same year, Kirchner *et al.* proposed a signed DMAS algorithm that can provide linear image reconstruction with increased image quality[102]. In 2022, Mulani *et al.* presented a high-order DMAS method, in which multi-term (e.g., three, four, or five terms) multiplication is used to replace two-term multiplication in the original DMAS algorithm[103].

Generally, the original DMAS algorithm can be mathematically written as[100]

$$S_{\text{DMAS}}(x, z) = \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} x_{ij}(t), \tag{30}$$

where S_{DMAS} is the reconstructed image, M is the total number of detectors, and $x_{ij}(t)$ is given by

$$x_{ij}(t) = \text{sign}[s_i(t_i)s_j(t_j)]\sqrt{|s_i(t_i)s_j(t_j)|}, \tag{31}$$

where sign($\cdot$) denotes the signum function. The workflow of the DMAS algorithm is illustrated in Fig. 9(b). If the center frequency of the photoacoustic signals $s_i(t)$ and $s_j(t)$ is f_0, the multiplication operation in Eq. (31) shifts the center frequency of the resultant signal to 0 and $2f_0$. As a result, the output signal S_{DMAS} needs to be filtered by a bandpass filter centered at $2f_0$ to extract the second harmonic component $2f_0$ while removing the direct current (DC) component. Therefore, DMAS is also called filtered-DMAS. Compared with DAS, DMAS uses both the amplitude and spatial correlation of delayed photoacoustic signals from different detectors for image reconstruction. For this reason, it can partially overcome the drawbacks of DAS and reconstruct images with improved spatial resolution and reduced side lobes[100].

A downside of DMAS is that the combinational multiplication in the algorithm increases the computational complexity. To solve this problem, in 2019, Jeon *et al.* proposed a DMAS algorithm with a modified CF, which avoids combinatorial multiplication in DMAS and significantly reduces the total number of multiplication operations[104]. In 2022, Paul *et al.* proposed a
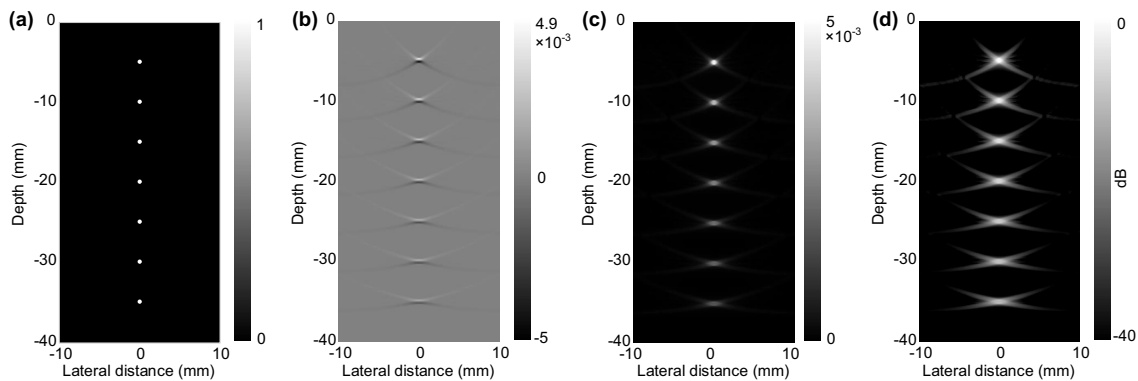

Fig. 11 An example showing DAS-based image reconstruction in PACT. (a) Ground truth. (b) Image reconstructed by DAS. (c) Envelope of (b). (d) Log transform of (c). The detector array is at the top of the image.

simplified-delay-multiply-and-standard-deviation (SDMASD) method[105], which is based on the measurement of the standard deviation of delayed and multiplied signals instead of normal delayed signals. Compared with native DAS and DMAS algorithms, the SDMASD algorithm can achieve real-time imaging using graphics processing units (GPUs) and produce improved image quality.

3.1.3 Short-lag spatial coherence

Short-lag spatial coherence (SLSC) is a beamforming technique that was initially developed by Lediju *et al.* for ultrasound imaging in 2011[106]. SLSC reconstructs an ultrasound image by calculating the spatial coherence of measured signals, and the reconstructed image is thus independent of the amplitude of the signals. As a result, SLSC can eliminate adverse effects caused by the different strengths of scatterers in ultrasound imaging. It has been demonstrated that compared with the conventional DAS beamforming algorithm, SLSC can achieve image reconstruction with considerable improvements in terms of resolution, contrast-to-noise ratio (CNR), and SNR.

In PACT, Muyinatu Bell *et al.* applied the SLSC algorithm to achieve imaging of prostate brachytherapy seeds[28,107] and demonstrated that the SLSC algorithm can enhance photoacoustic image quality compared with DAS, especially when the intensity of laser illumination is insufficient. Graham and Muyinatu Bell later developed a spatial coherence theory based on the van Cittert–Zernike theorem, a classical theorem in statistical optics, to explore the strengths and limitations of the SLSC algorithm[108,109]. In 2021, Mora *et al.* combined SLSC with DMAS and proposed a generalized spatial coherence algorithm for PACT, which can preserve relative signal magnitudes and improve the CNR and SNR of a reconstructed image[110].

Similar to DMAS, the photoacoustic signals in SLSC are also first delayed according to the TOF from the point of interest to each ultrasound detector and then combinatorically coupled and multiplied before summation. The SLSC algorithm can be formulated as[106]

$$\hat{R}_p(l) = \frac{1}{M-l} \sum_{i=1}^{M-l} \frac{\sum_{t=t_1}^{t_2} s_i(t) s_{i+l}(t)}{\sqrt{\sum_{t=t_1}^{t_2} s_i^2(t) \sum_{t=t_1}^{t_2} s_{i+l}^2(t)}}, \tag{32}$$

where $\hat{R}_p(l)$ is the normalized spatial coherence of the signals measured by a detector, M is the total number of detectors, and l represents the number of intervals between two detectors used to calculate the spatial coherence and is called the lag. The final SLSC image is obtained by summing all $\hat{R}_p(l)$ terms, i.e.,

$$S_{\text{SLSC}}(x, z) = \sum_{l=1}^{L} \hat{R}_p(l), \tag{33}$$

where L is the total number of lags. A large L helps improve the lateral resolution but decreases the CNR and SNR. Therefore, the value of L needs to be elaborately selected to achieve the best tradeoff between key image quality metrics, such as lateral spatial resolution, CNR, and SNR. The workflow of the SLSC algorithm is illustrated in Fig. 9(c).

3.1.4 Minimum variance

Minimum variance (MV) is a weighted DAS method that was originally devised by Capon in 1969 for narrowband signal processing applications such as sonar, radar, and wireless communication[111]. In 2002, Mann and Walker used a constrained adaptive beamformer (the MV method) for medical ultrasound imaging[112] and demonstrated its effectiveness in spatial resolution and contrast enhancement. Due to the remarkable improvement in spatial resolution, a large number of MV-based methods have been developed for ultrasound imaging[113–116].

MV-based image reconstruction has also been studied in PACT in recent years[30,117–120]. The MV reconstruction formula can be written as[114,116]

$$S_{\text{MV}}(x, z) = \sum_{i=1}^{M} w_i(t) s_i(t) = \mathbf{w}(t)^H \mathbf{s}(t), \tag{34}$$

where M is the total number of detectors, $w_i(t)$ is the optimal weight for the photoacoustic signal $s_i(t)$ measured by the ith detector, $\mathbf{s}(t)$ is a vector containing delayed photoacoustic signals from all detectors, $\mathbf{w}(t)$ is a vector containing optimal weights for the delayed photoacoustic signals in $\mathbf{s}(t)$, and the superscript H denotes the conjugate transpose. The workflow of the MV algorithm is illustrated in Fig. 9(d). The weight $\mathbf{w}(t)$ can be adaptively calculated by minimizing the variance of the output S_{MV} while maintaining the unit signal gain at the focal imaging point. The optimization problem can be mathematically written as[114]

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{R} \mathbf{w}, \quad \text{subject to } \mathbf{w}^H \mathbf{a} = 1, \tag{35}$$

where $\mathbf{R} = E[\mathbf{s}\mathbf{s}^H]$ is the spatial covariance matrix of the delayed photoacoustic signals $\mathbf{s}(t)$, E denotes the expectation, and $\mathbf{a}$ is the equivalent of the steering vector in narrowband applications[114,116]. When the photoacoustic signals of each detector are delayed, $\mathbf{a}$ is a simple all-one vector. The optimization problem in Eq. (35) can be solved by the method of Lagrange multipliers[121], which gives

$$\mathbf{w} = \frac{\mathbf{R}^{-1}\mathbf{a}}{\mathbf{a}^H \mathbf{R}^{-1}\mathbf{a}}. \tag{36}$$

By substituting Eq. (36) into Eq. (34), MV-based image reconstruction can be achieved.

To improve the robustness of the MV method, the covariance matrix $\mathbf{R}$ can be calculated based on a spatial smoothing strategy, where a detector array is divided into a group of overlapping subarrays, and the covariance matrices of all subarrays are calculated and averaged to form the final covariance matrix[114]. In this way, the covariance matrix $\mathbf{R}$ is given as[114]

$$\mathbf{R} = \frac{1}{N_{\text{d}} - L + 1} \sum_{l=1}^{N_{\text{d}}-L+1} \mathbf{s}_l(t) \mathbf{s}_l(t)^H, \tag{37}$$

where L is the number of detectors in a subarray and l is the index of the detector in the current subarray. To estimate the covariance matrix more accurately and enhance the contrast of MV, temporal averaging of multiple samples can be used together with spatial smoothing[116]. Once the covariance matrix $\mathbf{R}$ is estimated, the optimal weights $\mathbf{w}$ can be obtained by Eq. (36). The final image reconstructed by the MV algorithm can be written as[114]

$$S_{\text{MV}}(x, z) = \frac{1}{N_d - L + 1} \sum_{l=1}^{N_d - L + 1} \mathbf{w}(t)^H \mathbf{s}_l(t). \quad (38)$$

Although the MV method can produce narrower main lobes (i.e., the lobe located at the target) and higher spatial resolution than algorithms such as DAS and DMAS, its performance in terms of side lobe (i.e., the lobes adjacent to the main lobe) suppression and image contrast enhancement is limited. Therefore, many studies have been devoted to the development of MV-based hybrid beamforming algorithms. In 2008, Park *et al.* imposed additional CF weights on MV and achieved enhanced spatial resolution, contrast, and side lobe suppression[117]. In 2018, Mozaffarzadeh *et al.* developed an MV-based DMAS method[118] and an eigenspace-based MV method combined with DMAS for resolution improvement and side lobe reduction[119]. In 2021, Paul *et al.* proposed an adaptive-weighting-based MV to address the side lobe issue in MV beamformed images. It was demonstrated that the weighted MV approach can improve SNR while reducing main lobe width and side lobe intensity and has the potential for use in PACT imaging systems with a limited number of ultrasound detectors[30].

3.1.5 Coherence factor

The coherence factor (CF) is a pixel-level adaptive weighting factor that can improve the performance of DAS-based beamforming methods in side lobe suppression and spatial resolution improvement. CF was originally proposed by Mallart and Fink in 1994 for the evaluation of phase aberration correction techniques in scattering media[122] and was later used as an image quality metric for ultrasound imaging by Hollman *et al.*[123]. In 2003, Li *et al.* presented a generalized CF for ultrasound beamforming in heterogeneous media and showed that the combination of generalized CF and DAS-based beamformers could yield improved image quality[124].

In PACT, Liao *et al.* incorporated CF into DAS in 2004 and demonstrated the superiority of the CF-weighted DAS method in improving spatial resolution and SNR compared with DAS[29]. Wang and Li considered the local SNR in the formulation of CF and developed an SNR-dependent CF for ultrasound and photoacoustic imaging in 2014[125]. Wang *et al.* integrated CF with a focal-line-based image formation method to improve the contrast and elevational resolution of 3D photoacoustic imaging in 2016[126]. Mozaffarzadeh *et al.* proposed a high-resolution CF weighting technique and achieved improved resolution, SNR, and side lobe suppression in 2019[127]. Paul *et al.* considered the noise level variations of raw ultrasound data in the formulation of CF and achieved improvements in image resolution, side lobe reduction, SNR, and contrast in 2021[128]. In the same year, Mukaddim and Varghese extended CF from the spatial domain to the spatiotemporal domain[129]. This extension helps cancel out signals with low spatial and temporal coherence and results in higher background noise cancellation while preserving the main features of interest in reconstructed images.

Mathematically, the CF is defined as the ratio of the coherent sum of photoacoustic signals across detectors to the incoherent sum and can be formulated as[123]

$$\text{CF}(x, z) = \frac{\left| \sum_{i=1}^{M} s_i(t) \right|^2}{M \sum_{i=1}^{M} |s_i(t)|^2}, \quad (39)$$

where $s_i(t)$ is the photoacoustic signal measured by the ith detector at time t and M is the total number of detectors. The value of the CF ranges from zero to one. A large value means that the signals at that point are highly focused and that the point can be reconstructed with high quality. In contrast, a small CF value indicates that the signals are weakly focused and will result in lower image quality. The CF-weighted DAS method can be expressed as

$$S_{\text{CF-DAS}}(x, z) = \text{CF}(x, z) S_{\text{DAS}}(x, z), \quad (40)$$

where $S_{\text{CF-DAS}}(x, z)$ is the reconstructed image. The term $S_{\text{DAS}}(x, z)$ can be replaced by the outputs of other beamforming methods, such as DMAS, SLSC, and MV. The workflow of the CF algorithm is illustrated in Fig. 9(e).

Each beamforming method mentioned above has advantages and disadvantages. To improve the performance of image reconstruction in PACT, they can be combined to yield hybrid beamforming methods such as DMASD plus DAS/DMAS[105], SLSC plus filter DMAS[110], MV plus DMAS[118,119], CF plus DMAS[104,130], and CF plus MV[117,127,131].

Representative work on DAS-based image reconstruction in PACT is summarized in Table 3. For completeness, Table 3 also includes relevant work in ultrasound or microwave imaging.

3.2 Filtered Back Projection

DAS-type algorithms can achieve approximate photoacoustic image reconstruction and are inexact reconstruction techniques. To achieve exact image reconstruction, advanced algorithms are needed. Filtered back projection (FBP) is a class of algorithms that are based on the Radon transform (see Sec. 2.4). It achieves image reconstruction by first filtering measured photoacoustic signals and then back-projecting the filtered signals into the image domain. The back-projection operation in FBP is similar to the reconstruction procedure in DAS-type algorithms. Therefore, the native DAS algorithm can be regarded as a simplified version of FBP, which achieves image reconstruction by directly back-projecting original photoacoustic signals into the image domain.

3.2.1 Approximate filtered back projection

Early FBP algorithms were developed based on the condition of far-field approximation[35], which states that if the distance between a detector and a photoacoustic source is much greater than the size of the photoacoustic source itself (i.e., $|\mathbf{r}_s - \mathbf{r}_d| \gg d$, d: source size), the approximation $|\mathbf{r}_d - \mathbf{r}_s| \approx |\mathbf{r}_d - \mathbf{r}_s \cdot (\mathbf{r}_d / |\mathbf{r}_d|)|$ holds (see Fig. 12). Under the condition of far-field approximation, the photoacoustic integral over a spherical shell can be approximated by the integral over its tangential plane. Consequently, image reconstruction in PACT can be achieved by inverting the linear Radon transform. Kruger *et al.* proposed an approximate FBP algorithm in 1995[22], which is the first formal image reconstruction method in PACT. Xu and Wang developed approximate FBP algorithms from a more rigorous perspective in 2002[134,135] and 2003[136]. They deduced exact solutions to the image reconstruction problem and proposed approximate time-domain FBP algorithms for circular[134], spherical[135], planar, and cylindrical detection geometries[136].

The preceding approximate FBP algorithms were generally developed for specific detection geometries. In 2007, Burgholzer and Matt extended the approximate FBP algorithms

Table 3 Representative DAS-Type Algorithms Used for Image Reconstruction in PACT

Method	Author	Year	Variant	Source
DAS	Ma *et al.*	2020	Multiple DAS with enveloping	[132]
	Hoelen *et al.*	2000	Modified DAS	[97]
	Hoelen *et al.*	1998	Modified DAS	[25,26]
DMAS	Mulani *et al.*	2022	High-order DMAS	[103]
	Jeon *et al.*	2019	CF-weighted DMAS	[104]
	Mozaffarzadeh *et al.*	2018	Double-stage DMAS	[27]
	Kirchner *et al.*	2018	Signed DMAS	[102]
	Alshaya *et al.*	2016	Filter DMAS	[101]
	Lim *et al.*	2008	DMAS (microwave imaging)	[99]
SLSC	Graham *et al.*	2020	Photoacoustic spatial coherence theory for SLSC	[109]
	Bell *et al.*	2013	SLSC (for PACT)	[28]
	Lediju *et al.*	2011	SLSC (for ultrasound)	[106]
MV	Asl & Mahloojifar	2009	Modified MV (for ultrasound)	[116]
	Synnevag *et al.*	2007	MV (for ultrasound)	[114]
	Mann & Walker	2002	Constrained adaptive beamformer	[112]
CF	Mao *et al.*	2022	Spatial coherence + polarity coherence	[133]
	Mukaddim *et al.*	2021	Spatiotemporal CF	[129]
	Paul *et al.*	2021	Variational CF	[128]
	Wang *et al.*	2014	SNR-dependent CF	[125]
	Liao *et al.*	2004	CF-weighted DAS	[29]
	Li *et al.*	2003	Generalized CF	[124]
	Mallart & Fink	1994	CF	[122]
Hybrid	Paul *et al.*	2022	SDMASD + DAS/DMAS	[105]
	Mora *et al.*	2021	SLSC + Filter DMAS	[110]
	Mozaffarzadeh *et al.*	2019	CF + MV	[127,131]
	Mozaffarzadeh *et al.*	2018	CF + DMAS	[104,130]
	Mozaffarzadeh *et al.*	2018	MV + DMAS	[118,119]

for image reconstruction in arbitrarily closed detection geometries under the assumption of the far-field approximation. The extended FBP reconstruction formula can be written as[35]

$$p_0(\mathbf{r}_s) \cong \frac{1}{4\pi} \int_S b(\mathbf{r}_d, t) \delta\left(t - \frac{|\mathbf{r}_s - \mathbf{r}_d|}{v_0}\right) d\Omega, \tag{41}$$

where $p_0(\mathbf{r}_s)$ is the reconstructed photoacoustic image, δ is the Dirac delta function, $b(\mathbf{r}_d, t)$ is the back-projection term given by

$$b(\mathbf{r}_d, t) = -2t \frac{\partial p(\mathbf{r}_d, t)}{\partial t}, \tag{42}$$

and $d\Omega$ is the solid angle subtended by the element $d\sigma$ of a detection surface and can be calculated by

$$d\Omega = \frac{d\sigma}{|\mathbf{r}_s - \mathbf{r}_d|^2} \left(\mathbf{n}_d \cdot \frac{\mathbf{r}_s - \mathbf{r}_d}{|\mathbf{r}_s - \mathbf{r}_d|}\right). \tag{43}$$

Here $\mathbf{n}_d$ is the unit normal vector of the detector surface pointing to the region of interest (ROI).

The time-domain first derivative in Eq. (42) can be interpreted in the frequency domain as

$$\frac{\partial p(\mathbf{r}_d, t)}{\partial t} = \mathcal{F}^{-1}\{i\omega \mathcal{F} p(\mathbf{r}_d, t)\}, \tag{44}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the forward and inverse Fourier transforms, respectively; ω is a ramp filter, which suppresses the low-frequency contents of the measured photoacoustic signal $p(\mathbf{r}_d, t)$ and amplifies high-frequency information. Since the value of ω extends from $-\infty$ to ∞, the ramp filter is not integrable, and the inverse Fourier transform in Eq. (44) is undefined. To solve this problem, the ramp filter can be band-limited by a window function and Eq. (44) becomes
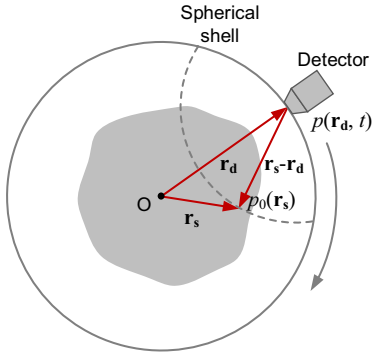
Fig. 12 Schematic diagram showing the signal detection and image reconstruction geometry in FBP. The forward problem and the image reconstruction problem in PACT correspond to the spherical Radon transform and its inverse, respectively. Under the condition of the far-field approximation, the integral over a spherical shell can be approximated by the integral over its tangential plane.

$$\frac{\partial p_0(\mathbf{r}_{\mathrm{d}}, t)}{\partial t} = \mathcal{F}^{-1}\{i\omega W(\omega)\mathcal{F} p_0(\mathbf{r}_{\mathrm{d}}, t)\}, \tag{45}$$

where $W(\omega)$ represents the window function. In practice, a smooth window such as the Hanning function is preferred over a box function because the latter may introduce undesirable ringing artifacts in the image domain.

One benefit of the far-field approximation in FBP is that it allows for simplified calculation of the solid angle $\mathrm{d}\Omega$. In other words, under the far-field approximation, $\mathrm{d}\Omega$ in Eq. (43) reduces to

$$\mathrm{d}\Omega \cong -\frac{\mathrm{d}\sigma}{|\mathbf{r}_{\mathrm{d}}|^2}\left(\mathbf{n}_{\mathrm{d}} \cdot \frac{\mathbf{r}_{\mathrm{d}}}{|\mathbf{r}_{\mathrm{d}}|}\right). \tag{46}$$

Compared with Eq. (43), Eq. (46) involves only the detector position $\mathbf{r}_{\mathrm{d}}$ when calculating $\mathrm{d}\Omega$ and is independent of the source position $\mathbf{r}_{\mathrm{s}}$, thereby reducing the computational complexity. For example, assuming that the size of a 3D image to be reconstructed is $N_x \times N_y \times N_z$ voxels ($N_x = N_y = N_z = n$) and the number of detectors used for imaging is M ($M = n \times n$), the computational complexity is $O(n^5)$ for Eq. (43) and only $O(n^2)$ for Eq. (46). Furthermore, if the detection geometry is spherical, $\mathrm{d}\Omega$ in Eq. (46) becomes a constant ($\mathrm{d}\Omega = \mathrm{d}\sigma/|\mathbf{r}_{\mathrm{d}}|^2$), indicating that the computational complexity is $O(1)$.

3.2.2 Exact filtered back projection

The approximate FBP algorithms are based on the condition of far-field approximation. This condition, however, may not be fully met in practice considering that photoacoustic signals may attenuate with propagation distance and that signal detection in space may be restricted. To solve this problem, in 2004, Finch *et al.* derived an exact FBP formula for image reconstruction in spherical detection geometries with odd dimensions ($n \geq 3$)[23]. In 2005, Xu and Wang presented a universal FBP formula for image reconstruction in infinite planar, infinite cylindrical, and closed spherical detection geometries[24]. The reconstruction formula is given as

$$p_0(\mathbf{r}_{\mathrm{s}}) = \int_\Omega b(\mathbf{r}_{\mathrm{d}}, t)\delta\left(t - \frac{|\mathbf{r}_{\mathrm{s}} - \mathbf{r}_{\mathrm{d}}|}{v_0}\right)\frac{\mathrm{d}\Omega}{\Omega}, \tag{47}$$

where the back-projection term

$$b(\mathbf{r}_{\mathrm{d}}, t) = 2\left[p(\mathbf{r}_{\mathrm{d}}, t) - t\frac{\partial p(\mathbf{r}_{\mathrm{d}}, t)}{\partial t}\right]. \tag{48}$$

Ω is the solid angle of the detection surface with respect to the reconstruction point and equals 2π for the infinite planar geometry and 4π for the spherical and cylindrical geometries. FBP is one of the most commonly used algorithms in PACT.

By comparing Eqs. (47) and (41), we find that the exact FBP algorithm is very close to the approximate FBP algorithm in the formulas. They both consist of three reconstruction steps: filtering, back projection (the δ function), and summation. Their major difference is in the back-projection terms. The back-projection term $b(\mathbf{r}_{\mathrm{d}}, t)$ in the exact FBP algorithm has an extra term $p(\mathbf{r}_{\mathrm{d}}, t)$, which is not present in the approximate FBP algorithm. The reasons are as follows. According to the forward solution of the photoacoustic wave equation [Eq. (18)], the amplitude of the photoacoustic signals received by a detector is proportional to the size of the source and inversely proportional to the detection distance, i.e., $p(\mathbf{r}_{\mathrm{d}}, t) \propto d/|\mathbf{r}_{\mathrm{d}} - \mathbf{r}_{\mathrm{s}}|$. Under the condition of the far-field approximation, i.e., $|\mathbf{r}_d - \mathbf{r}_{\mathrm{s}}| \gg d$, the amplitude of $p(\mathbf{r}_{\mathrm{d}}, t)$ is small enough compared with the derivative term $-t\partial p(\mathbf{r}_{\mathrm{d}}, t)/\partial t$ in Eq. (48) to be ignored without significant loss of reconstruction accuracy. Therefore, the back-projection term $b(\mathbf{r}_{\mathrm{d}}, t)$ in the approximate FBP algorithm does not involve the term $p(\mathbf{r}_{\mathrm{d}}, t)$.

Figure 13 presents an example demonstrating the principle of FBP-based image reconstruction. In this example, the photoacoustic source is a 5-mm-diameter uniform spherical absorber, and the photoacoustic signals generated from the source are recorded by a 40-mm-diameter circular detector array [Fig. 13(a)]. Figures 13(b) and 13(c) show the representative photoacoustic signal $p(\mathbf{r}_{\mathrm{d}}, t)$ measured by a detector and the corresponding back-projection signal $b(\mathbf{r}_{\mathrm{d}}, t)$, respectively. Figure 13(d) shows the projection images of detectors at different positions. Figures 13(e)–13(g) are the reconstruction results obtained by summing the projection images of four, 16, and 256 detectors, respectively. The results demonstrate that the FBP algorithm can effectively reconstruct the structure and amplitude information of the photoacoustic source. Note that the artifacts in the reconstructed images are not caused by the FBP algorithm but by the limited view angle of the circular detector array[137].

To evaluate the performance of the FBP algorithm for different detection geometries, a group of image reconstruction simulations was conducted. In the simulations, a multi-sphere phantom is used as the photoacoustic source. The phantom contains nine spherical absorbers with unit intensity, among which eight with diameters uniformly varying from 1 mm to 2 mm are evenly distributed on a circle with a diameter of 10 mm, while the ninth one, with a diameter of 1.6 mm, is seated at the origin. Figures 14(a)–14(c) show the relative positions of the multi-sphere phantom and a finite planar, finite cylindrical, and closed spherical detection surface. The planar detection surface is located 12 mm below the x–y plane, and the cylindrical and spherical detection surfaces are centered at the origin. The three detection surfaces have the same number (32768) of evenly distributed point-like detectors and approximately equal detection
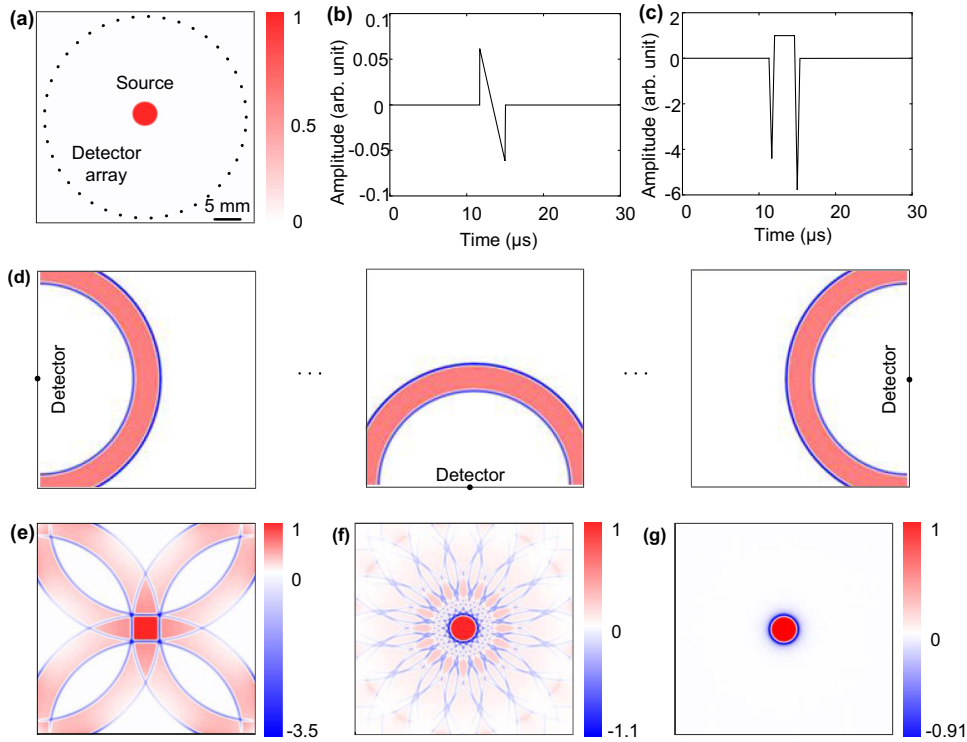
Fig. 13 Illustration of the principle of the FBP algorithm. (a) Schematic diagram showing a spherical photoacoustic source (diameter: 5 mm) and an array of point-like detectors uniformly distributed over a circle (diameter: 40 mm). (b) N-shaped photoacoustic signal recorded by a detector on the detection circle. (c) Back-projection signal [Eq. (48)]. (d) Projection images produced by the detectors at different positions. (e)–(g) Images reconstructed using 4, 16, and 256 detectors, respectively. Adapted from Ref. [137] with permission.

areas (plane: 150 mm × 150 mm; cylinder: 70 mm base diameter, 100 mm height; sphere: 84 mm diameter). The reconstructed images of the *x–z* and *x–y* cross sections of the phantom are displayed in Figs. 14(d)–14(f) and 14(g)–14(i), respectively. The results show that the FBP algorithm achieves stable image reconstruction for all three detection geometries. However, the image reconstructed in the closed spherical detection surface has much fewer artifacts than those reconstructed in the planar and cylindrical detection surfaces. This is because the spherical detection surface is closer to an ideal detection geometry than the finite planar and finite cylindrical surfaces.

To extend the application scenarios of FBP, many in-depth studies have also been carried out[138]. For example, in 2007, Kunyansky proposed a set of FBP-type formulas for image reconstruction in spherical detection geometries with arbitrary dimensions ($n \geq 2$)[139]. In the same year, Finch *et al.* also developed a set of FBP-type inverse formulas for spherical detection geometries with even dimensions[140]. In 2009, Nguyen derived a family of inverse formulas for thermoacoustic tomography, in which many previously known FBP inverse formulas can be obtained as special cases[141]. In addition, the exact FBP algorithms mentioned above are primarily used for planar, cylindrical, and spherical detection geometries with point-like detectors. In 2007, Burgholzer *et al.* developed a two-step FBP algorithm for integrating line detectors[142]. In 2012, Natterer proposed a novel FBP algorithm for elliptical detection geometries[143], which was further developed by Palamodov in 2012[144], Haltmeier in 2014[145,146], and Salman in 2014[147].

Moreover, the FBP algorithm is well suited for parallel computing. GPUs have been used to achieve real-time FBP image reconstruction in both 2D and 3D imaging[42,148–153]. A field-programmable gate array (FPGA) was also employed to accelerate image reconstruction in low-cost PACT systems[154].

Table 4 summarizes representative work on FBP-based image reconstruction in PACT.

3.3 Series Expansion

The basic principle of series expansion (SE) algorithms is to approximate the image to be reconstructed using mathematical series. Compared with other reconstruction algorithms in PACT, SE algorithms are simple and fast for specific detection geometries such as planar geometries because they can be implemented using FFT.

3.3.1 Series expansion for planar detection geometries

Norton *et al.* proposed exact SE methods for the reconstruction of acoustic reflectivity in circular geometries in 1980[156] and in planar, cylindrical, and spherical geometries in 1981[157]. Several groups have studied similar ideas for image reconstruction in PACT. Köstli *et al.* in 2001 proposed an exact SE algorithm for image reconstruction in planar detection geometries[31], and Xu *et al.* in 2002 reported a similar algorithm[158].

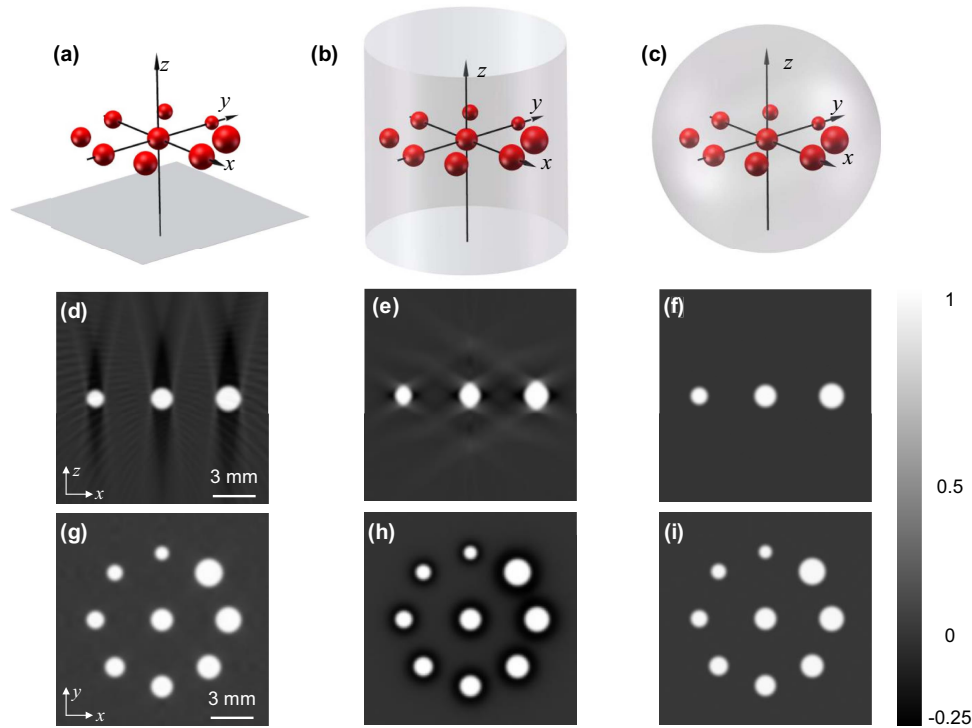Mathematically, SE-based PACT image reconstruction for planar detection geometries can be formulated as[31,83]

Fig. 14 Image reconstruction by FBP in three common detection geometries. (a)–(c) Schematic diagrams showing a multi-sphere phantom and planar, cylindrical, and spherical detection surfaces. The three detection surfaces have the same number of point-like detectors (32768) and approximately equal detection areas. Please refer to the text for more simulation settings. (d)–(f) Reconstructed images in the x–z plane. (g)–(i) Reconstructed images in the x–y plane. Adapted from Ref. [137] with permission.

Table 4 Representative Work on FBP-Based Image Reconstruction in PACT

Author	Year	Method	Detection Geometry	Dimension	Source
Haltmeier	2014	FBP	Elliptical	Arbitrary	[145,146]
Salman	2014	FBP	Elliptical	2D and 3D	[147]
Natterer	2012	FBP	Elliptical	3D	[143]
Palamodov	2012	FBP	Elliptical	Arbitrary	[144]
Nguyen	2009	FBP	Spherical	Arbitrary	[141]
Burgholzer & Matt	2007	Approximate FBP	Arbitrary	3D	[35]
Burgholzer *et al.*	2007	FBP	Arbitrary closed detection curve	2D or 3D	[142]
Kunyansky	2007	FBP	Spherical	Arbitrary	[139]
Finch *et al.*	2007	FBP	Spherical	Even	[140]
Xu & Wang	2005	FBP	Planar, cylindrical, and spherical	3D	[24]
Finch *et al.*	2004	FBP	Spherical	Odd	[23]
Xu & Wang	2003	Approximate FBP	Planar, cylindrical, and spherical	3D	[136]
Xu & Wang	2002	Approximate FBP	Circular	3D	[134]
Xu & Wang	2002	Approximate FBP	Spherical	3D	[135]
Xu *et al.*	2001	Approximate FBP	Circular	3D	[155]
Kruger *et al.*	1995	Approximate FBP	Circular	3D	[22]

$$P_0(k_x, k_y, \omega) = \frac{v_0^2 k_z}{2\omega} \mathcal{F}_{x,y,t}\{p(x, y, t)\}, \tag{49}$$

$$P_0(k_x, k_y, \omega) \xrightarrow{\omega^2 = v_0^2(k_x^2 + k_y^2 + k_z^2)} P_0(k_x, k_y, k_z), \tag{50}$$

$$p_0(x, y, z) = \mathcal{F}_{x,y,z}^{-1}\{P_0(k_x, k_y, k_z)\}, \tag{51}$$

where $p(x, y, t)$ is the photoacoustic signal measured by a planar surface at position (x, y) and time t; k_x, k_y, and k_z are the spatial wavenumber components in each Cartesian direction; ω is the temporal frequency; v_0 is the SOS; $\rightarrow$ represents the interpolation operation between the temporal and spatial frequencies ω and k_z; $p_0(x, y, z)$ and $P_0(k_x, k_y, k_z)$ are the initial photoacoustic pressure and its spatial Fourier transform; and $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent the forward and inverse Fourier transforms, respectively. The reconstruction procedure above involves one interpolation and two Fourier transforms and has a computational complexity of $O(n^3 \log n)$ for a 3D image with a size of $n \times n \times n$ voxels by use of FFT[159]. The numerical implementation of the SE algorithm is available in the k-Wave toolbox[83].

Figure 15 is a numerical example showing SE-based PACT image reconstruction using the k-Wave toolbox. The photoacoustic source in the xy plane contains nine spherical absorbers with unit intensity. Among them, eight absorbers with diameters uniformly varying from 1 mm to 2 mm are evenly distributed on a circle with a diameter of 10 mm, while the ninth absorber with a diameter of 1.6 mm is seated at the origin. The signal detection geometry is a planar surface 12 mm below the xy plane. The planar surface has 364×364 point detectors uniformly distributed spanning a physical size of 36 mm $\times$ 36 mm. Figures 15(b) and 15(c) are the xz and xy cross sections of the photoacoustic source, respectively, and Figs. 15(d) and 15(e) are the corresponding reconstructed images, which show that the SE algorithm can recover major structures of the source. Note that the distortion and blurring in the reconstructed images are due to the finite size of the planar detection surface.

3.3.2 Series expansions for circular, cylindrical, and spherical detection geometries

In 2002, the Wang group reported exact SE algorithms for PACT image reconstruction in cylindrical and spherical geometries[135,160]. However, the image reconstruction procedures in these two cases involve complicated mathematical calculations, preventing their implementations using FFT and are thus time-consuming[135,160]. Based on Norton's pioneering work on circular geometries[156], Haltmeier and Scherzer in 2007 proposed a 3D reconstruction algorithm for cylindrical detection geometries, where photoacoustic signals are recorded by linear integrating detectors[161]. The proposed algorithm has a computational complexity of $O(n^4)$ for a 3D image with a size of $n \times n \times n$ voxels, which is higher than that of the SE algorithm in planar geometries $[O(n^3 \log n)]$[159] but lower than that of FBP $[O(n^5)]$[24].

In 2012, Kunyansky proposed fast image reconstruction algorithms suitable for circular, cylindrical, and spherical detection geometries[33]. In Kunyansky's work, image reconstruction is based on the Fourier transforms of the initial photoacoustic pressure $p_0(\mathbf{r}_s)$ and the measured photoacoustic signals $p(\mathbf{r}_d, t)$, as summarized below. Suppose that the detection geometry is a circle, as shown in Fig. 7(c). By expanding the Fourier transform $P(\mathbf{r}_d, \omega)$ of the measured photoacoustic signal $p(\mathbf{r}_d, t)$ and the initial photoacoustic pressure $p_0(\mathbf{r}_s)$ in the Fourier series in φ and θ, we have[33]

$$P(\mathbf{r}_d, \omega) = \sum_{k=-\infty}^{\infty} P^k(\omega) e^{ik\varphi}, \qquad \mathbf{r}_d = (r_d \cos\varphi, r_d \sin\varphi), \tag{52}$$

$$p_0(\mathbf{r}_s) = \sum_{k=-\infty}^{\infty} p_0^k(r_s) e^{ik\theta}, \qquad \mathbf{r}_s = (r_s \cos\theta, r_s \sin\theta), \tag{53}$$

where i is the imaginary unit, ω is the angular frequency, $r_d = |\mathbf{r}_d|$ is the radius of the detection circle, $r_s = |\mathbf{r}_s|$ is the distance from the reconstruction point to the center of the circular
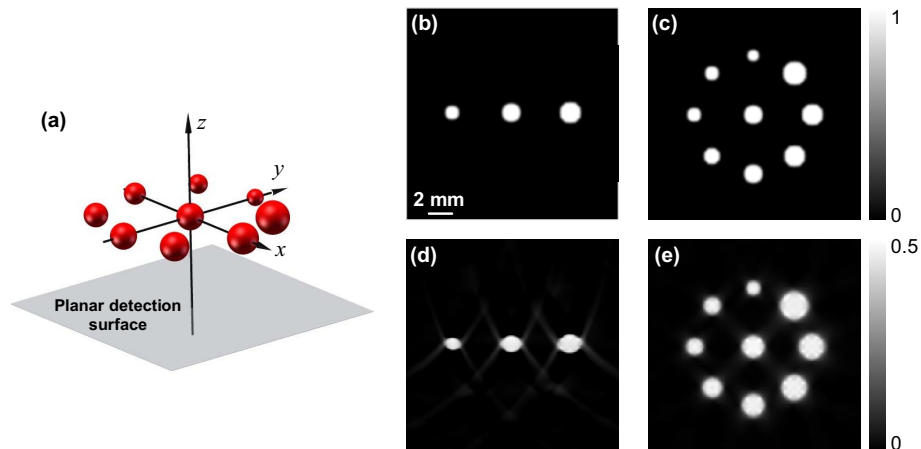

Fig. 15 Example of SE-based image reconstruction in PACT. (a) Schematic diagram of a planar detection geometry and a multi-sphere photoacoustic source. (b), (c) xz and xy cross sections of the source. (d), (e) xz and xy cross sections of the reconstructed source. Please refer to the text for the simulation settings.

detection geometry, and $P^k(\omega)$ and $p_0^k(r_s)$ are expansion coefficients given by[33]

$$P^k(\omega) = \frac{1}{2\pi} \int_0^{2\pi} P(\mathbf{r}_d, \omega) e^{-ik\varphi} d\varphi, \tag{54}$$

$$p_0^k(r_s) = \frac{2}{\pi} \int_0^\infty \frac{P^k(\omega)}{H_{|k|}^{(1)}(\omega R)} J_{|k|}(\omega r_s) d\omega, \tag{55}$$

where $H_{|k|}^{(1)}$ is the Hankel function of the first kind of order k and $J_{|k|}$ is the Bessel function of the first kind of order $|k|$[162]. Equation (55) is somewhat similar to an expression in Norton's work[156], where the Bessel function $J_{|k|}$ rather than the Hankel function $H_{|k|}^{(1)}$ is in the denominator and a term corresponding to the real part of $P^k(\omega)$ in the numerator.

By discretizing Eq. (55), we can obtain $p_0^k(r_s)$ and then the initial photoacoustic pressure $p_0(\mathbf{r}_s)$ [Eq. (53)]. However, direct discretization of Eq. (55) results in a computational complexity of $O(n^2)$ for each $p_0^k(r_s)$ and $O(n^3)$ for the whole reconstruction. To achieve fast reconstruction, substituting Eq. (55) into Eq. (53) yields[33]

$$p_0(\mathbf{r}_s) = \frac{1}{2\pi} \int_{\mathbb{R}^2} P_0(\Lambda) e^{ir_s \cdot \Lambda} d\Lambda, \qquad \Lambda = (\omega \cos \varphi, \omega \sin \varphi), \tag{56}$$

where $P_0(\Lambda)$ is given by

$$P_0(\Lambda) = \begin{cases} \frac{2}{\pi} \sum_{k=-\infty}^{\infty} \frac{(-i)^k p_0^k(\omega)}{\omega H_{|k|}^{(1)}(\omega R)} e^{ik\varphi}, & \Lambda \neq 0, \\ \int_0^\infty \frac{2P^{k=0}(\omega)}{\pi \omega H_0^{(1)}(\omega R)} RJ_1(\omega R) d\omega, & \Lambda = 0, \end{cases} \tag{57}$$

where R is the radius of the circular detection geometry. Equation (56) indicates that the initial photoacoustic pressure $p_0(\mathbf{r}_s)$ can be obtained by calculating the 2D inverse Fourier transform of $P_0(\Lambda)$, which has a computational complexity of $O(n^2 \log n)$, lower than that of the direct-discretization-based reconstruction [$O(n^3)$].

For cylindrical detection geometries with linear integrating detectors, image reconstruction can be readily realized by combining the fast image reconstruction procedure for 2D circular geometries and the 3D inverse Fourier transform[33]. This procedure can yield fast reconstruction with a computational complexity of $O(n^3 \log n)$. For spherical detection geometries, the derivation of the image reconstruction procedure is similar to those for circular geometries [Eqs. (52)–(57)]. A prominent difference is that $p_0(\mathbf{r}_s)$ and $P(\mathbf{r}_d, \omega)$ are expanded into spherical harmonics. In addition, the fast spherical harmonics (FSH) transform is adopted to achieve fast reconstruction[163], which results in a reconstruction complexity of $O(n^3 \log^2 n)$[33].

Wang and Anastasio in 2012 demonstrated that a mapping relationship exists between the spatial frequency components of initial photoacoustic pressure $p_0(\mathbf{r}_s)$ and the temporal frequency components of measured photoacoustic signals $p(\mathbf{r}_d, t)$[164]. They thus proposed a Fourier-transform-based image reconstruction algorithm whose computational complexity is $O(n^2 \log n)$ for 2D image reconstruction in circular geometries and $O(n^3 \log n)$ for 3D image reconstruction in spherical geometries[165]. The reconstruction formula does not involve

series expansion of special functions or multi-dimensional interpolation operations in Fourier space, which are commonly used in previous work.

3.3.3 Series expansion for general detection geometries

Kunyansky in 2007 presented a generalized SE method for image reconstruction in arbitrarily closed detection geometry provided that the eigenfunctions of the Dirichlet Laplacian are explicitly known[32].

Assuming that λ_k^2 and $u_k(\mathbf{r})$ are the eigenvalues and normalized eigenfunctions of the Dirichlet Laplacian in the interior Ω of a closed detection surface S, we have[32]

$$\nabla^2 u_k(\mathbf{r}) + \lambda_k^2 u_k(\mathbf{r}) = 0, \qquad \mathbf{r} \in \Omega, \qquad \Omega \subseteq \mathbb{R}^n$$

$$u_k(\mathbf{r}) = 0, \qquad \mathbf{r} \in S,$$

$$\|u_k\|_2^2 \equiv \int_\Omega |u_k(\mathbf{r})|^2 d\mathbf{r} = 1, \tag{58}$$

where n denotes the spatial dimension. Since the eigenfunctions $u_k(\mathbf{r})$ are orthogonal, the initial photoacoustic pressure $p_0(\mathbf{r})$ can be formulated in series form as

$$p_0(\mathbf{r}) = \sum_{k=0}^\infty \alpha_k u_k(\mathbf{r}), \tag{59}$$

where the coefficients α_k can be obtained from the measured projection data $p(\mathbf{r}_d, t)$, and $u_k(\mathbf{r})$ is the known eigenfunctions of the Dirichlet Laplacian determined by detection geometries. $u_k(\mathbf{r})$ is the solution of the Dirichlet problem for the Helmholtz equation with zero boundary conditions and has the following Helmholtz representation[32]:

$$u_k(\mathbf{r}) = \int_S \Phi_{\lambda_k}(|\mathbf{r} - \mathbf{r}_d|) \frac{\partial}{\partial \mathbf{n}_d} u_k(\mathbf{r}_d) dS, \qquad \mathbf{r} \in \Omega, \tag{60}$$

where Φ_{λ_k} is a free-space rotationally invariant Green's function[166], $\mathbf{n}_d$ is the unit normal vector of the detector surface pointing to the ROI, and $\mathbf{r}_d$ is the position of the detector. The coefficients α_k can be calculated by[32]

$$\alpha_k(\mathbf{r}) = \int_\Omega u_k(\mathbf{r}) p_0(\mathbf{r}) d\mathbf{r} = \int_S I(\mathbf{r}_d, \lambda_k) \frac{\partial}{\partial \mathbf{n}_d} u_k(\mathbf{r}_d) dS, \tag{61}$$

where $I(\mathbf{r}_d, \lambda_k)$ is given by

$$I(\mathbf{r}_d, \lambda_k) = \int_0^{\mathrm{diam}(\Omega)/v_0} g(\mathbf{r}_d, t) \Phi_{\lambda_k}(t) v_0 dt. \tag{62}$$

Here $\mathrm{diam}(\Omega)$ denotes the diameter of Ω, and $g(\mathbf{r}_d, t)$ is the integral over a spherical shell [Eq. (27)] centered at the detector position $\mathbf{r}_d$, which can be calculated from the measured projection data $p(\mathbf{r}_d, t)$. With the calculated coefficients α_k and eigenfunctions $u_k(\mathbf{r})$, the initial photoacoustic pressure $p_0(\mathbf{r})$ can be reconstructed [Eq. (59)].

If the eigenfunctions of the Dirichlet Laplacian of a detection geometry such as a sphere, a half-sphere, a finite cylinder, or a cube are explicitly known, the eigenfunction-based SE method can provide exact reconstruction for any closed detection geometry. In particular, when the detection geometry is a cube, the reconstruction procedure can be implemented using 3D FFT,

resulting in efficient computation with a complexity of $O(n^3 \log n)$[32]. Furthermore, unlike the FBP algorithm, this technique can provide theoretically exact reconstruction within the region enclosed by the detection geometry even if part of a photoacoustic source is outside the region.

The image reconstruction procedures discussed for planar, circular, cylindrical, spherical, and arbitrarily closed geometries are based on the assumption that the acoustic medium is homogeneous[32]. For the image reconstruction problem in heterogeneous media, Agranovsky and Kuchment proposed an analytic reconstruction formula that works for any closed detection geometry and variable SOS satisfying a non-trapping condition[167]. This algorithm can be regarded as a generalization of the eigenfunction-based SE method for arbitrary closed detection geometries [Eqs. (58)–(62)][167].

Table 5 lists representative work on SE-based fast image reconstruction algorithms and Table 6 summarizes representative work on SE-based image reconstruction in PACT.

3.4 Time Reversal

Time reversal (TR) is a type of algorithm that involves recovering initial photoacoustic pressure via numerically running a forward acoustic propagation model backward and re-transmitting the photoacoustic signals measured by each detector into the image domain in a temporally reversed order. TR algorithms can couple the acoustic properties of media such as SOS, density, dispersion, and absorption into the acoustic propagation model and can be used for image reconstruction in arbitrary closed detection geometry[170]. They are thus regarded as the 'least restrictive' image reconstruction algorithms in PACT[170,171].

3.4.1 Standard time reversal

In 2004, Xu and Wang proposed a TR algorithm in PACT and derived an analytical expression for image reconstruction based on the Green's function with a homogeneous Dirichlet boundary condition[34]. Under the condition of the far-field approximation [see Sec. 3.2.1], the derived analytical expression in[34] is reduced to the reconstruction formula in FBP [Eq. (47)]. Subsequent research on TR-based image reconstruction primarily focused on coupling the acoustic properties of heterogeneous media into the TR model to improve reconstruction performance[35,170–172]. For example, Hristova *et al.* studied TR-based image reconstruction in heterogeneous media, in even dimensions, and with partially closed detection surfaces[171]. Treeby *et al.* considered the effect of acoustic absorption of

Table 5 Representative Work on SE-Based Fast Image Reconstruction in PACT

Author	Year	Detection Geometry	Dimension	Complexity[a]	Source
Kunyansky	2012	Circular	2D	$O(n^2 \log n)$	[33]
		Spherical	3D	$O(n^3 \log^2 n)$	
		Cylindrical	3D	$O(n^3 \log n)$	
Wang *et al.*	2012	Circular	2D	$O(n^2 \log n)$	[165]
		Spherical	3D	$O(n^3 \log n)$	
Kunyansky	2007	Cubic	3D	$O(n^3 \log n)$	[32]
Xu *et al.*	2002	Planar	3D	$O(n^3 \log n)$	[160]
Köstli *et al.*	2001	Planar	3D	$O(n^3 \log n)$	[31]

[a]2D image size: $n \times n$; 3D image size: $n \times n \times n$.

Table 6 Representative Work on SE-Based Image Reconstruction in PACT

Author	Year	Detection Geometry	Detector	SOS	Source
Kunyansky	2012	Circular, spherical, and cylindrical	Point-like detectors for circles and spheres; linear integrating detectors for cylinder	Constant	[33]
Wang *et al.*	2012	Circular and spherical	Point-like detectors	Constant	[165]
Zangerl *et al.*	2009	Cylindrical	Circular integrating detectors	Constant	[168,169]
Haltmeier *et al.*	2007	Cylindrical	Linear integrating detectors	Constant	[161]
Kunyansky	2007	Arbitrary closed geometry	Point-like detectors	Constant	[32]
Agranovsky & Kuchment	2007	Arbitrary closed geometry	Point-like detectors	Constant or variable	[167]
Xu & Wang	2002	Spherical	Point-like detectors	Constant	[135]
Xu *et al.*	2002	Planar	Point-like detectors	Constant	[158]
Xu *et al.*	2002	Cylindrical	Point-like detectors	Constant	[160]
Köstli *et al.*	2001	Planar	Point-like detectors	Constant	[31]

media and proposed a TR algorithm for absorbing acoustic media[170].

The TR algorithms are based on the Huygens principle[171]. When the SOS of a medium is constant and the spatial dimension of an acoustic field is odd, the Huygens principle indicates that the acoustic wave field vanishes after a moment T[171] and TR algorithms enable exact image reconstruction for a sufficiently large T. When the SOS is variable or the spatial dimension is even, the Huygens principle no longer holds, and TR algorithms can only provide approximate reconstructions[171].

In a lossy medium, the photoacoustic wave equation can be formulated as an initial value problem by three coupled acoustic equations and corresponding initial conditions[170]. The coupled acoustic equations, including the linearized equation of motion, the linearized equation of continuity, and the adiabatic equation of state, are given as[170]

$$\frac{\partial}{\partial t}\mathbf{u}(\mathbf{r},t) = -\frac{1}{\rho_0(\mathbf{r})}\nabla p(\mathbf{r},t), \tag{63}$$

$$\frac{\partial}{\partial t}\rho(\mathbf{r},t) = -\rho_0(\mathbf{r})\nabla \cdot \mathbf{u}(\mathbf{r},t), \tag{64}$$

$$p(\mathbf{r},t) = v_0(\mathbf{r})^2\left\{1 + \tau(\mathbf{r})\frac{\partial}{\partial t}(-\nabla^2)^{y/2-1} + \eta(\mathbf{r})(-\nabla^2)^{(y+1)/2-1}\right\}\rho(\mathbf{r},t), \tag{65}$$

which are subject to the following initial conditions:

$$p(\mathbf{r},t)|_{t=0} = p_0(\mathbf{r}), \qquad \mathbf{u}(\mathbf{r},t)|_{t=0} = 0. \tag{66}$$

Here $p(\mathbf{r},t)$ is the acoustic pressure at time t and position $\mathbf{r}$ inside the imaging region, $\mathbf{u}(\mathbf{r},t)$ is the acoustic particle velocity, $\rho(\mathbf{r},t)$ is the acoustic density, $\rho_0(\mathbf{r})$ is the ambient density, and $\tau(\mathbf{r})$ and $\eta(\mathbf{r})$ are the absorption and dispersion proportionality coefficients given by

$$\tau(\mathbf{r}) = -2\alpha_0 v_0(\mathbf{r})^{y-1}, \qquad \eta(\mathbf{r}) = 2\alpha_0 v_0(\mathbf{r})^y \tan(\pi y/2), \tag{67}$$

where α_0 is the absorption coefficient (dB MHz^{-y} cm^{-1}) and y is a real constant representing the power law exponent.

TR-based image reconstruction is achieved by running the same photoacoustic forward model in temporally reversed order and thus can be formulated by the same set of equations [Eqs. (63)–(67)] except that the initial conditions in Eq. (66) are replaced with

$$p(\mathbf{r},t)|_{t=0} = 0, \qquad \mathbf{u}(\mathbf{r},t)|_{t=0} = 0,$$
$$p(\mathbf{r}_d,t) = p(\mathbf{r}_d, T-t). \tag{68}$$

Figure 16 illustrates the basic principle of TR-based image reconstruction.

Generally, there is no explicit analytical solution to the partial differential equations in Eqs. (63)–(65). Numerical methods are required to solve these equations. Finite element and finite difference are two commonly used methods for computing numerical solutions to partial differential equations due to their flexibility, accuracy, and rigorous mathematical foundations.
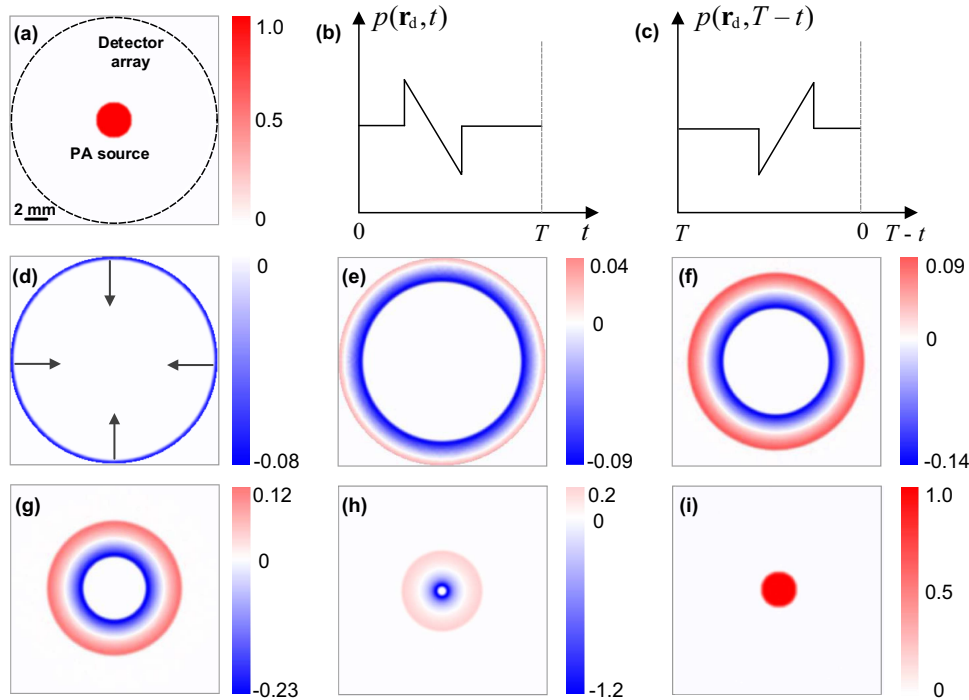

Fig. 16 Illustration of TR-based image reconstruction. (a) Cross section of a spherical absorber and a spherical detector array. (b) Photoacoustic signal measured by a detector. (c) Temporal reversion of the measured signal in (b). (d)–(i) Acoustic wave fields in the detection region at different moments during backward propagation of the time-reversed signal.

However, these approaches require many mesh points per wavelength and small time steps to achieve accurate calculations, which is computationally expensive, especially for high-frequency and large-scale imaging models.

To solve this problem, Cox *et al.* developed a k-space pseudospectral approach that can achieve accurate computations using fewer mesh points and more time steps[173]. The k-space-pseudospectral-method-based TR algorithm has been implemented in the k-Wave toolbox[83] and is based on the discrete form of the coupled acoustic equations [Eqs. (63)–(65)][170], which can be written as

$$\frac{\partial}{\partial \xi} p(\mathbf{r}, t) = \mathcal{F}^{-1}\{ik_\xi \kappa_k \mathcal{F}\{p(\mathbf{r}, t)\}\}, \tag{69}$$

$$\mathbf{u}_\xi(\mathbf{r}, t + \Delta t) = \mathbf{u}_\xi(\mathbf{r}, t) - \frac{\Delta t}{\rho_0(\mathbf{r})} \frac{\partial}{\partial \xi} p(\mathbf{r}, t), \tag{70}$$

$$\frac{\partial}{\partial \xi} \mathbf{u}_\xi(\mathbf{r}, t + \Delta t) = \mathcal{F}^{-1}\{ik_\xi \kappa_k \mathcal{F}\{\mathbf{u}_\xi(\mathbf{r}, t + \Delta t)\}\}, \tag{71}$$

$$\rho_\xi(\mathbf{r}, t + \Delta t) = \rho_\xi(\mathbf{r}, t) - \Delta t \rho_0(\mathbf{r}) \frac{\partial}{\partial \xi} \mathbf{u}_\xi(\mathbf{r}, t + \Delta t), \tag{72}$$

$$p(\mathbf{r}, t + \Delta t) = v_0(\mathbf{r})^2 \left[\sum_\xi \rho_\xi(\mathbf{r}, t + \Delta t) + \text{abs} + \text{disp} \right], \tag{73}$$

where i is the imaginary unit, ξ denotes the Cartesian coordinates [$\xi = x$ represents 1D space, $\xi = (x, y)$ represents 2D space, $\xi = (x, y, z)$ represents 3D space], k_ξ is the spatial wavenumber component, Δt is the time step, $\kappa_k = \text{sinc}[v_0(\mathbf{r})k\Delta t/2]$ is a k-space adjustment to the spatial derivative[170], and $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent the forward and inverse Fourier transforms, respectively. Equations (69) and (71) are equations for gradient calculation and Eqs. (70) and (72) are update equations. Equation (73) is an equation of state, including the absorption and dispersion of a medium, which is given by

$$\text{abs} = \tau(\mathbf{r}) \mathcal{F}^{-1}\left\{ (k_\xi \kappa_k)^{y-2} \mathcal{F}\left\{ \rho_0(\mathbf{r}) \sum_\xi \frac{\partial}{\partial \xi} \mathbf{u}_\xi(\mathbf{r}, t + \Delta t) \right\} \right\}, \tag{74}$$

$$\text{disp} = \eta(\mathbf{r}) \mathcal{F}^{-1}\left\{ (k_\xi \kappa_k)^{y-1} \mathcal{F}\left\{ \sum_\xi \rho_\xi(\mathbf{r}, t + \Delta t) \right\} \right\}. \tag{75}$$

The above equations are calculated iteratively and the entire acoustic field is updated for each time step.

One of the most prominent features of the TR algorithm is that it is well-suited for image reconstruction in acoustically heterogeneous media. To demonstrate this, a simulation is presented. Figure 17(a) shows an acoustically heterogeneous phantom consisting of a uniform background, several blood vessels, and a bone mimicking the cross section of a human finger. The background and blood vessels have a mass density of 1000 kg/m^3 and an SOS of 1500 m/s. In contrast, the acoustically heterogeneous region (i.e., the bone) has a density of 1850 kg/m^3 and an SOS of 3800 m/s. A 512-element full-ring detector array with a diameter of 50 mm enclosing the phantom is used for imaging. Figure 17(b) shows the image reconstructed by the TR algorithm under the assumption that the media are homogeneous. A typical SOS of 1505 m/s and a density of 1050 kg/m^3 are used. In this case, extensive image artifacts are present in the reconstructed image due to the inaccurate setting of the SOS and density for the TR model. As a comparison, Fig. 17(c) presents the image reconstructed by the TR algorithm coupling the true SOS and density of the heterogeneous media. The results show that by incorporating the acoustic properties of media, the TR algorithms can achieve improved image reconstruction in PACT.

Another important feature of TR is that it can address the image reconstruction problem in arbitrary closed detection geometry. Figures 18(a)–18(c) show an imaging example, where a 2D multi-disk phantom is enclosed by three different closed detection geometries, including a square, an octagon, and a circle. The octagon and the circle in Figs. 18(b) and 18(c) can be circumscribed by the square in Fig. 18(a). Figures 18(d)–18(f) show the corresponding images reconstructed by the TR algorithm, which shows that TR can produce good results for all three detection geometries.
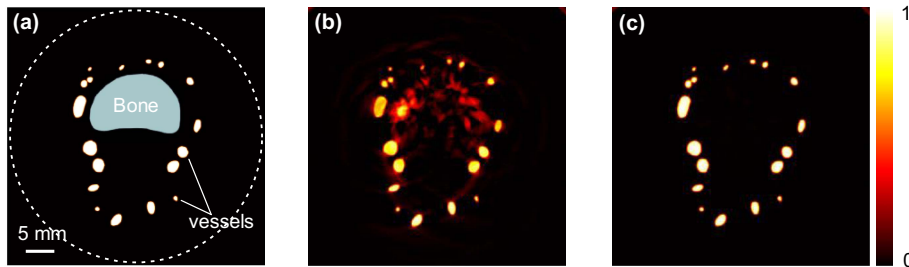

Fig. 17 TR-based image reconstruction in acoustically heterogeneous media. (a) A numerical phantom consisting of multiple blood vessels and a bone mimicking the cross section of a human finger. A 512-element full-ring detector array (dashed circle) with a diameter of 50 mm enclosing the phantom is used for imaging. (b) Image reconstructed by TR using a constant SOS (1505 m/s) and a constant density (1050 kg/m^3). (c) Image reconstructed by TR coupling the true SOS and density of the media. Please refer to the text for the simulation settings.
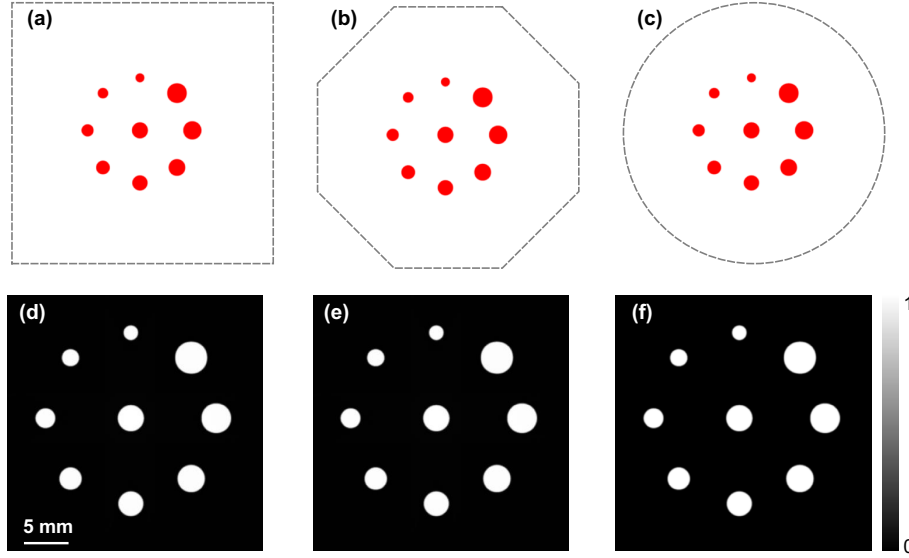
Fig. 18 TR-based image reconstruction under different detection geometries. (a)–(c) Schematic diagrams showing a phantom and three different detection geometries. The square detection geometry has a side length of 50 mm, the octagonal geometry has a side length of 20.7 mm, and the circular geometry has a diameter of 50 mm. (d)–(f) Corresponding images reconstructed by TR.

3.4.2 Modified time reversal based on the Neumann series

The standard TR algorithms discussed above are unable to find analytical solutions for the wave equations when the SOS is variable or the spatial dimension of an acoustic field is even. To address this problem, a Neumann-series-based TR algorithm was proposed by Stefanov and Uhlmann[174] and further developed by Qian *et al.*[175]. The basic idea of the Neumann series-based TR algorithm is to first modify the measurement data at $t = T$ using a harmonic extension operator, perform TR reconstruction, and then repeat the TR process[175]. The Neumann series-based TR algorithm can yield exact reconstruction if the detected pressure $p(\mathbf{r}_\mathrm{d}, t)$ is known on the whole boundary Ω_0 and the measurement time T is greater than a stability threshold[175].

In the Neumann-series-based TR algorithm, the initial conditions [Eq. (68)] for the coupled wave equations are modified to[63,175]

$$p(\mathbf{r}, t)|_{t=0} = P_\Omega p(\mathbf{r}, T), \qquad \mathbf{u}(\mathbf{r}, t)|_{t=0} = 0,$$
$$p(\mathbf{r}_\mathrm{d}, t) = p(\mathbf{r}_\mathrm{d}, T - t), \tag{76}$$

where P_Ω denotes the Poisson operator of the harmonic extension defined as $P_\Omega h(\mathbf{r}, T) = \phi(\mathbf{r})$, and $\phi(\mathbf{r})$ is the solution of the elliptic boundary value problem

$$\nabla^2 \phi(\mathbf{r}) = 0, \qquad \phi(\mathbf{r})|_{\mathbf{r} \in \Omega_0} = h(\mathbf{r}, T)|_{\mathbf{r} \in \Omega_0}, \tag{77}$$

where Ω is the domain defined by a detection surface and Ω_0 is the boundary of the domain.

By introducing a forward acoustic propagation operator $\mathcal{A}$ and a modified TR reconstruction operator $\mathcal{A}_\mathrm{modify}^\mathrm{TR}$, the photoacoustic forward problem and the modified TR reconstruction can be expressed as

$$p(\mathbf{r}_\mathrm{d}, t) = \mathcal{A} p_0(\mathbf{r}), \tag{78}$$

$$p_0(\mathbf{r}) = \mathcal{A}_\mathrm{modify}^\mathrm{TR} p(\mathbf{r}_\mathrm{d}, t). \tag{79}$$

The Neumann-series-based TR reconstruction can be expressed as[175]

$$p_0(\mathbf{r}) = \sum_{j=0}^{\infty} (I - \mathcal{A}_\mathrm{modify}^\mathrm{TR} \mathcal{A})^j \mathcal{A}_\mathrm{modify}^\mathrm{TR} p(\mathbf{r}_\mathrm{d}, t), \tag{80}$$

where I is an identity operator. Theoretically, Eq. (80) can provide an exact reconstruction if the variable j varies from zero to $+\infty$. However, in practice, j is always finite, and exact reconstruction is impossible. If j is cut off at a particular value m, the reconstruction formula in Eq. (80) can be written as

$$p_0^m(\mathbf{r}) = \sum_{j=0}^{m} (I - \mathcal{A}_\mathrm{modify}^\mathrm{TR} \mathcal{A})^j \mathcal{A}_\mathrm{modify}^\mathrm{TR} p(\mathbf{r}_\mathrm{d}, t). \tag{81}$$

Equation (81) can be further formulated in the form of iterative TR reconstruction[176], i.e.,

$$p_0^{m+1}(\mathbf{r}) = p_0^m(\mathbf{r}) + \mathcal{A}_\mathrm{modify}^\mathrm{TR}[p(\mathbf{r}_\mathrm{d}, t) - \mathcal{A} p_0^m(\mathbf{r})]. \tag{82}$$

The estimated initial photoacoustic pressure $p_0^m(\mathbf{r})$ usually contains pixels with negative intensity due to non-ideal imaging conditions, such as finite detector bandwidth and limited view angle. To improve reconstruction results, non-negative constraints can be imposed on the solutions during iteration. Moreover, the Poisson operator of harmonic extension can be neglected in some cases, e.g., when the spatial dimension of an acoustic field is even and the medium is homogeneous or when the time T is large enough to guarantee that acoustic waves inside a detection region decay sufficiently. Figure 19 shows an
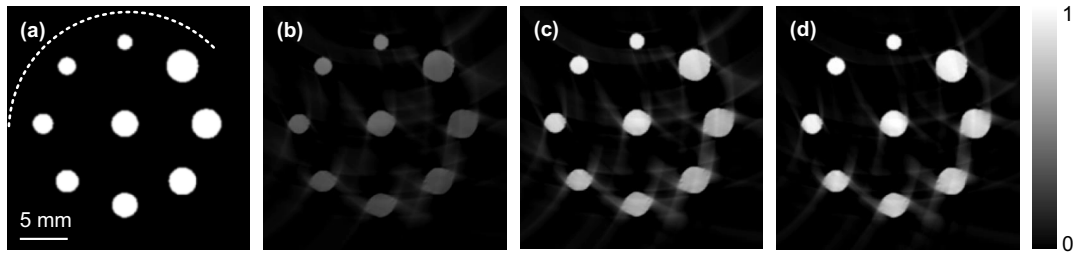
Fig. 19 Iterative TR-based image reconstruction for limited-view imaging. (a) Schematic diagram of a phantom and a limited-view detector array. The detector array has 455 detectors uniformly distributed over a 50-mm-diameter partial circle with a view angle of $3/4\pi$. (b)–(d) Images reconstructed by the Neumann-series-based TR algorithm after 3, 5, and 10 iterations.

example illustrating Neumann-series-based TR reconstruction in limited-view imaging.

Table 7 summarizes representative work on TR-based image reconstruction in PACT.

3.5 Iterative Reconstruction

The last class of conventional image reconstruction algorithms in PACT is iterative reconstruction (IR), which achieves image reconstruction through iteration. IR algorithms construct a set of linear equations based on a discrete photoacoustic imaging model and iteratively seek numerical solutions by minimizing the error between measured projection data and their estimates (see Fig. 20). Compared with the aforementioned DAS, FBP, SE, and TR algorithms, IR algorithms can yield improved photoacoustic image quality under non-ideal imaging scenarios, such as spare-view and limited-view imaging. They can also incorporate the physical models of an imaging system, such as transducer responses and acoustic heterogeneity, to achieve improved imaging. The downside of IR algorithms is that they are slow and computationally expensive due to the numerical optimization involved.

Table 7 Representative Work on TR-Based Image Reconstruction in PACT

Author	Year	Solution	Media	Source
Qian *et al.*	2011	Numerical	Heterogeneous (exact solution)	[175]
Treeby *et al.*	2010	Numerical	Heterogeneous, absorptive, and dispersive	[170]
Stefanov & Uhlmann	2009	Numerical	Heterogeneous (exact solution)	[174]
Hristova *et al.*	2008	Numerical	Heterogeneous	[171]
Burgholzer *et al.*	2007	Numerical	Heterogeneous	[35]
Xu & Wang	2004	Analytical	Heterogeneous	[34]
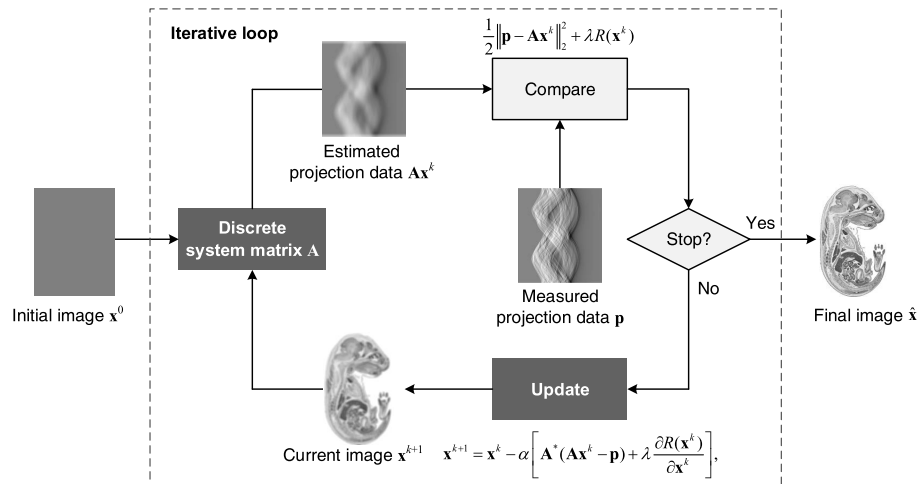
Fig. 20 Principle of the IR-based image reconstruction model.

3.5.1 Discrete forward imaging model

The basic idea of IR algorithms is to establish a mapping relationship between the photoacoustic image to be reconstructed and projection data using a set of linear equations and then solve them iteratively. To determine the relationship, a discrete photoacoustic imaging model was established and is illustrated in Fig. 21. In the discrete model, the 2D photoacoustic image is represented by $n \times n$ evenly distributed grid points, and the value of the jth grid point is x_j. The photoacoustic signal $p(\mathbf{r}_d, t)$ measured by a detector is discretely sampled, and the sampling length is K. The kth sampling point of the photoacoustic signal, denoted as p_k, relates to the spherical shell integral of the photoacoustic image over the kth detection shell, whose radius equals the TOF of the photoacoustic signal ($k\Delta t$, where Δt is the temporal sampling interval) multiplied by the SOS v_0. The goal of IR-based image reconstruction is to solve for x_j from the projection data p_k.

Based on the discrete imaging model, the relationship between the unknowns x_j and the projection data p_k can be described by the following set of linear equations:

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,N}x_N = p_1, \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,N}x_N = p_2, \\ \quad\quad\quad\vdots \\ a_{K,1}x_1 + a_{K,2}x_2 + \cdots + a_{K,N}x_N = p_K, \end{cases} \tag{83}$$

which can be written as

$$\sum_{j=1}^{N} a_{kj}x_j = p_k, \qquad k = 1, 2, \ldots, K, \tag{84}$$

where a_{kj} is a weighting factor representing the contribution of the jth grid point to the kth detection shell, $N = n \times n$ is the total number of grid points, and K is the total number of sampling points of a detector. If we consider the projection data measured by all M detectors and denote
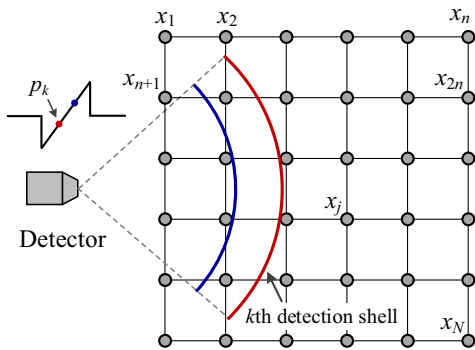
Fig. 21 Discrete photoacoustic imaging model in IR. The photoacoustic image is discretely represented by $n \times n$ evenly distributed grid points. The projection data p_k measured by a detector correspond to the spherical shell integral of the photoacoustic image over the kth detection shell.

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & & a_{2,N} \\ \vdots & \vdots & \vdots & \vdots \\ a_{K,1} & a_{K,2} & \cdots & a_{K,N} \\ a_{K+1,1} & a_{K+1,2} & \cdots & a_{K+1,N} \\ \vdots & \vdots & \vdots & \vdots \\ a_{MK,1} & a_{MK,2} & \cdots & a_{MK,N} \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_N \end{bmatrix},$$

$$\mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_K \\ p_{K+1} \\ \vdots \\ p_{MK} \end{bmatrix}, \tag{85}$$

Eq. (83) can be written in matrix form as

$$\mathbf{A}\mathbf{x} = \mathbf{p}, \tag{86}$$

where $\mathbf{A}$ is the system matrix that transforms the initial photoacoustic pressure $\mathbf{x}$ to the measured projection data $\mathbf{p}$. To solve Eq. (86), the system matrix $\mathbf{A}$ should be constructed first.

3.5.2 System matrix construction

A. System matrix construction for acoustically homogeneous media

If the imaging medium is acoustically homogeneous, the photoacoustic wave equations have an explicit analytical solution, as presented in Eq. (18). The solution describes how the measured photoacoustic signal $p(\mathbf{r}_d, t)$ relates to a photoacoustic source $p_0(\mathbf{r}_s)$. Discretizing the solution [Eq. (18)], the system matrix $\mathbf{A}$ can be constructed as

$$[\mathbf{p}]_{(m-1)K+k} = \left[\frac{1}{4\pi v_0^2} \frac{\partial}{\partial t} \int \frac{p_0(\mathbf{r}_s)}{|\mathbf{r}_s - \mathbf{r}_d^m|} \delta\left(t - \frac{|\mathbf{r}_s - \mathbf{r}_d^m|}{v_0} \right) d\mathbf{r}_s \right]_{t=k\Delta t}, \tag{87}$$

where $\mathbf{r}_d^m$ denotes the position of the mth detector, Δt is the temporal sampling interval, and $[\mathbf{p}]_{(m-1)K+k}$ is the photoacoustic signal measured by the mth detector at time $k\Delta t$. To calculate Eq. (87), $p_0(\mathbf{r}_s)$ can be expanded using a set of basis functions as[38]

$$p_0(\mathbf{r}_s) \approx \sum_{j=1}^{N} x(\mathbf{r}_s^j) \psi(\mathbf{r}_s^j), \tag{88}$$

where $\mathbf{r}_s^j$ represents the position of the jth grid point in the discrete image $\mathbf{x}$, and $\psi(\mathbf{r}_s^j)$ is the basis function at the jth grid point. Substituting Eq. (88) into Eq. (87), a discrete imaging model can be obtained as

$$[\mathbf{p}]_{(m-1)K+k}$$
$$= \left\{ \sum_{j=1}^{N} \left[\frac{1}{4\pi v_0^2} \frac{\partial}{\partial t} \int \frac{\psi(\mathbf{r}_s^j)}{|\mathbf{r}_s - \mathbf{r}_d^m|} \delta\left(t - \frac{|\mathbf{r}_s - \mathbf{r}_d^m|}{v_0} \right) d\mathbf{r}_s \right] x(\mathbf{r}_s^j) \right\}_{t=k\Delta t} .$$

$$(89)$$

Comparing Eqs. (84) and (89), the elements of the system matrix **A** can be obtained as

$$[\mathbf{A}]_{(m-1)K+k,j} = \left[\frac{1}{4\pi v_0^2} \frac{\partial}{\partial t} \int \frac{\psi(\mathbf{r}_s^j)}{|\mathbf{r}_s - \mathbf{r}_d^m|} \delta\left(t - \frac{|\mathbf{r}_s - \mathbf{r}_d^m|}{v_0} \right) d\mathbf{r}_s \right]_{t=k\Delta t} ,$$

$$(90)$$

where $[\mathbf{A}]_{(m-1)K+k,j}$ denotes the element in the $[(m-1)K+k]$th row and jth column of the system matrix **A**.

To construct a system matrix with sufficient accuracy, it is necessary to choose a suitable expansion function ψ. Several expansion functions are available for this purpose. Among them, a spherical-voxel function[36,177–179], Kaiser–Bessel functions[38], and linear interpolation functions[37,40,42,178,180] are the most commonly used. Figure 22 shows a schematic of the three expansion-function-based discrete grid models.

The expansion function ψ can help interpret the meaning of the system matrix **A**. Suppose that a photoacoustic image includes only a source defined by an expansion function ψ at grid point j. The element $[\mathbf{A}]_{(m-1)K+k,j}$ of the system matrix **A** is the response of the mth detector at time $k\Delta t$, as illustrated in Fig. 23. This implies that to construct the system matrix **A**, we can simply compute the photoacoustic signals measured by each detector and arrange the signals in a way according to Fig. 23. Since the system matrix **A** is only associated with the discrete photoacoustic imaging model (Fig. 21) and the expansion function ψ [Eq. (90)], it can be used for different experiments once constructed for a particular imaging system.

A1. Spherical-voxel-based system matrix

A spherical voxel function is defined as[181]

$$\psi(\mathbf{r}^j) = \begin{cases} 1, & \text{for } |\mathbf{r} - \mathbf{r}^j| \leq \varepsilon/2, \\ 0, & \text{for } |\mathbf{r} - \mathbf{r}^j| > \varepsilon/2, \end{cases} \qquad (91)$$

where ε is the spacing between two grid points, and $\mathbf{r}^j = (x^j, y^j, z^j)^T$ denotes the coordinate of the jth grid point. In a spherical-voxel-based discrete imaging model, each grid point in the image to be reconstructed is regarded as a uniform spherical source[36,177–179]. It is possible to directly construct the system matrix **A** based on the definition of the spherical voxel function. However, there may be problems. This is because a spherical voxel is typically small in radius. The photoacoustic signals are short in time and require a high sampling rate, which leads to an increase in computational complexity or aliasing
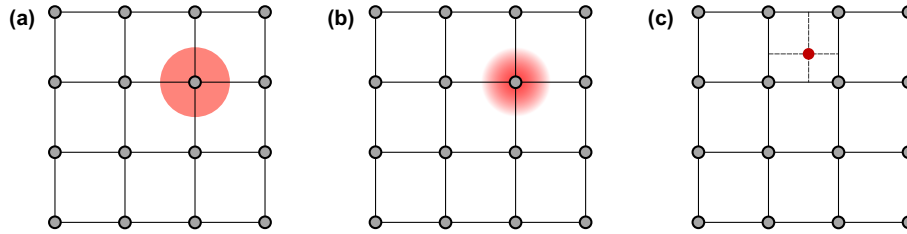

Fig. 22 Illustration of the discrete grid models based on different expansion functions. (a) Discrete grid model based on a spherical voxel. (b) Discrete grid model based on the Kaiser–Bessel function. (c) Discrete grid model based on the bilinear interpolation. The red dot in (c) represents the point to be interpolated.
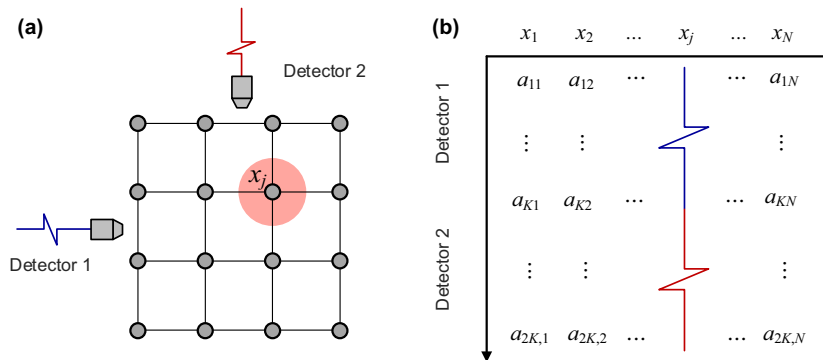

Fig. 23 Schematic diagram illustrating the meaning of the elements in a system matrix. (a) Spherical-voxel-function-based discrete imaging model. The photoacoustic signals generated from the spherical voxel at the jth grid point are recorded by detectors. (b) Elements of the system matrix. The jth column of the system matrix corresponds to the photoacoustic signal of the jth spherical voxel in (a).

artifacts[181]. To address this problem, a spherical-voxel-based system matrix can be constructed in the frequency domain[181].

Define $\mathbf{P}$ and $\tilde{\mathbf{A}}$ as the discrete Fourier transforms of the measured projection data $\mathbf{p}$ and the system matrix $\mathbf{A}$. The matrix form of the imaging model in Eq. (86) can be rewritten as[38,42]

$$\tilde{\mathbf{A}}\mathbf{x} = \mathbf{P}, \tag{92}$$

and the elements of the system matrix are given by[38,42]

$$[\tilde{\mathbf{A}}]_{(m-1)K+k,j} = P_0(f)\tilde{h}_{\mathrm{SIR}}^m(\mathbf{r}_s^j, f)|_{f=k\Delta f}, \tag{93}$$

where f denotes the frequency, Δf is the frequency sampling interval, $\tilde{h}_{\mathrm{SIR}}^m(\mathbf{r}_s^j, f)$ is the Fourier transform of the SIR of the mth detector and can be written as[38,42]

$$\tilde{h}_{\mathrm{SIR}}^m(\mathbf{r}_s^j, f) = \frac{1}{2\pi|\mathbf{r}_s^j - \mathbf{r}_d^m|}\exp\left(-i2\pi f\frac{|\mathbf{r}_s^j - \mathbf{r}_d^m|}{v_0}\right), \tag{94}$$

and $P_0(f)$ is the Fourier transform of the signal generated from a photoacoustic source defined by a spherical voxel and is given as[38,42]

$$P_0(f) = -i\frac{v_0}{f}\left[\frac{\varepsilon}{2v_0}\cos\left(\frac{\pi f\varepsilon}{v_0}\right) - \frac{1}{2\pi f}\sin\left(\frac{\pi f\varepsilon}{v_0}\right)\right]. \tag{95}$$

Combining Eqs. (93)–(95), the spherical-voxel-based system matrix $\mathbf{A}$ can be constructed in the frequency domain. This frequency domain approach can solve the sampling problem of high-frequency photoacoustic signals and make it easier to incorporate detector responses (see Sec. 3.5.3).

A2. Kaiser–Bessel-function-based system matrix

The Kaiser–Bessel (KB) functions are radially symmetric functions defined as[182,183]

$$b(r) = \begin{cases} \left(\sqrt{1-r^2/a^2}\right)^n \frac{I_n\left(\gamma\sqrt{1-r^2/a^2}\right)}{I_n(\gamma)}, & \text{if } 0 \le r \le a, \\ 0, & \text{otherwise,} \end{cases} \tag{96}$$

where I_n represents the modified Bessel function of the first kind of order n[184], a is the radius of support, and γ is a shape parameter governing the width of the radial profile. When $n = 0$, $a = \varepsilon/2$, and $\gamma = 0$, the KB function reduces to the spherical voxel function [Eq. (91)]. The KB-function-based expansion function can be expressed as

$$\psi(\mathbf{r}^j) = b(|\mathbf{r} - \mathbf{r}^j|), \tag{97}$$

where $\mathbf{r}^j = (x^j, y^j, z^j)^T$ denotes the coordinate of the jth grid point.

The elements of the KB-function-based system matrix have an analytical solution and can be calculated in the frequency domain[38]. The derivation is similar to that of the spherical-voxel-based system matrix described in Eqs. (92)–(94) except that Eq. (95) should be replaced with

$$P_0(f) = -\frac{i4\pi^2 f a^3}{v_0^2 I_n(\gamma)}\frac{j_{n+1}\left(\sqrt{4\pi^2 a^2 f^2/v_0^2 - \gamma^2}\right)}{(4\pi^2 a^2 f^2/v_0^2 - \gamma^2)^{(n+1)/2}}. \tag{98}$$

Here f denotes the frequency, $P_0(f)$ is the Fourier transform of the signal generated from the photoacoustic source defined by a KB function, and j_{n+1} is the spherical Bessel function of the first kind of order $n + 1$[184].

In the spherical-voxel-based imaging model, the expansion function ψ is a sphere with uniform intensity, which is not differentiable at the boundary. Therefore, generated photoacoustic signals have an infinite bandwidth and may induce numerical instabilities when calculating the system matrix using Eq. (90). In the KB-function-based imaging model, the expansion function ψ is also a sphere but with a smoothly varying intensity. It is differentiable at any position and thus can avoid numerical instabilities encountered by the spherical-voxel-based imaging model.

The shape of the KB function can be adjusted as needed. The parameter n affects the differentiability of the KB function and is typically set to a value greater than 2 ($n \ge 2$) to guarantee the continuity of derivatives at the boundary. The parameter a determines the effective size of a voxel and can be selected according to the desired spatial resolution. The parameter γ describes the smoothness of the KB function and impacts the width of the radial profile. It typically has a value of 10.4 for image reconstruction in PACT but may vary depending on application scenarios. Figure 24 illustrates the profiles of a family of KB functions for a given set of parameters. More details about the KB functions can be found in[185].
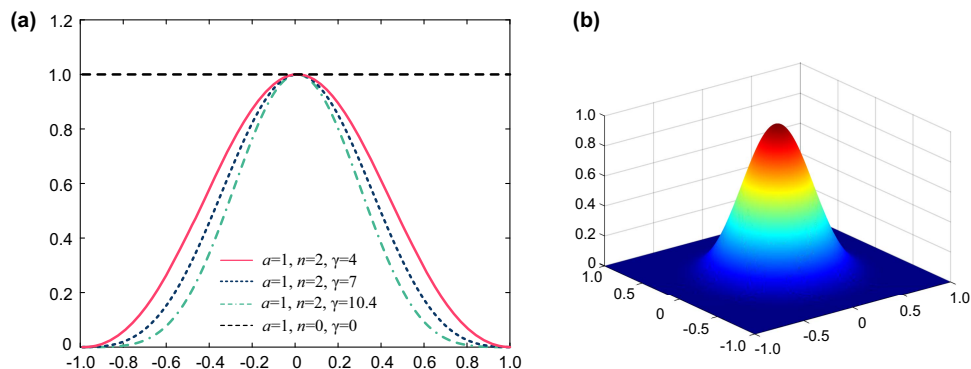
A3. Interpolation-function-based system matrix

Fig. 24 KB functions with different parameters. (a) Profiles of the KB functions with four groups of parameters. (b) 3D visualization of the KB function with $a = 1$, $n = 2$, and $\gamma = 10.4$.

In addition to the spherical-voxel- and KB-function-based approaches, the system matrix can also be constructed using linear interpolation functions. Several different interpolation functions can be used as the expansion function to describe the mapping relation between a continuous image and its discrete counterpart. Generally, a high-order interpolation function helps produce high-accuracy results, which, however, involves more calculations. Considering the trade-off between accuracy and efficiency, bilinear and trilinear interpolation functions are commonly used for 2D[43] and 3D[38,40,42] IR models, respectively. Specifically, when the trilinear interpolation method is employed in 3D space, the expansion function can be expressed as[42]

$$
\psi(\mathbf{r}^j) = \begin{cases} \left(1 - \dfrac{|x - x^j|}{\varepsilon}\right)\left(1 - \dfrac{|y - y^j|}{\varepsilon}\right)\left(1 - \dfrac{|z - z^j|}{\varepsilon}\right), & \text{for } |x - x^j|, |y - y^j|, |z - z^j| \leq \varepsilon, \\ 0, & \text{otherwise,} \end{cases}
\tag{99}
$$

where $\mathbf{r}^j = (x^j, y^j, z^j)^T$ denotes the coordinate of the jth grid point, and ε is the spacing between two grid points. The equation implies that the non-zero values of the expansion function ψ exist only in the ε-neighborhood of position $\mathbf{r}^j$. The image value at any position can be interpolated from its four neighbors in 2D space or eight neighbors in 3D space.

The interpolation-function-based system matrix $\mathbf{A}$ can be constructed in two steps as[42]

$$
\mathbf{p} = \mathbf{A}\mathbf{x} \equiv \mathbf{D}\mathbf{G}\mathbf{x},
\tag{100}
$$

where the matrix $\mathbf{G}$ represents the spherical Radon transform in 3D [Eq. (27)], and the matrix $\mathbf{D}$ represents the differential operation in Eq. (18). The matrix $\mathbf{G}$ can be obtained by accumulating all image grid points intersecting with an integrating spherical shell and satisfies the following relationship[42]:

$$
[\mathbf{G}\mathbf{x}]_{(m-1)K+k} = \varepsilon^2 \sum_{j=1}^{N} [\mathbf{x}]_j \sum_{p=1}^{N_p} \sum_{q=1}^{N_q} \psi(\mathbf{r}^j_{k,p,q}) \equiv [\mathbf{g}]_{(m-1)K+k},
\tag{101}
$$

where $[\mathbf{g}]_{(m-1)K+k} \approx g(\mathbf{r}^m_d, t = k\Delta t)$ is the approximation of the spherical Radon transform in Eq. (27), and N_p and N_q are the numbers of angular divisions over the polar and azimuth directions, respectively, on a local spherical coordinate system

centered at $\mathbf{r}^m_d$[42]. The differential matrix $\mathbf{D}$ can be obtained using a difference operation and satisfies the following relationship[42]:

$$
[\mathbf{D}\mathbf{g}]_{(m-1)K+k} = \frac{1}{8\pi v_0^2 \Delta t^2} \left[\frac{[\mathbf{g}]_{(m-1)K+k+1}}{k+1} - \frac{[\mathbf{g}]_{(m-1)K+k-1}}{k-1} \right]
$$
$$
\equiv [\mathbf{p}]_{(m-1)K+k}.
\tag{102}
$$

The preceding procedure implies that the system matrix $\mathbf{A}$ can be constructed via the two matrices $\mathbf{G}$ and $\mathbf{D}$, while the matrix elements in $\mathbf{A}$ do not need to be explicitly calculated.

In addition to the implicit method discussed above, the interpolation-function-based system matrix can also be constructed by explicitly calculating its elements via analytical methods[37,40,42,43]. The analytical methods typically have better computational stability due to more accurate calculations of derivatives.

B. System matrix construction for acoustically heterogeneous media

The system matrix discussed above is primarily constructed based on the analytical solution of the photoacoustic wave equations [Eq. (18)], where the acoustic medium is assumed to be homogeneous. When the acoustic media are heterogeneous, the photoacoustic discrete imaging model can still be described by the coupled photoacoustic wave equations [Eqs. (63)–(67)], which, however, have no analytical solutions. Based on the k-space pseudospectral method (see Sec. 3.4.1), Huang *et al.* proposed a method for constructing the system matrix $\mathbf{A}$ in acoustically heterogeneous media[41], which is briefly reviewed below.

Define a vector containing all photoacoustic field variables at the time step $k\Delta t$ as[41]

$$
\mathbf{v}_k = (\mathbf{u}_k^1, \mathbf{u}_k^2, \mathbf{u}_k^3, \boldsymbol{\rho}_k^1, \boldsymbol{\rho}_k^2, \boldsymbol{\rho}_k^3, \mathbf{p}_k)^T,
\tag{103}
$$

where $\mathbf{v}_k$ is a $7N \times 1$ vector, N is the total number of grid points in an image, $\mathbf{u}_k^i$ and $\boldsymbol{\rho}_k^i$ ($i = 1, 2, 3$) denote the particle velocity and acoustic density in the ith direction on a 3D Cartesian grid at
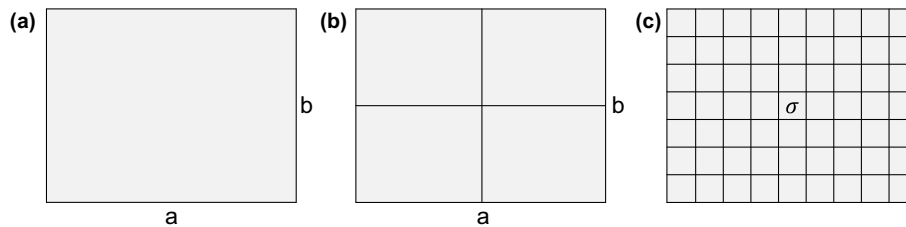

Fig. 25 Detector models for building SIR. (a) Detector model under the condition of far-field approximation. (b) Detector model based on patches ($n = 2$). (c) Detector model based on discrete detection elements.

the kth time step, $\mathbf{p}_k$ is the acoustic pressure at the grid points, and the superscript T denotes the matrix transpose. The update of the field variables from $t = 0$ to $t = (K - 1)\Delta t$ is described by

$$[\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{K-1}]^T = \mathbf{T}_{K-1}\mathbf{T}_{K-2}\cdots\mathbf{T}_1[\mathbf{v}_0, \mathbf{0}_{7N\times 1}, \cdots, \mathbf{0}_{7N\times 1}]^T, \tag{104}$$

where $\mathbf{T}$ is a $7NK \times 7NK$ matrix, and K is the total number of time steps. The vector $[\mathbf{v}_0, \mathbf{0}_{7N\times 1}, \cdots, \mathbf{0}_{7N\times 1}]$ represents the values of the field variables at $t = 0$ and can be calculated from the initial photoacoustic pressure $\mathbf{p}_0$ as

$$[\mathbf{v}_0, \mathbf{0}_{7N\times 1}, \cdots, \mathbf{0}_{7N\times 1}]^T = \mathbf{T}_0\mathbf{p}_0, \tag{105}$$

where $\mathbf{T}_0$ is a $7NK \times N$ matrix that maps the initial pressure $\mathbf{p}_0$ to the initial value of the field variables at time $t = 0$[41]. Since detectors generally do not coincide with image grid points, the measured photoacoustic projection data $\mathbf{p}$ can be related to the field quantities via interpolation as[41]

$$\mathbf{p} = \mathbf{S}[\mathbf{v}_0, \mathbf{v}_1, \cdots, \mathbf{v}_{K-1}]^T, \tag{106}$$

where $\mathbf{S}$ denotes the interpolation operation, such as the Delaunay triangulation-based interpolation. Combining Eqs. (104)–(106), we have

$$\mathbf{p} = \mathbf{S}\mathbf{T}_{K-1}\cdots\mathbf{T}_1\mathbf{T}_0\mathbf{p}_0. \tag{107}$$

By comparing Eq. (107) with Eq. (86), the system matrix can be obtained as

$$\mathbf{A} = \mathbf{S}\mathbf{T}_{K-1}\cdots\mathbf{T}_1\mathbf{T}_0. \tag{108}$$

As mentioned above, the jth column of system matrix $\mathbf{A}$ corresponds to the response of a detector to the source defined by an expansion function ψ at grid point j. Therefore, the system matrix $\mathbf{A}$ in heterogeneous media can be constructed by computing the response of a detector when the photoacoustic source defined by an expansion function ψ traverses all image grid points, which can be calculated using the k-Wave toolbox[186].

In a homogeneous medium, the construction of the system matrix $\mathbf{A}$ can be simpler. The response of a detector to an expansion-function-defined source at the first (reference) image grid point is computed and serves as the first column of $\mathbf{A}$. The remaining columns of $\mathbf{A}$ can be obtained by time-shifting the elements in the first column while taking into account the signal attenuation according to the relative position between the current and the reference grid points. This can greatly improve the construction speed of the system matrix $\mathbf{A}$.

3.5.3 Transducer response modeling

The IR model discussed above is established based on the assumption that an ultrasound detector is point-like in shape and has an infinite bandwidth. However, a real ultrasound detector always has a finite aperture size and a limited bandwidth, which makes the IR model less accurate. To address this problem, it is necessary to incorporate the SIR and EIR of an ultrasound detector (see Sec. 2.3) to make the system matrix $\mathbf{A}$ as realistic as possible.

A. EIR modeling

When the system matrix is constructed in the time domain, the discrete photoacoustic imaging model [Eq. (86)] incorporating the EIR of a detector can be written as[37]

$$\mathcal{F}^{-1}\tilde{\mathbf{E}}\mathcal{F}\mathbf{A}\mathbf{x} = \mathbf{p}, \tag{109}$$

where the matrix $\tilde{\mathbf{E}}$ denotes the Fourier transform of the EIR of a detector, and $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent the forward and inverse Fourier transforms, respectively. In contrast, when the system matrix is constructed in the frequency domain using the spherical voxel or the KB function, the EIR of a detector can be incorporated into a system matrix via[39]

$$[\tilde{\mathbf{A}}]_{(m-1)K+k} = P_0(f)\tilde{h}_{\text{SIR}}^m(r_s^j, f)\tilde{h}_{\text{EIR}}(f)|_{f=k\Delta f}. \tag{110}$$

Here, P_0 is the Fourier transform of the acoustic pressure generated from the photoacoustic source defined by the spherical voxel or the KB function, f is the frequency, Δf is the frequency sampling interval, k is the index of sampling points, and $\tilde{h}_{\text{SIR}}^m$ and $\tilde{h}_{\text{EIR}}$ are the Fourier transforms of the SIR and EIR of a detector, which can be modeled or measured in practice.

B. SIR modeling

B1. Far-field SIR

Several SIR models have been proposed and used for image reconstruction in PACT[187,188]. When a detector has a flat and rectangular surface with an area of $a \times b$ [see Fig. 25(a)] and the condition of far-field approximation[189] is met, the temporal Fourier transform of the SIR of a detector can be expressed as[187]

$$\tilde{h}_{\text{SIR}}^m(\mathbf{r}_s^j, f) = \frac{ab \exp\left(-i2\pi f \frac{|\mathbf{r}_s^j - \mathbf{r}_d^m|}{v_0}\right)}{2\pi|\mathbf{r}_s^j - \mathbf{r}_d^m|} \text{sinc}\left(\pi f \frac{ax_{\text{det}}^{jm}}{v_0|\mathbf{r}_s^j - \mathbf{r}_d^m|}\right)$$
$$\times \text{sinc}\left(\pi f \frac{by_{\text{det}}^{jm}}{v_0|\mathbf{r}_s^j - \mathbf{r}_d^m|}\right), \tag{111}$$

where x_{det}^{jm} and y_{det}^{jm} represent the coordinates of the jth grid in a local coordinate system at the mth detector. Other variables have the same meanings as those defined in Eq. (90). With Eq. (111), the system matrix $\mathbf{A}$ can be constructed according to Eq. (93).

B2. Patch-based SIR

When the far-field approximation condition does not hold, the far-field SIR model may lead to inaccurate reconstruction results. To address this problem, the surface of a detector can be divided into small patches [Fig. 25(b)]. For each small patch, the aforementioned far-field assumption is approximately met. The SIR of each small patch can be characterized by Eq. (111), and the SIR of the whole detector is obtained by summing the SIRs of all small patches. Suppose that a detector has an area of $a \times b$ and is divided into $n \times n$ patches. Denoting a' and b' as the length and width of each small patch (i.e., $a' = a/n$, $b' = b/n$), we have

$$\tilde{h}_{\text{SIR}}^m(\mathbf{r}_s^j, f) \approx \frac{1}{n^2}\sum_{l=1}^{n^2}\tilde{h}_{\text{SIR}}^{m,l}(\mathbf{r}_s^j, f). \tag{112}$$

Here $\tilde{h}_{\text{SIR}}^{m,l}(\mathbf{r}_s^j, f)$ is the SIR of the lth patch, which can be calculated using Eq. (111).

B3. Discrete-detection-element-based SIR

The far-field and patch-based SIR models are easy to incorporate into the system matrix based on the spherical voxel

function or the KB function but are difficult to incorporate into the system matrix based on linear interpolation or the k-space pseudospectral method. Huang *et al.*[41] and Ding *et al.*[40] solved this problem by dividing the entire detector surface into a set of detection elements with equal areas, as shown in Fig. 25(c). In this case, the acoustic pressure $p(\mathbf{r}_\mathrm{d}, t)$ recorded by a detector can be approximately written as[40]

$$p(r_\mathrm{d}, t) \approx \sum_{l=1}^{L} p(\mathbf{r}_\mathrm{d}^l, t)\sigma, \tag{113}$$

where L is the total number of detection elements, σ is the area of a detection element (dimensionless), and $\mathbf{r}_\mathrm{d}^l$ is the position of the lth detection element. $p(\mathbf{r}_\mathrm{d}^l, t)$ is the acoustic pressure recorded at position $\mathbf{r}_\mathrm{d}^l$ and time t and can be written in matrix form as [Eq. (86)]

$$\mathbf{A}_l \mathbf{x} = \mathbf{p}_l, \tag{114}$$

where $\mathbf{p}_l$ is the acoustic pressure recorded by the lth detection element, and $\mathbf{A}_l$ is the system matrix for the lth detection element. Substituting Eq. (114) into Eq. (113), we have

$$\sum_{l=1}^{L} \mathbf{A}_l \mathbf{x}\sigma \approx \mathbf{p}. \tag{115}$$

The system matrix $\mathbf{A}$ in the discrete imaging model in Eq. (86) is replaced with a weighted summation of L system matrices $\mathbf{A}_l$ in the detection-element-based SIR model. To achieve accurate representation, the size of the discrete detection element should be small, preferably much less than the acoustic wavelength[41]. Therefore, the discrete-detection-element-based SIR model involves more computations and is generally more time-consuming than the far-field- and patch-based SIR models.

3.5.4 Iterative reconstruction

Once the system matrix $\mathbf{A}$ of a PACT imaging system is constructed, the initial photoacoustic pressure $\mathbf{x}$ can be recovered from the measured projection data $\mathbf{p}$ by solving Eq. (86). In principle, the linear photoacoustic imaging model in Eq. (86) can be solved by the pseudo-inverse matrix method, i.e.,

$$\mathbf{x} = \mathbf{A}^\dagger \mathbf{p}, \tag{116}$$

where $\mathbf{A} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ is the Moore–Penrose pseudo-inverse matrix of $\mathbf{A}$, and the superscript H denotes the conjugate transpose. However, the pseudo-inverse method is not commonly used in practice because the system matrix $\mathbf{A}$ is generally too large to be inverted[61]. Alternatively, the imaging model can be solved iteratively. In this way, the image reconstruction problem is equivalent to solving the least-square problem

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\geq 0} \frac{1}{2}\|\mathbf{p} - \mathbf{A}\mathbf{x}\|_2^2, \tag{117}$$

where $\|\cdot\|_2$ represents the L_2-norm, and $\hat{\mathbf{x}}$ is the approximate solution of the least-square problem. Since the initial acoustic pressure is non-negative, a non-negativity constraint is imposed on Eq. (117). Generally, the optimization problem in Eq. (117) is ill-posed under non-ideal imaging conditions. To address this

problem, Eq. (117) is typically regularized by a penalty function and can be rewritten as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\geq 0} \frac{1}{2}\|\mathbf{p} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda R(\mathbf{x}), \tag{118}$$

where $\|\mathbf{p} - \mathbf{A}\mathbf{x}\|_2^2$ is the data fidelity term, $R(\mathbf{x})$ is a regularization term, and λ is a regularization coefficient controlling the weight of regularization, which can be automatically optimized by methods such as the generalized cross-validation (GCV) method[190] and the L-curve method[191,192].

One popular regularization technique used in Eq. (118) is Tikhonov regularization, which can be expressed as[63]

$$R(\mathbf{x}) = \|\mathbf{\Gamma}\mathbf{x}\|_2^2, \tag{119}$$

where $\mathbf{\Gamma}$ is the Tikhonov matrix. In many cases, the Tikhonov matrix is chosen as the identity matrix ($\mathbf{\Gamma} = \mathbf{I}$), giving preference to solutions with smaller norms. When the regularization term $R(\mathbf{x})$ is convex and differentiable, such as in Tikhonov regularization, the optimization problem in Eq. (118) can be solved by optimization methods such as stochastic gradient descent (SGD)[193,194], least-square QR (LSQR)[195], and conjugate gradient (CG)[196]. When a gradient descent method[197] is used, the optimization problem in Eq. (118) can be solved by

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \alpha\left[\mathbf{A}^*(\mathbf{A}\mathbf{x}^k - \mathbf{p}) + \lambda\frac{\partial R(\mathbf{x}^k)}{\partial \mathbf{x}^k}\right], \tag{120}$$

where k is the iteration number, α is the step size controlling the update rate, $\mathbf{A}^*(\mathbf{A}\mathbf{x}^k - \mathbf{p})$ is the gradient of the fidelity term $\|\mathbf{A}\mathbf{x} - \mathbf{p}\|_2^2$, and $\mathbf{A}^*$ is the adjoint of the matrix $\mathbf{A}$.

Tikhonov regularization emphasizes the smoothness of an image and tends to produce images with blurred edges. An alternative to Tikhonov regularization is sparse regularization, which is non-smooth and aims to find a solution with the minimum number of non-zero elements in certain transform domains[198–200]. L_1-norm-based regularization is one of the most commonly used sparse regularization methods and has the following form[63,201]:

$$R(\mathbf{x}) = \|\mathbf{\Phi}\mathbf{x}\|_1, \tag{121}$$

where $\mathbf{\Phi}$ is a transformation matrix that can be constructed using sparse basis functions, such as the wavelet function and finite difference operators. When a finite difference operator is employed, the L_1-norm regularization becomes total variation (TV) regularization[181], which can be written as

$$\begin{aligned} R(\mathbf{x}) &= \|\mathbf{x}\|_\mathrm{TV} \\ &= \sum_{j=1}^{N} \sqrt{([\mathbf{x}]_j - [\mathbf{x}]_{j_x-1})^2 + ([\mathbf{x}]_j - [\mathbf{x}]_{j_y-1})^2 + ([\mathbf{x}]_j - [\mathbf{x}]_{j_z-1})^2}, \end{aligned} \tag{122}$$

where $j_x - 1$, $j_y - 1$, and $j_z - 1$ are the neighboring grids of the jth grid along the x-axis, y-axis, and z-axis, respectively, and N is the total number of image grid points. Sparse regularization involves non-smooth cost functions and can be solved using proximal point algorithms, such as the iterative shrinkage thresholding

algorithm (ISTA) and its accelerated version fast ISTA (FISTA)[202], the iteratively reweighted least squares (IRLS)[203], and the alternating direction method of multipliers (ADMM)[204]. Specifically, the optimization problem in Eq. (118) can be solved by a proximal-gradient-descent scheme[45]:

$$\mathbf{x}^{k+1} = \text{prox}_{R,\lambda}[\mathbf{x}^k - \alpha\mathbf{A}^*(\mathbf{A}\mathbf{x}^k - \mathbf{p})], \quad (123)$$

where $\text{prox}_{R,\lambda}$ is a proximal operator related to the regularization term $R(\mathbf{x})$ and the parameter λ. From Eqs. (120) and (123), we know that the update process of an IR algorithm is related to the current reconstructed image $\mathbf{x}^k$, the regularization term $R(\mathbf{x})$, and the gradient term $\mathbf{A}^*(\mathbf{A}\mathbf{x}^k - \mathbf{p})$, as shown in Fig. 20.

Combining different regularization strategies may help achieve improved image quality compared with using only a single type of regularization[205]. When dual regularization is imposed, the regularization term can be written as

$$R(\mathbf{x}) = \lambda_1 R_1(\mathbf{x}) + \lambda_2 R_2(\mathbf{x}), \quad (124)$$

where $R_1(\mathbf{x})$ and $R_2(\mathbf{x})$ represent two types of regularization, and λ_1 and λ_2 are the corresponding coefficients. Based on this idea, Li *et al.* proposed non-local means and sparse-wavelet-based dual regularization and achieved image reconstruction with enhanced SNR and contrast[205].

Figure 26 shows an example of IR-based image reconstruction in PACT. Figure 26(a) shows a numerical blood vessel phantom. The photoacoustic signals generated from the phantom are received by a 50-mm-diameter full-ring detector array with 64 elements enclosing the phantom. Figure 26(b) is the image reconstructed by FBP [Eq. (47)]. Severe streak artifacts are present in the reconstructed image due to the limited number of ultrasound detectors. Figure 26(c) is the image reconstructed by TV-based IR. The results show that the IR algorithm can produce images with fewer artifacts, demonstrating its superior performance under non-ideal imaging scenarios. The intensity profiles of the reconstructed images indicated by the white arrow are shown in Fig. 26(d) for comparison of the results of IR and FBP.

Table 8 summarizes representative work on IR-based image reconstruction in PACT.

4 Deep Learning Approaches

The aforementioned conventional approaches can achieve high-quality image reconstruction under ideal imaging conditions. However, they may face challenges under non-ideal imaging conditions, such as limited detector bandwidth, finite detector aperture, limited view angle, sparse detector arrangement, and inhomogeneous acoustic media. Inspired by the successful applications of DL across industries such as computer vision[212], natural language processing[213], and biomedical engineering[214], DL-based image reconstruction in tomography has drawn considerable attention in recent years[215–217]. DL has been successfully used for image reconstruction in CT, magnetic resonance imaging (MRI), positron emission tomography (PET), ultrasound, and other imaging modalities[215,217,218]. It can achieve tomographic image reconstruction with improved image quality and computational efficiency. In PACT, DL has been used to address a range of image reconstruction problems[48,50], such as detector bandwidth expansion[51,52], resolution enhancement[53,54], low-power/cost imaging[219,220], artifact removal[55], reconstruction acceleration[57], and reconstruction enhancement in sparse-view and limited-view imaging[44,46,47,58]. Specifically, the applications of DL in PACT image reconstruction are mainly reflected in five aspects, including signal preprocessing in the data domain, image postprocessing in the image domain, hybrid-domain processing in both the data and the image domains, learned iterative reconstruction, and direct reconstruction.

4.1 Preprocessing in the Data Domain

Under non-ideal imaging conditions, the raw photoacoustic projection data measured by a detector array may be incomplete and contain noise. Data-domain signal preprocessing attempts to transform incomplete photoacoustic projection data to nearly complete projection data using neural networks before image reconstruction, as shown in Fig. 27.

Many related studies have emerged based on this idea[221]. Gutta *et al.* proposed a simple five-layer fully connected network to expand the bandwidth of measured photoacoustic projection data and achieved image reconstruction with improved contrast and quality[51]. Awasthi *et al.* developed a UNet-based model for super-resolution, denoising, and bandwidth enhancement of photoacoustic projection data in sparse-view imaging and achieved improved image reconstruction, as shown in Fig. 28[52]. Here, UNet is a specially designed U-shaped
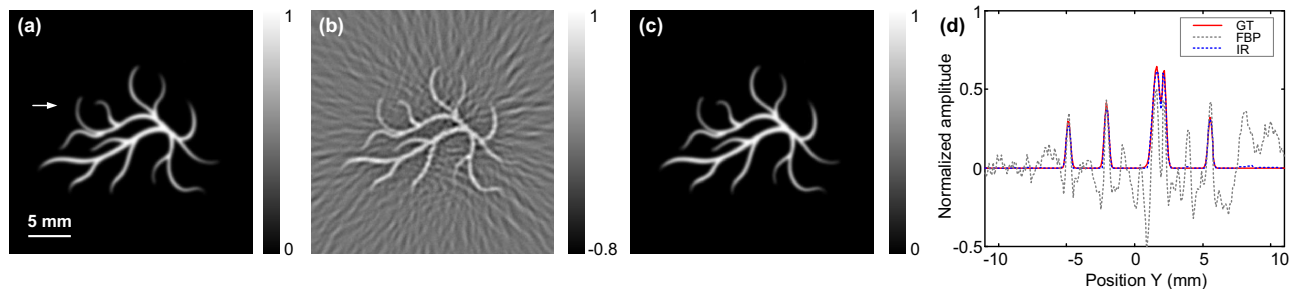

Fig. 26 Image reconstruction in sparse-view imaging by IR. (a) Numerical blood vessel phantom. The photoacoustic signals generated from the phantom are received by a full-ring detector array with 64 elements enclosing the phantom. (b) Image reconstructed by FBP [Eq. (47)]. (c) Image reconstructed by TV-based IR. (d) Intensity profiles of the reconstructed images indicated by the arrow in (a). GT: ground truth.

Table 8 Representative Work on IR-Based Image Reconstruction in PACT

Author	Year	System Matrix Model	Media	Detector Response	Regularization	Optimization	Dim	Source
Yalavarthy *et al.*	2021	k-space PSTD	Heterog	EIR	Non-local means + TV	IRLS	3D	[206]
Biton *et al.*	2019	Interpolation	Homog	–	Adaptive anisotropic TV	Chambolle-Pock	2D	[201]
Li *et al.*	2019	Interpolation	Homog	–	Non-local means + L_1 - norm	GD, FISTA	2D	[205]
Prakash *et al.*	2019	Interpolation	Homog	–	Second-order TV	CG	2D	[207]
Ding *et al.*	2017	Interpolation	Homog	SIR	–	LSQR	3D	[40]
Ding *et al.*	2016	Interpolation	Homog	–	–	LSQR	2D	[43]
Liu *et al.*	2016	Curve-driven	Homog	–	–	LSQR	2D	[208]
Javaherian& Holman	2016	k-space PSTD	Homog	–	TV	Multi-grid ISTA, FISTA	2D	[209]
Wang *et al.*	2014	KB function	Homog	EIR, SIR	Quadratic smoothness penalty/Tikhonov	Linear CG	3D	[38]
Mitsuhashi *et al.*	2014	Spherical voxel	Homog	EIR, SIR	Quadratic smoothness penalty	Fletcher-Reeves CG	3D	[189]
Huang *et al.*	2013	k-space PSTD	Heterog	EIR, SIR	TV	FISTA	2D	[41]
Wang *et al.*	2012	Spherical voxel	Homog	EIR, SIR	Quadratic smoothness Laplacian/TV	Fletcher-Reeves CG, FISTA	3D	[181]
Deán-Ben *et al.*	2012	Interpolation	Heterog (Weak)	–	Second-order Tikhonov	LSQR	2D	[210]
Deán-Ben *et al.*	2012	Interpolation	–	–	Tikhonov	LSQR	3D	[178]
Rosenthal *et al.*	2011	Interpolation	Homog	SIR	–	Pseudo-inverse, LSQR	2D	[211]
Wang *et al.*	2011	Spherical voxel	Homog	EIR, SIR	Quadratic smoothness penalty	Fletcher-Reeves CG	3D	[39]
Rosenthal *et al.*	2010	Interpolation	Homog	EIR	–	Pseudo-inverse, LSQR	2D	[37]
Provost & Lesage	2009	–	Homog	EIR	L_1-norm	LASSO	2D	[198]
Paltauf *et al.*	2002	Spherical voxel	Homog	–	–	Algebraic iteration	3D	[36]

Note: PSTD: pseudospectral time-domain; Homog: homogeneous; Heterog: heterogeneous; Dim: dimension; GD: gradient descent; LASSO: least absolute shrinkage and selection operator.
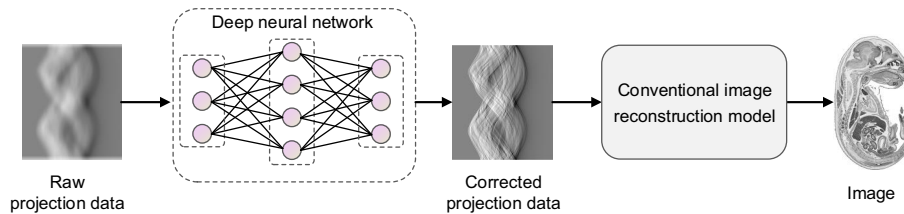
Fig. 27 DL-based projection data preprocessing in the data domain.

convolutional neural network, which has been widely used in various image processing tasks such as image denoising, image deconvolution, and image reconstruction. Zhang *et al*. designed a network based on a channel-wise attention mechanism to extract feature relations between sparse data channels and achieved full-channel projection data recovery from sparse input[222]. Zhang *et al*. proposed a UNet-based network to correct

photoacoustic projection data with skull-induced aberration in brain imaging and demonstrated that the aberration could be effectively removed after preprocessing[223]. These studies show that preprocessing projection data in the data domain can help enhance image reconstruction quality.

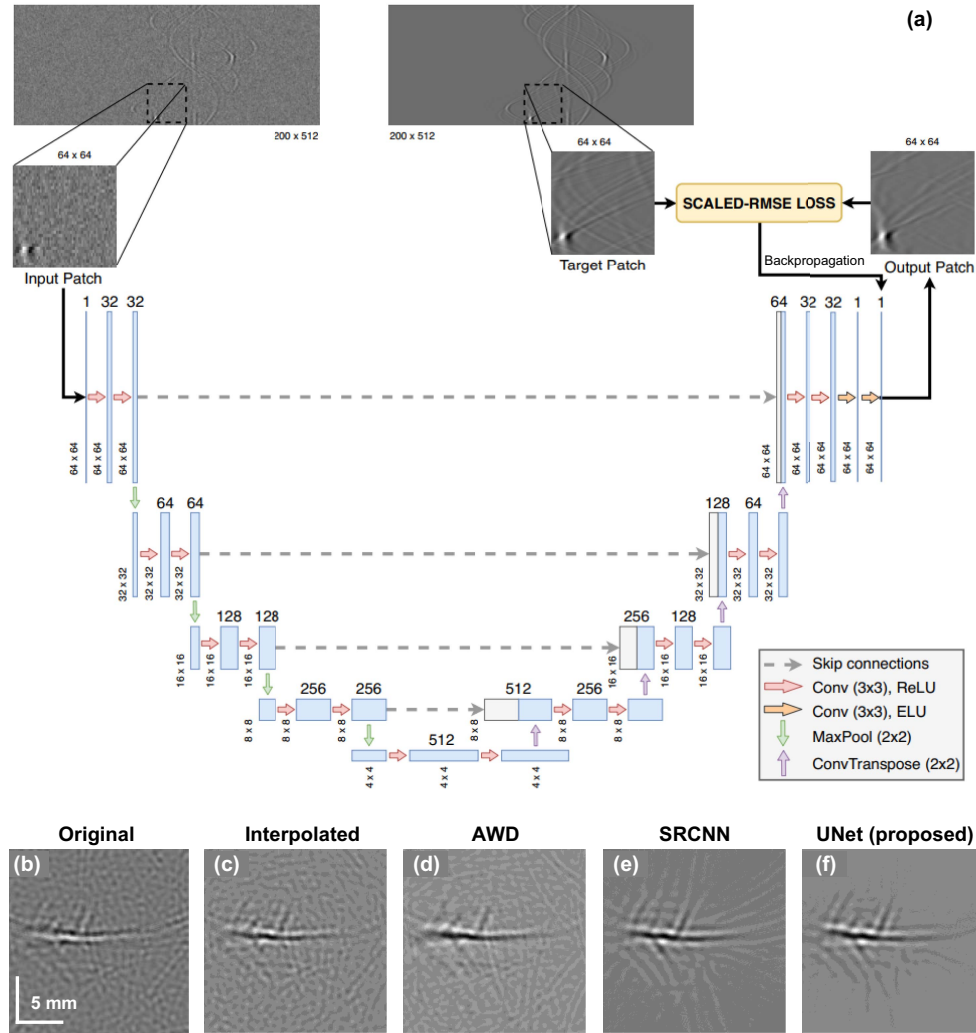Table 9 lists representative work on DL-based signal preprocessing in PACT.

Fig. 28 DL-based preprocessing helps correct photoacoustic projection data and enhances image reconstruction quality. The photoacoustic projection data in this case were recorded by sparsely distributed detectors with finite bandwidths. (a) Architecture of the UNet used for signal preprocessing. (b) Reconstructed image of a living rat brain using the raw bandwidth-limited projection data of 100 detectors. (c) Reconstructed image using the interpolated projection data of 200 detectors. (d) Reconstructed image using the interpolated projection data denoised by automated wavelet denoising (AWD). (e) Reconstructed image using the interpolated projection data (c) processed by the super-resolution CNN (SRCNN). (f) Reconstructed image using the interpolated projection data (c) processed by the UNet in (a). Adapted from Ref. [52] with permission.

Table 9 Representative Work on DL-Based Signal Preprocessing in PACT

Author	Year	Problem	Network	Dataset	Source
Zhang *et al.*	2023	Skull-induced aberration correction	UNet	Simulation	[223]
Zhang *et al.*	2022	Sparse-view imaging	Transformer	Simulation	[222]
Awasthi *et al.*	2020	Bandwidth expansion and sparse-view imaging	UNet	Simulation, phantom (test), *in vivo* (test)	[52]
Gutta *et al.*	2017	Detector bandwidth expansion	Five fully connected layers	Simulation, phantom (test)	[51]

4.2 Postprocessing in the Image Domain

In addition to signal preprocessing in the data domain, another application of DL in PACT is image postprocessing in the image domain, which indicates that deep neural networks can be used as postprocessing tools for image enhancement (see Fig. 29). This approach is especially useful for image artifacts removal in non-ideal imaging scenarios.

To mitigate the image artifacts that often appear in sparse-view or/and limited-view PACT images, a variety of DL-based postprocessing methods have been proposed. For example, Antholzer *et al.* proposed a UNet-based architecture to process images reconstructed by FBP in sparse-view imaging and demonstrated the effectiveness of the model in artifact removal[44]. Farnia *et al.* developed a UNet-based network to process images reconstructed by TR and achieved artifact suppression with limited projection data[224]. Davoudi *et al.* proposed a UNet-based framework for image quality recovery from sparse and limited projection data and demonstrated its performance with whole-body mouse imaging *in vivo*, as shown in Fig. 30[46].

In addition to the classic UNet, its variants are also commonly used as postprocessing tools in PACT. Guan *et al.* developed a fully dense UNet (FD-UNet) for image artifact removal in sparse-view PACT imaging and demonstrated its superiority over the conventional UNet[225]. Zhang *et al.* proposed a dual-domain UNet (DuDoUNet) with a specially designed information-sharing block, which takes both time-domain DAS images and k-space images as inputs, and verified its superiority with a public clinical database[226]. Guan *et al.* proposed a 3D modified UNet called dense dilation UNet (DD-UNet) and demonstrated its effectiveness in improving the imaging quality in 3D sparse-view and limited-view imaging[227]. Choi *et al.* reported a 3D progressive UNet for multi-parameter dynamic volume imaging and demonstrated that it could improve imaging speed and reduce image artifacts in limited-view PACT imaging, as shown in Fig. 31[47].

In addition to UNet and its variants, generative adversarial networks (GANs) are another deep neural network commonly used for image postprocessing in PACT[228]. GAN shows excellent image processing capabilities and has been successfully used in many applications, such as image style conversion, image enhancement, and image restoration. A GAN network consists of two sub-networks, i.e., a generator and a discriminator, which are adversarially trained until the discriminator cannot distinguish the result produced by the generator and the ground truth. Vu *et al.* explored the application of a Wasserstein GAN with gradient penalty (WGAN–GP) in limited-view and limited-bandwidth PACT imaging and achieved image artifact removal[229]. Lu *et al.* proposed a GAN-based DL approach (LV-GAN) to recover high-quality images in limited-view

imaging and achieved artifact removal and high recovery accuracy under a view angle of less than 60°[58]. Shahid *et al.* developed a residual-network-based GAN (ResGAN) to improve image quality by denoising and removing image artifacts in sparse-view imaging[230]. The deep neural networks in the above studies generally require supervision and need a large number of paired images for training, which are often difficult to obtain, especially in experiments. To address this problem, Lu *et al.* proposed an unsupervised cyclic GAN network (CycleGAN) that only needs unpaired training and successfully achieved artifact removal in sparse-view and limited-view PACT images[231].

In addition to solving the image enhancement problem in sparse-view and limited-view imaging, DL-based postprocessing has also been used to address a variety of other imaging problems, such as reflection artifact removal and aberration mitigation. For example, Shan *et al.* incorporated a deep neural network into conventional iterative algorithms and successfully corrected reflection artifacts caused by planar echogenic structures outside imaging tissues[232]. Jeon *et al.* developed a hybrid deep neural network based on SegNet and UNet and achieved SOS aberration mitigation, streak artifact removal, and temporal resolution improvement in sparse-view imaging[233]. Gao *et al.* proposed a modified encoder-decoder UNet to learn the mapping relationship between speckle patterns and target images in thick porous media and solved the scattering problem in transcranial photoacoustic brain imaging[234]. Shijo *et al.* proposed a shifted-window transformer to restore artifact-free images from artifact-heavy images reconstructed using a TR algorithm[235].

Deep neural networks have also been used to improve the spatial resolution of PACT images. For example, Rajendran and Pramanik employed a fully dense U-Net termed TARES to improve the spatially variant tangential resolution in circular scanning PACT imaging systems[53]. Zhang *et al.* exploited a fully dense UNet named Deep-E to improve the elevational resolution of linear-detector-array-based PACT imaging[54]. Based on this work, the same research group proposed a new Deep-E combining a 2D Deep-E and a 3D Deep-E and demonstrated its performance in elevational resolution improvement and deep vessel recovery in linear-detector-array-based PACT imaging[236].

In addition, DL can also be used to reduce noise in low-fluence light emitting diode (LED)-based PACT imaging. For example, Anas *et al.* designed a deep neural network consisting of a CNN and a recurrent neural network (RNN) to improve the SNR of noise-corrupted images in LED-based PACT imaging[219]. Hariri *et al.* proposed a multi-level wavelet CNN (MWCNN) model to enhance image contrast in low-fluence PACT imaging[220]. The above work demonstrates that DL-based postprocessing can enhance photoacoustic images in different
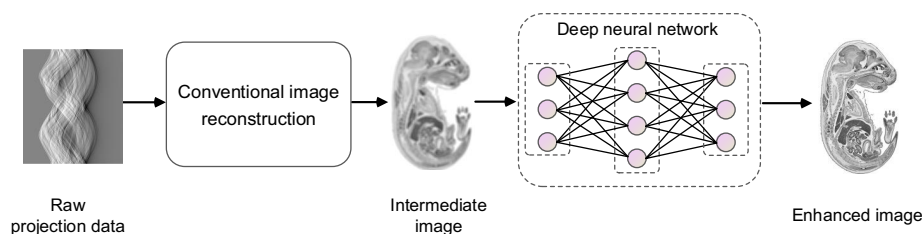

Raw projection data — Conventional image reconstruction — Intermediate image — Deep neural network — Enhanced image

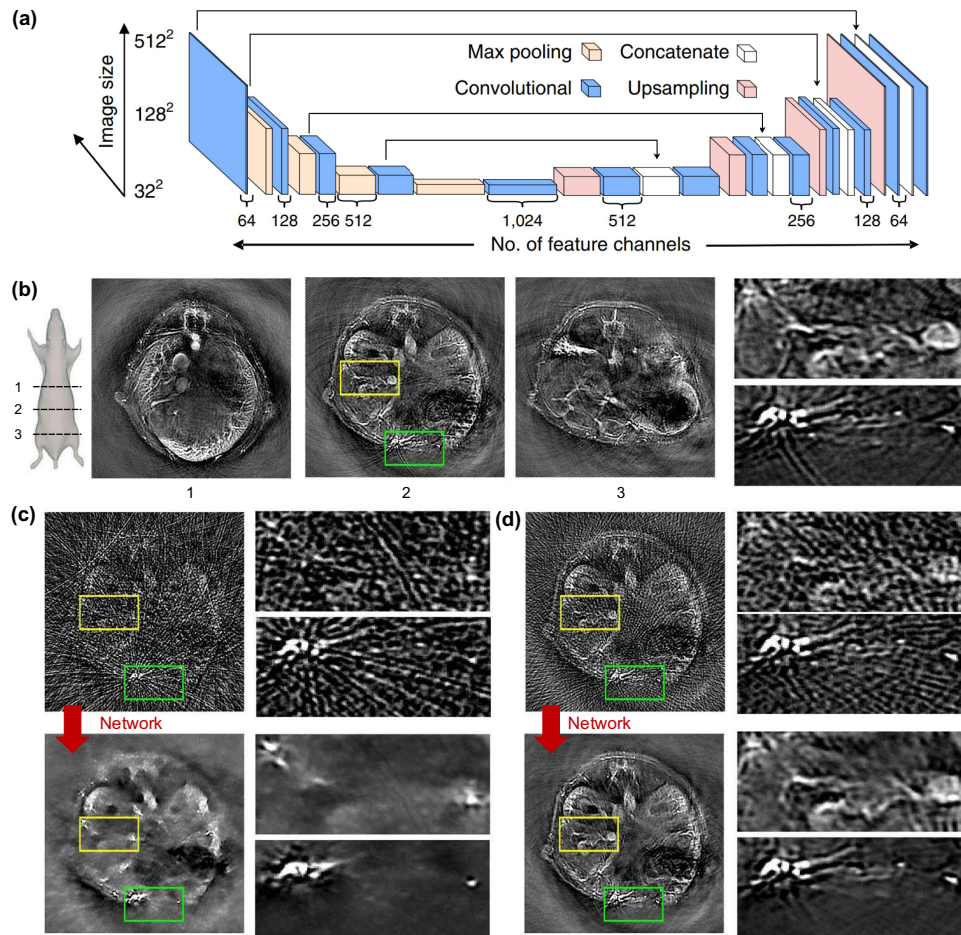**Fig. 29** DL-based image postprocessing in the image domain.

Fig. 30 DL-based postprocessing for image quality improvement in spare-view PACT. (a) Modified UNet for image postprocessing. (b) Reference images reconstructed using 512-channel projection data and their close-ups. (c) Images reconstructed with 32-channel projection data and their postprocessed versions by the modified UNet. (d) Images reconstructed with 128-channel projection data and their postprocessed versions by the modified UNet. Adapted from Ref. [46] with permission.

aspects, including artifact removal, contrast boost, resolution improvement, and aberration mitigation.

In addition to the above problems, DL-based postprocessing methods have also been used to address other problems in PACT, such as quantitative imaging[237–242], image fusion[243], image classification and segmentation[244–246], and band-frequency separation[247]. Table 10 summarizes representative work on DL-based image postprocessing in PACT.

4.3 Hybrid-Domain Processing

Data-domain preprocessing and image-domain postprocessing can only extract feature information from one domain. To achieve improved imaging results, hybrid-domain processing attempting to extract feature information from both the data domain and the image domain (Fig. 32) is also commonly used.

Based on this idea, Lan *et al*. proposed a hybrid deep neural network termed knowledge-infused GAN (Ki-GAN) for image enhancement in sparse-view PACT imaging[253]. The Ki-GAN

employs raw photoacoustic signals and DAS-reconstructed photoacoustic images as input and can produce images with better quality than conventional DAS-type algorithms and a UNet-based postprocessing method for both fully and sparsely sampled data. Based on this work, the authors further developed a new hybrid-domain DL model named Y-Net for image enhancement in limited-view PACT imaging[254]. The Y-Net consists of two encoders and one decoder. The two encoders are used to separately process raw photoacoustic signals and DAS-reconstructed photoacoustic images, while the decoder is employed to fuse the outputs of the two encoders to generate final images. Davoudi *et al*. developed a hybrid-domain CNN model using both time-resolved photoacoustic signals and reconstructed images as input to enhance the image quality in limited-view PACT imaging[255] and achieved excellent results as shown in Fig. 33. Moreover, Guo *et al*. proposed a hybrid-domain attention steered network (AS-Net) for image enhancement in sparse-view PACT imaging[256]. The AS-Net also takes raw photoacoustic signals and DAS-reconstructed images as inputs, but the photoacoustic signals are first transformed into
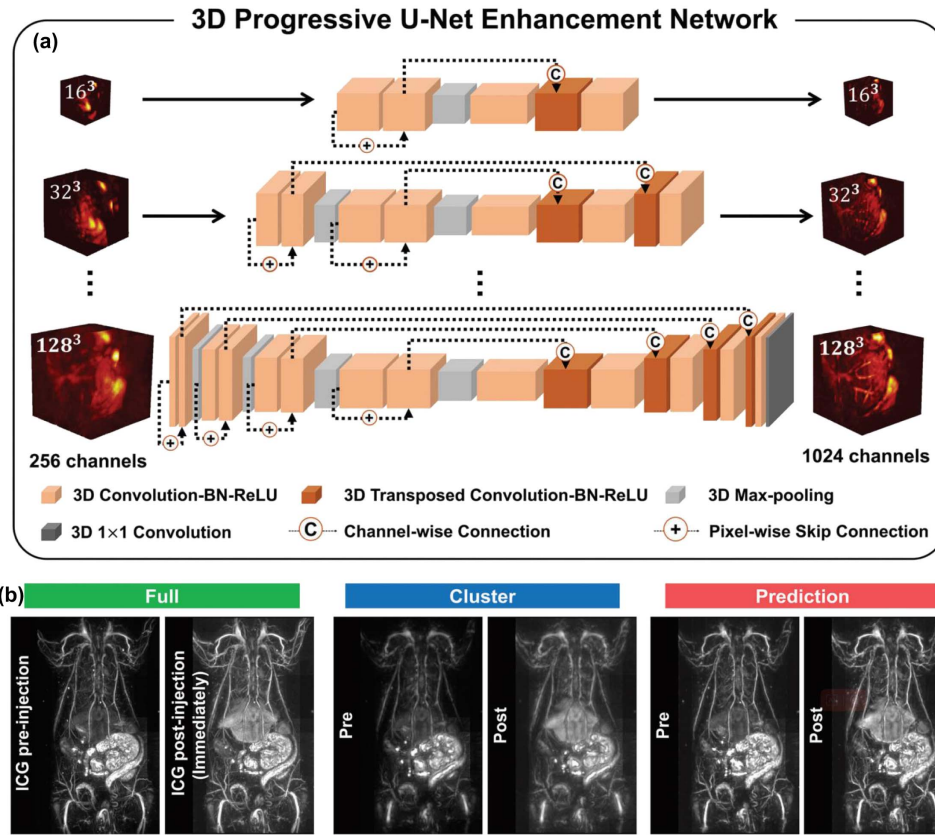
Fig. 31 DL-based postprocessing for imaging quality enhancement in limited-view 3D PACT. (a) Architecture of the 3D progressive UNet. (b) Reconstructed images in full-view imaging, cluster (limited) view imaging, and DL-enhanced cluster imaging. Adapted from Ref. [47] with permission.

Table 10 Representative Work on DL-Based Image Postprocessing in PACT

Author	Year	Problem	Network	Dataset	Source
		Image enhancement from incomplete projection data			
Choi *et al.*	2022	Limited-view imaging	3D progressive UNet	*In vivo*	[47]
Shahid *et al.*	2022	Sparse-view imaging	ResGAN	*In vivo* (public dataset)	[230]
Shahid *et al.*	2021	Sparse-view imaging	3-layer CNN, UNet, ResUNet	*In vivo*	[248]
Zhang *et al.*	2021	Sparse-view imaging	DuDoUNet	Simulation	[226]
Guan *et al.*	2021	Sparse-view and limited-view imaging	DD-UNet	Simulation	[227]
Godefroy *et al.*	2021	Limited-view and finite-bandwidth imaging	UNet	Simulation, phantom	[249]
Lu *et al.*	2021	Limited-view imaging	LV-GAN	Simulation, phantom, *in vivo* (test)	[58]
Lu *et al.*	2021	Sparse-view imaging	CycleGAN	Simulation, phantom, *in vivo*	[231]
Guan *et al.*	2020	Sparse-view imaging	FD-UNet	Simulation	[225]
Farnia *et al.*	2020	Sparse-view imaging	UNet	Simulation, *in vivo* (test)	[224]

(Table continued)

Vu *et al.*	2020	Limited-view and finite-bandwidth imaging	WGAN-GP	Simulation, phantom (test), *in vivo* (test)	[229]
Zhang *et al.*	2020	Sparse-view and limited-view imaging	RADL-Net(10-layer CNN)	Simulation, phantom (test), *in vivo* (test)	[250]
Antholzer *et al.*	2018	Sparse-view imaging	UNet	Simulation	[44]
Davoudi *et al.*	2019	Sparse-view and limited-view imaging	UNet	Simulation, phantom, *in vivo*	[46]

Inhomogeneous acoustic media

Gao *et al.*	2022	Thick porous media	UNet	Simulation, phantom, *ex vivo*	[234]
Jeon *et al.*	2020	SOS aberration	SegUNet	Simulation, phantom, *in vivo* (test)	[233]
Shan *et al.*	2019	Reflection artifact correction	UNet	Simulation	[232]

Resolution enhancement

Zheng *et al.*	2022	Elevational resolution enhancement	Deep-E (2D and 3D FD-UNet)	Simulation, phantom (test), *in vivo* (test)	[236]
Zhang *et al.*	2021	Elevational resolution enhancement	Deep-E (2D FD-UNet)	Simulation, phantom (test), *in vivo* (test)	[54]
Rajendran & Pramanik	2020	Tangential resolution enhancement	TARES (FD-UNet)	Simulation, phantom (test), *in vivo* (test)	[53]

Low-energy imaging

Hariri *et al.*	2020	Low-fluence imaging	Multi-level wavelet UNet	Phantom, *in vivo* (test)	[220]
Anas *et al.*	2018	LED imaging	RNN	Phantom, *in vivo* (test)	[219]

Image classification and segmentation

Lafci *et al.*	2021	Segmentation	UNet	*In vivo*	[246]
Chlis *et al.*	2020	Segmentation	Sparse UNet	*In vivo*	[245]
Zhang *et al.*	2019	Classification and segmentation	AlexNet and GoogLeNet	Simulation	[244]

Others

González *et al.*	2023	Band-frequency separation	FD-UNet	Simulation	[247]
Rajendran & Pramanik	2022	Imaging speed improvement	UNet	Simulation, phantom (test), *in vivo* (test)	[251]
Rajendran & Pramanik	2021	Radius calibration	RACOR-PAT(FD-UNet)	Simulation, phantom, *in vivo*	[252]
Awasthi *et al.*	2019	Image fusion	PA-Fuse (DeepFuse)	Simulation, phantom (test), *in vivo* (test)	[243]

a 3D square matrix before being fed into the network to ensure compatibility with the input of the AS-Net. The AS-Net adopts a self-attention mechanism for semantic feature extraction and fusion and is very efficient.

In addition to exploiting raw projection data directly, pre-processed projection data can also be used as the input of a hybrid-domain network. For example, Li *et al*. proposed a hybrid-domain CNN model called PRNet to improve the quality of sparse-view photoacoustic images[257]. The PRNet uses not only raw photoacoustic projection data but also their derivatives as input. Lan *et al*. proposed a joint feature fusion framework (JEFF-Net) to enhance photoacoustic images from limited-view projection data[258]. The JEFF-Net uses the sub-DAS images generated from each-channel raw projection data and the whole DAS images generated from all projection data as input, as shown in Fig. 34.

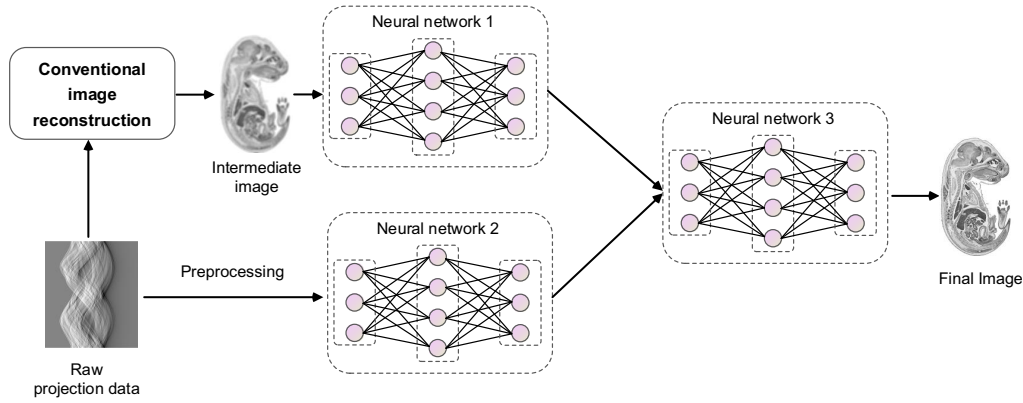Table 11 lists representative work on DL-based hybrid-domain processing in PACT.

Fig. 32 DL-based hybrid-domain processing. The neural networks 1 and 2 are used to extract feature information from the signal domain and the image domain, while the neural network 3 is used to fuse the outputs of the preceding two networks to generate the final images.

4.4 Learned Iterative Reconstruction

Conventional DL approaches generally employ deep neural networks as independent modules for signal preprocessing, image postprocessing, or hybrid processing. In contrast, learned iterative reconstruction (IR) algorithms attempt to fuse deep neural networks into an IR model to improve the quality and efficiency of image reconstruction. Specifically, learned IR algorithms employ deep neural networks to learn the regularization term or regularizer $R(\mathbf{x})$[259,260] or learn an entire iteration process[45,261], as shown in Fig. 35.

A regularizer is important for IR-based image reconstruction, especially for limited projection data. Conventional regularizers such as Tikhonov and TV may not necessarily be optimal. Data-driven DL-based regularizers were thus proposed to replace conventional regularizers. Li *et al.*[259] and Antholzer *et al.*[260] used a neural network to learn the Tikhonov regularizer to improve the image quality in sparse-view PACT imaging. However, the Tikhonov regularizer is trained before the IR loop and is not updated during iteration. Therefore, the learned Tikhonov regularizer may not be optimal for the image reconstruction problem in PACT. Wang *et al.* considered the entire IR process and proposed a learned IR model based on Eq. (120)[197], as shown in Fig. 36. The learned IR model takes the current reconstructed image $\mathbf{x}^k$ and the gradient term
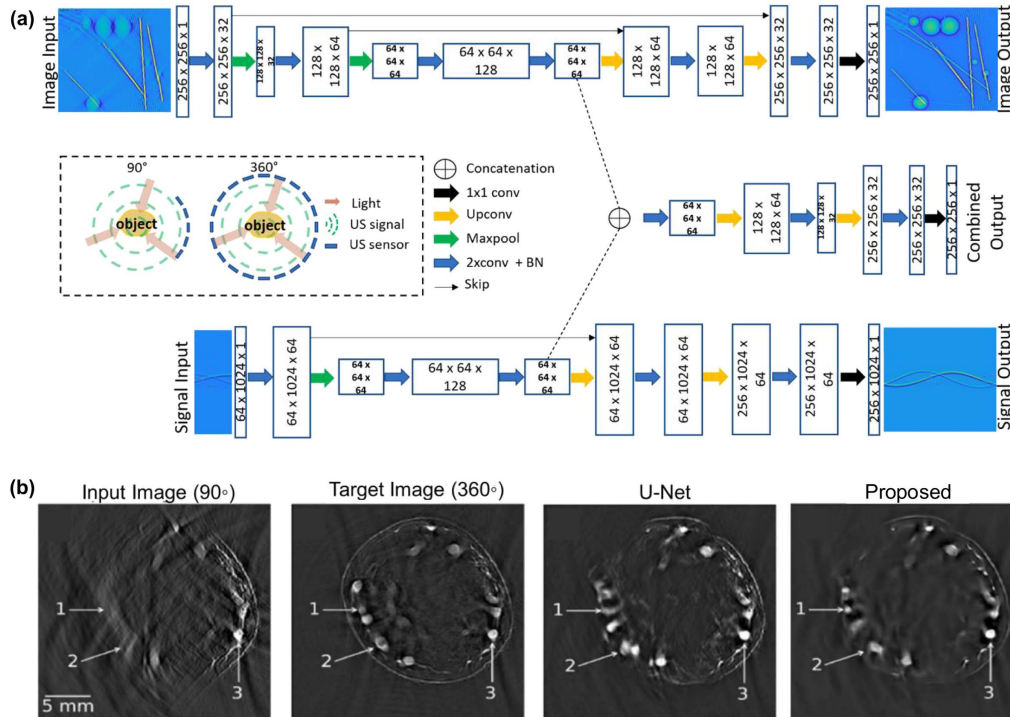
Fig. 33 DL-based hybrid-domain processing for enhancing the imaging quality in limited-view PACT imaging. (a) Architecture of the dual-domain CNN. (b) Reconstructed results of an *in vivo* human finger. Adapted from Ref. [255] with permission.
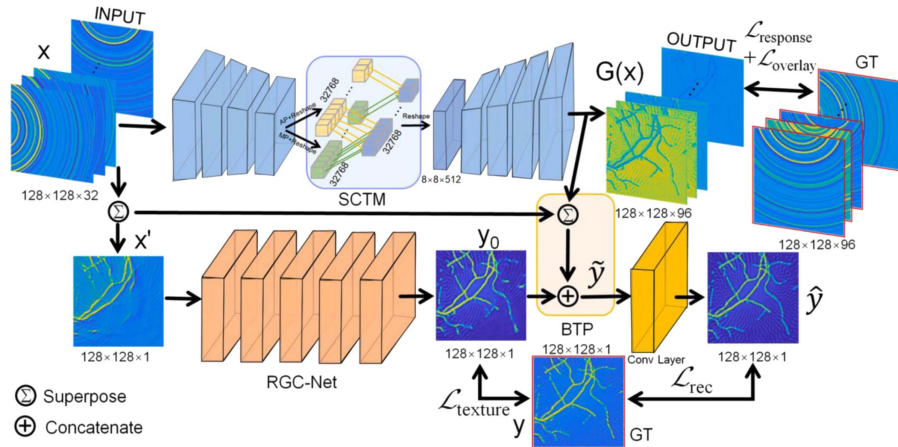
Fig. 34 JEFF-Net for image enhancement in limited-view PACT imaging. RGC-Net: residual global context subnetwork; SCTM: space-based calibration and transition module; GT: ground truth. Reprinted from Ref. [258] with permission.

Table 11 Representative Work on DL-Based Hybrid-Domain Processing in PACT

Author	Year	Problem	Network	Dataset	Source
Inputs: raw data + reconstructed image					
Guo *et al.*	2022	Sparse-view imaging	AS-Net	Simulation, *in vivo*	[256]
Davoudi *et al.*	2021	Limited-view imaging	CNN	*In vivo*	[255]
Lan *et al.*	2020	Limited-view imaging	Y-Net	Simulation, *ex vivo* (test), *in vivo* (test)	[254]
Lan *et al.*	2019	Sparse-view imaging	Ki-GAN	Simulation	[253]
Inputs: preprocessed raw data + reconstructed image					
Lan *et al.*	2023	Limited-view imaging	JEFF-Net	Simulation, *in vivo*	[258]
Li *et al.*	2021	Sparse-view imaging	PR-Net	Simulation, phantom	[257]

$\mathbf{A}^T(\mathbf{Ax}^k\mathbf{p})$ as inputs and updates the regularizer and the hyper-meters α and λ in each iteration. The experimental results show that the learned IR model is effective at improving the quality of sparse-view PACT images and has a faster convergence speed than conventional IR algorithms. Moreover, Lan *et al.* proposed a novel untrained CNN-based compressed sensing regularizer for sparse-view PACT imaging and achieved improved image quality compared with conventional IR algorithms[262].
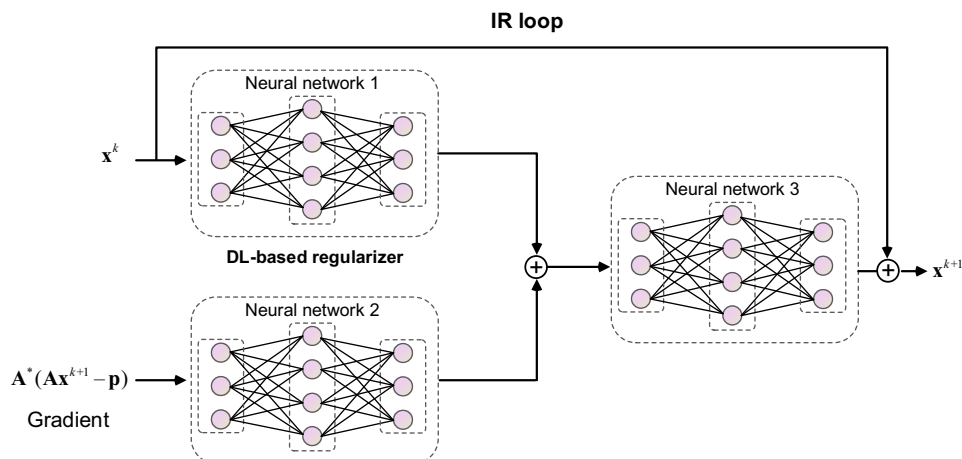
Fig. 35 Learned IR method in PACT.
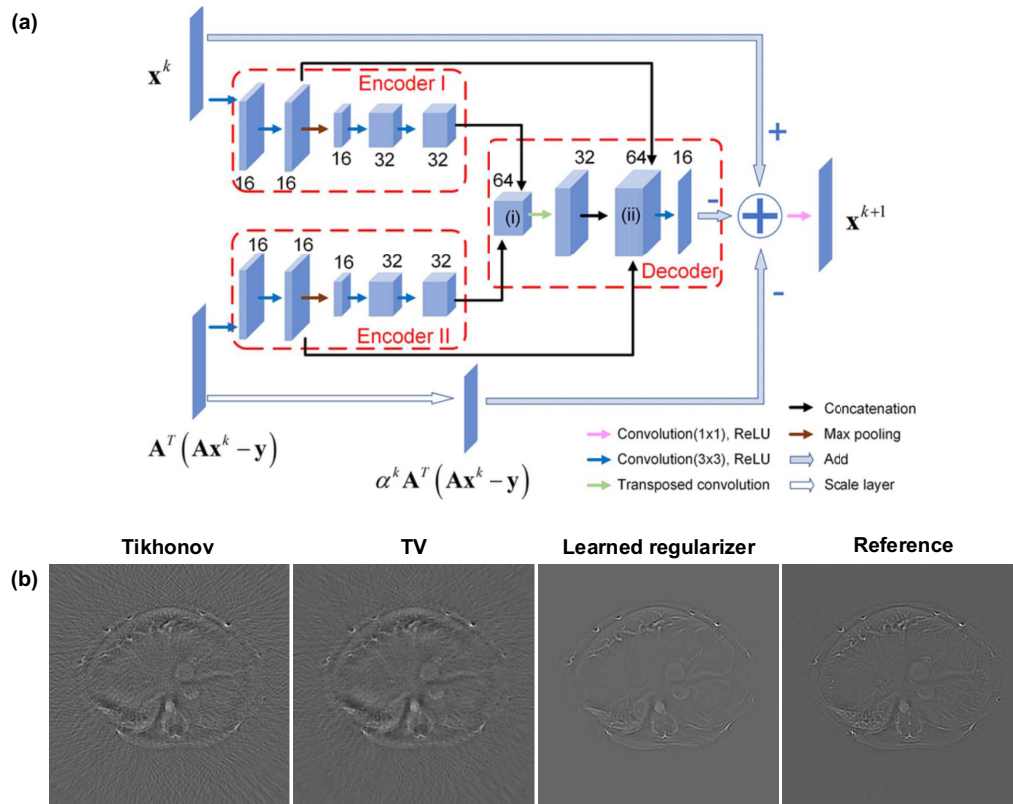
Fig. 36 Learning a regularizer. (a) Dual-path network for regularizer learning. (b) Cross-sectional images of a mouse reconstructed by IR with Tikhonov, TV, and learned regularizers. The learned regularizer has the highest image quality in terms of detail preservation and artifact suppression. Adapted from Ref. [197] with permission.

In addition to learning a regularizer, it is also possible to learn an entire iteration process [Eqs. (120) and (123)] using deep neural networks. For example, Hauptmann *et al.* proposed a deep gradient descent (DGD) algorithm to reconstruct high-resolution 3D images from restricted photoacoustic projection data in limited-view PACT imaging[45]. In the DGD algorithm, a CNN model was designed to learn the proximal operator in Eq. (123), which yields high-quality images at a fast convergence speed, as shown in Fig. 37. Similarly, Yang *et al.* proposed using recurrent inference machines (RIMs) to learn an iteration process in IR and achieved accelerated reconstruction[261]. Lan *et al.* developed a CNN network combining a variational model to learn an iteration process and achieved robust and accelerated reconstruction in limited-view PACT imaging[263]. Shan *et al.* proposed a simultaneous reconstruction network (SR-Net) to update the initial pressure **x** and the sound speed v_0 at each iteration of IR reconstruction for PACT imaging in acoustically heterogeneous media[264]. Moreover, Boink *et al.* proposed a learned primal-dual (L-PD) framework to learn an iteration process in IR and achieved simultaneous high-quality image reconstruction and segmentation in limited-view and sparse-view PACT imaging[265]. Recently, diffusion models in DL have attracted great attention in the field of biomedical image processing because of their powerful generative capabilities[266,267]. In PACT, some studies have incorporated diffusion models into iterative optimization frameworks, which show superior performance compared with conventional IR algorithms[268–271].

Generally, compared with conventional IR algorithms, learned IR methods can improve reconstruction efficiency by reducing the number of iterations. However, they are still time-consuming because the reconstruction model [Eqs. (120) and (123)] needs to be solved repeatedly. To accelerate the convergence of learned IR algorithms, Hsu *et al.* proposed a fast iterative reconstruction (FIRe) algorithm that simultaneously learns the discrete forward photoacoustic imaging model in Eq. (86) and an entire iteration process to reduce the reconstruction time[57]. Results showed that the FIRe algorithm can produce images with a quality comparable to that of learned IR and conventional IR algorithms but with reconstruction time reduced by nine and 620 times, respectively.

Table 12 lists representative work on DL-based IR in PACT.

4.5 Direct Reconstruction

In the aforementioned DL-based approaches, deep neural networks generally perform only certain functions, such as preprocessing, postprocessing, and regularizer learning, and cannot independently reconstruct images from raw projection data without the use of conventional algorithms. Nevertheless, deep neural networks can perform direct image reconstruction alone by learning the mapping relationship between raw photoacoustic projection data and reconstructed images (see Fig. 38).

In 2018, Waibel *et al.* used a UNet-like model to reconstruct images from limited-view photoacoustic projection data and demonstrated the feasibility of neural-network-based direct
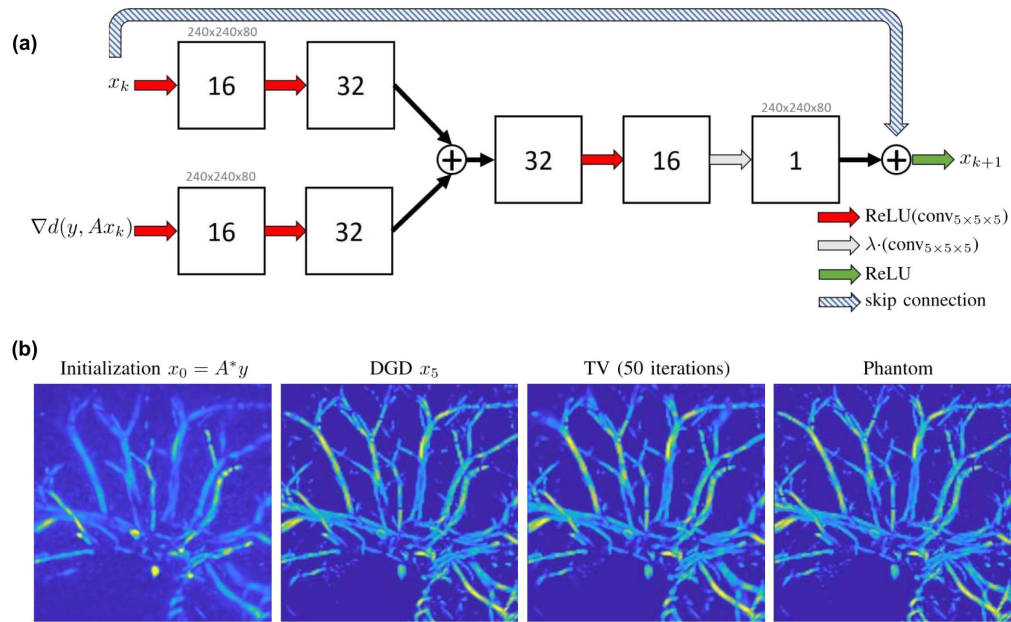
Fig. 37 Learning 3D IR based on DGD. (a) Structure of the CNN representing one iteration of DGD. (b) From left to right: initialization of the network, DGD result with five iterations, TV result with 50 iterations, and reference image. The proposed learned IR has a faster convergence speed than the conventional TV-based IR algorithm. Adapted from Ref. [45] with permission.

Table 12 Representative Work on DL-Based IR in PACT

| Author	Year	Problem	Network	Dataset	Source
Regularizer learning					
Wang *et al.*	2022	Sparse-view imaging	CNN	Simulation, *in vivo*	[197]
Lan *et al.*	2021	Sparse-view imaging	Untrained CNN (deep image prior)	Simulation (test), phantom (test), *in vivo* (test)	[262]
Antholzer *et al.*	2019	Sparse-view imaging	ResUNet	Simulation, phantom (test)	[260]
Li *et al.*	2018	Sparse-view imaging	Encoder-decoder	Simulation	[259]
Entire IR learning					
Hsu *et al.*	2023	Reconstruction acceleration	FIRe	Simulation	[57]
Lan *et al.*	2022	Limited-view imaging	CNN	Simulation, *in vivo* (public dataset)	[263]
Boink *et al.*	2020	Image reconstruction and segmentation	CNN	Simulation, phantom	[265]
Yang *et al.*	2019	Reconstruction acceleration	RIM	Simulation	[261]
Shan *et al.*	2019	Joint reconstruction	SR-Net	Simulation	[264]
Hauptmann *et al.*	2018	Limited view and acceleration	CNN	Simulation, phantom, *in vivo*	[45]
Diffusion model					
Guo *et al.*	2024	Limited-view imaging	UNet	Simulation, *in vivo*	[268]
Dey *et al.*	2024	Limited-view and sparse-view imaging	CNN	Simulation, *in vivo*	[270]
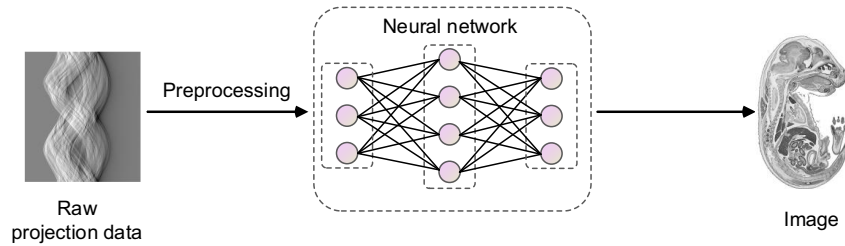
Fig. 38 DL-based direct image reconstruction in PACT.

image reconstruction through numerical simulation[272]. Lan *et al.* proposed a UNet-based deep neural network (DUNet) to reconstruct PACT images from multi-frequency photoacoustic projection data and demonstrated its superiority over conventional image reconstruction algorithms[273]. Feng *et al.* developed a residual-block-based end-to-end UNet (Res-UNet) to reconstruct PACT images from raw photoacoustic projection data and achieved high-quality image reconstruction with sharp edges and suppressed artifacts[274]. Lan *et al.* designed an encoder-decoder network to transform superimposed four-channel photoacoustic projection data into reconstructed images and achieved real-time PACT imaging with a single-channel data acquisition (DAQ) system[275]. Recently, Shen *et al.* developed a physics-driven DL-based filtered back projection (dFBP) framework for direct image reconstruction from raw projection

data[276]. The dFBP network is constructed based on the physical model of the analytical FBP algorithm [Eq. (47)] and consists of a filtering module, a back-projection module, and a fusion module, as shown in Fig. 39. The proposed dFBP is robust and flexible, can achieve direct signal-to-image transformation with enhanced accuracy, and reconstruct high-quality, artifact-suppressed images from sparse-view, limited-view, and acoustic heterogeneity-contaminated projection data. Moreover, Lan *et al.* designed a self-supervised deep-learning framework for image reconstruction from extremely sparse projection data. The success of the network can be mainly attributed to the adoption of a transformer-based self-attention mechanism[277].

To reduce the complexity of network training, raw photo-acoustic projection data can be preprocessed to produce high-level features before being fed into an image reconstruction
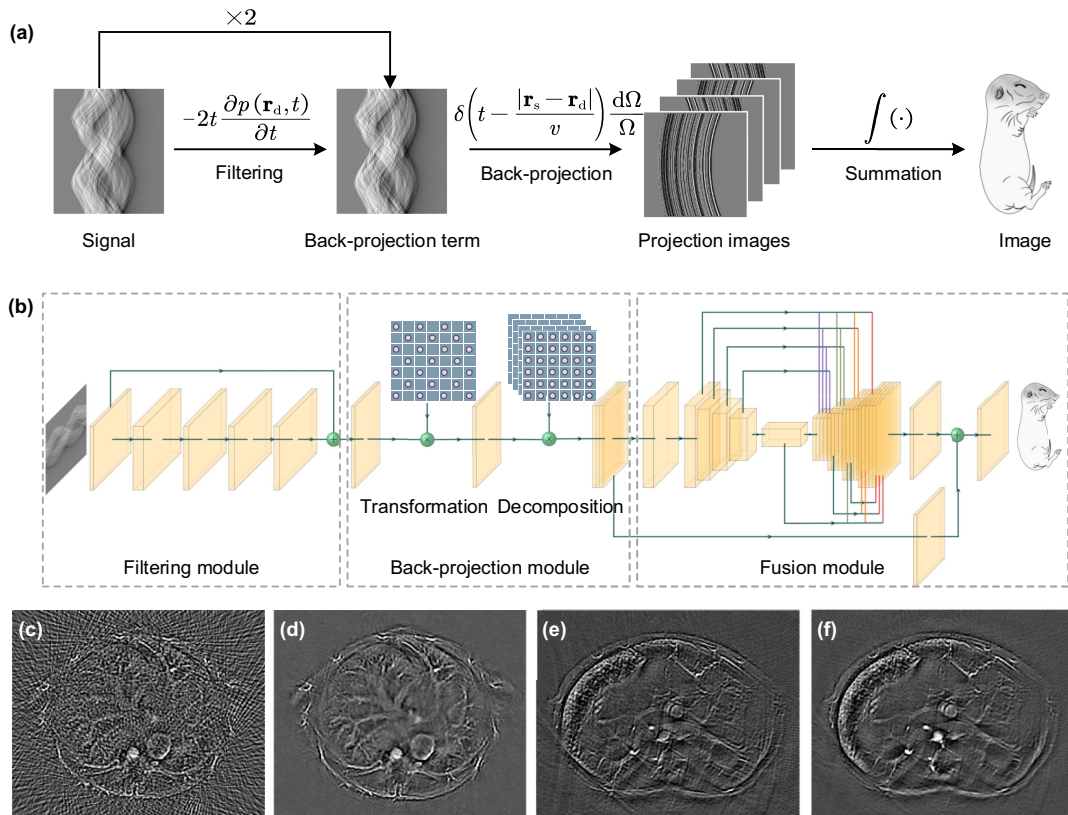

Fig. 39 Direct image reconstruction using dFBP. (a) Physical model of the analytical FBP [Eq. (47)]. (b) Architecture of the dFBP, which consists of a filtering module, a back-projection module, and a fusion module. (c), (d) Images separately reconstructed by FBP and dFBP using 128-channel photoacoustic projection data. (e), (f) Images separately reconstructed by FBP and dFBP from limited-view photoacoustic projection data (view angle: $3\pi/4$). Reprinted from Ref. [276] with permission.
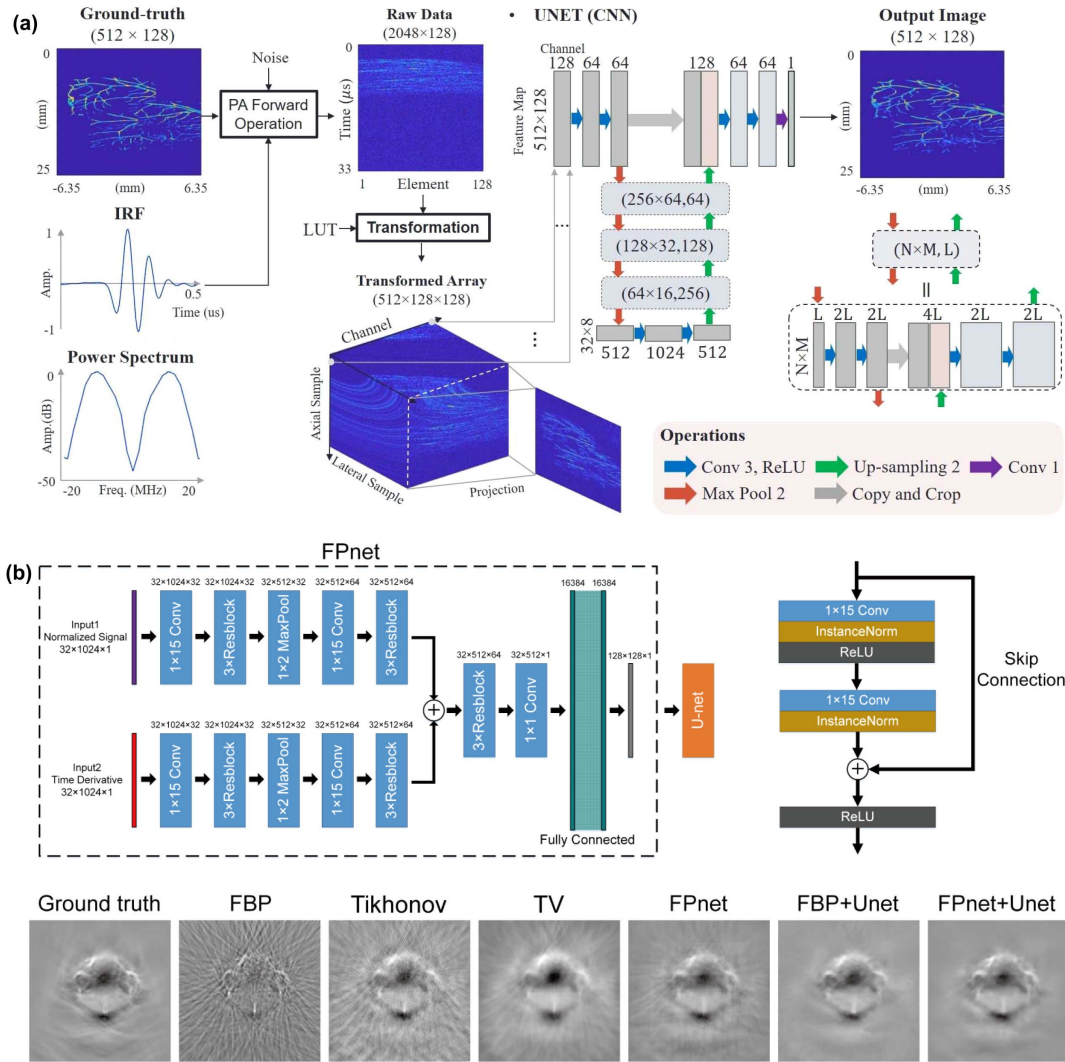
Fig. 40 DL-based direct image reconstruction using high-level features extracted from raw photoacoustic projection data. (a) Architecture of the upgUNet proposed by Kim and coworkers. Reprinted from Ref. [279] with permission. (b) Architecture of the FPNet + UNet proposed by Tong and coworkers (top) and representative reconstruction results (bottom). Adapted from Ref. [280] with permission.

network. Guan *et al*. proposed UNet-based pixel-wise deep learning (Pixel-DL) that uses pixel-interpolated data as input and realized direct image reconstruction in limited-view and sparse-view PACT imaging[278]. Similarly, Kim *et al*. developed a deep convolutional network (upgUNet), which takes multi-channel features extracted from raw photoacoustic projection data as input, and achieved direct image reconstruction in limited-view PACT imaging[279], as shown in Fig. 40(a). Inspired by the principle of FBP, Tong *et al*. designed a feature projection network (FPNet) that takes raw photoacoustic projection data and their first-order derivative as input [Fig. 40(b)] and achieved high reconstruction quality from limited-view data with sparse measurements[280]. Recently, Dehner *et al*. proposed a deep neural network named DeepMB for real-time IR image reconstruction with adjustable SOS[281,282]. DeepMB learns conventional IR algorithms using domain-transformed projection data as input and can reconstruct high-quality images 3000 times (less than 10 ms per image) faster than conventional IR algorithms.

Table 13 lists representative work on DL-based direct image reconstruction in PACT.

5 Performance Comparisons

Thus far, we have reviewed the principles of the six classes of image reconstruction algorithms in PACT, namely, DAS, FBP, SE, TR, IR, and DL. To choose the most appropriate algorithm in practice, it is necessary to understand the characteristics and performance of each method under different imaging scenarios. In this section, we discuss and compare the six classes of algorithms, focusing on image reconstruction quality and image reconstruction speed. Table 14 first summarizes the overall characteristics and performance of each algorithm when evaluated under different circumstances, which will be discussed in detail below.

Table 13 Representative Work on DL-Based Direct Signal-to-Image Reconstruction in PACT

Author	Year	Problem	Network	Dataset	Source
Input: raw data					
Shen *et al.*	2024	Sparse-view/limited-view imaging	Learned FBP	Simulation, *in vivo*	[276]
Lan *et al.*	2023	Channel data decoupling	Encoder-decoder	Simulation, *in vivo*	[275]
Feng *et al.*	2020	Direct image reconstruction	Res-UNet	Simulation, phantom (test)	[274]
Lan *et al.*	2019	Multi-frequency image reconstruction	DUNet	Simulation	[273]
Allman *et al.*	2018	Image reconstruction with source detection and reflection artifacts removal	Deep fully convolutional network + Fast R-CNN	Simulation, phantom	[283]
Waibel *et al.*	2018	Limited-view imaging	UNet	Simulation	[272]
Input: high-level features extracted from raw data					
Dehner *et al.*	2022	IR algorithm acceleration	DeepMB	Simulation, *in vivo* (test)	[281]
Guan *et al.*	2020	Sparse-view and limited-view imaging	FD-UNet	Simulation	[278]
Kim *et al.*	2020	Limited-view imaging	upgUNet	Simulation, phantom (test), *in vivo* (test)	[279]
Tong *et al.*	2020	Sparse-view & limited-view imaging	FPNet +UNet	Simulation, phantom, *in vivo*	[280]

5.1 Image Reconstruction Quality under Various Imaging Scenarios

5.1.1 Ideal imaging conditions

To achieve perfect image reconstruction or acoustic inversion in PACT, photoacoustic signal detection should be performed under ideal conditions: (1) the ultrasound detectors used for signal detection should have an infinite bandwidth; (2) the ultrasound detectors should have a point-like aperture; (3) the ultrasound detectors should be densely arranged in space; (4) the detector array formed by individual detectors should have a full view angle with respect to a sample; and (5) the

acoustic media should be homogeneous. When these conditions are satisfied, all six classes of image reconstruction algorithms are expected to produce high-quality images.

However, practical photoacoustic signal detection is unlikely to be ideal. In the following section, we discuss and compare the performances of the image reconstruction algorithms under non-ideal imaging scenarios.

5.1.2 Limited detector bandwidth

The ultrasound detectors used for photoacoustic signal detection should ideally have an infinite bandwidth so that they can

Table 14 Comparison of Different Image Reconstruction Algorithms in PACT[a]

Circumstance	DAS	FBP	SE	TR	IR	DL
Detector EIR modeling[b]	Difficult	Difficult	Difficult	Difficult	OK	OK
Detector SIR modeling[c]	Difficult	Difficult	Difficult	Difficult	OK	OK
Performance in sparse-view imaging	Poor	Poor	Poor	Poor	Good	Excellent
Performance in limited-view sampling	Poor	Poor	Poor	Poor	Good	Excellent
Media property coupling	Difficult	Difficult	Difficult	OK	OK	OK
Speed	Fast	Fast	Very fast	Slow	Very slow	Fast
Memory footprint	Very low	Very low	Low	High	Very high	High or very high

[a]The comparisons in this table are based on typical situations. The variants of some algorithms may have distinct characteristics. For example, modified FBP algorithms may be able to incorporate the SIR of a detector[284,285], modified SE algorithms can be used for image reconstruction in heterogeneous media[167], and iterative DL may be slow[45].

[b]The DAS, FBP, SE, and TR algorithms are generally not ready to incorporate the non-ideal EIR of a detector. However, the raw photoacoustic signals measured by a detector can be first deconvolved to correct the non-ideal EIR and then fed into the algorithms to obtain good results.

[c]The DAS, FBP, SE, and TR algorithms are generally not ready to incorporate the non-ideal SIR of a detector. However, the images reconstructed by these algorithms can be deconvolved to correct the non-ideal SIR.

respond to all frequency contents of a signal[4]. However, a practical detector always has a limited bandwidth and a non-ideal EIR, which distorts measured photoacoustic signals and blurs reconstructed images. To deal with this problem, a reconstruction algorithm should consider the non-ideal EIR. The six classes of algorithms reviewed in this paper have different characteristics in addressing this problem. It is generally difficult for DAS, FBP, SE, and TR algorithms to incorporate the non-ideal EIR of a detector into their reconstruction models but easy for IR- and DL-based algorithms, as summarized in Table 14. Specifically, IR algorithms can easily solve the non-ideal EIR problem by incorporating it into the system matrix $\mathbf{A}$ [Eq. (110)], and DL algorithms can solve the problem by training a reconstruction network with EIR-corrected signal-image datasets.

To understand how different image reconstruction algorithms behave in the case of non-ideal EIR, an example is given in Fig. 41. In this example, the photoacoustic source is a numerical blood vessel phantom distributed in the *x–y* plane [Fig. 41(a)]. The detector array used for imaging is a full ring with a diameter of 50 mm and has 512 evenly distributed elements, each of which has a point-like shape. To simulate the limited bandwidth case, the detector array is assigned a center frequency of 2 MHz and is assumed to have a Gaussian-like EIR with a fractional bandwidth of 80%. Figures 41(b)–41(d) show the images reconstructed by the FBP-, TR-, and EIR-corrected IR algorithms, respectively. Figures 41(e)–41(g) are the images of Figs. 41(b)–41(d) with negative values removed (negative values in a photoacoustic image have no physical meaning). The results show that the images reconstructed by FBP and TR have apparent artifacts, while the images reconstructed by IR have much fewer artifacts due to the correction of the detector EIR. Note that since DAS and SE typically have similar performance to FBP and are commonly adopted in PACT systems with linear and planar detector arrays, they are not compared in this example.
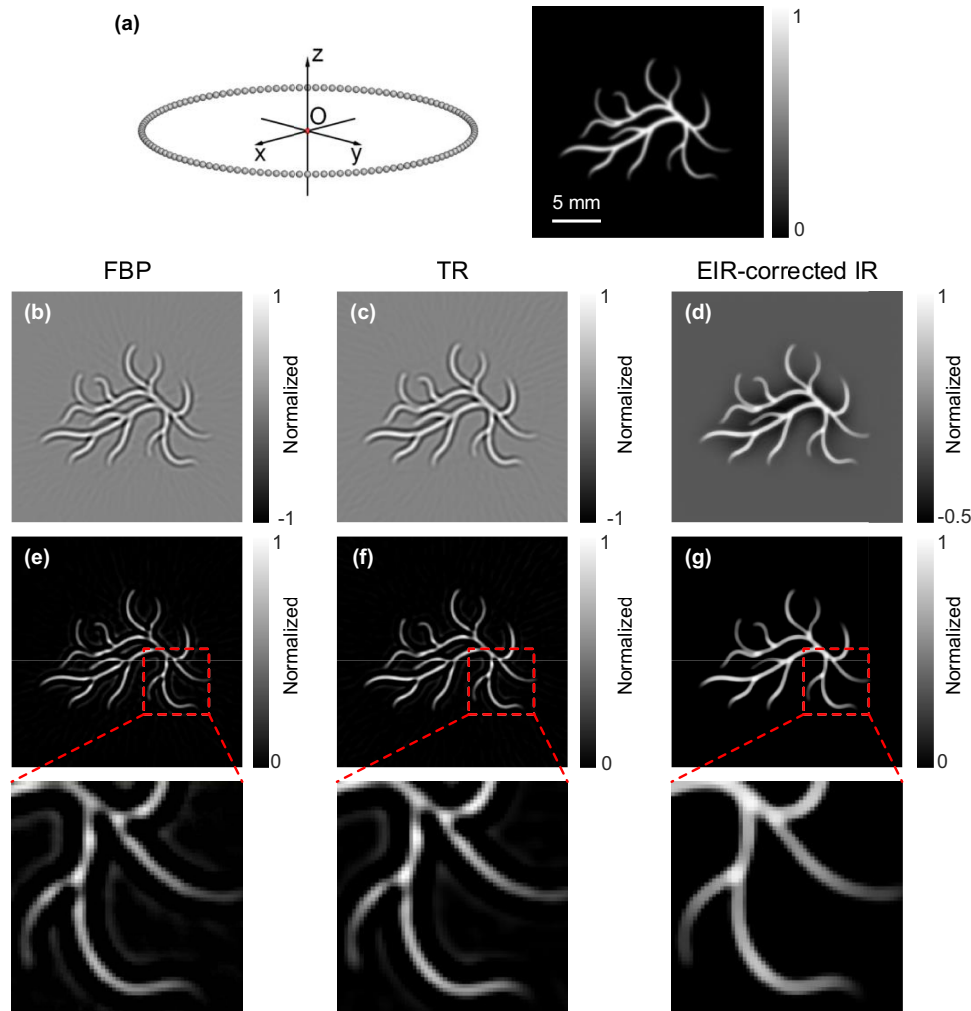

Fig. 41 Performance comparison of FBP, TR, and IR when the detectors used for signal detection have a limited bandwidth. (a) A full-ring detector array in 3D space and a numerical blood vessel phantom used for simulation. (b)–(d) Images reconstructed by FBP, TR, and EIR-corrected IR algorithms, respectively. (e)–(g) Images in (b)–(d) with negative values removed and close-up views of the images in the red dashed box. In this example, the detector array is assumed to have a Gaussian-like EIR with a bandwidth of 80%. Other simulation settings can be found in the text.
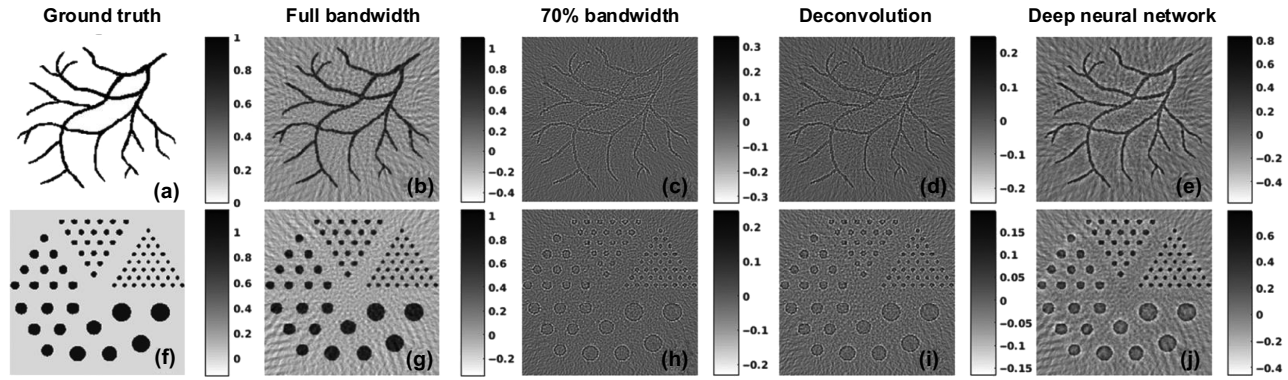
Fig. 42 Signal preprocessing helps correct the non-ideal EIR of a detector and improves image reconstruction quality. First row: (a) numerical blood vessel used for the test. (b), (c) Images reconstructed using full and limited-bandwidth photoacoustic signals, respectively. (d), (e) Images reconstructed using signals preprocessed by deconvolution and a deep neural network, respectively. Second row: (f)–(j) a Derenzo phantom and corresponding results. Detailed simulation settings can be found in the text. Adapted from Ref. [51] with permission.

Although DAS, FBP, SE, and TR algorithms are generally not ready to incorporate the EIR of a practical detector, their input signals can be preprocessed by deconvolution or DL to correct the non-ideal EIR. In deconvolution, the EIR of a detector is first experimentally measured using a point object and then incorporated into the signal detection model [Eq. (23)] for correction. In DL-based preprocessing, the EIR of a detector can be compensated for by training a neural network using EIR-corrected photoacoustic signals. Figure 42 shows an example demonstrating that signal preprocessing helps correct the non-ideal EIR of a detector and improves image reconstruction results[51]. In this example, the photoacoustic source is a numerical blood vessel, as shown in Fig. 42(a). Photoacoustic signals were measured by a full-ring detector array consisting of 100 elements with a radius of 37.02 mm, a center frequency of 2.25 MHz, and a fractional bandwidth of 70%. Figures 42(b) and 42(c) are images reconstructed using full and limited-bandwidth photoacoustic signals, respectively. Figures 42(d) and 42(e) are images reconstructed using signals separately preprocessed by deconvolution and a deep neural network, which were both designed to compensate for the EIR effect of the detector array. Figures 42(f)–42(j) show the corresponding results for a Derenzo phantom. The simulation results show that deconvolution and DL-based signal preprocessing can effectively correct non-ideal EIRs and improve image reconstruction quality.

5.1.3 Finite detector aperture size

In addition to having an infinite bandwidth, an ideal ultrasound detector used for photoacoustic signal detection should also have a point-like shape so that it can respond to photoacoustic signals from all directions[4]. However, a practical detector always has a finite aperture size and a non-ideal SIR, which distorts measured photoacoustic signals and blurs reconstructed images. To address this problem, a reconstruction algorithm should consider the non-ideal SIR. The six classes of algorithms reviewed in this paper have different characteristics in addressing this problem. It is generally difficult for DAS, FBP, SE, and TR algorithms to incorporate the non-ideal SIR of a detector into their reconstruction models but easy for IR- and DL-based algorithms, as summarized in Table 14. Specifically,

IR algorithms can easily solve the non-ideal SIR problem by incorporating it into the system matrix **A** [Eq. (93)], and DL algorithms can solve this problem by training a reconstruction neural network with SIR-corrected signal-image datasets.

To understand how different image reconstruction algorithms behave in the case of non-ideal SIRs, a simulation from[286] is adapted and shown in Fig. 43. In this example, the photoacoustic source has six-point absorbers located at different distances from the origin. Photoacoustic signals were recorded by a detector with a 6-mm-diameter aperture size and a center frequency of 5 MHz. The detector rotated around the photoacoustic source to record signals at different positions. Figures 43(a) and 43(b) show the images reconstructed by a model-based algorithm and DAS, respectively. The model-based algorithm uses a similar discrete imaging model [Eq. (86)] as IR and couples the SIR of the detector in its system matrix. The results show
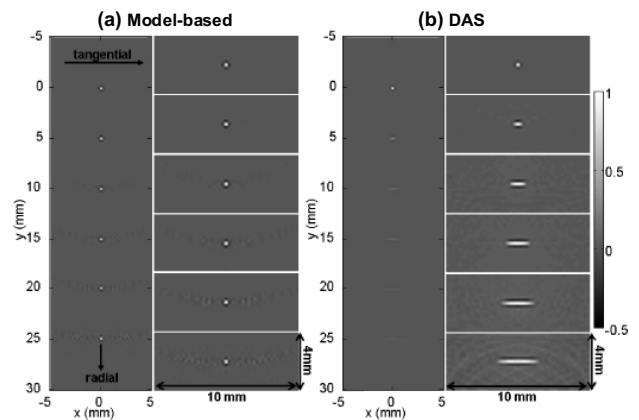

Fig. 43 Performance comparison of a model-based algorithm and DAS when the detector used for signal detection has a finite aperture size. (a), (b) Images reconstructed by the model-based algorithm and DAS, respectively. The model-based algorithm uses a similar discrete imaging model [Eq. (86)] as IR but couples the detector SIR in its system matrix. Reprinted from Ref. [286] with permission.
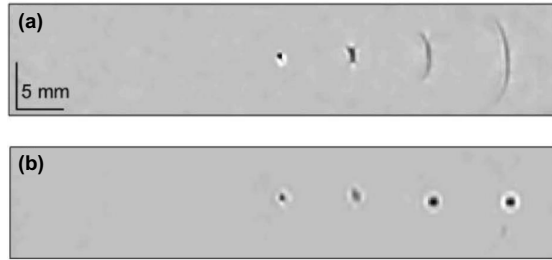
Fig. 44 Image postprocessing helps correct the non-ideal SIR of a detector and improves image quality. In this example, a detector with a finite aperture size rotates around the four-point sources for signal detection. (a), (b) Images reconstructed by DAS without and with DL-based postprocessing, respectively. Adapted from Ref. [53] with permission.

that the image reconstructed by DAS has significant distortions while the image recovered by the model-based algorithm has much fewer distortions due to the correction of the detector SIR.

Although DAS, FBP, SE, and TR algorithms are generally not ready to incorporate the SIR of a practical detector, their output images can be postprocessed by deconvolution or DL to correct the non-ideal SIR. In deconvolution, the SIR of a detector is first modeled using linear acoustics[39,287,288] and then incorporated into a non-blind image deconvolution model for SIR correction. Alternatively, the SIR can be unknown and a blind deconvolution algorithm is employed to correct the non-ideal SIR. In DL-based postprocessing, the SIR of a detector can be compensated for by training a neural network using SIR-corrected photoacoustic images. Figure 44 shows an example demonstrating that DL-based image postprocessing helps correct the non-ideal SIR of a detector and improves image quality[53]. In this example, the photoacoustic source contains four-point targets. The detector used for signal detection has a finite aperture size and a center frequency of 2.25 MHz. The detector rotates around the photoacoustic source to measure signals at different positions. Figures 44(a) and 44(b) are images reconstructed using DAS without and with DL-based image postprocessing, respectively. The results show that the DAS-reconstructed image without postprocessing has significant distortions, while the image with postprocessing has much fewer distortions. This indicates that DL-based image postprocessing can effectively correct non-ideal SIRs and improve image reconstruction quality.

Although here we independently discuss the EIR and SIR of a detector, they are never separable and together constitute the total impulse response (TIR) of a detector. To address the TIR of a detector, methods similar to those used for EIR and SIR correction can be used. In other words, the TIR of a detector can be experimentally measured and incorporated into IR or DL models for improved image reconstruction[289]. Alternatively, non-blind deconvolution, blind deconvolution, and DL-based image postprocessing techniques can also be adopted to correct the non-ideal TIR of a detector[290–292].

5.1.4 Sparse-view imaging

Ideally, the ultrasound detectors used for photoacoustic signal detection should be densely arranged in space to satisfy the spatial Nyquist sampling theorem[4]. However, the detectors in a practical detector array are often sparse due to high fabrication

costs, leading to the ill-posed problem of image reconstruction from sparse projection data. The six classes of algorithms reviewed in this paper have different performances when dealing with this problem. Generally, analytical algorithms such as DAS, FBP, SE, and TR have similar performance and require more projection data to reconstruct an image with reasonable image quality than do IR algorithms, which solve the image reconstruction problem in the sense of least squares; IR algorithms require more projection data than DL-based algorithms, which are data-driven and can realize complicated signal-to-image mapping using sparse projection data. In other words, DL-based algorithms have better performance than IR algorithms in sparse-view PACT imaging, and IR algorithms have better performance than DAS, FBP, SE, and TR algorithms, as summarized in Table 14.

To evaluate the performance of different image reconstruction algorithms in sparse-view imaging, an example is presented in Fig. 45. The simulation settings in this example are similar to those in Fig. 41 except that the circular detector array has a varying number of elements from 32 to 512 and all detector elements have infinite bandwidth. Figure 45 shows the images reconstructed by the FBP-, TR-, and TV-regularized IR algorithms using 32-, 128-, and 512-channel projection data. As expected, the image reconstruction quality of each algorithm improves with the increase of detector number. However, FBP and TR suffer from streak artifacts more significantly than TV-regularized IR in sparse-view cases (e.g., 32 and 128 views), although the three algorithms yield similar image quality when the detection views are dense (e.g., 512). Since DAS and SE typically have similar performance to FBP and are commonly adopted in PACT systems with linear and planar detector arrays, they are not compared in this example.

Compared with IR algorithms, DL-based algorithms may have better image reconstruction performance in sparse-view imaging[276,293]. To illustrate this, Fig. 46 shows an example comparing the performance of conventional algorithms such as TR, back projection, and TV-based IR, and two DL-based image reconstruction algorithms, namely, Pixel-DL[278] and model-based learning (MBLr)[45], for image reconstruction using sparsely sampled projection data[293]. In this example, the imaging target is mouse cerebral vasculature [Fig. 46(a)], and photoacoustic signals were received by an 8-mm-diameter full-ring array with 32 evenly distributed detectors. Figures 46(b)–46(f) are the images reconstructed by TR, back projection, TV-regularized IR, Pixel-DL, and MBLr, respectively. The results show that the TV-regularized IR algorithm has higher image reconstruction quality than the TR and back-projection algorithms, which suffer from severe artifacts due to the sparseness of the projection data, while the two DL-based reconstruction methods, especially MBLr, yield improved quality compared with all conventional algorithms in this case.

Although conventional image reconstruction algorithms suffer from artifacts in sparse-view imaging, they can be improved[221,294–296]. For example, Sandbichler *et al.* achieved enhanced image reconstruction using conventional algorithms by transforming sparse projection data into dense data using compressed sensing in sparse-view imaging[297]. Meng *et al.* developed a principal component analysis (PCA)-based method and achieved high-quality 3D image reconstruction with sparsely sampled data without involving an iterative process[298]. Hu *et al.* analyzed the aliasing problem caused by spatial undersampling and proposed a temporal low-pass-filtering and spatial
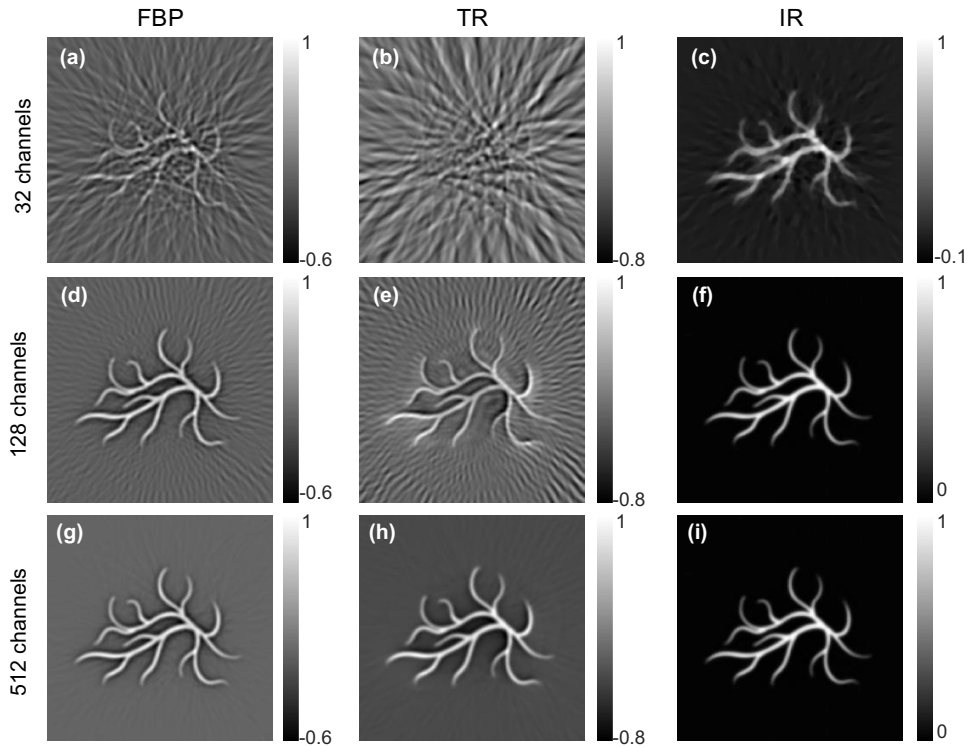
Fig. 45 Performance comparison of FBP, TR, and IR in sparse-view PACT imaging. First–third columns: images reconstructed by FBP, TR, and TV-regularized IR, respectively. First–third row: images reconstructed using (a)–(c) 32-, (d)–(f) 128-, and (g)–(i) 512-channel projection data, respectively. The simulation settings can be found in the text.

interpolation method for aliasing mitigation and artifact suppression[299,300]. Cai *et al.* analyzed the relationship between streak artifacts and sparse projection data and developed an adaptive FBP algorithm named contamination-tracing back-projection (CTBP) for image artifact suppression in sparse-view imaging[301]. Hakakzadeh *et al.* proposed a spatial-domain factor for sparse sampling circular-view PACT and achieved artifact suppression and resolution improvement compared with conventional algorithms[302]. Moreover, Wang *et al.* proposed an iterative scheme combining virtually parallel projecting and spatially adaptive filtering and achieved enhanced image reconstruction in sparse-view PACT imaging[303].
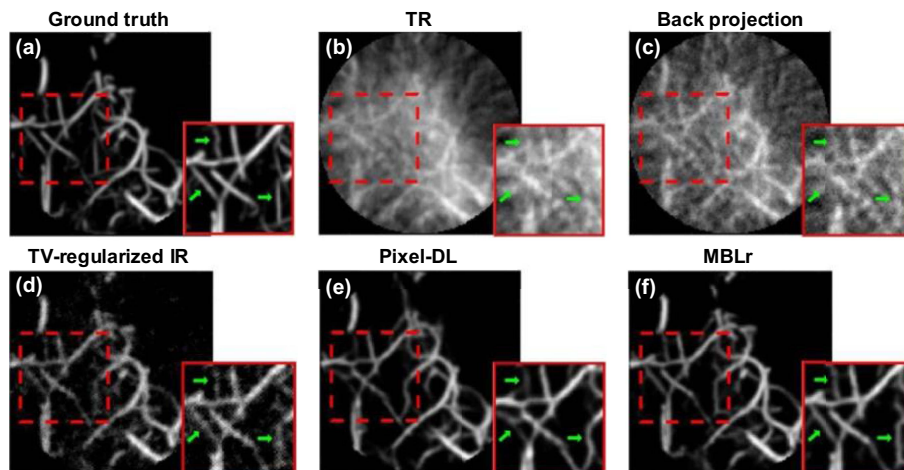

Fig. 46 Performance comparison of DL and conventional image reconstruction algorithms in sparse-view PACT imaging. (a) Mouse cerebral vasculature and local close-up image. (b)–(f) Images reconstructed by TR, back projection, TV-regularized IR, Pixel-DL, and MBLr, respectively. DL-based algorithms (MBLr and Pixel-DL) are superior to conventional algorithms in this case. The simulation settings can be found in the text. Adapted from Ref. [293] with permission.

5.1.5 Limited-view imaging

Ideally, the detector array used for PACT imaging should have a full view angle (4π steradians in 3D space) with respect to the sample being imaged[4]. However, a practical detector array always has a limited view angle, leading to the ill-posed problem of image reconstruction from limited-view projection data. The six classes of algorithms reviewed in this paper have different performances when dealing with this problem. Generally, analytical algorithms such as DAS, FBP, SE, and TR have similar performance and require a larger view angle to reconstruct an image with reasonable image quality than do IR algorithms, which solve the image reconstruction problem in the sense of least squares; IR algorithms require a larger view angle than DL-based algorithms, which are data-driven and can realize complicated signal-to-image mapping using limited-view projection data. In other words, DL-based algorithms have better performance than IR algorithms in limited-view PACT imaging and IR algorithms have better performance than DAS, FBP, SE, and TR algorithms, as summarized in Table 14.

To evaluate the performance of different image reconstruction algorithms in limited-view imaging, an example is presented in Fig. 47. The simulation settings in this example are similar to those used in Fig. 41 except that the circular detector arrays used for imaging have limited view angles ranging from $\pi/2$ (quarter circle) to 2π (full circle), and all detector elements have an infinite bandwidth. Figure 47 shows the images reconstructed by the FBP, TR, and TV-regularized IR algorithms when the view angles of the detector arrays are $\pi/2$, π, and 2π. As expected, the image reconstruction quality of each algorithm improves with the increase of the view angles of the detector arrays. However, FBP and TR suffer from artifacts more significantly than does TV-regularized IR in limited-view cases (e.g., $\pi/2$ and π), although the three algorithms yield similar image quality when the view angle is 2π.

Compared with conventional algorithms, DL-based methods may have better image reconstruction performance in limited-view imaging[276]. To illustrate this, Fig. 48 shows a simulation comparing the performance of FBP and a DL-based FBP algorithm (dFBP) for image reconstruction using limited-view projection data[276]. In this example, the imaging target is a numerical zebrafish, and photoacoustic signals are collected by a circular detector array, which has a diameter of 80 mm
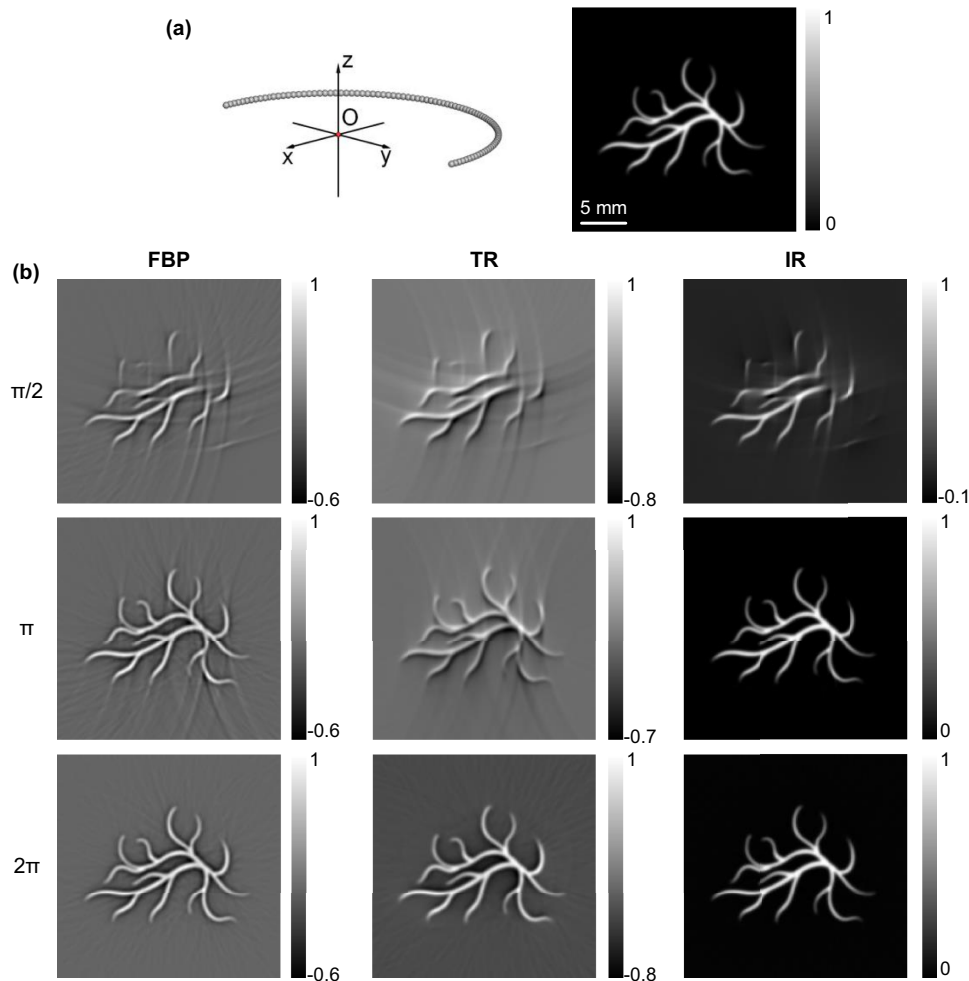

Fig. 47 Performance comparison of FBP, TR, and TV-regularized IR in limited-view PACT imaging. (a) A partial-ring detector array in 3D space and a numerical blood vessel phantom used for simulation. (b) Images reconstructed by FBP, TR, and TV-regularized IR under different imaging angles. The simulation settings can be found in the text.
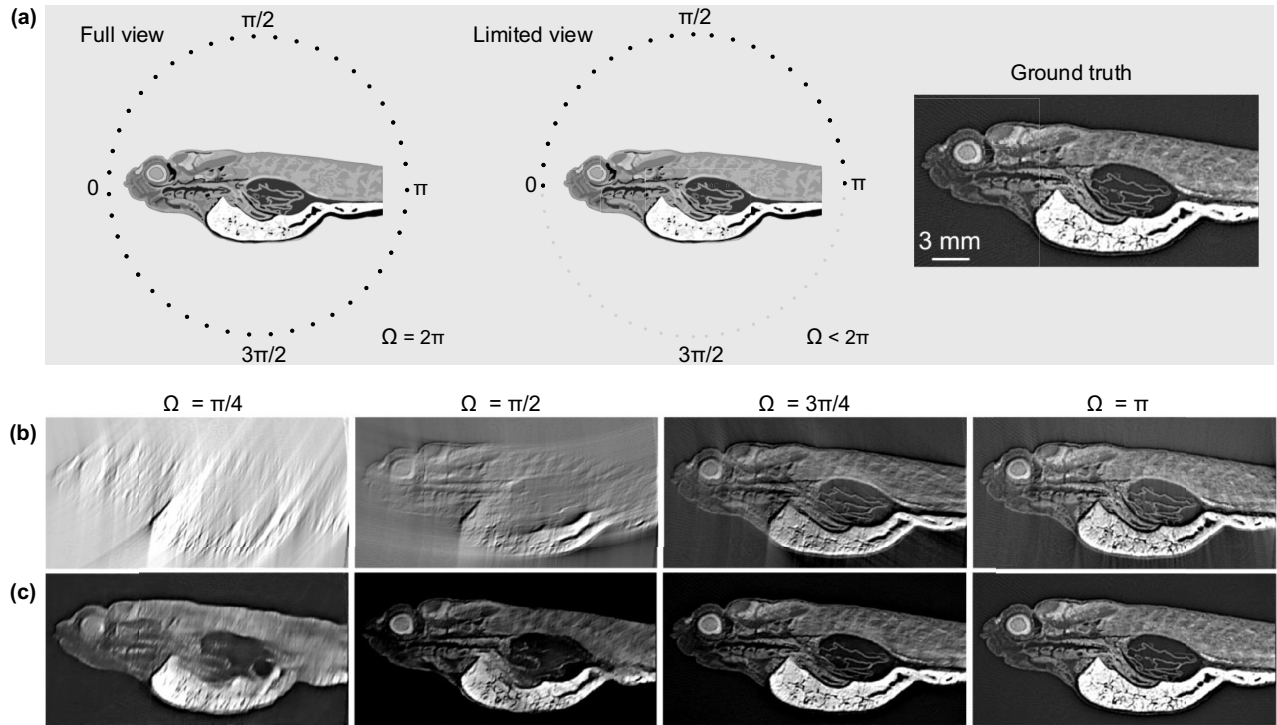
Fig. 48 Performance comparison of DL and conventional image reconstruction algorithms in limited-view PACT imaging. (a) Imaging configuration. (b), (c) Images reconstructed by FBP and dFBP under different view angles. Adapted from Ref. [276] with permission.

and 512 evenly distributed detectors, as shown in Fig. 48(a). To simulate limited-view imaging scenarios, the shape of the detector array is reduced from a full circle to partial circles with central angles of $\pi/4$, $\pi/2$, $3\pi/4$, and π. Figures 48(b) and 48(c) show the corresponding images reconstructed by FBP and dFBP under different view angles. The dFBP method can achieve better image reconstruction quality than FBP, which suffers from severe artifacts due to the incompleteness of the projection data.

Although image reconstruction in limited-view imaging is challenging, the reconstruction procedure can be improved. First, limited-view raw projection data can potentially be augmented to achieve improved reconstruction. In 2004, Patch derived the consistency conditions for projection and estimated missing data from measured data[304]. Gamelin *et al.* employed a single-stage Wiener optimal filter to augment measured projection data by interpolation between measurement locations[305]. Second, conventional algorithms can be modified to adapt to the image reconstruction problem. For example, Paltauf *et al.* proposed a modified FBP algorithm that uses a weight factor to improve the image reconstruction quality in limited-view imaging[306]. Third, DL-based methods can be used as postprocessing tools for image enhancement. For example, Lu *et al.* proposed a GAN-based image postprocessing method and achieved high-quality image recovery from limited-view photoacoustic images[58].

5.1.6 Heterogeneous media

Many image reconstruction algorithms in PACT, such as DAS and FBP, are derived based on the assumption that the acoustic media are homogeneous, lossless, and non-dispersive. However, this is not true for biological tissues, in which strong acoustic heterogeneities, such as bones and air cavities, may be present[296,307]. To address this problem, an image reconstruction algorithm should consider the properties of an acoustic medium. The six classes of algorithms reviewed in this paper have different characteristics in addressing this problem. DAS and FBP algorithms can employ dual SOS[308] or jointly reconstruct initial photoacoustic pressure and SOS[309] to reduce image artifacts caused by acoustic heterogeneity. Modified SE algorithms can incorporate variable SOS into reconstruction models to account for acoustic heterogeneity[167]. In comparison, TR-, IR-, and DL-based algorithms can more easily incorporate acoustic properties of a medium into their reconstruction models, as summarized in Table 14. Specifically, TR algorithms can solve the acoustic heterogeneity problem by incorporating the properties of acoustic media such as SOS, density, dispersion, and absorption into their acoustic propagation model (see Sec. 3.4.1). IR algorithms can solve this problem by incorporating the properties of acoustic media into the system matrix **A** by solving coupled photoacoustic wave equations (see Sec. 3.5.2). DL algorithms can solve this problem by training a reconstruction network with heterogeneity-corrected signal-image datasets. Certainly, incorporating the properties of acoustic media into a reconstruction model requires prior knowledge about the media.

To understand how different image reconstruction algorithms behave in the case of acoustic heterogeneity, an example is presented in Fig. 49. The simulation settings in this example are similar to those used in Fig. 41 except that the sound speeds in the background and the ROI [the white dashed box in Fig. 49(a)] are assumed to be 1500 and 1520 m/s, respectively, and all detector elements have an infinite bandwidth. Figure 49(b)
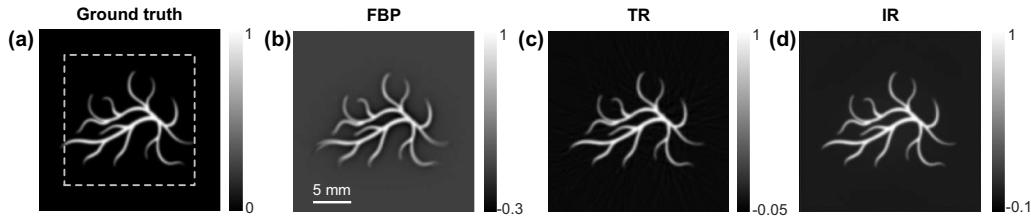
Fig. 49 Performance comparison of FBP, TR, and IR in acoustically heterogeneous media. (a) A numerical blood vessel phantom with a nonuniform SOS distribution. The SOSs in the background and the white dashed box are 1500 and 1520 m/s, respectively. (b) Image reconstructed by FBP with a constant SOS of 1505 m/s. (c), (d) Images reconstructed by TR and TV-regularized IR by coupling the actual SOS distribution into their reconstruction models. Detailed simulation settings can be found in the text.

shows the image reconstructed by FBP with a constant SOS of 1505 m/s, which is the optimal SOS for reconstruction. Figures 49(c) and 49(d) show the images reconstructed by TR- and TV-regularized IR by coupling the actual SOS distribution into their reconstruction models. The images reconstructed by TR and IR with a correct SOS distribution have better image quality than that of FBP, which contains splitting artifacts at the end of the vessel.

Although algorithms such as DAS, FBP, and SE are generally not ready to incorporate the properties of acoustic media into their reconstruction models, their output images can be postprocessed for image enhancement. Figure 50 shows a simulation demonstrating that image postprocessing helps correct acoustic heterogeneity and improves image reconstruction results[233]. The photoacoustic source in this example contains multiple elliptical and line absorbers and has three SOS regions with values of 1480, 1450, and 1575 m/s at different depths. The detector array used for imaging is a linear array located at the top of the image. Figure 50(a) shows the image reconstructed by a multistencil fast marching (MSFM) approach, which is regarded as the ideal result. The MSFM approach estimates the acoustic TOF based on the eikonal equation and the known SOS distribution and can achieve high-quality beamforming. Figure 50(b) shows the images reconstructed by a conventional Fourier beamformer using a constant SOS of 1540 m/s. Figure 50(c) shows the image in Fig. 50(b) processed by an automatic

SOS selection method, which attempts to maximize the sharpness of the photoacoustic image with an optimal SOS (1480 m/s in this case). Figure 50(d) shows the image of Fig. 50(b) processed by a deep neural network called SegUNet. The results show that the image produced by a conventional Fourier beamformer has significant distortions, which can be partially corrected by the automatic SOS selection method and fully corrected by the DL-based postprocessing method.

5.1.7 Other aspects

In addition to the qualitative reconstruction discussed above, quantitative image reconstruction is another important aspect to consider in certain imaging scenarios, such as quantitative photoacoustic imaging. Under ideal conditions, most algorithms can achieve accurate amplitude reconstruction of initial photoacoustic pressure. However, real imaging scenarios are never ideal, which makes quantitative image reconstruction very challenging. Generally, IR[41,64] and DL[293] algorithms can achieve more accurate image reconstruction compared to other methods and are more suitable for quantitative photoacoustic imaging. Additionally, negative artifacts that have no physical meanings often occur in reconstructed photoacoustic images under nonideal conditions[137]. In this case, IR algorithms with nonnegative constraints can be used to reconstruct photoacoustic images free from negative components[207].
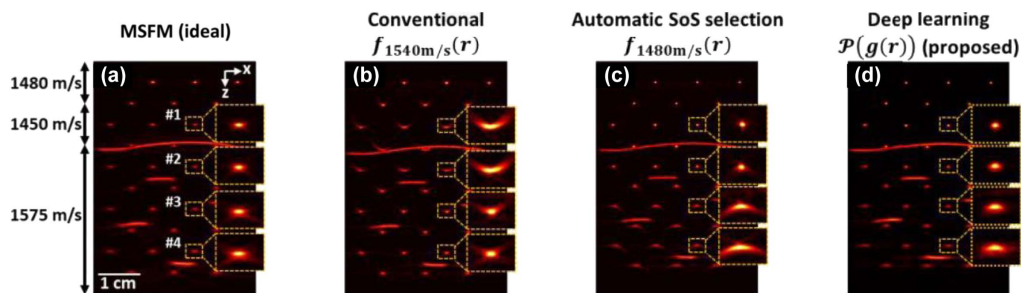

Fig. 50 DL-based postprocessing enhances image quality in acoustically heterogeneous media. In this example, the imaging media consist of three layers in the depth direction (*z* direction), and their SOSs are 1480, 1450, and 1575 m/s. (a) Image reconstructed using the MSFM method, which is regarded as the ideal result. (b) Image reconstructed by a conventional Fourier beamformer with a constant SOS of 1540 m/s. (c) Postprocessed image of (b) using an autofocus approach. (d) Postprocessed image of (b) using a DL-based method (SegUNet). Adapted from Ref. [233] with permission.
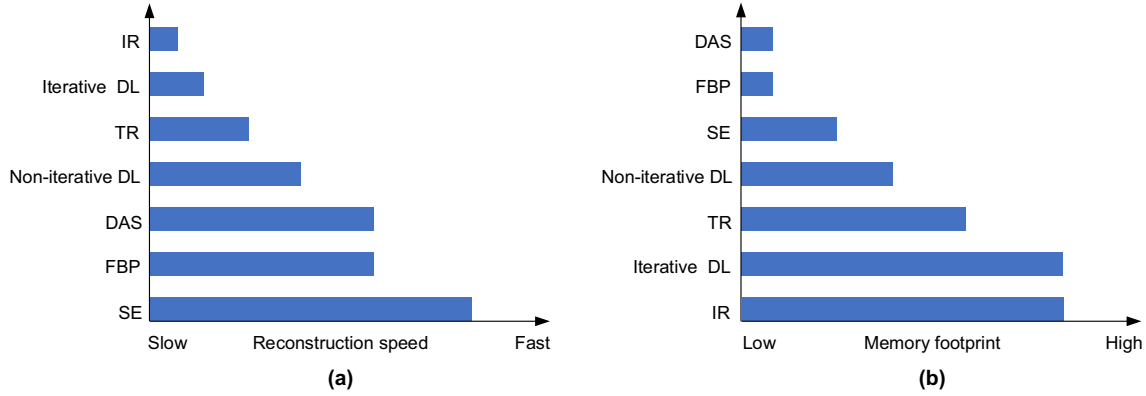
Fig. 51 Qualitative comparison of different image reconstruction algorithms in terms of (a) image reconstruction speed and (b) memory footprint.

5.2 Image Reconstruction Speed and Memory Footprint

5.2.1 Image reconstruction speed

In addition to image quality, image reconstruction speed and computational complexity are other critical indicators for measuring the performance of an algorithm. An ideal image reconstruction algorithm should be fast enough to allow a PACT imaging system to capture transient biodynamic processes such as heartbeat and blood flow in a living subject.

The computational complexity of image reconstruction algorithms in PACT varies from algorithm to algorithm. For non-iterative methods, such as DAS, FBP, SE, and TR, SE is the most computationally efficient due to the application of FFT. Assuming that a 3D image to be reconstructed has a size of $N_x \times N_y \times N_z$ ($N_x = N_y = N_z = n$) and the detector number $M = n \times n$[32], SE algorithms have a computational complexity of $O[n^3 \log(n)]$ for 3D image reconstruction and $O[n^2 \log(n)]$ for 2D image reconstruction for the specific detection geometries listed in Table 5. To perform the same reconstruction tasks, TR algorithms have computational complexities of $O[n^4 \log(n)]$ for 3D and $O[n^3 \log(n)]$ for 2D[310], while back-projection-type algorithms, such as DAS and FBP, have computational complexities of $O(n^5)$ for 3D and $O(n^3)$ for 2D[62,310]. One may notice that TR algorithms have lower complexity than back-projection-type algorithms for 3D image reconstruction but seem to be slower in practice. This is because TR algorithms need to compute an entire acoustic field [Eqs. (69)–(73)] step by step, which is time- and memory-consuming, especially for large-scale 3D imaging models. In contrast, back-projection-type algorithms can directly reconstruct the ROI instead of a whole imaging region and can be implemented with parallel computing. Compared with non-iterative methods, IR algorithms usually have a much greater computational complexity due to the repeated calculation of the image reconstruction model in Fig. 20.

The computational complexity of DL-based algorithms depends on the specific method. Generally, non-iterative DL algorithms, such as data-, image-, and hybrid-domain processing algorithms and direct reconstruction algorithms, can process an image in roughly the same amount of time as conventional non-iterative algorithms, such as FBP and TR. Iterative DL algorithms, such as learned IR, are generally faster than conventional IR algorithms because they normally require fewer iterations and because the discrete forward imaging model [Eq. (86)] can be learned by neural networks. Figure 51(a) shows a qualitative comparison of the speed of different image reconstruction algorithms.

The speed of an image reconstruction algorithm in PACT can be accelerated by GPUs[42,152,311]. For example, Wang *et al.* developed a comprehensive GPU-based framework for PACT image reconstruction that achieved an ~871-fold increase in speed when reconstructing extremely large-scale 3D images compared to a central processing unit (CPU)-based implementation[152]. Liu *et al.* reported an ultrafast GPU-based FBP implementation for PACT image reconstruction[311]. The implementation requires only 0.38 ms to reconstruct a 2D image with a size of 512 pixel × 512 pixel and 6.15 ms to reconstruct a 3D image with a size of 160 voxel × 160 voxel × 160 voxel. Wang *et al.* proposed GPU-based parallelization strategies to accelerate the FBP algorithm and a penalized least-square and interpolation-based IR (PLS-Int-IR) algorithm[42] and improved the image reconstruction speeds of FBP and PLS-Int-IR by factors of 1000 and 125, respectively. Although most image reconstruction algorithms can, in principle, be accelerated by GPUs, the speedup ratios may vary. Back-projection-type algorithms, such as DAS and FBP, inherently support parallel computing and thus can achieve a high speedup ratio. In contrast, TR and IR algorithms either need to compute an entire photoacoustic field [Eqs. (69)–(73)] or need to update the image reconstruction model (Fig. 20) step by step, thus offering limited acceleration performance.

5.2.2 Memory footprint

Memory footprint is also a critical indicator for measuring the performance of an algorithm. Generally, the memory footprint of an image reconstruction algorithm in PACT depends on the size of the image to be reconstructed $N_x \times N_y \times N_z$ and the size of the measured photoacoustic signals $M \times K$ (M: detector number, K: signal sampling length). For DAS, FBP, SE, and TR algorithms, the memory footprint can be approximately estimated as

$$\text{Memory}_1 \propto N_x N_y N_z + MK. \tag{125}$$

Similarly, the memory footprint of IR algorithms can be approximately estimated to be

$$\text{Memory}_2 \propto N_x N_y N_z MK + MK, \tag{126}$$

where $N_x N_y N_z MK$ is the size of the system matrix $\mathbf{A}$ in Eq. (86).

According to Eqs. (125)–(126), back-projection-type algorithms, such as DAS and FBP, require the least amount of memory, while IR algorithms require the most memory. Moreover, back-projection-type algorithms can reconstruct only a portion of an imaging region (i.e., ROI) rather than the whole region, which can further reduce the amount of memory needed. IR algorithms can also achieve direct image reconstruction of ROIs but the memory footprint is still high due to the extremely large size of the system matrix $\mathbf{A}$, which is sparse and can be compactly represented to reduce the memory footprint[36]. In contrast to that of conventional image reconstruction algorithms, the memory footprint of a DL algorithm depends on its network structure. In general, the memory footprint of non-iterative DL algorithms is greater than that of conventional non-iterative algorithms, but less than that of IR algorithms. The memory footprint of learned IR algorithms is comparable to that of conventional IR algorithms and is also very high. Figure 51(b) shows a qualitative comparison of the memory footprints of different image reconstruction algorithms.

6 Challenges and Discussion

The six classes of algorithms reviewed in this paper, namely, DAS, FBP, SE, TR, IR, and DL, can achieve high-quality image reconstruction under ideal imaging conditions. However, they may face challenges under non-ideal imaging conditions, especially in the following scenarios: (1) the bandwidths and aperture sizes of the detectors used for imaging are limited and finite; (2) the elements and view angle of the detector array used for imaging are sparse and limited; and (3) acoustic media are strongly heterogeneous.

The bandwidth and aperture size of an individual ultrasound detector in a detector array are important for high-quality photoacoustic image reconstruction. An ideal detector should have infinite bandwidth and a point-like aperture to respond to all frequency contents of a photoacoustic signal from all directions. However, practical ultrasound detectors always have limited bandwidths and finite aperture sizes, which distort measured photoacoustic signals and blur reconstructed images. Researchers typically first measure the impulse response of a detector and then couple it to IR algorithms[37,39] or use deconvolution approaches[291,292,312] to remove the impacts of detector bandwidth and aperture size. However, these methods lead to solving optimization problems, which are typically slow, ill-posed, and sensitive to noise. How to build more efficient models to compensate for the impact of detector bandwidth and aperture size is a subject worth studying.

The element number and view angle of a detector array are important for high-quality photoacoustic image reconstruction. Ideally, the element number of a detector array should be large enough to satisfy the spatial Nyquist sampling criteria for perfect sampling, and the view angle should be 4π steradian (full 3D view) to record complete photoacoustic signals in 3D space. However, practical detector arrays such as linear and planar arrays typically have a limited number of elements and a limited view angle, which eventually results in the challenging problem of image reconstruction from sparse-view and limited-view projections. IR and DL are two classes of algorithms that are commonly used to address image reconstruction problems in these scenarios[211,276,278,313]. However, these methods typically involve intensive calculations and/or require huge amounts of training data that are difficult to obtain in practice. How to develop more accurate and efficient algorithms to better address the image reconstruction problem in sparse-view and limited-view imaging is also worth studying.

The acoustic homogeneity of a medium is also important for high-quality photoacoustic image reconstruction. Most current image reconstruction algorithms used in PACT assume that the acoustic media are homogeneous. However, this does not hold in biological tissues, especially when strong heterogeneities, such as bones and air cavities, are present[307]. Current methods for image reconstruction in heterogeneous tissues include half-time[314] or partial-time[315] reconstruction, autofocus reconstruction by optimizing the SOS[316], joint reconstruction of optical absorption and SOS[317], and ultrasound-guided adaptive reconstruction[296]. However, these methods can only mitigate the acoustic heterogeneity problem to a certain extent and have limited improvements in reconstruction accuracy, especially in complex imaging scenarios such as transcranial imaging. An alternative method is full-waveform-based algorithms, such as TR[170] and IR[41], which are based on solutions to exact photoacoustic wave equations. These algorithms can consider the acoustic properties of the media such as SOS, absorption, dispersion, and density, and achieve accurate image reconstruction in acoustically heterogeneous media. However, these algorithms are not widely used thus far, especially in practical applications. The main reason is that they require prior information on media such as SOS maps, which can be measured using other imaging techniques such as ultrasound CT[318–320] but at the expense of introducing additional imaging facilities and computing resources. Therefore, how to better tackle the acoustic heterogeneity problem and achieve high-quality image reconstruction deserves further investigation.

In addition to image reconstruction quality, image reconstruction speed is also important for an algorithm. Currently, many algorithms such as DAS, FBP, SE, and non-iterative DL algorithms can achieve real-time reconstruction when accelerated with GPUs[33,104,152,279,311], which, however, is a great challenge for IR-based algorithms. The IR algorithms in PACT need to compute a large-scale system matrix $\mathbf{A}$ and perform image reconstruction iteratively. Therefore, they require huge amounts of computational resources, which makes real-time reconstruction difficult. This is especially true for 3D IR reconstruction, in which the computational time is typically on the order of hours[42]. DL-based IR reconstruction has been reported to help reduce the number of iterations[45] and improve computational efficiency[57]. Nevertheless, developing new strategies to improve the computational speed of IR algorithms and reduce the memory footprint still has important practical significance.

DL-based image reconstruction has superior performance over conventional image reconstruction algorithms and is an emerging technique in PACT. It has been used to solve a variety of non-ideal image reconstruction problems in PACT, such as detector bandwidth expansion and IR acceleration. Nonetheless, DL-based image reconstruction faces multiple challenges. First, the performance of DL methods heavily relies on the quality and quantity of training data. The training data are usually required to be paired, and the acquisition of large amounts of paired data is difficult in practice. Training neural networks using simulated datasets and/or small-scale experimental datasets may limit the

performance of the networks in terms of reconstruction accuracy, robustness, and generalizability in practical applications. Second, reference images such as full-bandwidth photoacoustic images for network training in DL methods are often difficult to obtain. In these situations, networks can only be trained using simulated datasets. Since experimental conditions are typically much more complex than simulations, training a network using only simulated datasets leads to degraded reconstruction performance and poor robustness. Third, there are currently few publicly available datasets for the comparison of the performances of different DL methods. Using private datasets to compare different DL methods may lead to biased conclusions. Finally, DL-based image reconstruction methods in PACT lack interpretability, which is also a long-standing problem in the field of DL. In response to the above issues, it is necessary to research the following aspects: (1) building large-scale publicly available datasets for the construction of networks with good generalization properties and performance comparisons of different networks; (2) developing unsupervised DL-based reconstruction methods that do not need paired data for training; (3) developing more powerful simulation platforms to generate reference data closely resembling real experimental data[321]; and (4) developing physics-informed networks to improve the interpretability of DL methods.

One important goal of image reconstruction from photoacoustic projections is disease diagnosis. The current diagnostic route from photoacoustic signals to photoacoustic images to knowledge is straightforward but can potentially lead to information loss and distortion during image reconstruction, which can decrease the accuracy of disease diagnosis. Artificial intelligence (AI) can mine knowledge from vast amounts of data and offer opportunities for disease diagnosis directly from raw data[322]. The use of raw-data-based diagnostic technology instead of image-based diagnostic technology may facilitate the development of fully automated scanning and diagnostic procedures and become an important direction in the future.

This review mainly discusses advanced image reconstruction algorithms in PAT for high-performance imaging. In addition to advanced algorithms, emerging technologies such as metamaterials and nanomaterials can be leveraged to enhance the imaging performance of PAT. For example, metamaterials can be appropriately designed to influence the propagation of electromagnetic waves or acoustic waves in a manner not observed in bulk materials. By tailoring the amplitude, phase, or polarization of light, optical metamaterials can emulate conventional optical lenses, waveplates, or holograms to achieve multidimensional light-field modulation in PAT[323–329]. Metamaterials can also be designed for photoacoustic signal sensing to address the sensitivity and bandwidth problem of conventional piezoelectric sensors[330–333] and provide a sensitive method for ultrasound detection in PAT.

Nanomaterials that often exhibit unique optical, electronic, or physicochemical properties are another emerging technology that can be used to enhance the imaging performance of PAT. Nanoparticles are a common type of nanomaterial used as molecular contrast agents in PAT[334] and can be generally divided into two categories: inorganic nanoparticles and organic nanoparticles. Inorganic nanoparticles, such as noble metal nanoparticles, iron oxide nanoparticles, semiconductor nanoparticles, magnetic nanoparticles, and carbon nanoparticles, possess versatile properties and have been investigated for various applications in biomedical imaging and therapeutics[335,336].

Among them, noble metal nanoparticles are the most commonly used contrast agents in PAT due to their high optical absorption cross section and biocompatibility[337]. Gold nanoparticles, such as gold nanospheres, nanorods, nanoshells, and nanostars, are particularly popular because of their high extinction coefficient in the near-infrared range and their high photoacoustic conversion efficiency[334,338,339]. In addition to inorganic nanoparticles, organic nanoparticles such as polymer nanoparticles and encapsulations are also increasingly being used in PAT[334,340,341]. Polymer nanoparticles usually have a molar extinction coefficient lower than that of gold nanoparticles but higher than that of small-molecule dyes and have an absorption peak in the near-infrared range. They typically possess high structural and functional flexibility, high biodegradability and biocompatibility, and have a high translational potential. Furthermore, polymer nanoparticles usually have stable surfaces that can be modified with specific targeting and therapeutic moieties for molecular imaging and targeted therapy, which enables the imaging of biological events and functionalities at multiple scales.

7 Conclusions

In this work, we systemically review the image formation problem in PACT over the past three decades, including the forward problem, conventional reconstruction methods, DL-based reconstruction methods, and the comparisons of different reconstruction methods.

The photoacoustic forward problem involves multiple physical processes, including photoacoustic signal generation, signal propagation, and signal detection. The generation of photoacoustic signals is based on the thermoelastic effect and should satisfy the conditions of thermal confinement and stress confinement. The propagation of photoacoustic signals is governed by the photoacoustic wave equation, which can be analytically solved by the method of Green's functions and numerically visualized by the k-Wave toolbox. The detection of photoacoustic signals typically involves the use of ultrasound detector arrays, whose properties such as detector bandwidth, aperture size, element number, and view angle impact detected photoacoustic signals and final images. The photoacoustic forward problem can be described by the spherical Radon transform. The task of image reconstruction in PACT can be regarded as finding the inverse of the spherical Radon transform.

Conventional reconstruction methods in PACT typically achieve image formation via analytical derivations or numerical computations and mainly include five classes of algorithms, i.e., DAS, FBP, SE, TR, and IR. The DAS-type algorithms reconstruct an image by summing the delayed photoacoustic signals of each detector and have many variants, such as native DAS, DMAS, SLSC, and MV. The DAS-type algorithms are simple in principle and fast in speed but lack mathematical rigor. FBP-type algorithms reconstruct an image by back-projecting the filtered photoacoustic signals of each detector to the image domain and can achieve accurate image reconstruction under ideal conditions. They are rigorously deduced from the photoacoustic wave equation and can be regarded as advanced versions of DAS. The SE-type algorithms reconstruct an image by representing it using a mathematical series. They are computationally efficient and super-fast for special detection geometries such as the planar geometry. The TR-type algorithms reconstruct an image by running a forward numerical acoustic propagation model backward and can be used for any closed detection geometry.

They are well suited for image reconstruction in heterogeneous media since the acoustic propagation model can couple acoustic properties of media, such as SOS, density, dispersion, and absorption of a medium. The IR-type algorithms reconstruct an image by minimizing the energy error between measured projection data and estimated projection data. They can incorporate various physical factors into reconstruction models, suppress image artifacts caused by incomplete projections, and are well-suited for non-ideal imaging scenarios.

DL achieves tomographic image reconstruction in PACT by learning the photoacoustic imaging model from big data using designed networks and thus is network-based, data-driven, and learning-oriented. DL techniques can be applied to photoacoustic image reconstruction from multiple aspects, including signal preprocessing in the data domain, image postprocessing in the image domain, hybrid-domain processing, direct signal-to-image reconstruction, and IR reconstruction learning. DL-based signal preprocessing employs a network to enhance raw photoacoustic projection data in the data domain and delivers the enhanced projection data to conventional reconstruction algorithms as input. DL-based image postprocessing employs a network to enhance photoacoustic images output from conventional reconstruction algorithms for artifacts and noise suppression. DL-based hybrid-domain processing employs a network to make full use of the information of raw projection data and reconstructed images and can yield images with enhanced quality. In contrast, DL-based direct reconstruction employs a network to form photoacoustic images directly from raw projection data and does not involve any conventional image reconstruction algorithms. Training such a direct transformation network, however, requires large-scale datasets. DL-based IR reconstruction employs a network to learn the reconstruction process of conventional IR algorithms and can also achieve direct signal-to-image reconstruction. This method typically has better robustness and generalizability but consumes more time and memory than other DL methods.

The image reconstruction algorithms discussed in this review have distinct characteristics. In terms of image reconstruction quality, most image reconstruction algorithms work well under ideal imaging conditions and can yield high-quality images. However, they behave differently under non-ideal imaging scenarios, such as sparse-view imaging, limited-view imaging, imaging with bandwidth-limited and/or finite-aperture detectors, and imaging in the presence of acoustic heterogeneity. Generally, the DAS, FBP, and SE algorithms have relatively poor performance in these scenarios. The TR and IR algorithms can yield improved results because they can incorporate the physical models of an imaging system, such as the SOS of media, the EIR, and the SIR of a detector. The DL algorithms can produce surprisingly good results that are unattainable for conventional algorithms due to their network-based data-driven nature. In terms of image reconstruction speed and memory footprint, DAS, FBP, SE, TR, and non-iterative DL algorithms are generally faster than IR and iterative DL methods. TR, IR, and iterative DL algorithms generally require more memory than DAS, FBP, and SE methods.

This review is expected to help general readers better understand the image formation problem in PACT, provide a self-contained reference guide for beginners as well as specialists, and facilitate further developments and applications of novel image reconstruction algorithms.

References

1. A. G. Bell, "On the production and reproduction of sound by light," *Am. J. Sci.* **s3-20**, 305 (1880).
2. L. V. Wang, "Multiscale photoacoustic microscopy and computed tomography," *Nat. Photonics* **3**, 503 (2009).
3. L. V. Wang and S. Hu, "Photoacoustic tomography: in vivo imaging from organelles to organs," *Science* **335**, 1458 (2012).
4. C. Tian *et al.*, "Spatial resolution in photoacoustic computed tomography," *Rep. Prog. Phys.* **84**, 036701 (2021).
5. C. Li and L. V. Wang, "Photoacoustic tomography and sensing in biomedicine," *Phys. Med. Biol.* **54**, R59 (2009).
6. S.-L. Chen and C. Tian, "Recent developments in photoacoustic imaging and sensing for nondestructive testing and evaluation," *Vis. Comput. Ind. Biomed. Art.* **4**, 1 (2021).
7. C. Tian *et al.*, "Noninvasive chorioretinal imaging in living rabbits using integrated photoacoustic microscopy and optical coherence tomography," *Opt. Express* **25**, 15947 (2017).
8. S. Li *et al.*, "Photoacoustic imaging of peripheral vessels in extremities by large-scale synthetic matrix array," *J. Biomed. Opt.* **29**, S11519 (2024).
9. S. Liu *et al.*, "Validation of photoacoustic/ultrasound dual imaging in evaluating blood oxygen saturation," *Biomed. Opt. Express* **13**, 5551 (2022).
10. M. Yang *et al.*, "Synovial oxygenation at photoacoustic imaging to assess rheumatoid arthritis disease activity," *Radiology* **306**, 220 (2023).
11. S. Liu *et al.*, "On the imaging depth limit of photoacoustic tomography in the visible and first near-infrared windows," *Opt. Express* **32**, 5460 (2024).
12. S. Liu *et al.*, "In vivo photoacoustic sentinel lymph node imaging using clinically-approved carbon nanoparticles," *IEEE Trans. Biomed. Eng.* **67**, 2033 (2019).
13. W. Pang *et al.*, "Direct monitoring of whole-brain electrodynamics via high-spatiotemporal-resolution photoacoustics with voltage-sensitive dye," *Laser Photonics Rev.*, 2400165 (2024).
14. T. Bowen, "Radiation-induced thermoacoustic soft tissue imaging," in *Ultrasonics Symposium* (IEEE, 1981), p. 817.
15. T. Bowen *et al.*, "Some experimental results on the thermoacoustic imaging of tissue equivalent phantom materials," in *Ultrasonics Symposium* (IEEE, 1981), p. 823.
16. T. Bowen, "Radiation-induced thermoacoustic imaging," *U.S. patent* 4,385,634 (1981).
17. A. A. Oraevsky, S. L. Jacques, and F. K. Tittel, "Determination of tissue optical properties by piezoelectric detection of laser-

induced stress waves," in *Laser-Tissue Interaction IV* (SPIE, 1993), p. 86.

18. A. A. Oraevsky *et al.*, "Laser-based optoacoustic imaging in biological tissues," in *Laser-Tissue Interaction V and Ultraviolet Radiation Hazards* (SPIE, 1994), p. 122.

19. A. A. Oraevsky *et al.*, "Lateral and z-axial resolution in laser optoacoustic imaging with ultrasonic transducers," in *Optical Tomography, Photon Migration, and Spectroscopy of Tissue and Model Media: Theory, Human Studies, and Instrumentation* (SPIE, 1995), p. 198.

20. R. A. Kruger, "Photoacoustic ultrasound," *Med. Phys.* **21**, 127 (1994).

21. R. A. Kruger and P. Liu, "Photoacoustic ultrasound: pulse production and detection in 0.5% liposyn," *Med. Phys.* **21**, 1179 (1994).

22. R. A. Kruger *et al.*, "Photoacoustic ultrasound (PAUS) reconstruction tomography," *Med. Phys.* **22**, 1605 (1995).

23. D. Finch, S. K. Patch, and R. Rakesh, "Determining a function from its mean values over a family of spheres," *SIAM J. Math. Anal.* **35**, 1213 (2004).

24. M. Xu and L. V. Wang, "Universal back-projection algorithm for photoacoustic computed tomography," *Phys. Rev. E* **71**, 016706 (2005).

25. C. G. A. Hoelen *et al.*, "Three-dimensional photoacoustic imaging of blood vessels in tissue," *Opt. Lett.* **23**, 648 (1998).

26. C. G. A. Hoelen *et al.*, "Photoacoustic blood cell detection and imaging of blood vessels in phantom tissue," in *Optical and Imaging Techniques for Biomonitoring III* (SPIE, 1998), p. 142.

27. M. Mozaffarzadeh *et al.*, "Double-stage delay multiply and sum beamforming algorithm: application to linear-array photoacoustic imaging," *IEEE Trans. Biomed. Eng.* **65**, 31 (2018).

28. M. A. Lediju Bell *et al.*, "Short-lag spatial coherence beamforming of photoacoustic images for enhanced visualization of prostate brachytherapy seeds," *Biomed. Opt. Express* **4**, 1964 (2013).

29. C.-K. Liao, M.-L. Li, and P.-C. Li, "Optoacoustic imaging with synthetic aperture focusing and coherence weighting," *Opt. Lett.* **29**, 2506 (2004).

30. S. Paul, A. Thomas, and M. S. Singh, "Delay-and-sum-to-delay-standard-deviation factor: a promising adaptive beamformer," *Opt. Lett.* **46**, 4662 (2021).

31. K. P. Köstli *et al.*, "Temporal backward projection of optoacoustic pressure transients using Fourier transform methods," *Phys. Med. Biol.* **46**, 1863 (2001).

32. L. A. Kunyansky, "A series solution and a fast algorithm for the inversion of the spherical mean Radon transform," *Inverse Probl.* **23**, S11 (2007).

33. L. Kunyansky, "Fast reconstruction algorithms for the thermoacoustic tomography in certain domains with cylindrical or spherical symmetries," *Inverse Prob. Imaging* **6**, 111 (2012).

34. Y. Xu and L. V. Wang, "Time reversal and its application to tomography with diffracting sources," *Phys. Rev. Lett.* **92**, 033902 (2004).

35. P. Burgholzer *et al.*, "Exact and approximative imaging methods for photoacoustic tomography using an arbitrary detection surface," *Phys. Rev. E* **75**, 046706 (2007).

36. G. Paltauf *et al.*, "Iterative reconstruction algorithm for optoacoustic imaging," *J. Acoust. Soc. Am.* **112**, 1536 (2002).

37. A. Rosenthal, D. Razansky, and V. Ntziachristos, "Fast semi-analytical model-based acoustic inversion for quantitative optoacoustic tomography," *IEEE Trans. Med. Imaging* **29**, 1275 (2010).

38. K. Wang *et al.*, "Discrete imaging models for three-dimensional optoacoustic tomography using radially symmetric expansion functions," *IEEE Trans. Med. Imaging* **33**, 1180 (2014).

39. K. Wang *et al.*, "An imaging model incorporating ultrasonic transducer properties for three-dimensional optoacoustic tomography," *IEEE Trans. Med. Imaging* **30**, 203 (2011).

40. L. Ding, X. L. Deán-Ben, and D. Razansky, "Efficient 3-D model-based reconstruction scheme for arbitrary optoacoustic acquisition geometries," *IEEE Trans. Med. Imaging* **36**, 1858 (2017).

41. C. Huang *et al.*, "Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media," *IEEE Trans. Med. Imaging* **32**, 1097 (2013).

42. K. Wang *et al.*, "Accelerating image reconstruction in three-dimensional optoacoustic tomography on graphics processing units," *Med. Phys.* **40**, 023301 (2013).

43. L. Ding, X. L. Deán-Ben, and D. Razansky, "Real-time model-based inversion in cross-sectional optoacoustic tomography," *IEEE Trans. Med. Imaging* **35**, 1883 (2016).

44. S. Antholzer, M. Haltmeier, and J. Schwab, "Deep learning for photoacoustic tomography from sparse data," *Inverse Probl. Sci. Eng.* **27**, 987 (2019).

45. A. Hauptmann *et al.*, "Model-based learning for accelerated, limited-view 3-D photoacoustic tomography," *IEEE Trans. Med. Imaging* **37**, 1382 (2018).

46. N. Davoudi, X. L. Deán-Ben, and D. Razansky, "Deep learning optoacoustic tomography with sparse data," *Nat. Mach. Intell.* **1**, 453 (2019).

47. S. Choi *et al.*, "Deep learning enhances multiparametric dynamic volumetric photoacoustic computed tomography in vivo (DL-PACT)," *Adv. Sci.* **10**, e2202089 (2022).

48. A. Hauptmann and B. Cox, "Deep learning in photoacoustic tomography: current approaches and future directions," *J. Biomed. Opt.* **25**, 112903 (2020).

49. H. Deng *et al.*, "Deep learning in photoacoustic imaging: a review," *J. Biomed. Opt.* **26**, 040901 (2021).

50. P. Rajendran, A. Sharma, and M. Pramanik, "Photoacoustic imaging aided with deep learning: a review," *Biomed. Eng. Lett.* **12**, 155 (2022).

51. S. Gutta *et al.*, "Deep neural network-based bandwidth enhancement of photoacoustic data," *J. Biomed. Opt.* **22**, 1 (2017).

52. N. Awasthi *et al.*, "Deep neural network-based sinogram super-resolution and bandwidth enhancement for limited-data photoacoustic tomography," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **67**, 2660 (2020).

53. P. Rajendran and M. Pramanik, "Deep learning approach to improve tangential resolution in photoacoustic tomography," *Biomed. Opt. Express* **11**, 7311 (2020).

54. H. Zhang *et al.*, "Deep-E: a fully-dense neural network for improving the elevation resolution in linear-array-based photoacoustic tomography," *IEEE Trans. Med. Imaging* **41**, 1279 (2021).

55. C. Dehner *et al.*, "Deep-learning-based electrical noise removal enables high spectral optoacoustic contrast in deep tissue," *IEEE Trans. Med. Imaging* **41**, 3182 (2022).

56. H. Zhao *et al.*, "Deep learning enables superior photoacoustic imaging at ultralow laser dosages," *Adv. Sci.* **8**, 2003097 (2021).

57. K. T. Hsu, S. Guan, and P. V. Chitnis, "Fast iterative reconstruction for photoacoustic tomography using learned physical model: Theoretical validation," *Photoacoustics* **29**, 100452 (2023).

58. T. Lu *et al.*, "LV-GAN: a deep learning approach for limited-view optoacoustic imaging based on hybrid datasets," *J. Biophotonics* **14**, e202000325 (2021).

59. X. Zhang *et al.*, "Sparse-sampling photoacoustic computed tomography: deep learning vs. compressed sensing," *Biomed. Signal Process. Control* **71**, 103233 (2022).

60. P. Kuchment and L. Kunyansky, "Mathematics of thermoacoustic tomography," *Eur. J. Appl. Math.* **19**, 191 (2008).

61. A. Rosenthal, V. Ntziachristos, and D. Razansky, "Acoustic inversion in optoacoustic tomography: a review," *Curr. Med. Imaging Rev.* **9**, 318 (2013).

62. C. Lutzweiler and D. Razansky, "Optoacoustic imaging and tomography: reconstruction approaches and outstanding challenges in image performance and quantification," *Sensors* **13**, 7345 (2013).

63. J. Poudel, Y. Lou, and M. A. Anastasio, "A survey of computational frameworks for solving the acoustic inverse problem in

three-dimensional photoacoustic computed tomography," *Phys. Med. Biol.* **64**, 14TR01 (2019).

64. X. L. Deán-Ben and D. Razansky, "A practical guide for model-based reconstruction in optoacoustic imaging," *Front. Phys.* **10**, 1057 (2022).

65. C. Yang *et al.*, "Review of deep learning for photoacoustic imaging," *Photoacoustics* **21**, 100215 (2021).

66. H. B. Yedder, B. Cardoen, and G. Hamarneh, "Deep learning for biomedical image reconstruction: a survey," *Artif. Intell. Rev.* **54**, 215 (2021).

67. A. DiSpirito III *et al.*, "Sounding out the hidden data: a concise review of deep learning in photoacoustic imaging," *Exp. Biol. Med.* **246**, 1355 (2021).

68. J. Gröhl *et al.*, "Deep learning for biomedical photoacoustic imaging: a review," *Photoacoustics* **22**, 100241 (2021).

69. B. Cox *et al.*, "Quantitative spectroscopic photoacoustic imaging: a review," *J. Biomed. Opt.* **17**, 061202 (2012).

70. X. Tang, J. Fu, and H. Qin, "Microwave-induced thermoacoustic imaging with functional nanoparticles," *J. Innov. Opt. Health Sci.* **16**, 2230014 (2023).

71. Q. Liu *et al.*, "Biomedical microwave-induced thermoacoustic imaging," *J. Innov. Opt. Health Sci.* **15**, 2230007 (2022).

72. Z. Liang *et al.*, "Study on response of metal wire in thermoacoustic imaging," *J. Innov. Opt. Health Sci.* **15**, 2250015 (2022).

73. X. Liang *et al.*, "Investigation of artifacts by mapping SAR in thermoacoustic imaging," *J. Innov. Opt. Health Sci.* **14**, 2150011 (2021).

74. L. V. Wang and H. Wu, *Biomedical Optics: Principles and Imaging* (John Wiley & Sons, 2007).

75. B. T. Cox and P. C. Beard, "Fast calculation of pulsed photoacoustic fields in fluids using k-space methods," *J. Acoust. Soc. Am.* **117**, 3616 (2005).

76. American National Standards Institute, *American National Standard for Safe Use of Lasers ANSI Z136.1-2007* (Laser Institute of America, 2007).

77. A. C. Tam, "Applications of photoacoustic sensing techniques," *Rev. Mod. Phys.* **58**, 381 (1986).

78. B. Cox and P. C. Beard, "Modeling photoacoustic propagation in tissue using k-space techniques," in *Photoacoustic Imaging and Spectroscopy* (CRC Press, 2017), p. 25.

79. H. Jiang, "Fundamentals of photoacoustic tomography," in *Photoacoustic Tomography* (CRC Press, 2015), p. 1.

80. L. V. Wang, "Tutorial on photoacoustic microscopy and computed tomography," *IEEE J. Sel. Top. Quantum Electron.* **14**, 171 (2008).

81. C. Guo and C. S. Singh, *Handbook of Laser Technology and Applications: Laser Applications: Medical, Metrology and Communication (volume four)* (CRC Press, 2021).

82. C. Tian *et al.*, "Impact of system factors on the performance of photoacoustic tomography scanners," *Phys. Rev. Appl.* **13**, 014001 (2020).

83. B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *J. Biomed. Opt.* **15**, 021314 (2010).

84. B. E. Treeby *et al.*, "Modeling nonlinear ultrasound propagation in heterogeneous media with power law absorption using a k-space pseudospectral method," *J. Acoust. Soc. Am.* **131**, 4324 (2012).

85. N. N. Bojarski, "The k-space formulation of the scattering problem in the time domain," *J. Acoust. Soc. Am.* **72**, 570 (1982).

86. M. Tabei, T. D. Mast, and R. C. Waag, "A k-space method for coupled first-order acoustic propagation equations," *J. Acoust. Soc. Am.* **111**, 53 (2002).

87. Z. Chenxi, T. Zhijian, and T. Chao, "Point spread function modeling for photoacoustic tomography–I: three-dimensional detection geometries," *Opt. Express* **32**, 1063 (2024).

88. C. Zhang, C. Chen, and C. Tian, "Point spread function modeling for photoacoustic tomography–II: Two-dimensional detection geometries," *Opt. Express* **32**, 1088 (2024).

89. S. R. Deans, *The Radon Transform and Some of its Applications* (Courier Corporation, 2007).

90. A. C. Kak and M. Slaney, "Algorithms for reconstruction with nondiffracting sources," in *Principles of Computerized Tomographic Imaging* (IEEE, 2001), p. 49.

91. N. J. Redding and G. N. Newsam, "Radon transform and its inverse," in *Inverting the Circular Radon Transform* (DSTO Electronics and Surveillance Research Laboratory, 2002), p. 2.

92. N. J. Redding and T. M. Payne, "Inverting the spherical Radon transform for 3D SAR image formation," in *Proceedings of the International Conference on Radar (IEEE Cat. No. 03EX695)* (IEEE, 2003), p. 466.

93. K. E. Thomenius, "Evolution of ultrasound beamformers," in *Ultrasonics Symposium* (IEEE, 1996), p. 1615.

94. J. C. Somer, "Electronic sector scanning for ultrasonic diagnosis," *Ultrasonics* **6**, 153 (1968).

95. R. E. McKeighen and M. P. Buchin, "New techniques for dynamically variable electronic delays for real time ultrasonic imaging," in *Ultrasonics Symposium* (IEEE, 1977), p. 250.

96. V. Perrot *et al.*, "So you think you can DAS? a viewpoint on delay-and-sum beamforming," *Ultrasonics* **111**, 106309 (2021).

97. C. G. A. Hoelen and F. F. de Mul, "Image reconstruction for photoacoustic scanning of tissue structures," *Appl. Opt.* **39**, 5872 (2000).

98. D. Feng *et al.*, "Microwave-induced thermoacoustic tomography: reconstruction by synthetic aperture," *Med. Phys.* **28**, 2427 (2001).

99. H. B. Lim *et al.*, "Confocal microwave imaging for breast cancer detection: delay-multiply-and-sum image reconstruction algorithm," *IEEE Trans. Biomed. Eng.* **55**, 1697 (2008).

100. G. Matrone *et al.*, "The delay multiply and sum beamforming algorithm in ultrasound B-mode medical imaging," *IEEE Trans. Med. Imaging* **34**, 940 (2015).

101. A. Alshaya *et al.*, "Spatial resolution and contrast enhancement in photoacoustic imaging with filter delay multiply and sum beamforming technique," in *IEEE International Ultrasonics Symposium (IUS)* (IEEE, 2016), p. 1.

102. T. Kirchner *et al.*, "Signed real-time delay multiply and sum beamforming for multispectral photoacoustic imaging," *J. Imaging* **4**, 121 (2018).

103. S. Mulani, S. Paul, and M. S. Singh, "Higher-order correlation based real-time beamforming in photoacoustic imaging," *J. Opt. Soc. Am. A* **39**, 1805 (2022).

104. S. Jeon *et al.*, "Real-time delay-multiply-and-sum beamforming with coherence factor for in vivo clinical photoacoustic imaging of humans," *Photoacoustics* **15**, 100136 (2019).

105. S. Paul *et al.*, "Simplified delay multiply and sum based promising beamformer for real-time photoacoustic imaging," *IEEE Trans. Instrum. Meas.* **71**, 1 (2022).

106. M. A. Lediju *et al.*, "Short-lag spatial coherence of backscattered echoes: Imaging characteristics," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **58**, 1377 (2011).

107. M. A. Lediju Bell *et al.*, "In vivo visualization of prostate brachytherapy seeds with photoacoustic imaging," *J. Biomed. Opt.* **19**, 126011 (2014).

108. M. T. Graham and M. A. L. Bell, "Theoretical application of short-lag spatial coherence to photoacoustic imaging," in *IEEE International Ultrasonics Symposium (IUS)* (IEEE, 2017), p. 1.

109. M. T. Graham and M. A. L. Bell, "Photoacoustic spatial coherence theory and applications to coherence-based image contrast and resolution," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **67**, 2069 (2020).

110. J. Tordera Mora *et al.*, "Generalized spatial coherence reconstruction for photoacoustic computed tomography," *J. Biomed. Opt.* **26**, 046002 (2021).

111. J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE* **57**, 1408 (1969).

112. J. Mann and W. Walker, "A constrained adaptive beamformer for medical ultrasound: Initial results," in *Ultrasonics Symposium* (IEEE, 2002), p. 1807.

113. M. Sasso and C. Cohen-Bacrie, "Medical ultrasound imaging using the fully adaptive beamformer," in *IEEE International Conference on Acoustics, Speech, and Signal Processing* (IEEE, 2005), p. ii/489.

114. J. F. Synnevag, A. Austeng, and S. Holm, "Adaptive beamforming applied to medical ultrasound imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **54**, 1606 (2007).

115. I. K. Holfort, F. Gran, and J. A. Jensen, "Broadband minimum variance beamforming for ultrasound imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **56**, 314 (2009).

116. B. M. Asl and A. Mahloojifar, "Minimum variance beamforming combined with adaptive coherence weighting applied to medical ultrasound imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **56**, 1923 (2009).

117. S. Park *et al.*, "Adaptive beamforming for photoacoustic imaging," *Opt. Lett.* **33**, 1291 (2008).

118. M. Mozaffarzadeh *et al.*, "Linear-array photoacoustic imaging using minimum variance-based delay multiply and sum adaptive beamforming algorithm," *J. Biomed. Opt.* **23**, 026002 (2018).

119. M. Mozaffarzadeh *et al.*, "Eigenspace-based minimum variance combined with delay multiply and sum beamformer: application to linear-array photoacoustic imaging," *IEEE J. Sel. Top. Quantum Electron* **25**, 1 (2018).

120. R. Al Mukaddim, R. Ahmed, and T. Varghese, "Improving minimum variance beamforming with sub-aperture processing for photoacoustic imaging," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2021), p. 2879.

121. O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE* **60**, 926 (1972).

122. R. Mallart and M. Fink, "Adaptive focusing in scattering media through sound-speed inhomogeneities: the van cittert zernike approach and focusing criterion," *J. Acoust. Soc. Am.* **96**, 3721 (1994).

123. K. W. Hollman, K. Rigby, and M. O'donnell, "Coherence factor of speckle from a multi-row probe," in *Ultrasonics Symposium* (IEEE, 1999), p. 1257.

124. P.-C. Li and M.-L. Li, "Adaptive imaging using the generalized coherence factor," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **50**, 128 (2003).

125. Y.-H. Wang and P.-C. Li, "SNR-dependent coherence-based adaptive imaging for high-frame-rate ultrasonic and photoacoustic imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **61**, 1419 (2014).

126. D. Wang *et al.*, "Coherent-weighted three-dimensional image reconstruction in linear-array-based photoacoustic tomography," *Biomed. Opt. Express* **7**, 1957 (2016).

127. M. Mozaffarzadeh *et al.*, "Image improvement in linear-array photoacoustic imaging using high resolution coherence factor weighting technique," *BMC Biomed. Eng.* **1**, 10 (2019).

128. S. Paul, S. Mandal, and M. S. Singh, "Noise adaptive beamforming for linear array photoacoustic imaging," *IEEE Trans. Instrum. Meas.* **70**, 1 (2021).

129. R. A. Mukaddim and T. Varghese, "Spatiotemporal coherence weighting for in vivo cardiac photoacoustic image beamformation," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **68**, 586 (2021).

130. M. Mozaffarzadeh *et al.*, "Enhanced linear-array photoacoustic beamforming using modified coherence factor," *J. Biomed. Opt.* **23**, 026005 (2018).

131. S. Shamekhi *et al.*, "Eigenspace-based minimum variance beamformer combined with sign coherence factor: application to linear-array photoacoustic imaging," *Ultrasonics* **108**, 106174 (2020).

132. X. Ma *et al.*, "Multiple delay and sum with enveloping beamforming algorithm for photoacoustic imaging," *IEEE Trans. Med. Imaging* **39**, 1812 (2020).

133. Q. Mao *et al.*, "Improving photoacoustic imaging in low signal-to-noise ratio by using spatial and polarity coherence," *Photoacoustics* **28**, 100427 (2022).

134. M. Xu and L. V. Wang, "Pulsed-microwave-induced thermoacoustic tomography: filtered backprojection in a circular measurement configuration," *Med. Phys.* **29**, 1661 (2002).

135. M. Xu and L. V. Wang, "Time-domain reconstruction for thermoacoustic tomography in a spherical geometry," *IEEE Trans. Med. Imaging* **21**, 814 (2002).

136. M. Xu, Y. Xu, and L. V. Wang, "Time-domain reconstruction algorithms and numerical simulations for thermoacoustic tomography in various geometries," *IEEE Trans. Biomed. Eng.* **50**, 1086 (2003).

137. K. Shen *et al.*, "Negativity artifacts in back-projection based photoacoustic tomography," *J. Phys. D: Appl. Phys.* **54**, 074001 (2020).

138. R. Gao *et al.*, "Restoring the imaging quality of circular transducer array-based PACT using synthetic aperture focusing technique integrated with 2nd-derivative-based back projection scheme," *Photoacoustics* **32**, 100537 (2023).

139. L. A. Kunyansky, "Explicit inversion formulae for the spherical mean Radon transform," *Inverse Probl.* **23**, 373 (2007).

140. D. Finch, M. Haltmeier, and Rakesh, "Inversion of spherical means and the wave equation in even dimensions," *SIAM J. Appl. Math.* **68**, 392 (2007).

141. L. V. Nguyen, "A family of inversion formulas in thermoacoustic tomography," *Inverse Prob. Imaging* **3**, 649 (2009).

142. P. Burgholzer *et al.*, "Temporal back-projection algorithms for photoacoustic tomography with integrating line detectors," *Inverse Probl.* **23**, S65 (2007).

143. F. Natterer, "Photo-acoustic inversion in convex domains," *Inverse Prob. Imaging* **6**, 1 (2012).

144. V. P. Palamodov, "A uniform reconstruction formula in integral geometry," *Inverse Probl.* **28**, 065014 (2012).

145. M. Haltmeier, "Exact reconstruction formula for the spherical mean Radon transform on ellipsoids," *Inverse Probl.* **30**, 105006 (2014).

146. M. Haltmeier, "Universal inversion formulas for recovering a function from spherical means," *SIAM J. Math. Anal.* **46**, 214 (2014).

147. Y. Salman, "An inversion formula for the spherical mean transform with data on an ellipsoid in two and three dimensions," *J. Math. Anal. Appl.* **420**, 612 (2014).

148. X. L. Deán-Ben, A. Ozbek, and D. Razansky, "Volumetric real-time tracking of peripheral human vasculature with GPU-accelerated three-dimensional optoacoustic tomography," *IEEE Trans. Med. Imaging* **32**, 2050 (2013).

149. J. Yuan *et al.*, "Real-time photoacoustic and ultrasound dual-modality imaging system facilitated with graphics processing unit and code parallel optimization," *J. Biomed. Opt.* **18**, 86001 (2013).

150. X. L. Deán-Ben, H. López-Schier, and D. Razansky, "Optoacoustic micro-tomography at 100 volumes per second," *Sci. Rep.* **7**, 6850 (2017).

151. Y. Zhang and L. Wang, "Video-rate ring-array ultrasound and photoacoustic tomography," *IEEE Trans. Med. Imaging* **39**, 4369 (2020).

152. Y. Wang and C. Li, "Comprehensive framework of GPU-accelerated image reconstruction for photoacoustic computed tomography," *J. Biomed. Opt.* **29**, 066006 (2024).

153. Y. Zhang *et al.*, "Video-rate dual-modal wide-beam harmonic ultrasound and photoacoustic computed tomography," *IEEE Trans. Med. Imaging* **41**, 727 (2021).

154. Z. Gao *et al.*, "Implementation and comparison of three image reconstruction algorithms in FPGA towards palm-size photoacoustic tomography," *IEEE Sens. J.* **23**, 8605 (2023).

155. M. Xu, G. Ku, and L. V. Wang, "Microwave-induced thermoacoustic tomography using multi-sector scanning," *Med. Phys.* **28**, 1958 (2001).

156. S. J. Norton, "Reconstruction of a two-dimensional reflecting medium over a circular domain: exact solution," *J. Acoust. Soc. Am.* **67**, 1266 (1980).

157. S. J. Norton and M. Linzer, "Ultrasonic reflectivity imaging in three dimensions: exact inverse scattering solutions for plane, cylindrical, and spherical apertures," *IEEE Trans. Biomed. Eng.* **BME-28**, 202 (1981).

158. Y. Xu, D. Feng, and L. V. Wang, "Exact frequency-domain reconstruction for thermoacoustic tomography. I. planar geometry," *IEEE Trans. Med. Imaging* **21**, 823 (2002).

159. M. Haltmeier, O. Scherzer, and G. Zangerl, "A reconstruction algorithm for photoacoustic imaging based on the nonuniform FFT," *IEEE Trans. Med. Imaging* **28**, 1727 (2009).

160. Y. Xu, M. Xu, and L. V. Wang, "Exact frequency-domain reconstruction for thermoacoustic tomography. II. cylindrical geometry," *IEEE Trans. Med. Imaging* **21**, 829 (2002).

161. M. Haltmeier *et al.*, "Thermoacoustic tomography and the circular Radon transform: exact inversion formula," *Math. Models Methods Appl. Sci.* **17**, 635 (2007).

162. D. L. Colton, R. Kress, and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Vol. **93** (Springer, 1998).

163. R. Suda and M. Takami, "A fast spherical harmonics transform algorithm," *Math. Comput.* **71**, 703 (2002).

164. M. A. Anastasio *et al.*, "Application of inverse source concepts to photoacoustic tomography," *Inverse Probl.* **23**, S21 (2007).

165. K. Wang and M. A. Anastasio, "A simple Fourier transform-based reconstruction formula for photoacoustic computed tomography with a circular or spherical measurement geometry," *Phys. Med. Biol.* **57**, N493 (2012).

166. G. Beylkin, "On representations of the Helmholtz Green's function," *Appl. Comput. Harmon. Anal.* **70**, 101633 (2024).

167. M. Agranovsky and P. Kuchment, "Uniqueness of reconstruction and an inversion procedure for thermoacoustic and photoacoustic tomography with variable sound speed," *Inverse Probl.* **23**, 2089 (2007).

168. G. Zangerl, O. Scherzer, and M. Haltmeier, "Circular integrating detectors in photo and thermoacoustic tomography," *Inverse Probl. Sci. Eng.* **17**, 133 (2009).

169. G. Zangerl, O. Scherzer, and M. Haltmeier, "Exact series reconstruction in photoacoustic tomography with circular integrating detectors," *Commun. Math. Sci.* **7**, 665 (2008).

170. B. E. Treeby, E. Z. Zhang, and B. T. Cox, "Photoacoustic tomography in absorbing acoustic media using time reversal," *Inverse Probl.* **26**, 115003 (2010).

171. Y. Hristova, P. Kuchment, and L. Nguyen, "Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media," *Inverse Probl.* **24**, 055006 (2008).

172. B. T. Cox and B. E. Treeby, "Artifact trapping during time reversal photoacoustic imaging for acoustically heterogeneous media," *IEEE Trans. Med. Imaging* **29**, 387 (2010).

173. B. T. Cox *et al.*, "K-space propagation models for acoustically heterogeneous media: application to biomedical photoacoustics," *J. Acoust. Soc. Am.* **121**, 3453 (2007).

174. P. Stefanov and G. Uhlmann, "Thermoacoustic tomography with variable sound speed," *Inverse Probl.* **25**, 075011 (2009).

175. J. Qian *et al.*, "An efficient neumann series–based algorithm for thermoacoustic and photoacoustic tomography with variable sound speed," *SIAM J. Imaging Sci.* **4**, 850 (2011).

176. S. R. Arridge *et al.*, "On the adjoint operator in photoacoustic tomography," *Inverse Probl.* **32**, 115012 (2016).

177. A. Rosenthal, V. Ntziachristos, and D. Razansky, "Optoacoustic methods for frequency calibration of ultrasonic sensors," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **58**, 316 (2011).

178. X. L. Dean-Ben *et al.*, "Accurate model-based reconstruction algorithm for three-dimensional optoacoustic tomography," *IEEE Trans. Med. Imaging* **31**, 1922 (2012).

179. K. Wang *et al.*, "Investigation of iterative image reconstruction in optoacoustic tomography," in *Photons Plus Ultrasound: Imaging and Sensing* (SPIE, 2012), p. 379.

180. J. Zhang *et al.*, "Effects of different imaging models on least-squares image reconstruction accuracy in photoacoustic tomography," *IEEE Trans. Med. Imaging* **28**, 1781 (2009).

181. K. Wang *et al.*, "Investigation of iterative image reconstruction in three-dimensional optoacoustic tomography," *Phys. Med. Biol.* **57**, 5399 (2012).

182. R. M. Lewitt, "Multidimensional digital image representations using generalized kaiser–bessel window functions," *J. Opt. Soc. Am. A* **7**, 1834 (1990).

183. M. Schweiger and S. R. Arridge, "Image reconstruction in optical tomography using local basis functions," *J. Electron. Imaging* **12**, 583 (2003).

184. G. B. Arfken, H. J. Weber, and F. E. Harris, *Mathematical Methods for Physicists: a Comprehensive Guide* (Academic press, 2011).

185. S. Matej and R. M. Lewitt, "Practical considerations for 3-D image reconstruction using spherically symmetric volume elements," *IEEE Trans. Med. Imaging* **15**, 68 (1996).

186. C. B. Shaw *et al.*, "Least squares QR-based decomposition provides an efficient way of computing optimal regularization parameter in photoacoustic tomography," *J. Biomed. Opt.* **18**, 080501 (2013).

187. P. R. Stepanishen, "Transient radiation from pistons in an infinite planar baffle," *J. Acoust. Soc. Am.* **49**, 1629 (1971).

188. J. C. Lockwood and J. G. Willette, "High-speed method for computing the exact solution for the pressure variations in the near-field of a baffled piston," *J. Acoust. Soc. Am.* **53**, 735 (1973).

189. K. Mitsuhashi, K. Wang, and M. A. Anastasio, "Investigation of the far-field approximation for modeling a transducer's spatial impulse response in photoacoustic computed tomography," *Photoacoustics* **2**, 21 (2014).

190. R. P. K. Jagannath and P. K. Yalavarthy, "Minimal residual method provides optimal regularization parameter for diffuse optical tomography," *J. Biomed. Opt.* **17**, 106015 (2012).

191. P. C. Hansen, "The L-curve and its use in the numerical treatment of inverse problems," in *Computational Inverse Problems in Electrocardiology* (WIT Press, 2001), p. 119.

192. D. Calvetti *et al.*, "Tikhonov regularization and the l-curve for large discrete ill-posed problems," *J. Comput. Appl. Math.* **123**, 423 (2000).

193. L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proceedings of COMPSTAT'2010* (Springer, 2010), p. 177.

194. L. Bottou, "Stochastic gradient descent tricks," in *Neural Networks: Tricks of the Trade* (Springer, 2012), p. 421.

195. C. C. Paige and M. A. Saunders, "LSQR: an algorithm for sparse linear equations and sparse least squares," *ACM Trans. Math. Software* **8**, 43 (1982).

196. O. Axelsson, "A generalized conjugate gradient, least square method," *Numer. Math.* **51**, 209 (1987).

197. T. Wang *et al.*, "Learned regularization for image reconstruction in sparse-view photoacoustic tomography," *Biomed. Opt. Express* **13**, 5721 (2022).

198. J. Provost and F. Lesage, "The application of compressed sensing for photo-acoustic tomography," *IEEE Trans. Med. Imaging* **28**, 585 (2009).

199. Z. Guo *et al.*, "Compressed sensing in photoacoustic tomography in vivo," *J. Biomed. Opt.* **15**, 021311 (2010).

200. F. Lucka *et al.*, "Enhancing compressed sensing 4D photoacoustic tomography by simultaneous motion estimation," *SIAM J. Imaging Sci.* **11**, 2224 (2018).

201. S. Biton *et al.*, "Optoacoustic model-based inversion using anisotropic adaptive total-variation regularization," *Photoacoustics* **16**, 100142 (2019).

202. A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.* **2**, 183 (2009).

203. I. Daubechies *et al.*, "Iteratively reweighted least squares minimization for sparse recovery," *Commun. Pure Appl. Math.* **63**, 1 (2010).

204. Y. Wang, W. Yin, and J. Zeng, "Global convergence of ADMM in nonconvex nonsmooth optimization," *J. Sci. Comput.* **78**, 29 (2019).

205. X. Li *et al.*, "Model-based optoacoustic tomography image reconstruction with non-local and sparsity regularizations," *IEEE Access* **7**, 102136 (2019).

206. P. K. Yalavarthy *et al.*, "Non-local means improves total-variation constrained photoacoustic image reconstruction," *J. Biophotonics* **14**, e202000191 (2021).

207. J. Prakash *et al.*, "Maximum entropy based non-negative optoacoustic tomographic image reconstruction," *IEEE Trans. Biomed. Eng.* **66**, 2604 (2019).

208. H. Liu *et al.*, "Curve-driven-based acoustic inversion for photoacoustic tomography," *IEEE Trans. Med. Imaging* **35**, 2546 (2016).

209. A. Javaherian and S. Holman, "A multi-grid iterative method for photoacoustic tomography," *IEEE Trans. Med. Imaging* **36**, 696 (2016).

210. X. L. Dean-Ben, V. Ntziachristos, and D. Razansky, "Acceleration of optoacoustic model-based reconstruction using angular image discretization," *IEEE Trans. Med. Imaging* **31**, 1154 (2012).

211. A. Buehler *et al.*, "Model-based optoacoustic inversions with incomplete projection data," *Med. Phys.* **38**, 1694 (2011).

212. A. Voulodimos *et al.*, "Deep learning for computer vision: a brief review," *Comput. Intell. Neurosci.* **2018**, 7068349 (2018).

213. D. W. Otter, J. R. Medina, and J. K. Kalita, "A survey of the usages of deep learning for natural language processing," *IEEE Trans. Neural Netw. Learn. Syst.* **32**, 604 (2021).

214. S. S. Nisha and N. Meeral, "Applications of deep learning in biomedical engineering," in *Handbook of Deep Learning in Biomedical Engineering Techniques and Applications* (Academic Press, 2020), p. 245.

215. G. Wang, J. C. Ye, and B. De Man, "Deep learning for tomographic image reconstruction," *Nat. Mach. Intell.* **2**, 737 (2020).

216. G. Wang, "A perspective on deep imaging," *IEEE Access* **4**, 8914 (2016).

217. G. Wang *et al.*, "Deep tomographic image reconstruction: yesterday, today, and tomorrow—editorial for the 2nd special issue 'Machine Learning for Image Reconstruction'," *IEEE Trans. Med. Imaging* **40**, 2956 (2021).

218. D. Liang *et al.*, "Deep magnetic resonance image reconstruction: inverse problems meet neural networks," *IEEE Signal Process Mag.* **37**, 141 (2020).

219. E. M. A. Anas *et al.*, "Enabling fast and high quality LED photoacoustic imaging: a recurrent neural networks based approach," *Biomed. Opt. Express* **9**, 3852 (2018).

220. A. Hariri *et al.*, "Deep learning improves contrast in low-fluence photoacoustic imaging," *Biomed. Opt. Express* **11**, 3360 (2020).

221. M. Yamakawa and T. Shiina, "Artifact reduction in photoacoustic images by generating virtual dense array sensor from hemispheric sparse array sensor using deep learning," *J. Med. Ultrason.* **51**, 169 (2024).

222. J. Zhang *et al.*, "PAFormer: photoacoustic reconstruction via transformer with mask mechanism (IUS)," in *IEEE International Ultrasonics Symposium* (IEEE, 2022), p. 1.

223. F. Zhang *et al.*, "Photoacoustic digital brain and deep-learning-assisted image reconstruction," *Photoacoustics* **31**, 100517 (2023).

224. P. Farnia *et al.*, "High-quality photoacoustic image reconstruction based on deep convolutional neural network: towards intra-operative photoacoustic imaging," *Biomed. Phys. Eng. Express* **6**, 045019 (2020).

225. S. Guan *et al.*, "Fully dense UNet for 2-D sparse photoacoustic tomography artifact removal," *IEEE J. Biomed. Health. Inf.* **24**, 568 (2020).

226. J. Zhang *et al.*, "Limited-view photoacoustic imaging reconstruction with dual domain inputs under mutual information constraint," arXiv:2011.06147 (2020).

227. S. Guan *et al.*, "Dense dilated UNet: deep learning for 3D photoacoustic tomography image reconstruction," arXiv:2104.03130 (2021).

228. A. Creswell *et al.*, "Generative adversarial networks: an overview," *IEEE Signal Process Mag.* **35**, 53 (2018).

229. T. Vu *et al.*, "A generative adversarial network for artifact removal in photoacoustic computed tomography with a linear-array transducer," *Commun. Pure Appl. Math.* **245**, 597 (2020).

230. H. Shahid *et al.*, "Feasibility of a generative adversarial network for artifact removal in experimental photoacoustic imaging," *Ultrasound Med. Biol.* **48**, 1628 (2022).

231. M. Lu *et al.*, "Artifact removal in photoacoustic tomography with an unsupervised method," *Biomed. Opt. Express* **12**, 6284 (2021).

232. H. Shan, G. Wang, and Y. Yang, "Accelerated correction of reflection artifacts by deep neural networks in photoacoustic tomography," *Appl. Sci.* **9**, 2615 (2019).

233. S. Jeon *et al.*, "A deep learning-based model that reduces speed of sound aberrations for improved in vivo photoacoustic imaging," *IEEE Trans. Image Process.* **30**, 8773 (2021).

234. Y. Gao *et al.*, "Deep learning-based photoacoustic imaging of vascular network through thick porous media," *IEEE Trans. Med. Imaging* **41**, 2191 (2022).

235. V. Shijo *et al.*, "SwinIR for photoacoustic computed tomography artifact reduction," in *IEEE International Ultrasonics Symposium (IUS)* (IEEE, 2023), p. 1.

236. W. Zheng *et al.*, "Deep-E enhanced photoacoustic tomography using three-dimensional reconstruction for high-quality vascular imaging," *Sensors* **22**, 7725 (2022).

237. C. Cai *et al.*, "End-to-end deep neural network for optical inversion in quantitative photoacoustic imaging," *Opt. Lett.* **43**, 2752 (2018).

238. I. Olefir *et al.*, "Deep learning-based spectral unmixing for optoacoustic imaging of tissue oxygen saturation," *IEEE Trans. Med. Imaging* **39**, 3643 (2020).

239. C. Bench, A. Hauptmann, and B. T. Cox, "Toward accurate quantitative photoacoustic imaging: learning vascular blood oxygen saturation in three dimensions," *J. Biomed. Opt.* **25**, 085003 (2020).

240. C. Yang and F. Gao, "EDA-Net: dense aggregation of deep and shallow information achieves quantitative photoacoustic blood oxygenation imaging deep in human breast," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019* (Springer, 2019), p. 246.

241. C. Yang *et al.*, "Quantitative photoacoustic blood oxygenation imaging using deep residual and recurrent neural network," in *IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* (IEEE, 2019), p. 741.

242. Z. Wang *et al.*, "Extractor-attention-predictor network for quantitative photoacoustic tomography," *Photoacoustics* **38**, 100609 (2024).

243. N. Awasthi *et al.*, "PA-fuse: a deep supervised approach for fusion of photoacoustic images with distinct reconstruction characteristics," *Biomed. Opt. Express* **10**, 2227 (2019).

244. J. Zhang *et al.*, "Photoacoustic image classification and segmentation of breast cancer: a feasibility study," *IEEE Access* **7**, 5457 (2019).

245. N.-K. Chlis *et al.*, "A sparse deep learning approach for automatic segmentation of human vasculature in multispectral optoacoustic tomography," *Photoacoustics* **20**, 100203 (2020).

246. B. Lafci *et al.*, "Deep learning for automatic segmentation of hybrid optoacoustic ultrasound (OPUS) images," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **68**, 688 (2021).

247. M. G. González, M. Vera, and L. J. R. Vega, "Combining band-frequency separation and deep neural networks for optoacoustic imaging," *Opt. Lasers Eng.* **163**, 107471 (2023).

248. H. Shahid *et al.*, "A deep learning approach for the photoacoustic tomography recovery from undersampled measurements," *Front. Neurosci.* **15**, 598693 (2021).

249. G. Godefroy, B. Arnal, and E. Bossy, "Compensating for visibility artefacts in photoacoustic imaging with a deep learning approach providing prediction uncertainties," *Photoacoustics* **21**, 100218 (2021).

250. H. Zhang *et al.*, "A new deep learning network for mitigating limited-view and under-sampling artifacts in ring-shaped photoacoustic tomography," *Comput. Med. Imaging Graph.* **84**, 101720 (2020).

251. P. Rajendran and M. Pramanik, "High frame rate (approximately 3 Hz) circular photoacoustic tomography using single-element ultrasound transducer aided with deep learning," *J. Biomed. Opt.* **27**, 066005 (2022).

252. P. Rajendran and M. Pramanik, "Deep-learning-based multi-transducer photoacoustic tomography imaging without radius calibration," *Opt. Lett.* **46**, 4510 (2021).

253. H. Lan *et al.*, "Ki-GAN: knowledge infusion generative adversarial network for photoacoustic image reconstruction in vivo," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference* (Springer, 2019), p. 273.

254. H. Lan *et al.*, "Y-Net: hybrid deep learning image reconstruction for photoacoustic tomography in vivo," *Photoacoustics* **20**, 100197 (2020).

255. N. Davoudi *et al.*, "Deep learning of image and time-domain data enhances the visibility of structures in optoacoustic tomography," *Opt. Lett.* **46**, 3029 (2021).

256. M. Guo *et al.*, "AS-Net: fast photoacoustic reconstruction with multi-feature fusion from sparse data," *IEEE Trans. Comput. Imaging* **8**, 215 (2022).

257. W. Li *et al.*, "Deep learning reconstruction algorithm based on sparse photoacoustic tomography system," in *IEEE International Ultrasonics Symposium (IUS)* (IEEE, 2021), p. 1.

258. H. Lan, C. Yang, and F. Gao, "A jointed feature fusion framework for photoacoustic image reconstruction," *Photoacoustics* **29**, 100442 (2023).

259. H. Li *et al.*, "NETT: solving inverse problems with deep neural networks," *Inverse Probl.* **36**, 065005 (2020).

260. S. Antholzer *et al.*, "NETT regularization for compressed sensing photoacoustic tomography," in *Photons Plus Ultrasound: Imaging and Sensing*, (SPIE, 2019), p. 272.

261. C. Yang, H. Lan, and F. Gao, "Accelerated photoacoustic tomography reconstruction via recurrent inference machines," in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (IEEE, 2019), p. 6371.

262. H. Lan *et al.*, "Compressed sensing for photoacoustic computed tomography based on an untrained neural network with a shape prior," *Biomed. Opt. Express* **12**, 7835 (2021).

263. H. Lan, J. Gong, and F. Gao, "Deep learning adapted acceleration for limited-view photoacoustic image reconstruction," *Opt. Lett.* **47**, 1911 (2022).

264. H. Shan, G. Wang, and Y. Yang, "Simultaneous reconstruction of the initial pressure and sound speed in photoacoustic tomography using a deep-learning approach," *Proc. SPIE* **11105**, 1110504 (2019).

265. Y. E. Boink, S. Manohar, and C. Brune, "A partially-learned algorithm for joint photo-acoustic reconstruction and segmentation," *IEEE Trans. Med. Imaging* **39**, 129 (2020).

266. J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Adv. Neural Inf. Process. Syst.* **33**, 6840 (2020).

267. Z. Luo *et al.*, "Image restoration with mean-reverting stochastic differential equations," arXiv:2301.11699 (2023).

268. K. Guo *et al.*, "Score-based generative model-assisted information compensation for high-quality limited-view reconstruction in photoacoustic tomography," *Photoacoustics* **38**, 100623 (2024).

269. S. Tong *et al.*, "Score-based generative models for photoacoustic image reconstruction with rotation consistency constraints," arXiv:2306.13843 (2023).

270. S. Dey *et al.*, "Score-based diffusion models for photoacoustic tomography image reconstruction," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, 2024), p. 2470.

271. X. Song *et al.*, "Sparse-view reconstruction for photoacoustic tomography combining diffusion model with model-based iteration," *Photoacoustics* **33**, 100558 (2023).

272. D. Waibel *et al.*, "Reconstruction of initial pressure from limited view photoacoustic images using deep learning," in *Photons Plus Ultrasound: Imaging and Sensing* (SPIE, 2018), p. 196.

273. H. Lan *et al.*, "Reconstruct the photoacoustic image based on deep learning with multi-frequency ring-shape transducer array," in *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE, 2019), p. 7115.

274. J. Feng *et al.*, "End-to-end Res-Unet based reconstruction algorithm for photoacoustic imaging," *Biomed. Opt. Express* **11**, 5321 (2020).

275. H. Lan *et al.*, "Deep learning enabled real-time photoacoustic tomography system via single data acquisition channel," *Photoacoustics* **22**, 100270 (2021).

276. K. Shen *et al.*, "Physics-driven deep learning photoacoustic tomography," *Fundam. Res.* (2024).

277. H. Lan *et al.*, "Masked cross-domain self-supervised deep learning framework for photoacoustic computed tomography reconstruction," *Neural Netw.* **179**, 106515 (2024).

278. S. Guan *et al.*, "Limited-view and sparse photoacoustic tomography for neuroimaging with deep learning," *Sci. Rep.* **10**, 8510 (2020).

279. M. Kim *et al.*, "Deep-learning image reconstruction for real-time photoacoustic system," *IEEE Trans. Med. Imaging* **39**, 3379 (2020).

280. T. Tong *et al.*, "Domain transform network for photoacoustic tomography from limited-view and sparsely sampled data," *Photoacoustics* **19**, 100190 (2020).

281. C. Dehner *et al.*, "DeepMB: deep neural network for real-time model-based optoacoustic image reconstruction with adjustable speed of sound," arXiv:2206.14485 (2022).

282. C. Dehner *et al.*, "Deep model-based optoacoustic image reconstruction (DeepMB)," in *Photons Plus Ultrasound: Imaging and Sensing* (SPIE, 2024), p. 66.

283. D. Allman, A. Reiter, and M. A. L. Bell, "Photoacoustic source detection and reflection artifact removal enabled by deep learning," *IEEE Trans. Med. Imaging* **37**, 1464 (2018).

284. X. Luo *et al.*, "Fast correction of "finite aperture effect" in photoacoustic tomography based on spatial impulse response," *Photonics* **8**, 356 (2021).

285. B. Wang *et al.*, "Approximate back-projection method for improving lateral resolution in circular-scanning-based photoacoustic tomography," *Med. Phys.* **48**, 3011 (2021).

286. M.-L. Li, Y.-C. Tseng, and C.-C. Cheng, "Model-based correction of finite aperture effect in photoacoustic tomography," *Opt. Express* **18**, 26285 (2010).

287. V. G. Andreev *et al.*, "Detection of optoacoustic transients with a rectangular transducer of finite dimensions," in *Biomedical Optoacoustics III* (SPIE, 2002), p. 153.

288. S. A. Ermilov *et al.*, "Development of laser optoacoustic and ultrasonic imaging system for breast cancer utilizing handheld array probes," in *Photons Plus Ultrasound: Imaging and Sensing* (SPIE, 2009), p. 28.

289. K. B. Chowdhury *et al.*, "A synthetic total impulse response characterization method for correction of hand-held optoacoustic images," *IEEE Trans. Med. Imaging* **39**, 3218 (2020).

290. L. Qi *et al.*, "Photoacoustic tomography image restoration with measured spatially variant point spread functions," *IEEE Trans. Med. Imaging* **40**, 2318 (2021).

291. D. Xie *et al.*, "Spatially-variant image deconvolution for photoacoustic tomography," *Opt. Express* **31**, 21641 (2023).

292. W. Dong *et al.*, "Image restoration for ring-array photoacoustic tomography system based on blind spatially rotational deconvolution," *Photoacoustics* **38**, 100607 (2024).

293. K. T. Hsu, S. Guan, and P. V. Chitnis, "Comparing deep learning frameworks for photoacoustic tomography image reconstruction," *Photoacoustics* **23**, 100271 (2021).

294. T. Wang *et al.*, "Sparse-view photoacoustic image quality enhancement based on a modified U-Net," *Laser Optoelectron. Prog.* **59**, 0617022 (2022).

295. T. Wang *et al.*, "Streak artifact suppressed back projection for sparse-view photoacoustic computed tomography," *Appl. Opt.* **62**, 3917 (2023).

296. Y. Zhao *et al.*, "Ultrasound-guided adaptive photoacoustic tomography," *Opt. Lett.* **47**, 3960 (2022).

297. M. Sandbichler *et al.*, "A novel compressed sensing scheme for photoacoustic tomography," *SIAM J. Appl. Math.* **75**, 2475 (2015).

298. J. Meng *et al.*, "High-speed, sparse-sampling three-dimensional photoacoustic computed tomography in vivo based on principal component analysis," *J. Biomed. Opt.* **21**, 076007 (2016).

299. P. Hu *et al.*, "Spatiotemporal antialiasing in photoacoustic computed tomography," *IEEE Trans. Med. Imaging* **39**, 3535 (2020).

300. P. Hu, L. Li, and L. V. Wang, "Location-dependent spatiotemporal antialiasing in photoacoustic computed tomography," *IEEE Trans. Med. Imaging* **42**, 1210 (2022).

301. C. Cai *et al.*, "Streak artifact suppression in photoacoustic computed tomography using adaptive back projection," *Biomed. Opt. Express* **10**, 4803 (2019).

302. S. Hakakzadeh *et al.*, "A spatial-domain factor for sparse-sampling circular-view photoacoustic tomography," *IEEE J. Sel. Top. Quantum Electron.* **29**, 1 (2022).

303. Y. Wang *et al.*, "Enhancing sparse-view photoacoustic tomography with combined virtually parallel projecting and spatially adaptive filtering," *Biomed. Opt. Express* **9**, 4569 (2018).

304. S. K. Patch, "Thermoacoustic tomography-consistency conditions and the partial scan problem," *Phys. Med. Biol.* **49**, 2305 (2004).

305. J. K. Gamelin, A. Aguirre, and Q. Zhu, "Fast, limited-data photoacoustic imaging for multiplexed systems using a frequency-domain estimation technique," *Med. Phys.* **38**, 1503 (2011).

306. G. Paltauf, R. Nuster, and P. Burgholzer, "Weight factors for limited angle photoacoustic tomography," *Phys. Med. Biol.* **54**, 3303 (2009).

307. T. Wang, W. Liu, and C. Tian, "Combating acoustic heterogeneity in photoacoustic computed tomography: a review," *J. Innov. Opt. Health Sci.* **13**, 2030007 (2020).

308. Y. Zhang and L. Wang, "Adaptive dual-speed ultrasound and photoacoustic computed tomography," *Photoacoustics* **27**, 100380 (2022).

309. C. Cai *et al.*, "Feature coupling photoacoustic computed tomography for joint reconstruction of initial pressure and sound speed in vivo," *Biomed. Opt. Express* **10**, 3447 (2019).

310. J. Wei, K. Shen, and C. Tian, "Comparisons of filtered back-projection and time reversal algorithms in photoacoustic tomography," in *Second Conference on Biomedical Photonics and Cross-Fusion (BPC 2023)* (SPIE, 2023), p. 68.

311. S. Liu *et al.*, *Ultrafast Filtered Back-Projection for Photoacoustic Computed Tomography* (2024).

312. D. Van de Sompel *et al.*, "Comparison of deconvolution filters for photoacoustic tomography," *PLoS One* **11**, e0152597 (2016).

313. K. Shen *et al.*, "Dual-domain neural network for sparse-view photoacoustic image reconstruction," *Chin. J. Lasers* **49**, 0507208 (2022).

314. M. A. Anastasio *et al.*, "Half-time image reconstruction in thermoacoustic tomography," *IEEE Trans. Med. Imaging* **24**, 199 (2005).

315. J. Poudel *et al.*, "Mitigation of artifacts due to isolated acoustic heterogeneities in photoacoustic computed tomography using a variable data truncation-based reconstruction method," *J. Biomed. Opt.* **22**, 041018 (2017).

316. B. E. Treeby *et al.*, "Automatic sound speed selection in photoacoustic image reconstruction using an autofocus approach," *J. Biomed. Opt.* **16**, 090501 (2011).

317. T. P. Matthews *et al.*, "Parameterized joint reconstruction of the initial pressure and sound speed distributions for photoacoustic computed tomography," *SIAM J. Imaging Sci.* **11**, 1560 (2018).

318. J. Xia *et al.*, "Enhancement of photoacoustic tomography by ultrasonic computed tomography based on optical excitation of elements of a full-ring transducer array," *Opt. Lett.* **38**, 3140 (2013).

319. E. Merčep *et al.*, "Transmission–reflection optoacoustic ultrasound (TROPUS) computed tomography of small animals," *Light Sci. Appl.* **8**, 1 (2019).

320. Z. Hu *et al.*, "Sound speed imaging of small animal organs by ultrasound computed tomography," *JUSTC* **52**, 8 (2022).

321. J. Li *et al.*, "Deep learning-based quantitative optoacoustic tomography of deep tissues in the absence of labeled experimental data," *Optica* **9**, 32 (2022).

322. B. He *et al.*, "From signal to knowledge: the diagnostic value of rawdata in artificial intelligence prediction of human data for the first time," *Engineering* **34**, 60 (2022).

323. W. Fu *et al.*, "Ultracompact meta-imagers for arbitrary all-optical convolution," *Light Sci. Appl.* **11**, 62 (2022).

324. A. M. Shaltout, V. M. Shalaev, and M. L. Brongersma, "Spatiotemporal light control with active metasurfaces," *Science* **364**, eaat3100 (2019).

325. Q. Jiang, G. Jin, and L. Cao, "When metasurface meets hologram: principle and advances," *Adv. Opt. Photonics* **11**, 518 (2019).

326. S. Sun *et al.*, "Electromagnetic metasurfaces: physics and applications," *Adv. Opt. Photonics* **11**, 380 (2019).

327. Y. Zhao *et al.*, "Ultraviolet metalens for photoacoustic microscopy with an elongated depth of focus," *Opt. Lett.* **48**, 3435 (2023).

328. W. Song *et al.*, "Ultraviolet metasurface-assisted photoacoustic microscopy with great enhancement in DOF for fast histology imaging," *Photoacoustics* **32**, 100525 (2023).

329. A. Barulin *et al.*, "Dual-wavelength UV-visible metalens for multispectral photoacoustic microscopy: a simulation study," *Photoacoustics* **32**, 100545 (2023).

330. C. Tian *et al.*, "Non-contact photoacoustic imaging using a commercial heterodyne interferometer," *IEEE Sens. J.* **16**, 2079 (2016).

331. B. Dong, C. Sun, and H. F. Zhang, "Optical detection of ultrasound in photoacoustic imaging," *IEEE Trans. Biomed. Eng.* **64**, 4 (2016).

332. G. Wissmeyer *et al.*, "Looking at sound: optoacoustics with all-optical ultrasound detection," *Light Sci. Appl.* **7**, 53 (2018).

333. V. V. Yakovlev *et al.*, "Ultrasensitive non-resonant detection of ultrasound with plasmonic metamaterials," *Adv. Mater.* **25**, 2351 (2013).
334. J. Weber, P. C. Beard, and S. E. Bohndiek, "Contrast agents for molecular photoacoustic imaging," *Nat. Methods* **13**, 639 (2016).
335. W. Li and X. Chen, "Gold nanoparticles for photoacoustic imaging," *Nanomed.* **10**, 299 (2015).
336. T.-H. Shin *et al.*, "Recent advances in magnetic nanoparticle-based multi-modal imaging," *Chem. Soc. Rev.* **44**, 4501 (2015).
337. Y. Mantri and J. V. Jokerst, "Engineering plasmonic nanoparticles for enhanced photoacoustic imaging," *ACS Nano* **14**, 9408 (2020).
338. C. Tian *et al.*, "Plasmonic nanoparticles with quantitatively controlled bioconjugation for photoacoustic imaging of live cancer cells," *Adv. Sci.* **3**, 1600237 (2016).
339. V. P. Nguyen *et al.*, "Chain-like gold nanoparticle clusters for multimodal photoacoustic microscopy and optical coherence tomography enhanced molecular imaging," *Nat. Commun.* **12**, 34 (2021).
340. K. Pu *et al.*, "Semiconducting polymer nanoparticles as photoacoustic molecular imaging probes in living mice," *Nat. Nanotechnol.* **9**, 233 (2014).
341. S. Y. Ong *et al.*, "Recent advances in polymeric nanoparticles for enhanced fluorescence and photoacoustic imaging," *Angew. Chem. Int. Ed.* **60**, 17797 (2021).