Imaging Heart Dynamics With Ultrafast Cascaded-Wave Ultrasound

Yang Zhang, *Member, IEEE*, He Li, and Wei-Ning Lee, *Member, IEEE*

Abstract—The heart is an organ with highly dynamic complexity, including cyclic fast electrical activation, muscle kinematics, and blood dynamics. Although ultrafast cardiac imaging techniques based on pulsed-wave ultrasound (PUS) have rapidly emerged to permit mapping of heart dynamics, they suffer from limited sonographic signal-to-noise ratio (SNR) and penetration due to insufficient energy delivery and inevitable attenuation through the chest wall. We hereby propose ultrafast cascaded-wave ultrasound (uCUS) imaging to depict heart dynamics in higher SNR and larger penetration than conventional ultrafast PUS. To solve the known tradeoff between the length of transmitted ultrasound signals and spatial resolution while achieving ultrafast frame rates (>1000 Hz), we develop a cascaded synthetic aperture (CaSA) imaging method. In CaSA, an array probe is divided into subapertures; each subaperture transmits a train of diverging waves. These diverging waves are weighted in both the aperture (i.e., spatial) and range (i.e., temporal) directions with a coding matrix containing only $+1$ and -1 polarity coefficients. A corresponding spatiotemporal decoding matrix is designed to recover backscattered signals. The decoded signals are thereafter beamformed and coherently compounded to obtain one high-SNR beamformed image frame. For CaSA with M subapertures and N cascaded diverging waves, sonographic SNR is increased by $10 \times \log_{10}(N \times M)$ (dB) compared with conventional synthetic aperture (SA) imaging. The proposed uCUS with CaSA was evaluated with conventional SA and Hadamard-encoded SA (H-SA) methods in a calibration phantom for B-mode image quality and an *in vivo* human heart in a transthoracic setting for the quality assessment of anatomical, myocardial motion, and chamber blood power Doppler images. Our results demonstrated that the proposed uCUS with CaSA (4 subapertures, 32 cascaded waves) improved SNR ($+20.46$ dB versus SA, $+14.83$ dB versus H-SA) and contrast ratio ($+8.44$ dB versus SA, $+7.81$ dB versus H-SA) with comparable spatial resolutions to and at the same frame rates as benchmarks.

Index Terms—Cascaded wave, diverging wave, doppler, hadamard, myocardial motion, signal-to-noise ratio (SNR), ultrafast, ultrasound.

I. Introduction

IMAGING techniques for quantifying and mapping heart dynamics have paved the way for the diagnosis of cardiovascular diseases (CVDs), which account for approximately 31% of global deaths and among which myocardial infarction

Y. Zhang and H. Li are with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong.

W.-N. Lee is with the Department of Electrical and Electronic Engineering and the Biomedical Engineering Programme, The University of Hong Kong, Hong Kong (e-mail: wnlee@eee.hku.hk).

is the major form [1]. A recent World Health Organization (WHO) publication not only projects an increase of 4.7 million in the annual number of CVD deaths by 2030 but also reports a high incidence ($>75\%$) of CVD deaths in underprivileged regions and groups [2]. Accessible low-cost technologies are thus deemed crucial for CVD management.

Medical ultrasound imaging remains the predominant clinical cardiac imaging modality because of its affordability and real-time visualization of morphology (i.e., B-mode) and chamber blood dynamics (i.e., Doppler mode) of the beating heart. Beyond these two conventional modes, in past decades, significant technological advances in cardiac ultrasound imaging have been made to study myocardial: 1) kinematics by tissue Doppler imaging [3]–[5] and strain (rate) imaging [6]–[8]; 2) mechanical properties by natural pulse [9], acoustic radiation force (ARF) impulse imaging [10] and ARF-based shear wave imaging [11]–[13]; and 3) perfusion and hemodynamics by ultrafast Doppler and contrast-enhanced ultrasound imaging [14]–[17]. These ultrasound developments are envisioned to be of paramount importance in primary health care for the monitoring and prognosis of heart function.

The core technique that has facilitated the aforementioned technological advances is cardiac ultrasound image acquisition [18], [19] in high frame rate or ultrafast regimes, such as composite methodologies [20], [21], multiline transmission [22], [23], and diverging wave transmission [24]–[27], instead of conventional real-time line-by-line scanning. The ultrafast image acquisition permits assessment of critical instants of the highly dynamic and complex heart. Pulsed-wave ultrasound (PUS), instead of cascaded-wave ultrasound (CUS), is conventionally used in the high frame rate or ultrafast paradigm to trade the length of the transmitted ultrasound signals (i.e., acoustic intensity) for spatial resolution. Conventional ultrafast PUS can be achieved in either focused or unfocused wave transmission modes. One example in the category of focused PUS is multiline transmit imaging [22]. Multiple focused beams are transmitted simultaneously to increase the frame rate, but the crosstalk between the transmitted beams likely causes image artifacts. In the case of unfocused PUS, synthetic aperture (SA) imaging is widely accepted [27], [28]; a subarray diverging wave is transmitted by a single element [28] or a virtual-source-based subaperture [27] to obtain one low-quality image with a large field-of-view (FOV). Multiple low-quality images yielded from different elements or subapertures are coherently summed to obtain a high-quality image. Since the SA uses a single element or subaperture in each transmission, its signal-to-noise ratio (SNR) is lower than that of the full-aperture diverging wave transmission [24].

In order to further increase the SNR, a Hadamard-encoded SA (H-SA) imaging [29], [30] is proposed to fully utilize the spatial domain of the array elements. Unlike SA imaging, which only activates partial array elements in each transmission, H-SA utilizes full aperture in the transmission event with Hadamard coefficients applied to each element or subaperture. In reception, a decoding process is realized to increase the SNR without compromising other primary image characteristics, such as spatial resolution. Even though the aforementioned PUS-based ultrafast imaging methods fully utilize the spatial domain of the array, their sonographic SNR and penetration remain limited in cardiac imaging. It is mainly because of the insufficient energy delivery by the short ultrasound pulses (microseconds) and attenuated ultrasound waves through the chest wall [31], [32] to reach the heart. In conventional real-time ultrasound imaging, coded excitation techniques have been applied to increase the SNR by transmitting long pulses, such as the Chirp and Golay code [33], [34]. However, the Chirp coding and decoding rely on the transducer bandwidth and the prior knowledge of the signal propagation properties of the medium. The Golay coding leads to a reduction of frame rate, which is not optimal for ultrafast imaging.

In this study, we propose an ultrafast CUS (uCUS) for imaging and mapping the heart dynamics with cascaded diverging waves in the SA scheme. This is termed as cascaded SA (CaSA) imaging to tackle the aforementioned dilemma. The proposed method encodes the transmitted ultrasound waves and decodes the backscattered signals in both the spatial and temporal domains to approach the physical limits in terms of SNR. The coding–decoding scheme in the temporal domain of diverging waves in the phased array configuration is developed based on our recent work for the linear array [35]. The spatial domain employs H-SA but in a phased array scenario. The performance of uCUS in both B-mode image quality and quantitative mapping of myocardial motion and chamber blood is evaluated on the human heart *in vivo*.

II. Theory

A theoretical framework is introduced to describe several previous and our proposed models in this study for the transmission, propagation, and reception of signals, and the corresponding SNR in array-based medical imaging scenarios.

In order to analyze the SNR of array-based medical imaging scenarios, we built a general transmission–propagation–reception model as shown in Fig. 1. We divided the array into K subarrays for transmission but used full-array for the reception; there were M transmission–reception events to obtain a final image. In general, the number of transmission–reception events is equal to the number of subarrays, so let us assume $M = K$ for all subsequent analysis. In the mth transmission–reception event, we have the transmission matrix $X_m(t)$ as

$$X_m(t) = [f_{m1}(s_1(t)) \ f_{m2}(s_2(t)) \ \cdots \ f_{mk}(s_K(t))] \quad (1)$$

where $s_k(t)$ is the source signal (i.e., basic transmitted pulse) from the kth subarray, $f_{mk}(s_k(t))$ is the final transmitted
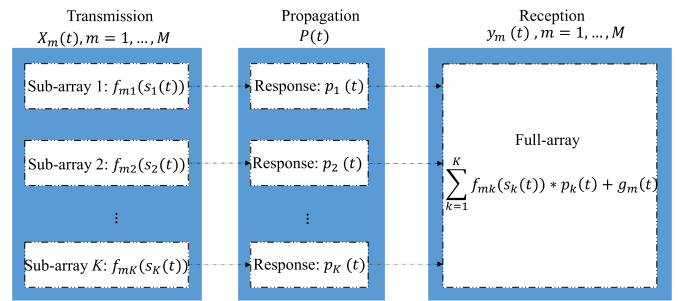

Fig. 1. Transmitter and receiver model. $s_k(t)$ is the source signal (i.e., basic transmitted pulse) from the kth subarray, $f_{mk}(s_k(t))$ is the transmitted signal for the kth subarray for the mth transmission–reception event. $p_k(t)$ is the spatial impulse response from the kth subarray to the full-array. $k = 1, 2, \ldots, K$, $m = 1, 2, \ldots, M$. $g_m(t)$ is the noise in the received signals for the mth transmission–reception event.

signal (output) from the kth subarray in the mth transmission–reception event, and $m = 1, \ldots, M$, $k = 1, \ldots, K$. The function $f_{mk}(t)$ is different for different imaging models. The transmitted signals propagate in the medium, and the signals (i.e., echoes) are received by the full-array. The propagation response matrix $P(t)$ from each subarray to the full-array is

$$P(t) = [p_1(t) \ p_2(t) \ \cdots \ p_k(t)] \quad (2)$$

where $p_k(t)$ is the spatial impulse response from the kth subarray to the full-array. Then, we can obtain the received signals, $y_m(t)$, for the mth transmission–reception event as

$$y_m(t) = \sum_{k=1}^{K} f_{mk}(s_k(t)) * p_k(t) + g_m(t) \quad (3)$$

where $*$ is the convolution operation, and $g_m(t)$ is the noise in the received signals for the mth transmission–reception event. The noise mainly comes from the thermal noise of the electronic acquisition system and the disturbance of the environment noise. We assumed that the noise was additive white Gaussian noise (AWGN) throughout the analysis.

A. Synthetic Aperture Imaging Model

In the conventional SA imaging model [28], the transmission array is divided into K subarrays as shown in Fig. 1. In each transmission–reception event, a source signal from one subarray is transmitted. Thus, K source signals $s_1(t), s_2(t), \ldots, s_K(t)$ are sequentially transmitted in the M transmission–reception events. Since the number of subarrays is equal to the number of transmission–reception events, we have $M = K$ and $s_K(t) = s_M(t)$. Notice that we use the subscript S to represent the SA model, so the transmissions can be described as

$$X_S = \begin{bmatrix} X_{S1}(t) \\ X_{S2}(t) \\ \vdots \\ X_{SM}(t) \end{bmatrix} = I_{M \times M} \cdot S_{K \times 1}$$

$$= \begin{bmatrix} s_1(t) & 0 & \cdots & 0 \\ 0 & s_2(t) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & s_K(t) \end{bmatrix} \quad (4)$$

where $X_{S1}(t)$, $X_{S2}(t)$, ..., $X_{SM}(t)$ represent the transmitted signals for the first, second, ..., Mth transmission–reception events in the SA model as described in (1), $I_{M \times M}$ is the identity matrix and $S_{K \times 1}$ is the matrix of source signals $s_1(t)$, $s_2(t)$, ..., $s_K(t)$, $K = M$. According to the propagation response matrix $P(t)$ for the propagation of the signal in the medium, the matrix of received signals Y_S can be described as

$$Y_S = \begin{bmatrix} y_{S1}(t) \\ y_{S2}(t) \\ \vdots \\ y_{SM}(t) \end{bmatrix} = \begin{bmatrix} s_1(t) * p_1(t) + g_{S1}(t) \\ s_2(t) * p_2(t) + g_{S2}(t) \\ \vdots \\ s_K(t) * p_K(t) + g_{SM}(t) \end{bmatrix} \quad (5)$$

where $y_{S1}(t)$, $y_{S2}(t)$, ..., $y_{SM}(t)$ represent the received signals for the first, second, ..., Mth transmission–reception events in the SA model, respectively; $g_{S1}(t)$, $g_{S2}(t)$, ..., $g_{SM}(t)$ are the respective noises for reception signals in the SA model, and $P_K(t) = P_M(t)$.

Suppose that the noise-free part of the Y_S is Y_O as

$$Y_O = \begin{bmatrix} y_{O1}(t) \\ y_{O2}(t) \\ \vdots \\ y_{OM}(t) \end{bmatrix} = \begin{bmatrix} s_1(t) * p_1(t) \\ s_2(t) * p_2(t) \\ \vdots \\ s_K(t) * p_K(t) \end{bmatrix}. \quad (6)$$

Thus, the SNR of the received signals for each transmission–reception event can be written as

$$\text{SNR}_S = \begin{bmatrix} \mu_{y_{O1}}^2 / \sigma^2 \\ \mu_{y_{O2}}^2 / \sigma^2 \\ \vdots \\ \mu_{y_{OM}}^2 / \sigma^2 \end{bmatrix} \quad (7)$$

where μ_{yO1}, μ_{yO2}, ..., μ_{yOM} are the mean values of the received signals from the first, second, ..., Mth transmission–reception events, and σ^2 is the variance of the noise for each transmission–reception event.

B. Hadamard-Encoded Synthetic Aperture Imaging Model

In order to increase the SNR of received signals without increasing the amplitude or the number of transmission–reception events, spatial coding with Hadamard matrix was utilized as the H-SA imaging model for medical imaging [29], [30]. Unlike the SA model in which a source signal from one subarray is transmitted for each transmission–reception event, K source signals from K subarrays (full-aperture) are transmitted simultaneously with the Hadamard matrix-encoded coefficients for each transmission–reception event in the H-SA model. The number of transmission–reception events is the same as the number of subarrays, and then we have $M = K$ and $H_{M \times K} = H_{M \times M}$. Thus, an $M \times M$ Hadamard matrix is given by

$$H_{M \times M} = \begin{bmatrix} a_{11} & \cdots & a_{1K} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MK} \end{bmatrix} \quad (8)$$

where a_{11}, a_{12}, ..., a_{MK} represent the coefficients of the signals with a given value of -1 or $+1$ from the Hadamard matrix. The row elements of the Hadamard matrix can be used as the coefficients of source signals for each transmission–reception event. Note that we use the subscript H to represent the H-SA model. The signals for M transmission–reception events and K subarrays as shown in Fig. 1 are given by

$$X_H = \begin{bmatrix} X_{H1}(t) \\ X_{H2}(t) \\ \vdots \\ X_{HM}(t) \end{bmatrix} = H_{M \times M} \cdot S_{K \times 1}$$

$$= \begin{bmatrix} a_{11} \cdot s_1(t) & a_{12} \cdot s_2(t) & \cdots & a_{1K} \cdot s_K(t) \\ a_{21} \cdot s_1(t) & a_{22} \cdot s_2(t) & \cdots & a_{2K} \cdot s_K(t) \\ \vdots & \vdots & \vdots & \vdots \\ a_{M1} \cdot s_1(t) & a_{M2} \cdot s_2(t) & \cdots & a_{MK} \cdot s_K(t) \end{bmatrix} \quad (9)$$

where $X_{H1}(t)$, $X_{H2}(t)$, ..., $X_{HM}(t)$ represent the transmitted signals for the first, second, ..., Mth transmission–reception events in H-SA model and $S_{K \times 1}$ is the matrix of source signals $s_1(t)$, $s_2(t)$, ..., $s_K(t)$, $K = M$. Similar to the SA model and according to (3) and (6), the H-SA model would formulate the received signals from each transmission–reception event (Fig. 1) as

$$Y_H = \begin{bmatrix} y_{H1}(t) \\ y_{H2}(t) \\ \vdots \\ y_{HM}(t) \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{K} a_{1k} \cdot y_{Ok}(t) + g_{H1}(t) \\ \sum_{k=1}^{K} a_{2k} \cdot y_{Ok}(t) + g_{H2}(t) \\ \vdots \\ \sum_{k=1}^{K} a_{Mk} \cdot y_{Ok}(t) + g_{HM}(t) \end{bmatrix} \quad (10)$$

where $g_{H1}(t)$, $g_{H2}(t)$, ..., $g_{HM}(t)$ are the noises for respective reception signals $y_{H1}(t)$, $y_{H2}(t)$, ..., $y_{HM}(t)$. According to (6), $y_{Ok}(t) = s_k(t) * p_k(t)$, $K = M$. The received signals consist of the signals from individual source signals. In order to reconstruct the image for each subarray transmission, a decoding procedure needs to be designed to recover the equivalent signals to the received signals from the SA model.

We can write the matrix of received signals as

$$Y_H = H_{M \times M} \cdot Y_O + G_H \quad (11)$$

where G_H is an $M \times 1$ matrix of noise $g_{H1}(t)$, $g_{H2}(t)$, ..., $g_{HM}(t)$ for individual received signals in the H-SA model. In the decoding procedure, the decoded received signals (i.e., equivalent signals as in the SA model) can be obtained by the multiplication of the inverse of the Hadamard matrix [29], [30] as

$$Y_{Hd} = H_{M \times M}^{-1} \cdot Y_H = Y_O + H_{M \times M}^{-1} \cdot G_H. \quad (12)$$

Notice that we use the subscript Hd to represent the decoded signals of H-SA model. It is known that the inverse of $H_{M \times M}$ is itself with the multiplication of a factor [36] as

$$H_{M \times M}^{-1} = \frac{1}{M} \cdot H_{M \times M}. \quad (13)$$

Decoding of the received signals can be formulated as

$$
Y_{Hd} = Y_O + \frac{1}{M} \cdot H_{M \times M} \cdot G_H
$$

$$
= \begin{bmatrix} y_{O1}(t) + \frac{1}{M} \sum_{m=1}^{M} a_{1m} \cdot g_{Hm}(t) \\ y_{O2}(t) + \frac{1}{M} \sum_{m=1}^{M} a_{2m} \cdot g_{Hm}(t) \\ \vdots \\ y_{OM}(t) + \frac{1}{M} \sum_{m=1}^{M} a_{Mm} \cdot g_{Hm}(t) \end{bmatrix}. \tag{14}
$$

Since the noise is AWGN, the variance of the noise for each decoded received signal is written as

$$
\sigma_{Hd} = \begin{bmatrix} \sigma^2/M \\ \sigma^2/M \\ \vdots \\ \sigma^2/M \end{bmatrix}. \tag{15}
$$

Thus, the SNR of the received signals for each transmission–reception event is

$$
\text{SNR}_{Hd} = M \cdot \text{SNR}_S = \begin{bmatrix} M \cdot \mu_{yO1}^2 \cdot \sigma^2 \\ M \cdot \mu_{yO2}^2 / \sigma^2 \\ \vdots \\ M \cdot \mu_{yOM}^2 / \sigma^2 \end{bmatrix}. \tag{16}
$$

The SNR of the decoded received signals in the H-SA is M times higher than that of the SA model. The underlying principle of the increase in SNR is that the decoded received signals are equivalent to the M times repetition of the transmissions in the SA model.

C. Proposed Cascaded Synthetic Aperture Imaging Model

In a previous work, we proposed a cascaded dual-polarity wave (CDW) model, in which N cascaded waves are temporally transmitted. The previously designed CDW model is in the time domain for plane wave imaging [35]. In this study, we extend the CDW model into the spatiotemporal domain as the coined CaSA imaging model. We use the subscript C to represent the CaSA model. The signals for M transmission–reception events can be described as

$$
X_C = \begin{bmatrix} X_{C1}(t) \\ X_{C2}(t) \\ \vdots \\ X_{CM}(t) \end{bmatrix}
$$

$$
= \begin{bmatrix} x_{11}(t) & x_{12}(t) & \cdots & x_{1K}(t) \\ x_{21}(t) & x_{22}(t) & \cdots & x_{2K}(t) \\ \vdots & \vdots & \vdots & \vdots \\ x_{M1}(t) & x_{M2}(t) & \cdots & x_{MK}(t) \end{bmatrix} \tag{17}
$$

where $X_{C1}, X_{C2}, \ldots, X_{CM}$ represent the transmitted signals for the first, second, ... , Mth transmission–reception events,

and $x_{mk}(t)$ represent the transmitted signals in the mth transmission and kth subarray, $m = 1, \ldots, M$, $k = 1, \ldots, K$, $M = K$. In the CaSA model, each subarray transmission is a pulse train with CDWs in each transmission–reception event. It is given by

$$
x_{mk}(t) = \sum_{n=1}^{N} c_{\text{in}} \cdot s_k(t) * \delta(t - (n-1) \cdot \tau) \tag{18}
$$

where $s_k(t)$ is the source signal of the kth subarray as defined in (1), τ is the preset delay added between two successive cascaded waves, and c_{in} is the polarity coefficient of each cascaded wave. The index n indicates the nth cascaded wave, and $n = 1, 2, \ldots, N$. The other index i depends on the transmission–reception event m and the subarray k, which will be defined later (24) according to the designed coding matrices. In order to decode the received signals, we need to design a new matrix as the coefficients of the spatiotemporal cascaded waves.

1) Design of the Coding Matrices: The previously designed CDW matrix is a $2 \times N$ matrix with $N = 2^Q$, where Q is an integer [35]. The CDW matrix is a coding matrix with two polarities: $+1$ and -1, and it has more columns than the square Hadamard matrix of order 2. For cascaded-wave imaging with full-aperture plane waves, N CDWs with the polarities determined from the CDW matrix are transmitted for each reception. The cascaded wave with designed matrix $C_{2 \times N}$ is not only for plane wave but also for any other types of waves, such as diverging wave. In our previous plane wave transmission work [35], we applied the matrix in the time domain. In this presented SA setting, both the spatial and temporal domains of the transmitted waves are considered. Specifically, we extend the time-domain $C_{2 \times N}$ matrix to a new spatiotemporal matrix $C_N H_M$. The $C_N H_M$ matrix has more than two rows (i.e., spatial domain) and is designed based on the Hadamard matrix of order M (i.e., $H_{M \times M}$, where $M = 2^Z$, and Z is an integer). The design of the matrix $C_N H_M$ from $C_{2 \times N}$ and $H_{M \times M}$ is shown in Fig. 2. For example, the design of $C_8 H_2$ from $C_{2 \times 8}$ and $H_{2 \times 2}$ is provided as follows.

1) As shown in our previous work [35], the $C_{2 \times 8}$ CDW matrix can be obtained from the $C_{2 \times 4}$ CDW matrix, which is similar to the $C_{2 \times 4}$ CDW matrix derived from the $H_{2 \times 2}$ Hadamard matrix by repeating and rearranging $H_{2 \times 2}$ elements. The previously designed $C_{2 \times 8}$ matrix is

$$
C_{2 \times 8} = \begin{bmatrix} +1 & +1 & +1 & -1 & +1 & -1 & +1 & +1 \\ +1 & +1 & -1 & +1 & +1 & -1 & -1 & -1 \end{bmatrix}. \tag{19}
$$

2) Four block representations of the $C_{2 \times 8}$ matrix can be written as

$$
\begin{cases} C_{1L} = [+1 \ +1 \ +1 \ -1] \\ C_{1R} = [+1 \ -1 \ +1 \ +1] \\ C_{2L} = [+1 \ +1 \ -1 \ +1] \\ C_{2R} = [+1 \ -1 \ -1 \ -1] \end{cases} \tag{20}
$$

where C_{1L} and C_{1R} are, respectively, the block matrices of the left and the right halves of the first row, and C_{2L}
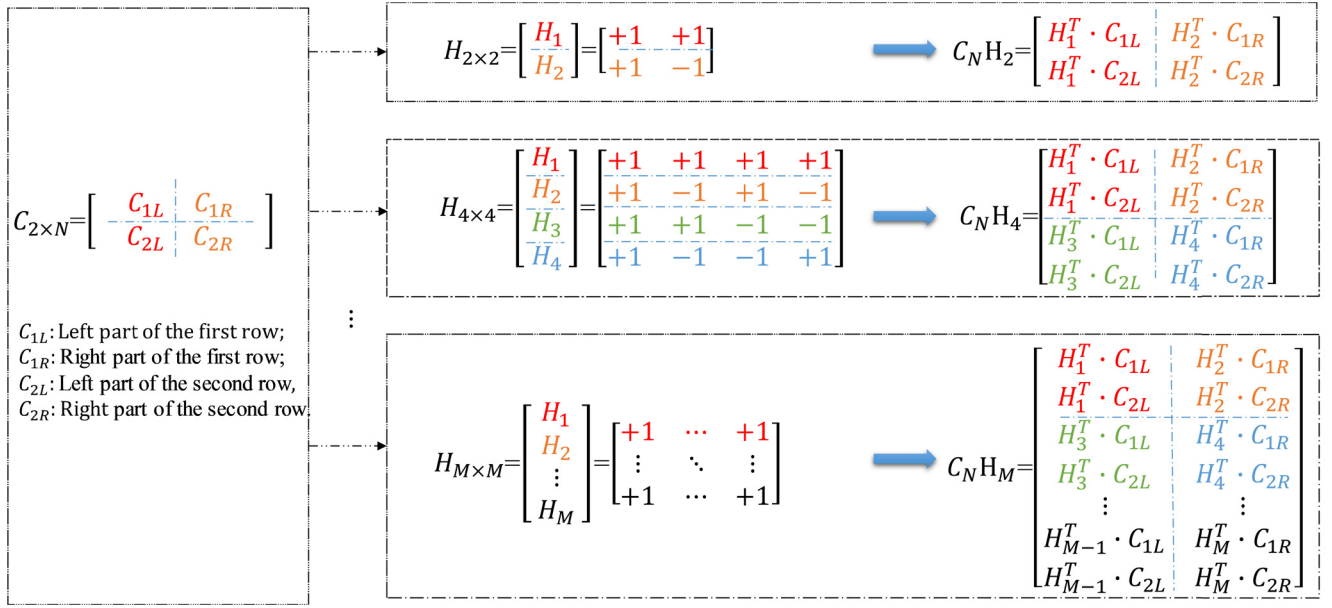
Fig. 2. Design of the spatiotemporal coding matrix $C_N H_M$. $C_{2\times N}$ is the previously designed CDW matrix. $H_{M\times M}$ is the Hadamard matrix.

and C_{2R} are, respectively, the block matrices of the left and right halves of the second row.

3) We then multiply the transpose of H_1, H_2 and that of C_{1L}, C_{1R}, C_{2L}, C_{2R} as shown in Fig. 2 and provided below

$$\begin{cases} H_1^T \cdot C_{1L} = \begin{bmatrix} +1 & +1 & +1 & -1 \\ +1 & +1 & +1 & -1 \end{bmatrix} \\ H_2^T \cdot C_{1R} = \begin{bmatrix} +1 & -1 & +1 & +1 \\ -1 & +1 & -1 & -1 \end{bmatrix} \\ H_1^T \cdot C_{2L} = \begin{bmatrix} +1 & +1 & -1 & +1 \\ +1 & +1 & -1 & +1 \end{bmatrix} \\ H_2^T \cdot C_{2R} = \begin{bmatrix} +1 & -1 & -1 & -1 \\ -1 & +1 & +1 & +1 \end{bmatrix} \end{cases} \tag{21}$$

where H_1 and H_2 are the two block representations of the Hadamard matrix as shown in Fig. 2.

4) The designed $C_8 H_2$ is

$$\begin{aligned} C_8 H_2 &= \begin{bmatrix} H_1^T \cdot C_{1L} & H_2^T \cdot C_{1R} \\ H_1^T \cdot C_{2L} & H_2^T \cdot C_{2R} \end{bmatrix} \\ &= \begin{bmatrix} +1 & +1 & +1 & -1 & +1 & -1 & +1 & +1 \\ +1 & +1 & +1 & -1 & -1 & +1 & -1 & -1 \\ +1 & +1 & -1 & +1 & +1 & -1 & -1 & -1 \\ +1 & +1 & -1 & +1 & -1 & +1 & +1 & +1 \end{bmatrix}. \end{aligned} \tag{22}$$

Similar to the design of $C_8 H_2$ matrix, the $C_8 H_4$ matrix can be obtained from $C_{2\times 8}$ and $H_{4\times 4}$ matrices. Based on the procedure illustrated in Fig. 2, an arbitrary $C_N H_M$ matrix can

be obtained by

$$C_N H_M = \begin{bmatrix} H_1^T \cdot C_{1L} & H_2^T \cdot C_{1R} \\ H_1^T \cdot C_{2L} & H_2^T \cdot C_{2R} \\ H_3^T \cdot C_{1L} & H_4^T \cdot C_{1R} \\ H_3^T \cdot C_{2L} & H_4^T \cdot C_{2R} \\ \vdots & \vdots \\ H_{M-1}^T \cdot C_{1L} & H_M^T \cdot C_{1R} \\ H_{M-1}^T \cdot C_{2L} & H_M^T \cdot C_{2R} \end{bmatrix} \tag{23}$$

where H_1^T, H_2^T, ..., H_M^T are the transpose of the rows of Hadamard matrix $H_{M\times M}$. The dimension of the $C_N H_M$ matrix is $M^2 \times N$.

2) Application of the Designed Matrix to the Model: According to the designed matrix $C_N H_M$ and the proposed cascaded model, the polarity coefficients of the cascaded waves c_{in} are obtained from the ith row and nth column element of the $C_N H_M$ matrix. The index i of the c_{in} in (18) is

$$i = (m-1) \cdot M + k. \tag{24}$$

We take two subarrays and eight cascaded waves ($M = 2$, $N = 8$) as an example. In the first transmission–reception event ($m = 1$), the coefficients of the first subarray ($k = 1$) cascaded waves are from the first row of the $C_8 H_2$ matrix, and the coefficients of the second subarray ($k = 2$) cascaded waves are from the second row of the $C_8 H_2$ matrix. In the second transmission–reception event ($m = 2$), the coefficients of the first subarray ($k = 1$) cascaded waves are from the third row of the $C_8 H_2$ matrix, and the coefficients of the second subarray ($k = 2$) cascaded waves are from the fourth row of the $C_8 H_2$ matrix. Similar to the H-SA model,
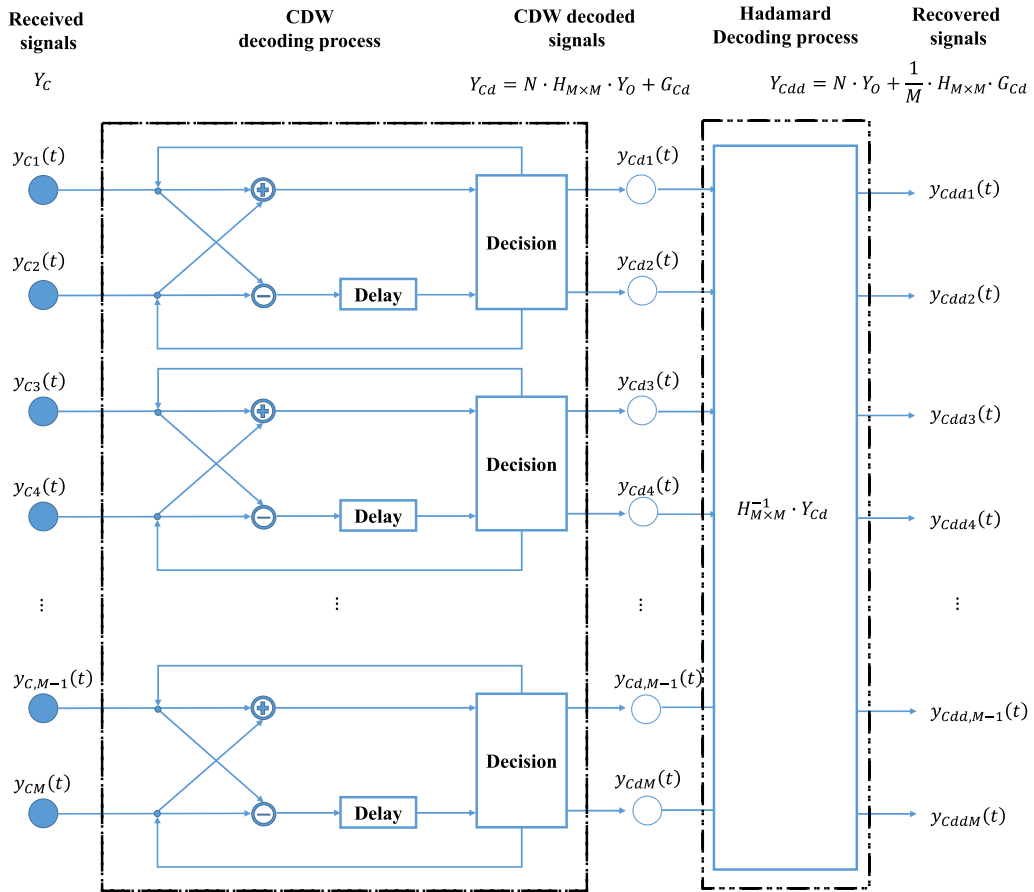
Fig. 3.   Design of the decoding process in the proposed CaSA imaging model. Y_C is the matrix of the received signals, Y_{Cd} is the matrix of the CDW decoded signals, Y_{Cdd} is the matrix of the finally decoded signals, Y_O is the matrix of the noise-free component of the received signals in the SA model, G_{Cd} is the matrix of the noise, $H_{M \times M}$ is the M Hadamard matrix, M is the number of transmission–reception events, and N is the number of cascaded waves.

the received signals from each transmission–reception event as shown in Fig. 1 in the CaSA model can be obtained as

$$Y_C = \begin{bmatrix} y_{C1}(t) \\ y_{C2}(t) \\ \vdots \\ y_{CM}(t) \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{K} x_{1k}*p_k(t) + g_{C1}(t) \\ \sum_{k=1}^{K} x_{2k}*p_k(t) + g_{C2}(t) \\ \vdots \\ \sum_{k=1}^{K} x_{Mk}*p_k(t) + g_{CM}(t) \end{bmatrix} \quad (25)$$

where $g_{C1}(t)$, $g_{C2}(t)$, ..., $g_{CM}(t)$ are the noises for each reception signal $y_{C1}(t)$, $y_{C2}(t)$, ..., $y_{CM}(t)$. The received signals consist of the signals from individual source signals.

According to (6), (18), and (25), we can obtain (26), as shown at the bottom of the page. The received signals consist of the signals from individual source signals and the delay of source signals. Next, a decoding procedure needs to be designed to recover the signals which correspond to the received signals from the SA model.

3) Design of the Decoding Process of the Model: The decoding process can be divided into two steps as shown in Fig. 3. After receiving the signals $y_{C1}(t)$,

$$Y_C = \begin{bmatrix} y_{C1}(t) \\ y_{C2}(t) \\ \vdots \\ y_{CM}(t) \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{K}\left(\sum_{n=1}^{N} c_{kn} \cdot y_{Ok}(t)*\delta(t-(i-1)\cdot\tau)\right) + g_{C1}(t) \\ \sum_{k=1}^{K}\left(\sum_{n=1}^{N} c_{(M+k)n} \cdot y_{Ok}(t)*\delta(t-(i-1)\cdot\tau)\right) + g_{C2}(t) \\ \vdots \\ \sum_{k=1}^{K}\left(\sum_{n=1}^{N} c_{((M-1)\cdot M+k)n} \cdot y_{Ok}(t)*\delta(t-(i-1)\cdot\tau)\right) + g_{CM}(t) \end{bmatrix} \quad (26)$$

$y_{C2}(t), \ldots, y_{CM}(t)$ from M transmission–reception events, each pair of them is utilized to obtain the decoded signals $y_{Cd1}(t), y_{Cd2}(t), \ldots, y_{CdM}(t)$ based on the previously designed CDW decoding process. The CDW decoder consists of addition, subtraction, and delay operations. These operations may require several iterations and are illustrated by closed loops as shown in Fig. 3. In each iteration, the echoes from the two adjacent transmissions are summed or subtracted and then delayed to obtain the processed echoes. The decider is used to send the processed echoes for the next iteration or the final decoded output echoes. We use the subscript Cd to represent the decoded signals from the CDW decoding process of the CaSA model. Let the noise be AWGN, and the CDW decoded signal is thus

$$
Y_{Cd} = \begin{bmatrix} y_{Cd1}(t) \\ y_{Cd2}(t) \\ \vdots \\ y_{CdM}(t) \end{bmatrix} = N \cdot H_{M \times M} \cdot Y_O + G_{Cd}
$$

$$
= \begin{bmatrix} N \cdot \sum_{k=1}^{K} a_{1k} \cdot y_{Ok}(t) + \sum_{n=1}^{N} g_{1n}(t) \\ N \cdot \sum_{k=1}^{K} a_{1k} \cdot y_{Ok}(t) + \sum_{n=1}^{N} g_{2n}(t) \\ \vdots \\ N \cdot \sum_{k=1}^{K} a_{1k} \cdot y_{Ok}(t) + \sum_{n=1}^{N} g_{Mn}(t) \end{bmatrix} \tag{27}
$$

where G_{Cd} represents the matrix of summation of AWGN noise, and N is the number of cascaded waves.

Then, the final decoded signals can be obtained by the Hadamard decoding process as in the H-SA. We use the subscript Cdd to represent the final decoded signals from the CDW and Hadamard decoding process of the CaSA model. The final decoded signals are

$$
Y_{Cdd} = \begin{bmatrix} y_{Cdd1}(t) \\ y_{Cdd2}(t) \\ \vdots \\ y_{CddM}(t) \end{bmatrix} = N \cdot Y_O + \frac{1}{M} \cdot H_{M \times M} \cdot G_{Cd}
$$

$$
= \begin{bmatrix} N \cdot y_{O1}(t) + \frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} g_{mn}(t) \\ N \cdot y_{O2}(t) + \frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} g_{mn}(t) \\ \vdots \\ N \cdot y_{OM}(t) + \frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} g_{mn}(t) \end{bmatrix}. \tag{28}
$$

Since the noise is AWGN, the variance of the noise for each decoded received signal is written as

$$
\sigma_{Cdd} = \begin{bmatrix} N \cdot \sigma^2/M \\ N \cdot \sigma^2/M \\ \vdots \\ N \cdot \sigma^2/M \end{bmatrix}. \tag{29}
$$

The SNR of the received signals for each transmission–reception event is thus

$$
\text{SNR}_{Cdd} = M \cdot N \cdot \text{SNR}_S = \begin{bmatrix} M \cdot N \cdot \mu_{yO1}^2/\sigma^2 \\ M \cdot N \cdot \mu_{yO2}^2/\sigma^2 \\ \vdots \\ M \cdot N \cdot \mu_{yOM}^2/\sigma^2 \end{bmatrix}. \tag{30}
$$

In theory, the proposed CaSA model can increase the SNR by $10 \cdot \log_{10}(N \times M)$ compared with the conventional SA model and can increase the SNR by $10 \cdot \log_{10}(N)$ compared with the H-SA model.

D. Exemplary Cascaded Synthetic Aperture Imaging Sequence

In order to elucidate the proposed CaSA method for medical ultrasound imaging with M transmission–reception events and N cascaded waves, two ($M = 2$) transmission–reception events and eight ($N = 8$) cascaded waves are exemplified as shown in Fig. 4. In transmission, the array is equally divided into two subapertures; each subaperture transmits eight cascaded diverging waves with a guaranteed short time interval between two successive sets of cascaded waves. T1-R1 and T2-R2 indicate the first and second transmission–reception events, respectively. Therefore, in each transmission–reception event, a group of 16 diverging waves coded in both the spatial (columns) and temporal (rows) domains is transmitted. The positive [shown in Fig. 4 (yellow)] and negative [shown in Fig. 4 (blue)] coefficients are obtained from the designed coefficient matrix $C_8 H_2$ as explained in Section II-C (22). The coefficients of the waves for T1-R1 and T2-R2 events are obtained from the first to second rows, and third to fourth rows of the coefficient matrix $C_8 H_2$, respectively.

In reception, the full aperture of the array is deployed to receive the backscattered signals in each transmission–reception event. A two-stage decoding process is directly applied to the received signals R1 and R2 as shown in Fig. 4(b)–(h). Fig. 4(b)–(g) shows the results of each step of the CDW decoding process (temporal decoding) with addition, subtraction, and delay operations. The intensity shown in Fig. 4(g) has an eightfold increase compared with that shown in Fig. 4(a) because of eight cascaded waves. The final step of the decoding process is the Hadamard decoding (spatial decoding), by which we obtain Fig. 4(h) from Fig. 4(g). Finally, we obtain a 16-fold increase in the intensity shown in Fig. 4(h) when compared with the original shown in Fig. 4(a). The decoded signals are lastly beamformed and coherently compounded to obtain one high-SNR beamformed image frame. Note that Fig. 4(g) and (h) corresponds to the H-SA and SA models, respectively. In this example, in theory, CaSA can increase the SNR by $10 \cdot \log_{10}(16)$ compared with the conventional SA model and by $10 \cdot \log_{10}(8)$ compared with the H-SA model.

III. METHODS

The proposed uCUS imaging sequence with CaSA was tested on a multi-purpose multi-tissue ultrasound phantom (CIRS Model 040GSE, attenuation 0.5 dB/cm/MHz) and an
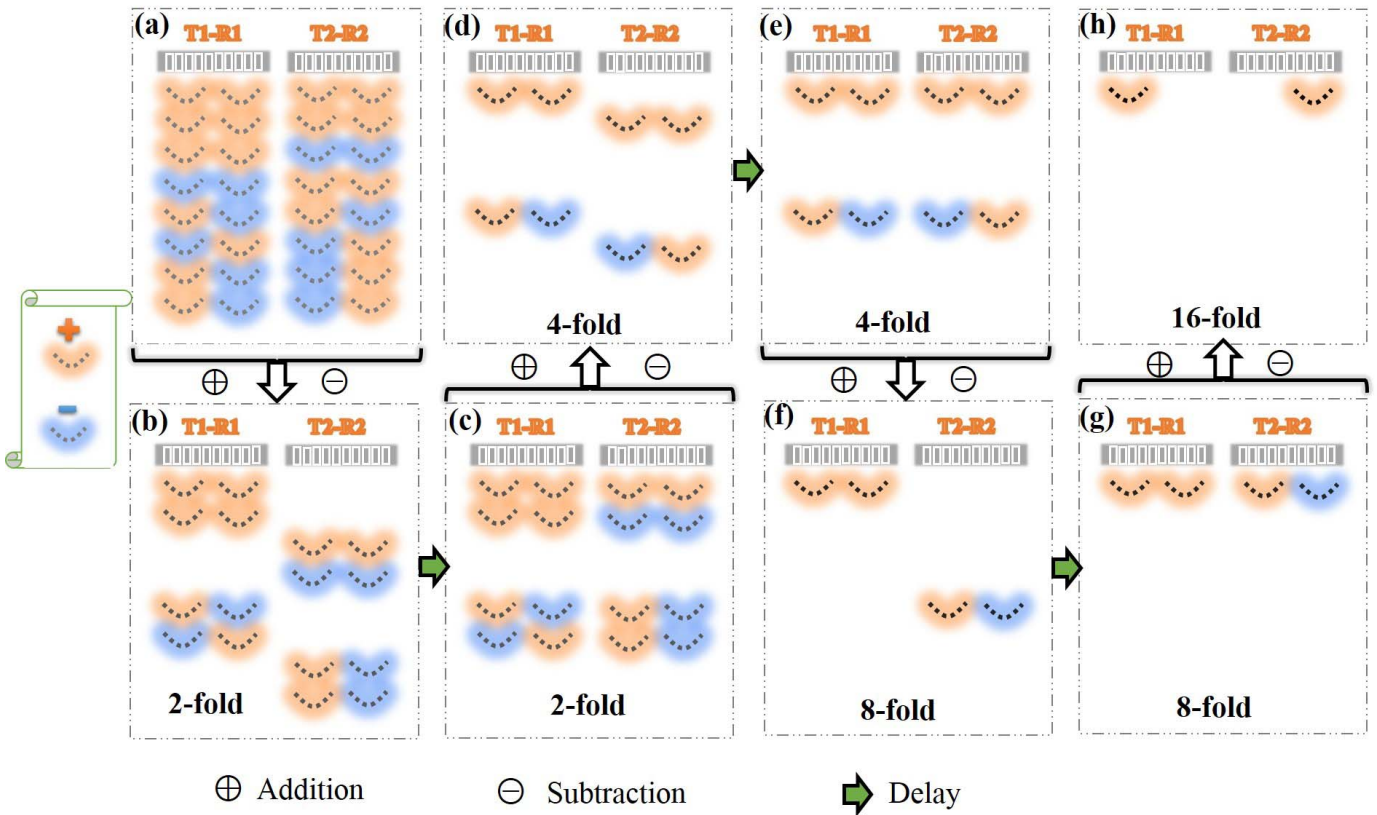
Fig. 4. Illustration of the CaSA imaging mode with two transmission–reception events and eight cascaded diverging waves ($M = 2$, $N = 8$). T1-R1 and T2-R2 indicate the first and second transmission–reception events, respectively. (a) Transmission settings with two subapertures and eight cascaded waves. (b)–(g) CDW decoding process with addition, subtraction, and delay operations. (h) Final step of the decoding process as Hadamard decoding from (g).

in vivo beating human heart for B-mode imaging, myocardial motion, and power Doppler imaging. The data were acquired by a Verasonics Vantage 256 system (Verasonics, Kirkland, WA, USA) with a phased array probe P4-2 (64 array elements) whose center frequency was 2.5 MHz. The transmitted center frequency was the same as the probe center frequency. The phased array was divided into four subapertures ($M = 4$) with 16 elements each. The distance between the virtual source and the array was half of the subaperture. The number of cascaded waves is 32 ($N = 32$) with 0.6-μs short time interval for every two adjacent cascaded waves. The base of the transmitted diverging wave was a short burst of one cycle. We designed an interleaved sequence for all three methods (SA, H-SA, and CaSA) on our Verasonics system to ensure the same frame rate and the same region of interest (ROI) at the same time for a fair comparison. In the interleaved sequence, since each of the three imaging methods obtained one compounded image from every four transmission–reception events, we sequentially performed four SA transmissions, four H-SA transmissions, and four CaSA transmissions; this process was then repeated. The acquisition frame rate for each transmission–reception event was 4000 frame per second (fps). The driving voltage settings on the Vantage 256 system in this study for B-mode and Doppler images were 1.6 and 16.0 V, respectively. Note that this study focuses on the methodology, and we did not optimize the parameters and measure the corresponding acoustic pressures. The experimental protocol of the *in vivo* beating

human heart was approved prior to use by the Institutional Review Board of The University of Hong Kong (UW13-566). Informed consent was given and signed.

A. B-Mode Imaging

The B-mode (brightness) images were obtained by the log compression of the envelope of the beamformed (pixel-oriented delay-and-sum) echoes. In the calibration phantom study, the quantitative evaluation metrics, such as resolution, SNR, and contrast, were computed from 100 acquired images in the same scanning region [35]. The contrast ratio (CR) was calculated as $CR = |\mu_s - \mu_c|$, where μ_s and μ_c are the mean intensities of the regions of background and anechoic cyst region, respectively. The contrast-to-noise ratio (CNR) was calculated using $CNR = |\mu_s - \mu_c|/(\sigma_s^2 + \sigma_c^2)^{1/2}$, where σ_s and σ_c are the intensity standard deviations of the background and anechoic cyst, respectively. In the *in vivo* beating human heart study, a human heart (male, 26 years old) was scanned in the apical four-chamber view. The electrocardiography (ECG) signals were also recorded to identify cardiac phases.

B. Myocardial Motion Imaging

An in-house RF-based speckle tracking method based on cross correlation with a 2-D matching kernel was employed to estimate the axial displacement of the myocardium [37], [38]. Spline interpolation was first performed between RF signals in
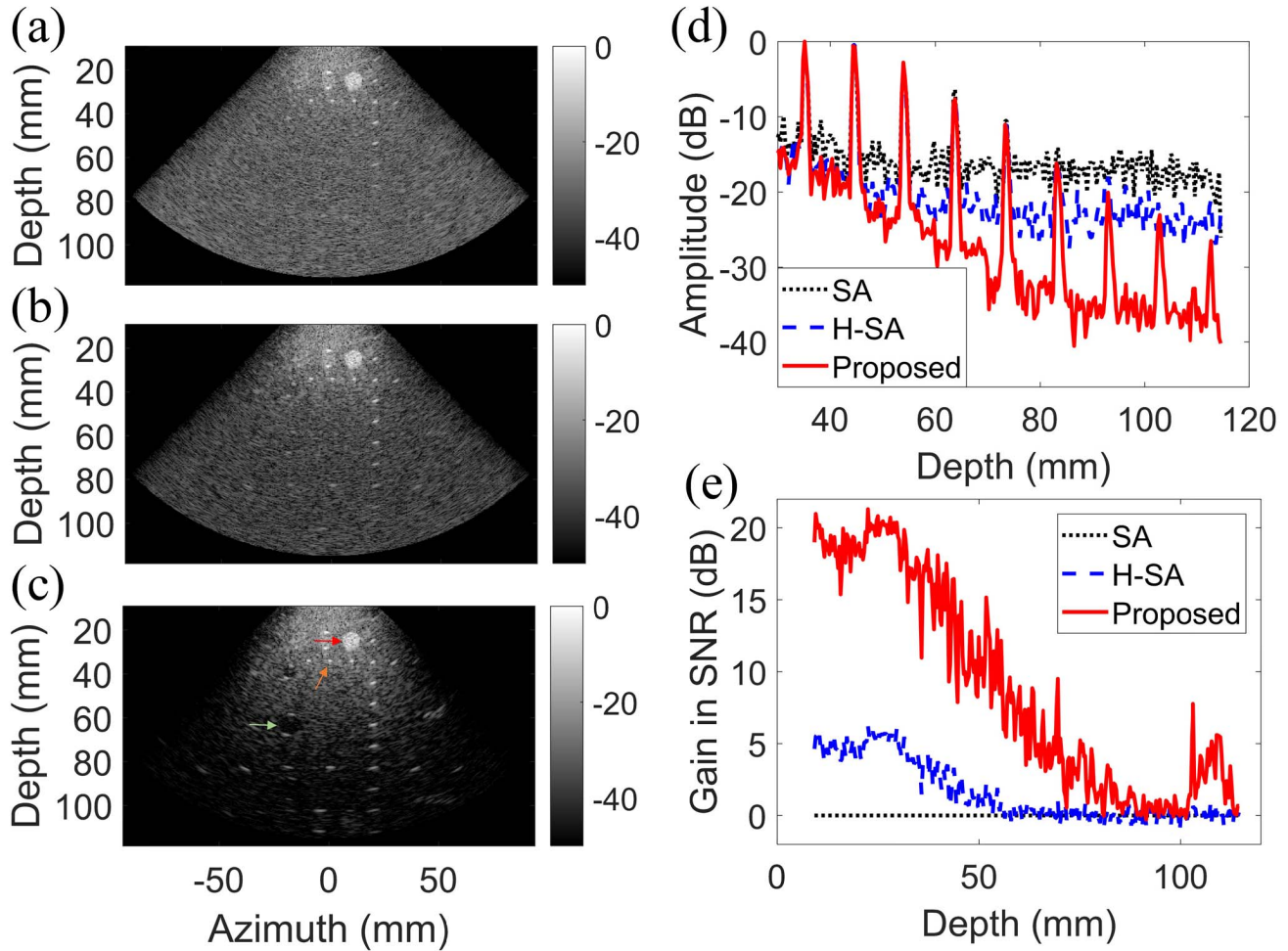
Fig. 5. B-mode images of a commercial calibration phantom by (a) SA, (b) H-SA, and (c) proposed CaSA. (a)–(c) Point target (indicated by an orange arrow) and the cyst (indicated by red and green arrows). (d) Axial profiles along the strong reflectors at azimuth around 20 mm. (e) Gain in SNR of H-SA and the proposed CaSA compared with that of SA at azimuth around 10 mm.

the postdisplaced frame to increase lateral sampling. The 2-D cross correlation was then performed between the predisplaced and postdisplaced RF frames in a predefined 2-D search region to generate a normalized cross correlation (NCC) map, which was interpolated to find the peak on the interpolated NCC map to obtain the displacement to a subsample scale. In this study, only the axial displacement was estimated and presented to compare the estimation quality from RF frames reconstructed by SA, H-SA, and CaSA methods. The entire septum was chosen as the ROI to compare the estimation quality of the axial displacement among the three image reconstruction methods. We calculated the displacement SNR as the ratio between the mean values and standard deviations of the displacements in the entire septum region. This ratio reflected the quality of the estimated displacement image.

C. Power Doppler Imaging

In addition to B-mode, power Doppler imaging was performed for the detection of intracavity blood of the beating human heart *in vivo*. In all, 70 postcompounded images were used to generate 1 power Doppler image without sliding

window. The tissue clutter filter is similar to that described in [17], in which a singular value decomposition (SVD) filter was applied for spatial–temporal processing of the blood and tissue information. First, the acquired image data set was spatiotemporally organized to convert a 3-D data set into a 2-D data set. Then, we performed the truncated SVD of the 2-D data set by processing several large singular values (cutoff value was 60). Due to the fast change of blood motion, blood signals had much lower coherence than the tissue signals. Thus, we could extract the blood signals from the 2-D data set by removing the first several singular value data which were associated with the tissue signals. We then integrated the energy of the signals as the power Doppler image.

IV. RESULTS

A. Calibration Phantom Study

Fig. 5 shows the B-mode images, axial profiles, and SNR curves of the SA, H-SA, and proposed CaSA method realized in the phantom which contains multiple strong reflectors and hypoechoic cysts. In the SA image [Fig. 5(a)], the strong

TABLE I
COMPARISONS OF RESOLUTION, CONTRAST, AND SNR

	SA	H-SA	CaSA	CaSA vs. SA	CaSA vs. H-SA
Axial resolution (mm)	1.06 ± 0.081	1.05 ± 0.047	1.07 ± 0.006	(- 0.93 %)	(- 1.86 %)
Lateral resolution (mm)	1.52 ± 0.557	1.42 ± 0.033	1.43 ± 0.006	(+ 6.29 %)	(- 0.70 %)
CR (dB)	0.70 ± 0.769	1.33 ± 0.790	9.14 ± 0.657	(+ 8.44)	(+ 7.81)
CNR (dB)	0.11 ± 0.119	0.20 ± 0.118	1.16 ± 0.068	(+ 1.05)	(+ 0.96)
SNR (dB)	20.26	25.89	40.72	(+ 20.46)	(+ 14.83)

reflectors and hypoechoic cysts were veiled by the background noise, especially at the depth deeper than 60 mm. In the H-SA image [Fig. 5(b)], the strong reflectors were revealed up to 80-mm depth, and the hypoechoic cysts remained obscured. In contrast, the proposed CaSA method [Fig. 5(c)] was capable of highlighting the strong reflectors even in the deep zone and clearly displayed two cysts located at the depths of 40 and 63 mm at −17 mm azimuthally. Besides, the axial profiles around azimuth of 20 mm show significant suppression of the background noise by the proposed CaSA. Fig. 5(e) shows that the gains in SNR of H-SA and CaSA with respect to SA were approximately 5 and 20 dB, respectively.

Table I lists the spatial (i.e., axial and lateral) resolutions, contrasts (CR and CNR), and SNR of the SA, H-SA, and proposed CaSA imaging methods realized at the same frame rate. The spatial resolutions were calculated at the strong reflector at 0.8-mm azimuthally and depth of 34.5 mm (orange arrow) as an example. The proposed CaSA shows comparable spatial resolutions with SA and H-SA imaging. The CR and CNR were calculated from the anechoic cyst region at the depth of 63 mm (green arrow). The H-SA method shows slight CR and CNR improvements by the SA method. The proposed CaSA method had greater improvement of CR (+8.44 dB versus SA, +7.81 dB versus H-SA) and CNR (+1.05 versus SA, +0.96 versus H-SA). SNR calculated in the middle of the grayscale round target at a depth of 26.18 mm (red arrow) by H-SA was improved by 5.63 dB compared with SA imaging. The proposed CaSA imaging had experimental improvement of SNR (+20.46 dB versus SA, + 14.83 dB versus H-SA), which was close to the theoretical prediction ($10 \times \log_{10}(4 \times 32) = 21.07$ dB and $10 \times \log_{10}(32) = 15.05$ dB).

B. In Vivo Transthoracic B-Mode Images of the Heart

Fig. 6 shows the B-mode images of the *in vivo* human heart in the diastolic phase in the apical four-chamber view. The chamber walls were best visualized and delineated from the proposed CaSA method because of the SNR improvement. After normalization, the SA image had the lowest SNR, and the heart wall structure was not clearly seen. The H-SA image was better than the SA image because of the enhanced SNR. The CaSA image exhibited the lowest noise level and the best myocardial wall visualization. The axial profile [Fig. 6(d), taken at the red vertical arrow in (c)] was obtained at azimuth 29.2 mm through the left ventricle and left atrium. CaSA produced the lowest noise level inside the heart chambers. The lateral profile [Fig. 6(e), taken at the red horizontal arrow in (c)] was obtained at a depth of 78.2 mm through

both ventricles. After normalization, the noise levels in the left ventricle at 78 mm in depth and 29.2 mm at azimuth (crossing point of two profiles indicated by the two arrows) were approximately −35 dB in SA imaging, −40 dB in H-SA imaging, and −52 dB in CaSA imaging. The lateral profile obtained by CaSA showed three peaks and clearer myocardial walls. Supplementary movies (see PowerPoint file🖿) show the B-mode images reconstructed by SA, H-SA, and the proposed CaSA, respectively, in a complete cardiac cycle.

C. In Vivo Myocardial Motion Imaging

Fig. 7 shows the images of axial displacements estimated by our in-house RF speckle tracking technique [37], [38] from RF frames of the same human myocardium shown in Fig. 6 reconstructed by SA, H-SA, and CaSA methods. Positive and negative axial displacements represent upward and downward myocardial motion, respectively (Fig. 7). The superior quality of the axial displacements estimated from the proposed CaSA method was demonstrated not only in the temporal profiles of displacement SNR during diastole [Fig. 7(d)] and systole [Fig. 7(h)] but also in the spatial maps [Fig. 7(c) and (g) and supplementary movies in the PowerPoint file] in comparison with SA [Fig. 7(a) and (e) and supplementary movies in the PowerPoint file] and H-SA [Fig. 5(b) and (f) and supplementary movies in the PowerPoint file].

D. In Vivo Power Doppler Imaging

Fig. 8 shows power Doppler images of the blood flow in the apical four-chamber view of the heart. The power Doppler images [(Fig. 8(a)–(c)] show that the proposed CaSA imaging method achieved the best contrast between chamber blood and myocardium. The axial [Fig. 8(d), taken at the white vertical arrow in (c)] and lateral profiles [Fig. 8(e), taken at the white horizontal arrow in (c)] further show that the myocardial region had much lower intensity than the heart chamber region.

V. DISCUSSION

This study aims at deploying CUS for ultrafast imaging of the heart dynamics with enhanced SNR and without compromising spatiotemporal resolution. To solve the tradeoff between the axial resolution and the length of transmitted pulses, we propose a new imaging sequence CaSA with a newly designed spatiotemporal coding matrix for each segment of the cascaded wave. The advantage of the CaSA method is that it strives to approach the physical limits of SNR in
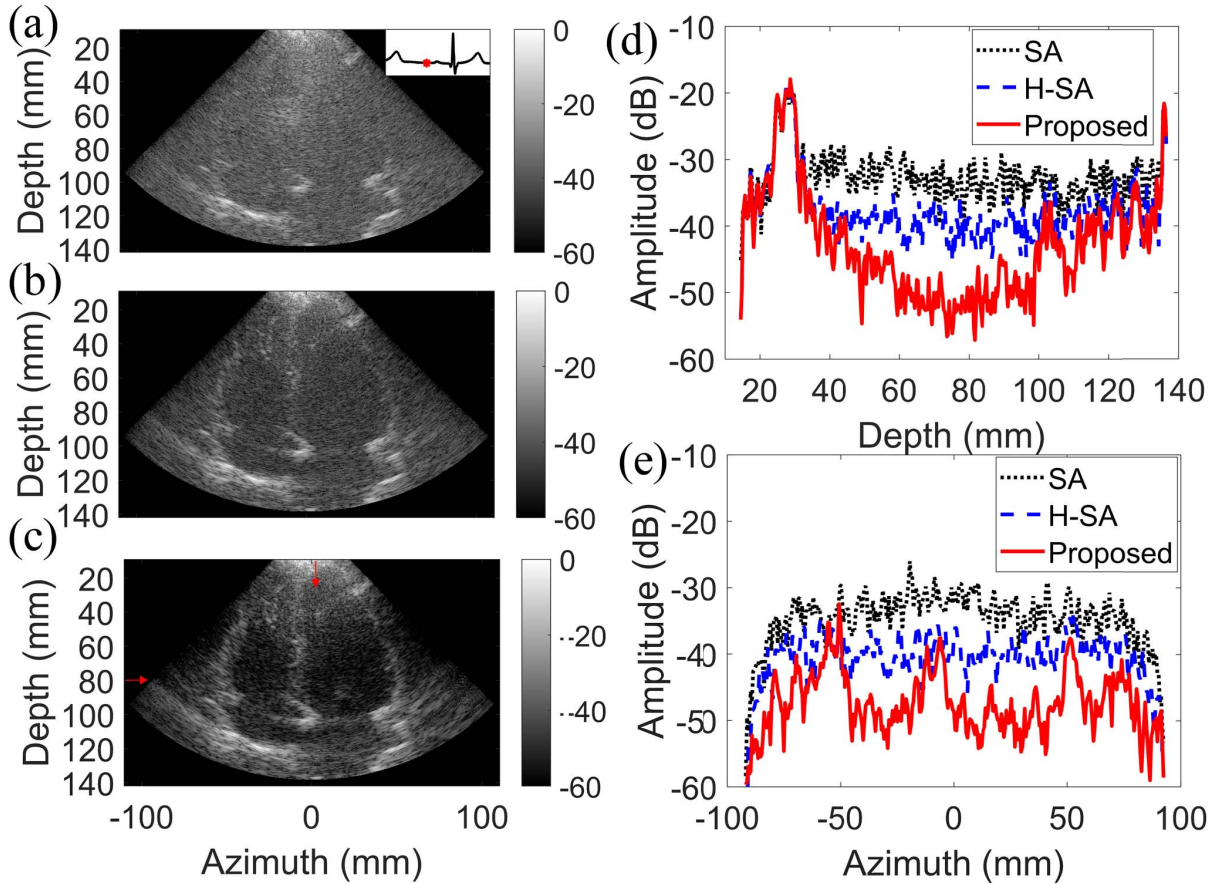
Fig. 6. B-mode images of an *in vivo* human heart in the apical four-chamber view. (a) SA. (b) H-SA. (c) Proposed CaSA. (d) Axial profile at azimuth 29 mm (red axial arrow). (e) Lateral profile at a depth of 78 mm (red lateral arrow). The phase is shown as the red point in the ECG signal.

ultrafast ultrasound imaging and shows enhanced quality for the imaging of heart dynamics, including myocardial motion (Fig. 7) and blood dynamics (Fig. 8), than currently available SA and H-SA imaging methods.

The tradeoff between the resolution and the length of transmitted waves is resolved by the newly designed spatiotemporal coding matrix $C_N H_M$. Based on the $C_N H_M$ matrix, we can fully utilize both the spatial and temporal domains of the array aperture to achieve the maximal SNR available. In the spatial domain, M subaperture diverging waves coded with the spatial coefficients are transmitted simultaneously, and the received signals for each subaperture diverging wave can then be separated by the spatial decoding process, without lateral resolution degradation and cross talks from each subaperture diverging wave. In the temporal domain, the N cascaded diverging waves are transmitted as the CUS with the temporal coding coefficients from the $C_N H_M$ matrix, and the received signals from each wave are separated from the temporal decoding process, without sacrificing the axial resolution. As detailed in Section II, the SNR can be improved by $10 \times \log_{10}(N \times M)$ with respect to SA imaging, and $10 \times \log_{10}(N)$ compared with H-SA imaging.

We evaluated the performance of the CaSA method with four subapertures ($M = 4$) and 32 cascaded waves ($N = 32$). Because of the designed CDW transmissions and

corresponding decoding process, the CaSA imaging achieved the increase in SNR and contrast with comparable spatial resolution as evidenced in the calibration phantom results. The improvement of contrast in deep regions stemmed from the improvement of SNR. In the case of the *in vivo* human heart, the electronic noise complicated the delineation of the myocardium wall in both SA and H-SA images. The heart wall was better delineated from the CaSA image because of the SNR increase. Since the sonographic SNR is known to impact ultrasound speckle tracking [38], the proposed CaSA shows the highest quality of myocardial displacement image [Fig. 7(c), (d) and (g), (h)] than that of SA and H-SA ones. In the case of blood dynamics, the clutter filtering process can profoundly suppress the echoes from the myocardium, but the electronic noise remains in the SA images. Since the echoes from the blood are weak and interfered by the electronic noise, the separation of the blood in the heart chamber from the heart wall necessitates noise suppression. The H-SA images show the better delineation of the blood than SA ones because of the fourfold increase in SNR by spatial coding. In the case of CaSA, the electronic noise level is more significantly suppressed, and the blood dynamics in the ventricle can thus be highlighted.

In theory, the CaSA method can be extended to any $N = 2^Q$, $M = 2^Z$, where Q and Z are integers. The number
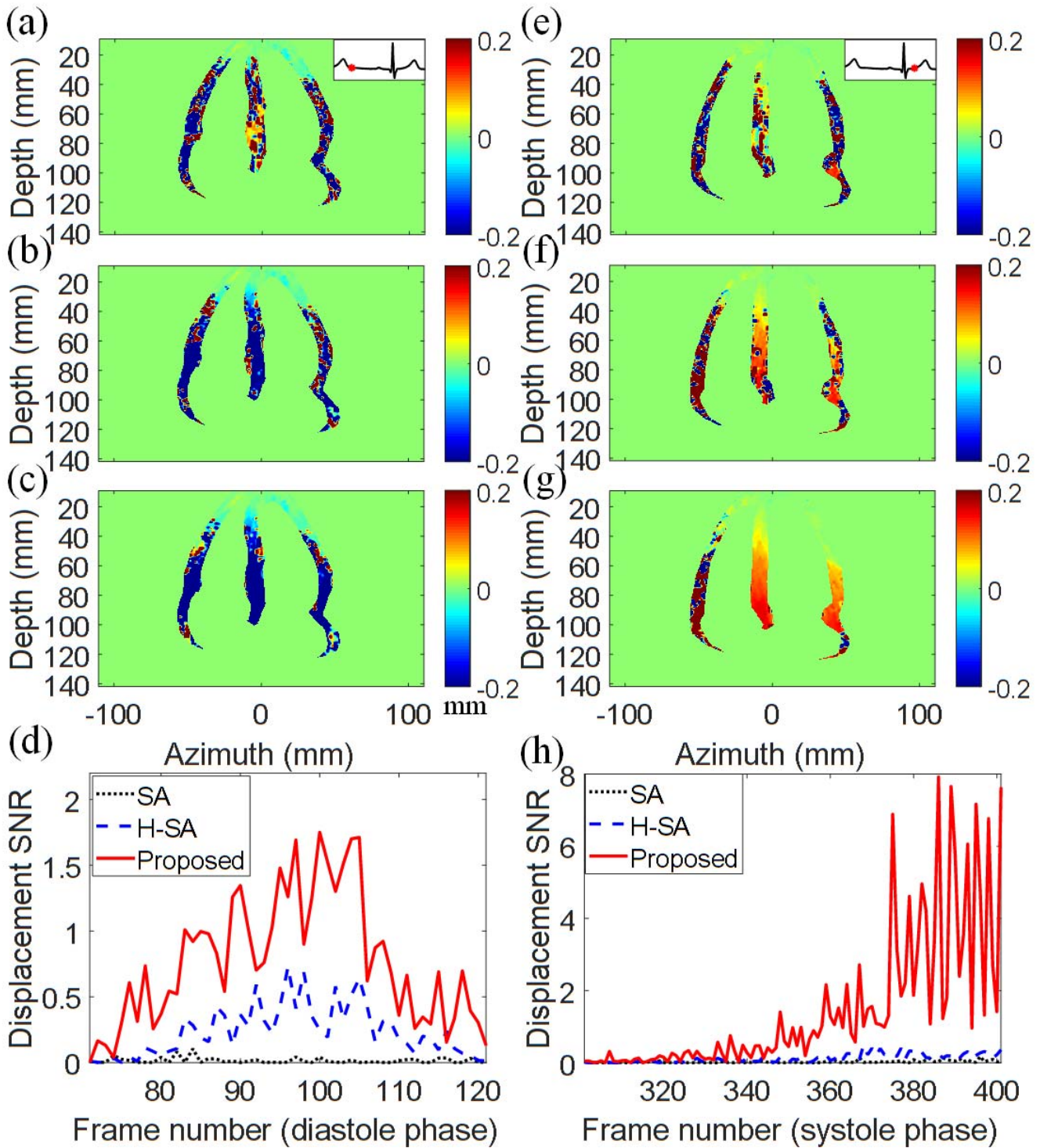
Fig. 7. Displacement maps of the myocardium in the apical four-chamber view based on (a) SA, (b) H-SA, and (c) proposed CaSA reconstructed images at a diastolic phase and (e) SA, (f) H-SA, and (g) proposed CaSA at a systolic phase. (d) Temporal profiles of displacement SNR in the septum during diastole. (h) Displacement SNR in the septum region during systole. The phase is shown as the red point in the ECG signal in (a) and (e).

of transmission–reception events (M) and the number of cascaded waves (N) can be determined by users according to applications. If the frame rate is much more important, we use a smaller M; if the image quality (i.e., image contrast) is much more important, we choose a larger M. If the SNR is much more important, we should consider a larger N.

However, we should use a smaller N if the ROI is in the vicinity of the probe surface. One of the factors that limit the number of transmission–reception events and the number of cascaded waves is the motion susceptibility of the method. The tissue motion will lead to the misalignment of the received signals from different transmission–reception events in the
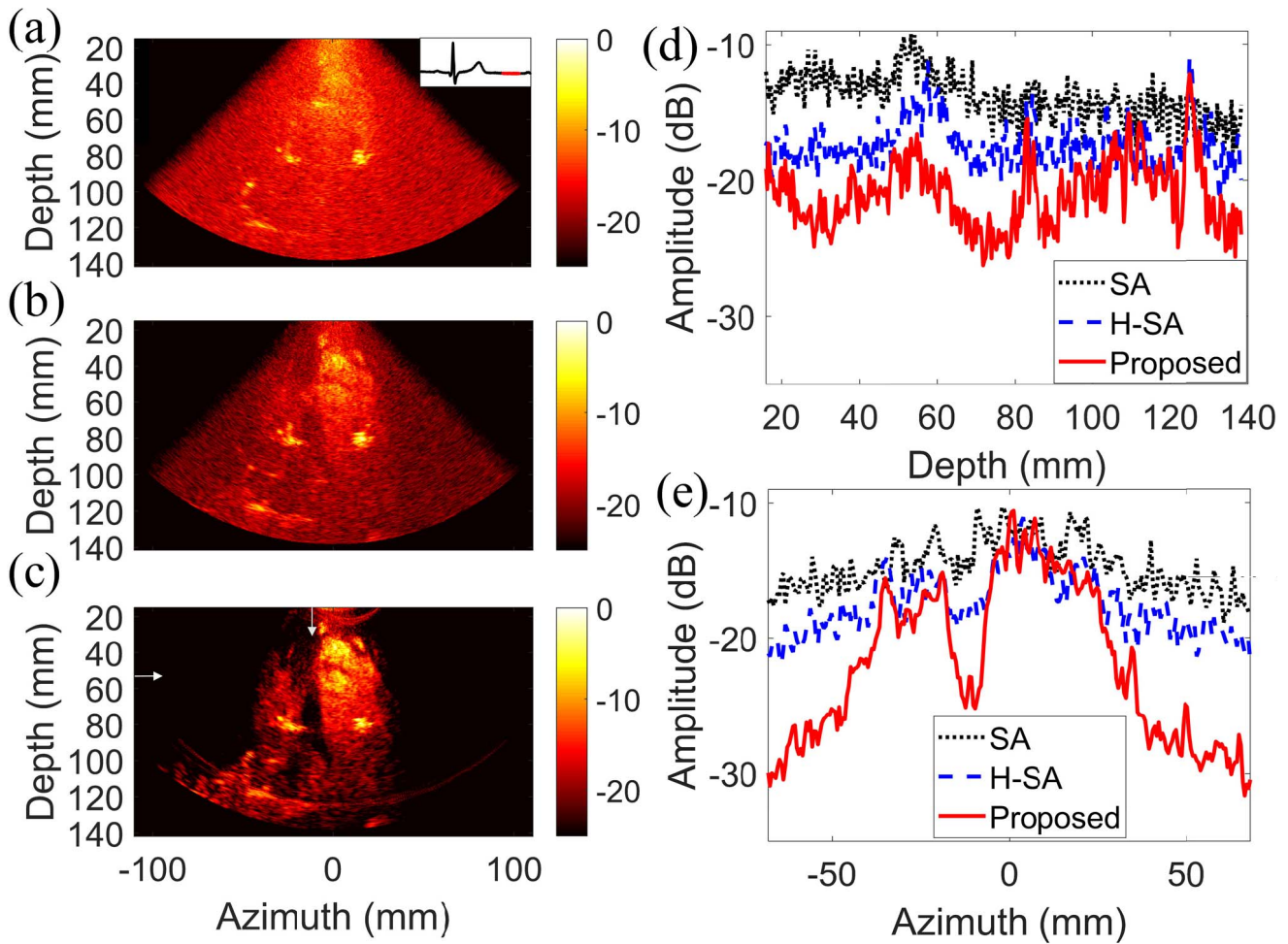
Fig. 8. Power Doppler images of the *in vivo* human heart during diastole by (a) SA, (b) H-SA, and (c) proposed CaSA. (d) Axial profiles at azimuth −12 mm (white axial arrow). (e) Lateral profiles at a depth of 55 mm (white lateral arrow). The examined cardiac phase is highlighted as the red region in the ECG signal in (a).

decoding process. The longer the pulses and the larger number of subaperture transmissions, the more visible the motion artifacts because of imperfect decoding. For imaging fast moving targets (i.e., the heart), we suggest a small number of subapertures and a limited number of cascaded waves (i.e., $M = 4$, and $N = 32$ in the presented example). However, for imaging slow moving targets (i.e., perfusion imaging), we can increase the number of subapertures and cascaded waves to achieve higher image contrast and SNR. Another factor that limits the number of cascaded waves is the dead zone in the near field when the transducer does not receive the signals during transmissions. The range of the dead zone is half the length of the transmitted cascaded waves. For example, in our experiments, the duration of a single pulse was 0.4 μs, and the short time interval between two successive cascaded waves was 0.6 μs. The duration of 32 cascaded waves was approximately 32 μs. The range of the dead zone was estimated to be around 24.64 mm given the sound velocity of 1540 m/s. However, for cardiac imaging, the ROI is generally beyond the estimated size of the dead zone except for the apical region shown in Fig. 6.

In the experiments, we ensured that the three methods used the same voltage output from the Vantage 256 system to drive the phased array probe for the generation of ultrasound pulses. The transmitted voltage in the Verasonics system is directly proportional to the transmitted ultrasound wave pressure level or the peak intensity of the transmitted ultrasound signals. We performed the acoustic measurement in a separate study, but in this study, we mainly focus on the methods and did not measure the acoustic pressure for the calculation of the mechanical index (MI) and the spatial peak temporal average intensity (I_{SPTA}). Nevertheless, we ensured that the three methods (SA, H-SA, CaSA) were under the same transmission settings, including the transmitted voltage, frequency, apodization, subaperture, and so on. The only difference among CaSA, SA, and H-SA was the total length of the transmitted signals. As we used the same transmission settings, three methods in theory had the same acoustic pressure level and MI. The I_{SPTA} values in theory were different among the three methods because we used different lengths of the transmitted signals. Although longer signals were transmitted in CaSA than SA and H-SA, the total signal length and therefore I_{SPTA} in CaSA remained way below the FDA limit in the current settings.

There are several limitations of the current experiments. We used the low transmitted voltage (1.6 V for B-mode, and 16 V for Doppler imaging) for the transmissions in

the experiments. Note that when the SNR is high enough in certain applications, even though we improve the SNR by the CaSA method, we may not observe a significant enhancement of image quality (especially the image contrast). For example, the improvement of CaSA for image quality will not be such significant as compared to SA and H-SA when we use a very high voltage (i.e., 30 V) for transmission. In this case, for the B-mode images, SNR is not the dominant factor of image quality. Instead, the artifacts from the side lobes and axial lobes will affect the image contrast. This study thus aimed to demonstrate that CaSA could increase the SNR of ultrafast ultrasound imaging, and that CaSA could achieve acceptable SNR when we used the lowest available voltage to drive the array probe. In addition, for the experimental settings, the pulselength was the same for B-mode and power Doppler. We understand that Doppler pulses are typically longer than those of B-mode for the increase of backscattering intensity. In order to be consistent with the calibration phantom study and for the convenience of *in vivo* experiment comparison, we used the same transmission settings, except for the voltage, to obtain B-mode images and Doppler images. This is the limitation of the study where we did not redesign the transmitted waveform of CaSA based on long basis pulses of SA and H-SA for Doppler imaging. Moreover, given the nonlinear propagation of ultrasonic signals, the second harmonic signals of the positive and negative waves cannot be canceled in the decoding process, and imperfect signal cancellation may cause the crosstalk of different subapertures. However, this kind of crosstalk could be suppressed by applying a bandpass filter for the received signals.

Most importantly, at the lowest available voltage that drives the array probe, CaSA can achieve acceptable SNR. Medical ultrasound imaging may exert adverse mechanical or thermal effects on the living organism if the MI, thermal index (TI), and acoustic intensity parameters exceed the upper limits set by the FDA guidelines. The MI and TI of the CaSA method are less than the upper limits in the experimental settings. The proposed CaSA method can abate the transmitted voltage by the square root of $N \times M$ to obtain ultrasound images with a comparable SNR to conventional methods. In addition, the decoding process of CaSA only consists of several delay, addition, and subtraction operations, and its resultant extra computational load is much lower than the image formation process itself. The extra time by CaSA is negligible in terms of real-time display. We have implemented CaSA in our Verasonics Vantage 256 system for real-time display at 40 Hz. We envision that the enhanced image qualities and ultrafast frame rates offered by CaSA imaging beget great potential for various novel imaging applications.

VI. Conclusion

Central to healthy aging and reduction in CVD mortality are affordable and safe health care technologies that can monitor cardiac function. Biomedical ultrasound is the superior choice because it is cost-effective, portable, real time, and free from ionizing radiation. Imaging the heart at high frame rates (>1000 Hz) permits detailed quantification of heart dynamics, including morphological alterations, muscle contraction, and

blood dynamics, at instants of interest in a cardiac cycle. However, degraded SNR and limited penetration are traded for high frame rate. This study presents an unprecedented method—uCUS with a straightforward design of spatiotemporal coding–decoding of ultrasound waves—that overcomes the tradeoffs for the assessment of heart dynamics as a potential low-cost, risk-free diagnostic tool in cardiology.

References

[1] World Health Organization, *World Health Statistics 2018: Monitoring Health for the Sustainable Development Goals (SDGs)*. Geneva, Switzerland: World Health Organization, 2018.

[2] *Hearts: Technical Package for Cardiovascular Disease Management in Primary Health Care*. Geneva, Switzerland: World Health Organization, 2016.

[3] L. N. Bohs and G. E. Trahey, "A novel method for angle independent ultrasonic imaging of blood flow and tissue motion," *IEEE Trans. Biomed. Eng.*, vol. 38, no. 3, pp. 280–286, Mar. 1991.

[4] W. N. McDicken, G. R. Sutherland, C. M. Moran, and L. N. Gordon, "Colour Doppler velocity imaging of the myocardium," *Ultrasound Med. Biol.*, vol. 18, nos. 6–7, pp. 651–654, 1992.

[5] K. Miyatake *et al.*, "New method for evaluating left ventricular wall motion by color-coded tissue Doppler imaging: In vitro and *in vivo* studies," *J. Amer. College Cardiol.*, vol. 25, no. 3, pp. 717–724, Mar. 1995.

[6] E. Konofagou *et al.*, "Two-dimensional ultrasonic strain rate measurement of the human heart *in vivo*," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 49, no. 2, pp. 281–286, Feb. 2002.

[7] E. E. Konofagou, J. D'hooge, and J. Ophir, "Myocardial elastography—A feasibility study *in vivo*," *Ultrasound Med. Biol.*, vol. 28, no. 4, pp. 475–482, Apr. 2002.

[8] A. Heimdal, A. Støylen, H. Torp, and T. Skjærpe, "Real-time strain rate imaging of the left ventricle by ultrasound," *J. Amer. Soc. Echocardiography*, vol. 11, no. 11, pp. 1013–1019, Nov. 1998.

[9] H. Kanai, "Propagation of spontaneously actuated pulsive vibration in human heart wall and *in vivo* viscoelasticity estimation," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 52, no. 11, pp. 1931–1942, Nov. 2005.

[10] S. J. Hsu, R. R. Bouchard, D. M. Dumont, P. D. Wolf, and G. E. Trahey, "*In vivo* assessment of myocardial stiffness with acoustic radiation force impulse imaging," *Ultrasound Med. Biol.*, vol. 33, no. 11, pp. 1706–1719, Nov. 2007.

[11] M. Couade *et al.*, "*In vivo* quantitative mapping of myocardial stiffening and transmural anisotropy during the cardiac cycle," *IEEE Trans. Med. Imag.*, vol. 30, no. 2, pp. 295–305, Feb. 2011.

[12] M. W. Urban, C. Pislaru, I. Z. Nenadic, R. R. Kinnick, and J. F. Greenleaf, "Measurement of viscoelastic properties of *in vivo* swine myocardium using Lamb wave dispersion ultrasound vibrometry (LDUV)," *IEEE Trans. Med. Imag.*, vol. 32, no. 2, pp. 247–261, Feb. 2013.

[13] R. R. Bouchard, S. J. Hsu, P. D. Wolf, and G. E. Trahey, "*In vivo* cardiac, acoustic-radiation-force-driven, shear wave velocimetry," *Ultrason. Imag.*, vol. 31, no. 3, pp. 201–213, Jul. 2009.

[14] B.-F. Osmanski, D. Maresca, E. Messas, M. Tanter, and M. Pernot, "Transthoracic ultrafast Doppler imaging of human left ventricular hemodynamic function," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 61, no. 8, pp. 1268–1275, Aug. 2014.

[15] D. Posada *et al.*, "Staggered multiple-PRF ultrafast color Doppler," *IEEE Trans. Med. Imag.*, vol. 35, no. 6, pp. 1510–1521, Jun. 2016.

[16] M. Toulemonde *et al.*, "Cardiac imaging with high frame rate contrast enhanced ultrasound: In-vivo demonstration," in *Proc. IEEE Int. Ultrason. Symp. (IUS)*, Sep. 2016, pp. 1–4.

[17] D. Maresca *et al.*, "Noninvasive imaging of the coronary vasculature using ultrafast ultrasound," *JACC, Cardiovascular Imag.*, vol. 11, no. 6, pp. 798–808, Jun. 2018.

[18] M. Cikes, L. Tong, G. R. Sutherland, and J. D'hooge, "Ultrafast cardiac ultrasound imaging: Technical principles, applications, and clinical benefits," *JACC: Cardiovascular Imag.*, vol. 7, no. 8, pp. 812–823, Aug. 2014.

[19] J. Provost *et al.*, "3D ultrafast ultrasound imaging *in vivo*," *Phys. Med. Biol.*, vol. 59, no. 19, pp. L1–L13, Sep. 2014.

[20] S. Wang, W.-N. Lee, J. Provost, J. Luo, and E. E. Konofagou, "A composite high-frame-rate system for clinical cardiovascular imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 55, no. 10, pp. 2221–2233, Oct. 2008.

[21] C. Moore *et al.*, "Live high-frame-rate echocardiography," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 62, no. 10, pp. 1779–1787, Oct. 2015.

[22] L. Tong, A. Ramalli, R. Jasaityte, P. Tortoli, and J. D'hooge, "Multi-transmit beam forming for fast cardiac imaging—Experimental validation and *in vivo* application," *IEEE Trans. Med. Imag.*, vol. 33, no. 6, pp. 1205–1219, Jun. 2014.

[23] G. Matrone, A. Ramalli, A. S. Savoia, P. Tortoli, and G. Magenes, "High frame-rate, high resolution ultrasound imaging with multi-line transmission and filtered-delay multiply and sum beamforming," *IEEE Trans. Med. Imag.*, vol. 36, no. 2, pp. 478–486, Feb. 2017.

[24] H. Hasegawa and H. Kanai, "High-frame-rate echocardiography using diverging transmit beams and parallel receive beamforming," *J. Med. Ultrason.*, vol. 38, no. 3, pp. 129–140, 2011.

[25] J. Porée, D. Posada, A. Hodzic, F. Tournoux, G. Cloutier, and D. Garcia, "High-frame-rate echocardiography using coherent compounding with Doppler-based motion-compensation," *IEEE Trans. Med. Imag.*, vol. 35, no. 7, pp. 1647–1657, Jul. 2016.

[26] M. Correia, J. Provost, S. Chatelin, O. Villemain, M. Tanter, and M. Pernot, "Ultrafast harmonic coherent compound (UHCC) imaging for high frame rate echocardiography and shear-wave elastography," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 63, no. 3, pp. 420–431, Mar. 2016.

[27] C. Papadacci, M. Pernot, M. Couade, M. Fink, and M. Tanter, "High-contrast ultrafast imaging of the heart," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 61, no. 2, pp. 288–301, Feb. 2014.

[28] J. A. Jensen, S. I. Nikolov, K. L. Gammelmark, and M. H. Pedersen, "Synthetic aperture ultrasound imaging," *Ultrasonics*, vol. 44, pp. e5–e15, Dec. 2006.

[29] P. Gong, M. C. Kolios, and Y. Xu, "Delay-encoded transmission and image reconstruction method in synthetic transmit aperture imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 62, no. 10, pp. 1745–1756, Oct. 2015.

[30] R. Y. Chiao, L. J. Thomas, and S. D. Silverstein, "Sparse array imaging with spatially-encoded transmits," in *Proc. IEEE Ultrason. Symp. Int. Symp.*, Oct. 1997, pp. 1679–1682.

[31] T. D. Mast, L. M. Hinkelman, L. A. Metlay, M. J. Orr, and R. C. Waag, "Simulation of ultrasonic pulse propagation, distortion, and attenuation in the human chest wall," *J. Acoust. Soc. Amer.*, vol. 106, no. 6, pp. 3665–3677, Dec. 1999.

[32] E. G. Moros, X. Fan, and W. L. Straube, "Ultrasound power deposition model for the chest wall," *Ultrasound Med. Biol.*, vol. 25, no. 8, pp. 1275–1287, Oct. 1999.

[33] T. Misaridis and J. A. Jensen, "Use of modulated excitation signals in medical ultrasound. Part I: Basic concepts and expected benefits," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 52, no. 2, pp. 177–191, Feb. 2005.

[34] T. Misaridis and J. A. Jensen, "Use of modulated excitation signals in medical ultrasound. Part II: Design and performance for medical imaging applications," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 52, no. 2, pp. 192–207, Feb. 2005.

[35] Y. Zhang, Y. Guo, and W.-N. Lee, "Ultrafast ultrasound imaging with cascaded dual-polarity waves," *IEEE Trans. Med. Imag.*, vol. 37, no. 4, pp. 906–917, Apr. 2018.

[36] S. S. Agaian, *Hadamard Matrices and their Applications*. Princeton, NJ, USA: Princeton Univ. Press, 2012.

[37] H. Li, Y. Guo, and W.-N. Lee, "Systematic performance evaluation of a cross-correlation-based ultrasound strain imaging method," *Ultrasound Med. Biol.*, vol. 42, no. 10, pp. 2436–2456, 2016.

[38] H. Li and W.-N. Lee, "Effects of tissue mechanical and acoustic anisotropies on the performance of a cross-correlation-based ultrasound strain imaging method," *Phys. Med. Biol.*, vol. 62, no. 4, p. 1456, Jan. 2017.

Yang Zhang (M'19) received the B.Eng. degree in information engineering from the South China University of Technology, Guangzhou, China, in 2012, the M.Eng. degree in signal and information processing from Peking University, Beijing, China, in 2015, and the Ph.D. degree in electrical and electronic engineering from the University of Hong Kong, Hong Kong, in 2018.

He is currently a Postdoctoral Fellow with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong. His current research interests include developing mathematical and physical tools for addressing medical and health issues.

He Li received the M.S. degree in biomedical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2013, and the Ph.D. degree from the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, in 2017.

He is currently a Postdoctoral Fellow with the Department of Electrical and Electronic Engineering, The University of Hong Kong. His current research interests include ultrasound motion estimation and strain imaging in cardiovascular applications and ultrasound contrast imaging.

Wei-Ning Lee (M'14) received the B.S. and M.S. degrees in electrical engineering from National Taiwan University, Taipei, Taiwan, and the Ph.D. degree in biomedical engineering from Columbia University, New York, NY, USA.

She worked as a Postdoctoral Fellow at Institut Langevin, ESPCI ParisTech, Paris, France, from 2010 to 2012. She then joined as Assistant Professor with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong, where she is currently Associate Professor. Her current research interests focus on functional ultrafast ultrasound imaging methods to investigate microstructure, mechanics, and hemodynamics of cardiovascular and musculoskeletal tissues.

Dr. Lee was the recipient of the New Investigator Award at the American Institute of Ultrasound in Medicine (AIUM) Annual Convention in 2009 and an Early Career Award from the Hong Kong Research Grands Council in 2013. She currently serves as an Academic Editor for *PLoS One* and a member for International Advisory Board of *Physics in Medicine and Biology* and on the Advisory Editorial Board of *Ultrasound in Medicine and Biology*.