Noise Suppression of SAFT Ultrasound Video Using Eigen Filters

CHENGYANG HUANG and FRANCESCO LANZA DI SCALEA

ABSTRACT

In ultrasound scanning videos, noise and sub-stationary artifacts can severely impair the detection of the intended targets. This paper advances an unsupervised learning framework for noise suppression in ultrasound videos based on eigen filter concept. Specifically, the study focuses on the removal of artifacts in Synthetic Aperture Focusing Technique (SAFT) beamformed videos created by a moving transducer array in a scanning mode. The specific application is in-motion imaging of internal flaws in railroad tracks using a transducer array mounted in a Rolling Search Unit (RSU). By employing Singular Value Decomposition (SVD) analysis, an ultrasound video is decomposed into principal components through their time-space coherence/correlation. Artifacts can be eliminated by projecting out the associated principal components from the video (i.e. eigen filtering). A novel recursive algorithm with signal rectification is proposed to allow SVD filters to better capture the subtle movement of artifacts in ultrasound tomography. As opposed to the prior knowledge of a stopband in conventional SVD filters, the recursive SVD filter only removes the primary eigen mode until convergence. The proposed approach is validated through experiments using an RSU hosting a 25-element linear transducer array in defected rail sections in the laboratory and in the field. The results demonstrate a remarkable filtering ability, where artifacts (e.g. reflections at the wheel-rail interface, reverberations within the wheel, reflections from the bottom flange of the rail head) are successfully eliminated and only the internal flaws are imaged with high accuracy. This approach is applicable to any ultrasound imaging test where “loud” artifacts are present in addition to the intended targets.

Chengyang Huang; Francesco Lanza di Scalea, Experimental Mechanics & NDE Laboratory, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, U.S.A.

INTRODUCTION

Ultrasonic NDE is often challenged by strong clutter artifacts and noise that obscure flaw signals. In particular, heterogeneous materials and irregular geometries produce structural response/noise that can completely submerge defect echoes, greatly limiting detectability. A common artifact-removal strategy is baseline subtraction, where a reference (undamaged) waveform is subtracted from the current measurement. Under ideal conditions this procedure cancels static reflections and cleanly reveals new flaw echoes. However, any mismatch between the baseline and inspection conditions (e.g. due to temperature changes, loading, or slight sensor repositioning) produces significant residuals. In practice, these uncertainties often cause baseline subtraction to fail at isolating defect signals, motivating the need for more robust clutter-removal techniques.

Consequently, advanced filtering techniques are needed to improve the signal-to-noise ratio and reveal subtle flaws. Principal component analysis (PCA) is one such adaptive unsupervised learning approach that has been applied for clutter suppression in ultrasound. In PCA, the data are re-expressed on an orthogonal basis determined from the signal covariance; projecting the data onto its principal modes then allows dominant correlated clutter to be identified and rejected, thereby enhancing weaker defect signals [1]. In medical ultrasound imaging, clutter suppression is often achieved via singular value decomposition (SVD) filtering [2]. Biological tissue echoes tend to be highly coherent and dominated by a few low-rank modes. For example, the first few singular vectors of quasi-static tissue often share similar frequency content, reflecting the low-dimensional nature of the background [3]. The SVD filtering framework exploits this property by adaptively weighting or rejecting SVD modes to remove stationary clutter. The low-order singular components capture the bulk of the tissue (static) echo, allowing dynamic flow or contrast signals to be separated by discarding those modes.

By contrast, engineered structures often violate these assumptions. Structural elements (e.g. holes, welds, fasteners) and guided-wave multipathing produce complex echoes that lack the uniformity of soft-tissue signals. Frequency components of flaws and clutter can overlap, so that singular modes do not correspond neatly to background vs. defect. Consequently, directly applying SVD filters designed for medical imaging is problematic in NDE, since the singular spectrum does not cleanly separate static and dynamic components. This study presents a way of rejecting sub-static artifact signals from Ultrafast Synthetic Aperture Focusing Technique (SAFT) videos via a recursive SVD filtering algorithm. Considerable improvements in the image/video quality are shown by separating quasi-static “baseline” reflections from rail flaws.

PROBLEM STATEMENT

The ultrasound video of a railway track was recorded using ultrafast capability of enhanced SAFT [4]. As shown in Figure 1, the RSU wheel was mounted on a test rail section in the UCSD rail defect farm by designing an aluminum frame that allowed longitudinal movement for scanning along the rail. The orientation of the 25-element linear transducer array inside the RSU was rotated at 16 degrees from the horizontal plane to create inclined transmission/reception beam directions so as to maximize the reflections from transverse-type defects (TDs). As shown in Figure 1(a), the current study considers a 2D image in the x-y vertical plane in the rail head. The enhanced

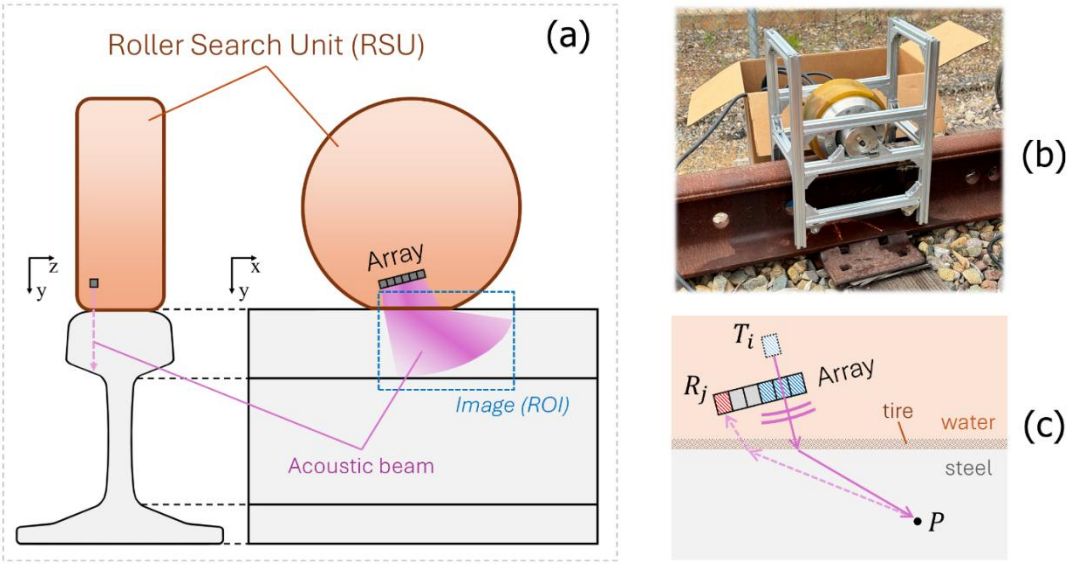

Figure 1. (a) Schematic of the Roller Search Unit (RSU) for the ultrasonic imaging of rail flaws. (b) The RSU scanning a rail section. A mounting frame helps to position the RSU on the track. (c) Ray path connecting a transmitting virtual element T_i, a focus point P, and a receiving element R_j.

SAFT approach beamforms an ultrasonic image (a frame) by transmitting a subaperture in each firing and receiving in the whole aperture. As schematized in Figure 1(c), a subaperture is fired in a designated delay pattern to mimic a virtual transmit element T_i. Each subaperture uses 11 elements and fires at a 30-degree angular range, which increases the SNR of transmitted beam while preserving the resolution. A multiplexer records the Full-Matrix Capture (FMC) dataset, which can be seen as in inter-element response between the 25-element virtual transmission array and the 25-element physical reception array. The Time-of-Flight (TOF) in each inter-element response can be found by considering wave refraction at the tire/steel interface through ray tracing algorithms.

An ultrafast SAFT acquisition consists of a stack of beamformed ultrasound intensity images and can be represented under the positive valued variable s(x, y, t), where t stands for discrete time that is sampled at a frequency noted as frame rate. As pointed out in Figure 2(a), a typical ultrasound SAFT image of a rail head has the following major contributions: 1) reflection from the water/tire interface, 2) reflection from the tire/steel interface, 3) scattering from water that is illuminated by the

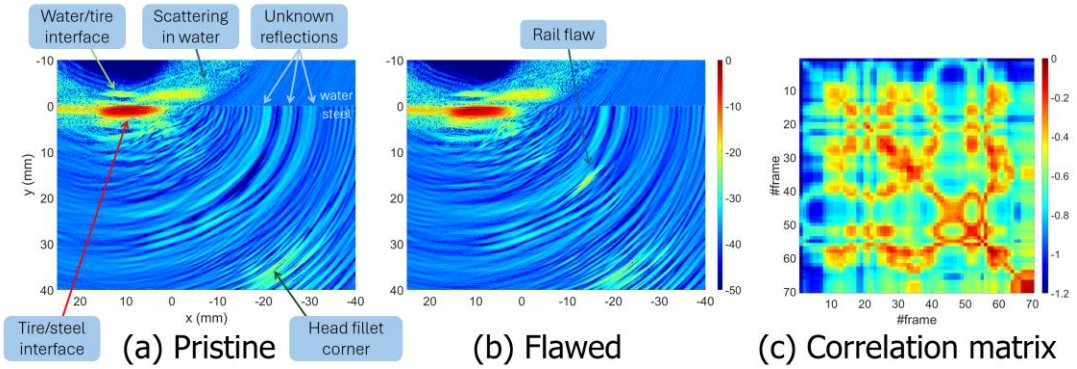

Figure 2. Raw ultrasonic video beamformed by the enhanced SAFT approach. (a) A typical frame showing no present rail flaws (pristine case). (b) A frame showing possible indications of a rail flaw. (c) The temporal correlation matrix of the raw ultrasonic video.

transmitted acoustic beam, 4) multiple bounces of reflection from the assembly parts within the water in the RSU that is incorrectly beamformed in the region of steel (denoted as unknown reflections) and 5) reflection from the rail head fillet corner. These principal components are coherent in s(x, y, t) over space and time. For the targets of interest (rail flaws), they only occur in very few frames in the ultrasound video as shown in Figure 2(b) and thus do not contribute to the principal components. It is also worth noting that for one particular rail flaw, its spatial information within neighboring frames can be very distinct since the RSU is moving along the rail track. In other words, with sufficient recording time t (that covers a long enough distance along the rail track), the rail flaws have a sparse distribution over space and time in s(x, y, t). The example rail flaw shown in Figure 2(b) is a reflection from a side-drilled hole (SDH) that has an intensity of -20 dB with respect to the strongest tire/steel reflection. Given that the noise floor is around -35 dB, the SNR of the SDH is actually 15 dB. For weaker reflectors such as natural rail flaws (caused by rolling contact fatigue), the SNR of the target can be much smaller [5]. This paper aims to find a post-processing scheme on the ultrasonic video s(x, y, t) that suppresses the principal components while increasing the SNR of the target signal (rail flaws).

METHODS AND RESULTS

Let us now consider the spatiotemporal matrix s(x, y, t) that corresponds to a set of (N_x, N_y, N_t) discrete samples. The raw data matrix can be reshaped under a Casorati matrix form by compressing 2D images (x, y) into 1D column such that the spatiotemporal matrix has a 2D dimension of ($N_x \times N_y$, N_t). Without prior knowledge of the principal signal components, one of the most frequently used Blind Source Separation (BSS) techniques is Singular Value Decomposition (SVD), a generalization of eigen decomposition that decomposes the Casorati matrix C as

$$C = U\Sigma V^{\mathrm{H}} \tag{1}$$

where U and V are orthonormal matrices with respective dimensions ($N_x \times N_y$, $N_x \times N_y$) and (N_t, N_t), and Σ is a non-square ($N_x \times N_y$, N_t) diagonal matrix containing the singular values. If the singular modes are properly sorted (typically from largest to smallest according to the singular values in Σ), then each i^{th} column in U and V corresponds to the left and right singular vectors of the i^{th} singular mode whose singular value is the i^{th} (largest) term in Σ. The SVD of the Casorati matrix can be seen as decomposing C into a weighted, ordered sum of orthogonal matrices

$$s(x, y, t) = \sum_i \sigma_i I_i(x, y)\, V_i(t) \tag{2}$$

where each left singular vector U_i with dimension of $N_x \times N_y$ becomes a 2D image I_i with dimension of (N_x, N_y) that is denoted as an "eigen image", and each right singular vector V_i with dimension of N_t is denoted as an "eigen history". In other words, (2) shows that an SAFT video is a weighted sum of eigen images behaving in their eigen history.

From the principal component analysis of ultrasonic tomography video via SVD, the baseline to be removed from the raw video are expected to exist mainly in the first few singular modes, since their high spatiotemporal coherence ensures that a large

number of spatial pixels exhibit the same time profile (outer product of eigen image and eigen history). Thus, a way to eliminate the baseline can be a direct subtraction of the contribution of the dominant singular modes

$$s_{filtered}(x,y,t) = s(x,y,t) - \sum_{i=1}^{N_m} \sigma_i I_i(x,y) V_i(t) \quad (3)$$

where the subtracted collection of dominant singular modes starts from the 1st (strongest singular value) to N_m that is determined by a threshold. The determination of threshold for BSS filtering strategies has been a topic of open research. There are three major approaches to select the appropriate subset of principal components: empirically per application, experimentally optimized per acquisition, or adaptively based on features of the components (data-driven). The current study only considers trail-and-error to find an optimum threshold based on experimental data. Figure 3 shows two frames of the filtering results using (3) with a different threshold N_m are performed on an experimental SAFT video with a dimension of (N_x, N_y, N_t) = (400, 300, 70). The conventional SVD filter can subtract the baseline fairly well when N_m is smaller than 10. However, none of these results show a perfectly cleansed image as the SNR of the rail flaw with respect to the background noise remains almost the same. This is because the dominant singular modes fails to capture subtle movements of the noise. When the number of cutoff modes N_m is large, the target signal can be eliminated as highlighted by the red boxes. The comparison of Figure 3(a) and (b) shows that these target signals can be embedded at different orders of the singular modes. Thus, the choice of N_m should be carefully considered (from empirical knowledge of the Casorati matrix) to preserve these signals after the filtering.

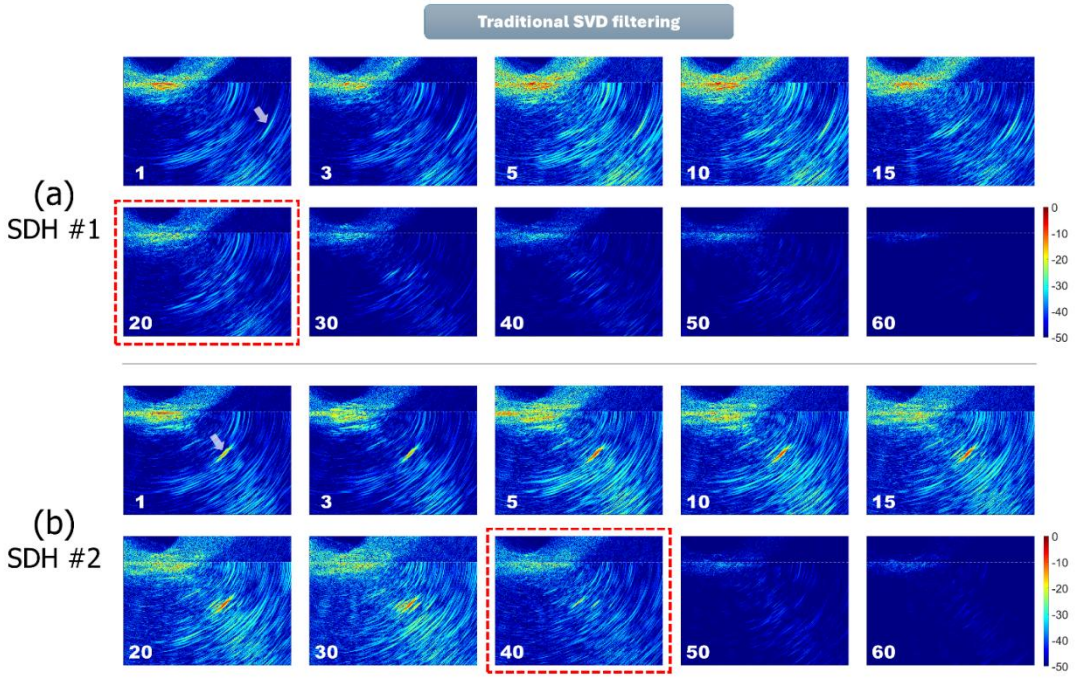

Figure 3. Inefficiency of implementing traditional SVD filtering on the ultrasonic video for rail flaw detection. Different choices of the cut-off singular mode are compared. Filtering result for (a) frame SDH #1: when the SDH is in the right of the ROI, and (b) frame SDH #2: when the SDH is centered in the ROI. The red box highlights the lowest cut-off mode that removes the target signal for each frame.

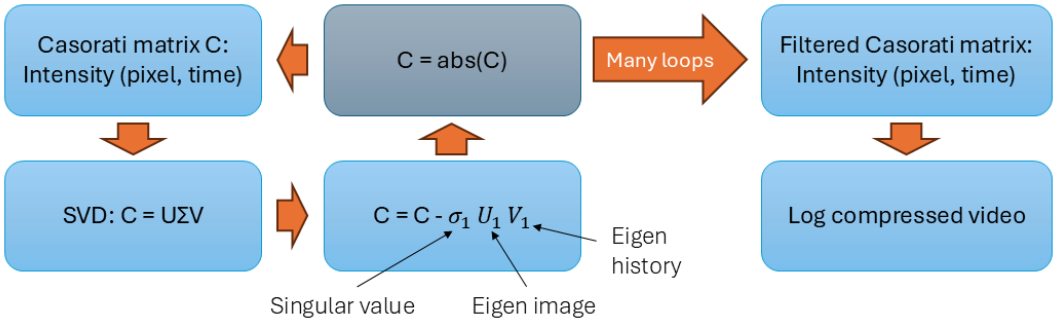

Figure 4. Flowchart of the proposed recursive SVD filter.

The authors are now developing a novel BSS filter technique based on recursive SVD as schematized in Figure 4. Since the SVD of the Casorati matrix is purely numerical, the eigen images and the subtracted Casorati matrix can have considerable number of pixels that exhibit non-physical negative values. To address such limitations, the key features of the proposed recursive SVD filter consists of 1) only rejecting the 1st dominant singular mode per SVD, and 2) taking the absolute of the remainder (filtered) Casorati matrix before feeding to the next SVD. The reason to force rectification in the absolute value of each pixel is because the raw Casorati matrix for ultrasound video is an intensity map (energy distribution) that only allows positive values. When performing BSS filter in a recursive manner, after each subtraction there can be pixels where the value becomes negative and thus loses their physical meaning. A side benefit of forcing the Casorati matrix to be positive in each iteration is that the clusters with phase uncertainty disturbance are eliminated as the loop runs. This results from taking the norm of each pixel that masks spatial regions of semi-consistent response over time, which is equivalent to calculating an envelope of an ultrasound time-series signal.

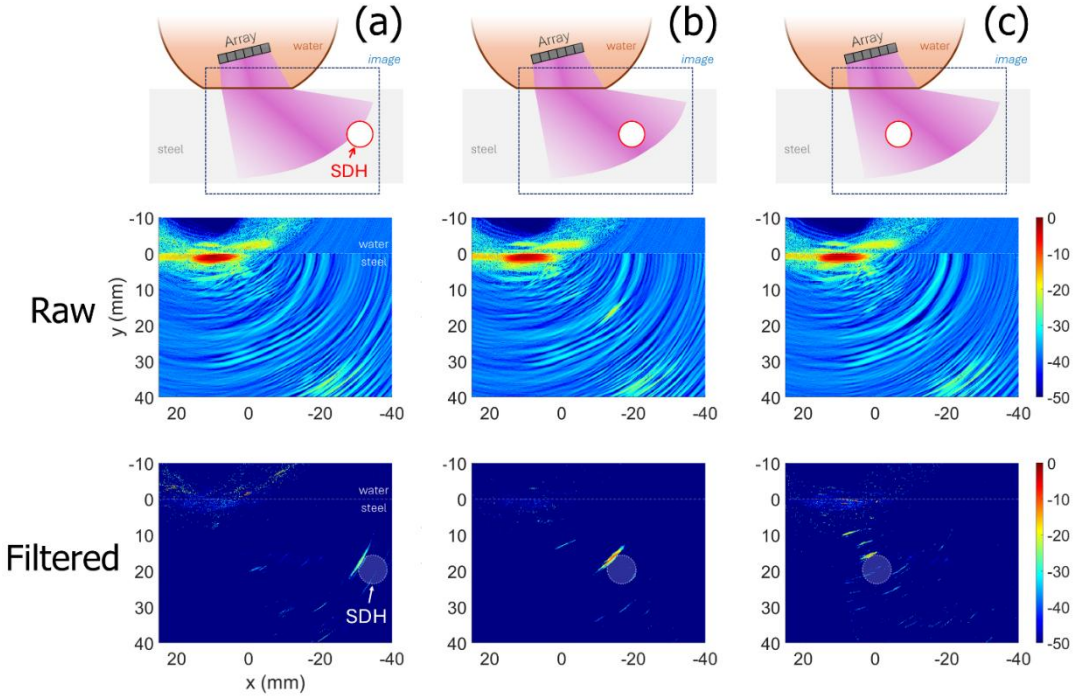

Figure 5. Dramatic noise suppression effect of recursive SVD filter. The signal of interest is an 8-mm Side-drilled Hole (SDH). The RSU moves to the right from (a) to (c).

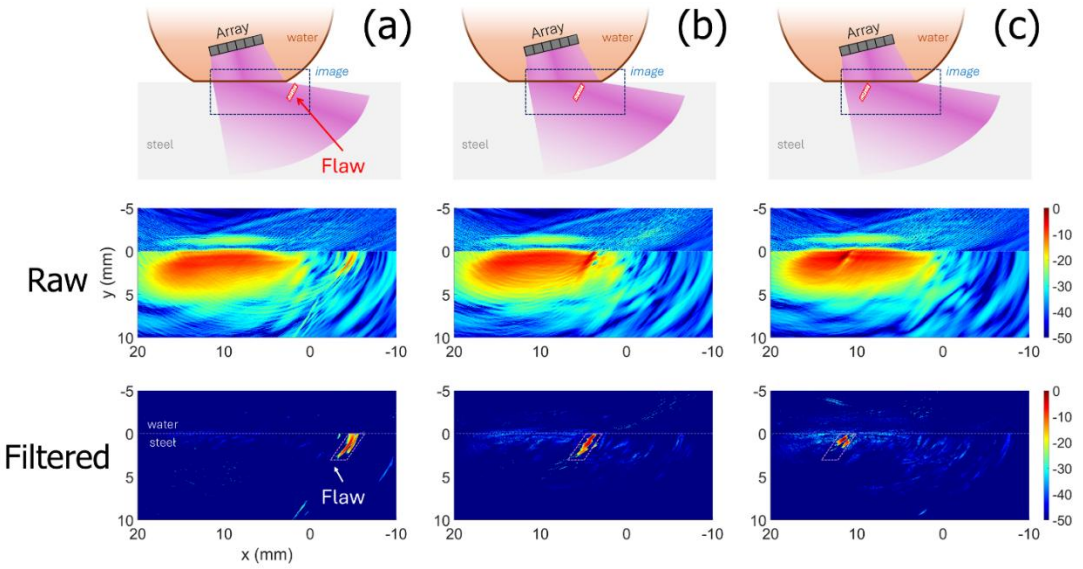

Figure 6. Recursive SVD filtering on a rail sample with a natural rail flaw. Potential flawed region is highlighted by white dashed boxes. The filtered video highlights the target signals (rail flaw echoes) that overlap with one of the principal components (tire/steel interface reflections) in the raw video.

Figure 5 shows the dramatic noise suppression effect of the proposed recursive SVD filter on the same SAFT video as discussed above. Three key frames with the responses from the SDH is shown, and the raw SAFT images are compared to the filtered result. Most of the pixels of the principal components are suppressed below the displayed 50 dB dynamic range. The result is most impressive in the frames shown in Figure 5(a) and (c) where the SDH is located off the broadside of the array, which naturally sees smaller reflection intensity. Finally, Figure 6 shows the cleansing power of the purposed recursive SVD filter on another SAFT video containing a natural rail flaw. The rail flaw is a vertical reflector originating on the rail surface into the depth of roughly 3 mm. The filtered video successfully preserves the response from the rail flaw, even when the tire/steel interface response masks the target signals. The filtered response underestimates the size of the rail flaw because the rail flaw is placed off-axis in the transmission beam. The lack of beam steering in this angular range and the vertical orientation of the rail flaw limits the possibility of accurate ultrasound detection.

CONCLUSION

This paper presents a recursive SVD filtering technique to perform baseline subtraction of ultrasonic SAFT videos on railway flaws. The recursive filter captures space-temporal coherence adaptively to the ultrasound video, and the absolute value rectification enforces elimination of sub-stationary principal signals without prior knowledge of the cut-off rank. Experimental results on SDHs and natural transverse defects show dramatic improvement in dynamic range and cleansing performance, compared to traditional SVD filtering as used in medical ultrasound. Future investigations will be focused on applications on other NDE modalities, such as 1-D ultrasound signals, or thermal infrared videos.

ACKNOWLEDGMENT

This research was funded by the U.S. Office of Naval Research (Agreement No. N00014-24-1-2481), and by MxV Rail through a University Grand Challenge Contract no. 24-0312-008622.

REFERENCES

1. Candès, E. J., X. Li, Y. Ma, and J. Wright. 2011. "Robust Principal Component Analysis?" *J. ACM*, 58(3):1-37.
2. Demené, C., T. Deffieux, M. Pernot, B. F. Osmanski, V. Biran, J. L. Gennisson, L. A. Sieu, A. Bergel, S. Franqui, J. M. Correas, I. Cohen, O. Baud, and M. Tanter. 2015. "Spatiotemporal Clutter Filtering of Ultrafast Ultrasound Data Highly Increases Doppler and fUltrasound Sensitivity." *IEEE Trans. Med. Imaging*, 34(11):2271-2285.
3. Wildeboer, R. R., F. Sammali, R. J. G. Van Sloun, Y. Huang, P. Chen, M. Bruce, C. Rabotti, S. Shulepov, G. Salomon, B. C. Schoot, and H. Wijkstra. 2020. "Blind Source Separation for Clutter and Noise Suppression in Ultrasound Imaging: Review for Different Applications." *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 67(8):1497-1512.
4. Huang, C., and F. Lanza di Scalea. 2024. "Application of Sparse Synthetic Aperture Focusing Techniques to Ultrasound Imaging in Solids Using a Transducer Wedge." *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 71(2):280-294.
5. Huang, C., and F. Lanza di Scalea. 2024. "Rail Flaw Imaging Prototype Based on Improved Ultrasonic Synthetic Aperture Focus Method." *Mater. Eval.*, 82(1):51-59.