

A Portable Ultrasound System Integrating a CMUT Sensor Array with an FPGA-Based Delay-and-Sum Beamformer

Mert Doğan¹, Can Ufuk Ertenli^{2,3}, Mehmet Oğulcan Güngör⁴, Berkay Karacaer²

¹Department of Electrical and Electronics Engineering, METU, Ankara, Turkey

²K AND K Müh. Tek. Ltd. Şti., Ankara, Turkey

³Department of Computer Engineering, METU, Ankara, Turkey

⁴Department of Micro and Nanotechnology, METU, Ankara, Turkey

brk@knktechs.com

Summary:

A portable ultrasound imaging system combining a custom CMUT array, fabricated via a novel SOI-based MEMS process under patent review, with a real-time FPGA delay-and-sum beamformer is presented. The 64-channel array is scanned in eight sequential groups of 8 via analog multiplexing, with on-chip receive beamforming at 100 MHz yielding ~33 ms per focal depth within a 7.9 W total system power budget. Phantom experiments at 1 and 10 MHz validate point-target resolution and tissue-mimicking contrast, confirming feasibility for handheld point-of-care ultrasound.

Keywords: CMUT, FPGA beamforming, portable ultrasound, MEMS, point-of-care imaging

Background, Motivation and Objective

Point-of-care ultrasound (POCUS) demands transducer technology and signal-processing electronics that are simultaneously high-performance, compact, and power-efficient. Capacitive Micromachined Ultrasonic Transducers (CMUTs) are a compelling alternative to conventional piezoelectric arrays, offering batch-fabrication compatibility, wide acoustic bandwidth, and potential for monolithic CMOS integration [1], but translating them into functional imaging systems requires careful co-design of the sensor and downstream beamforming hardware.

FPGAs support parallel receive processing at deterministic latency and far lower power than GPU or CPU-based pipelines [2], and storing all working data in on-chip block RAM eliminates the PCB area and power overhead of external memory. We present a complete sensor-to-image pipeline in which a custom CMUT array, fabricated via a novel Silicon-on-Insulator based MEMS process, is interfaced with an FPGA running a Verilog delay-and-sum beamformer.

Description of the System

The imaging system (Fig. 1a) comprises three subsystems: a CMUT array front-end, a high-voltage pulser and analog receive chain, and an FPGA-based digital beamformer.

The CMUT array is fabricated utilizing custom Silicon-on-Insulator wafers. The process

achieves well-controlled membrane geometry: across a 15-wafer lot, the device layer thickness was $1.93 \pm 0.09 \mu\text{m}$ wafer-to-wafer (interwafer CoV 4.5%), confirming consistent membrane definition across the lot. The assembled sensor module comprising the CMUT die wirebonded to a custom PCB with acoustic matching and isolation layers is shown in Fig. 1b, alongside optical microscope images of individual array elements and cells. Devices target a resonant frequency in the 1–10 MHz diagnostic range and operate in pre-collapsed mode to maximize coupling efficiency. Specific array geometry, element count, and process-layer details are withheld pending patent review.

Transmit pulses are produced by a high-voltage pulser board (EV28J28A and HV7351) with programmable bipolar excitation. Received echoes are conditioned by a low-noise amplifier and anti-aliasing filter before digitization via an AD9671 analog front-end IC. A custom multiplexer board enables element addressing, reducing the simultaneous digital channel count to eight which is a practical power and area trade-off at the prototype stage.

The beamformer is implemented in Verilog HDL on a Nexys Video FPGA (Artix 7- XC7A200T) development board clocked at 100 MHz. The full 64 channel array is scanned in eight sequential groups of 8 channels via the analog multiplexer, with each group digitized and receive-beamformed independently. For each output

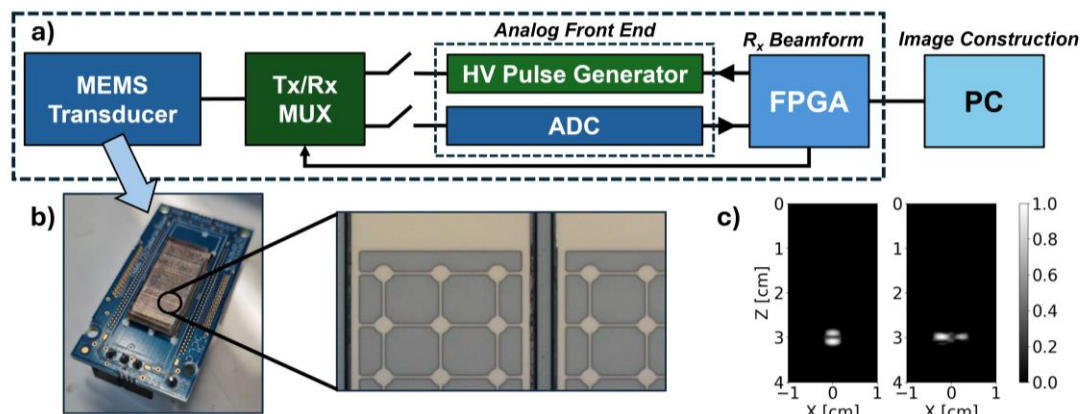


Fig. 1. (a) System block diagram of the signal chain from MEMS transducer to PC reconstruction. (b) Assembled sensor module and optical microscope image of the CMUT array elements. (c) B-mode images of wire targets at the same axial (left) and lateral (right) positions.

pixel, the round-trip delay to each element is computed from the element positions with tissue sound speed of 1540 m/s. The corresponding sample is selected from a per-channel BRAM buffer and all eight contributions are summed. Delay arithmetic uses fixed-point integers to minimize DSP tile usage and log compression is applied via a look-up table. All working data, reside entirely in on-chip BRAM, eliminating external DDR memory. Beamformed frames from all eight groups are streamed over USB to a host PC, where they are coherently combined and processed for envelope detection, dynamic range selection, scan conversion, and display. At 100 MHz, a single frame requires 3.3 M cycles per group, corresponding to 33 ms per focal depth per group. Table 1 summarizes the implementation parameters.

Results

Phantom imaging experiments (Fig. 1c) were conducted using a wire-target and a tissue-mimicking gelatin phantom at two operating frequencies, characterising the resolution–penetration trade-off across the 1–10 MHz bandwidth of the transducer. Point spread function measurements were obtained from 0.1 mm-diameter wire targets, with lateral and axial resolution extracted at the –6 dB level. At 10 MHz, axial resolution remains stable between 1.13 and 1.46 mm over 1–9 cm depth, degrading

to 2.06 mm at 10 cm due to frequency-dependent tissue attenuation. Lateral resolution degrades from 0.51 mm at 1 cm to 1.42 mm at 10 cm (0.10 mm/cm), reflecting the tighter beam at higher frequency. At 1 MHz, axial resolution spans 1.23–2.50 mm across the full range, while lateral resolution degrades more steeply (0.92 mm/cm) to 18.6 mm at 20 cm, consistent with the longer wavelength and fixed single-focus transmit configuration. Tissue-mimicking phantom experiments confirm imaging of hypo- and hyperechoic inclusions at both frequencies, with no significant grating-lobe artefacts observed.

The total system power consumption is 7.9 W. The analog front-end draws 5.2 W during active transmission (3.72 W pulser + 1.44 W ADC and peripherals across 8 channels), and the FPGA digital back-end consumes 2.7 W. These results confirm the feasibility of the sensor–FPGA co-integration approach for portable operation; dynamic transmit focusing is identified as the primary next step toward improved lateral resolution at depth.

Acknowledgements

This work was supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) under grant no. 7230798.

References

- [1] G. Caliano et al., “Design, fabrication and characterization of a capacitive micromachined ultrasonic probe for medical imaging,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 52, no. 12, pp. 2259–2269, 2005; doi: 10.1109/TUFFC.2005.1561630
- [2] E. Boni et al., “ULA-OP 256: A 256-channel open scanner for development and real-time implementation of new ultrasound methods,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 63, no. 10, pp. 1488–1495, 2016; doi: 10.1109/TUFFC.2016.2566458

Tab. 1: FPGA implementation summary.

Parameter	Value / Description
Samples / channel	32,768 × 16-bit
Beamformed image	100×100 × 32 bit ≈ 400 KB
Total memory	≈4.4 MB (BRAM)
Beamforming cycles	~3.3 M (100×100×32,768)
Frame time	~33 ms ≈ 30 fps theoretical
Delay arithmetic	Fixed-point integer
Log compression	LUT-based