



# Real-Time Codebook-based Speech Enhancement Now Possible With *n*VIDIA GPU's

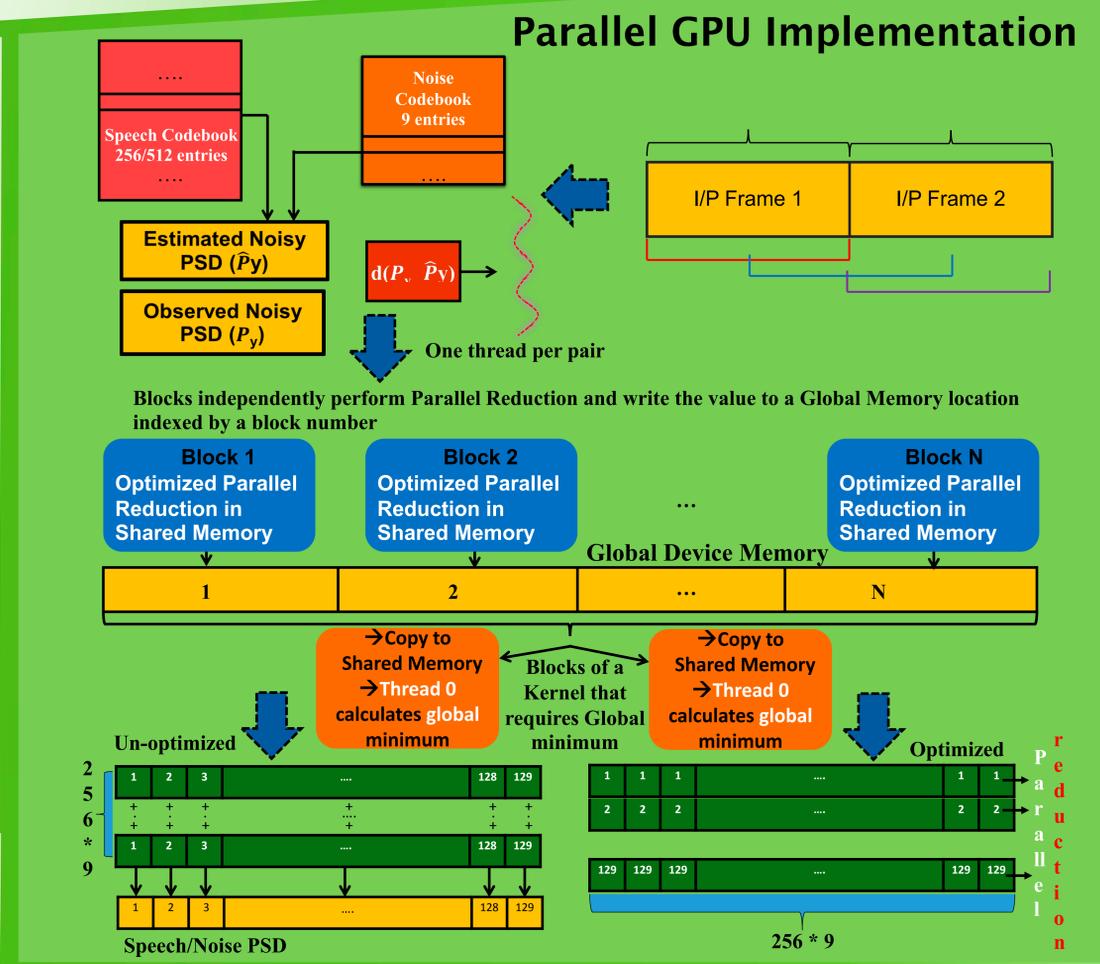
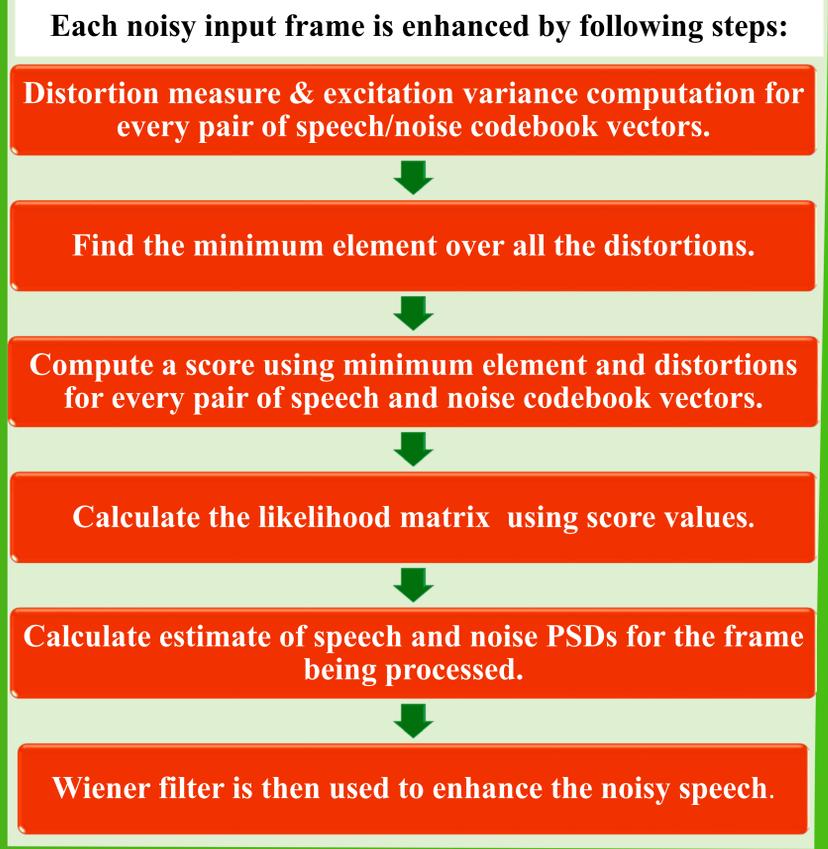
AN. Sai Prasanna\*, Iyer Chandrashekar G\*, D. H. R. Naidu, Pallav K Baruah  
Sri Sathya Sai Institute of Higher Learning, Prasanthi Nilayam, India.



### ABSTRACT

- ❖ The advent of ubiquitous mobile communication has posed a lot of challenges, one of them being suppression of nonstationary background noise.
- ❖ Codebook-based speech enhancement (CBSE) is one of the most effective technique for handling nonstationary noise.
- ❖ CBSE is inapplicable in real-time speech enhancement scenarios due to its compute intensive nature.
- ❖ Speedup of 30 and 54 for speech codebooks with 256 and 512 entries is achieved resulting in a speech transmission delay which are well within the limits of realizing real-time speech enhancement.
- ❖ CBSE can be used in a variety of applications demanding real-time speech processing.

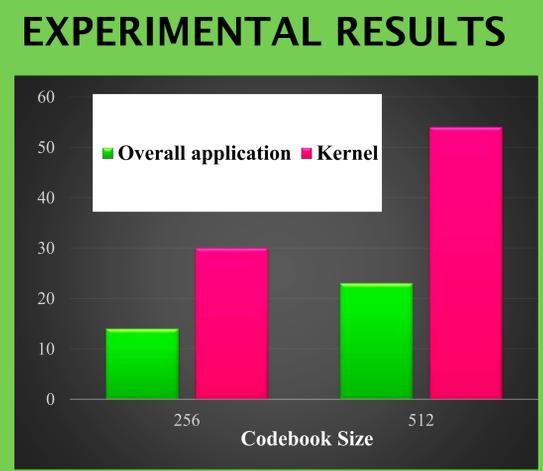
### CBSE ALGORITHM



- ### EXPERIMENTAL SETUP
- NVIDIA TESLA K20c GPU card.
  - Used MEX-Interfaced MATLAB to write the parallel CUDA version code.
  - 10 noisy input utterances with signal-to-noise ratio of 10 dB and length approximately 4.5 to 5.5 sec.
  - Frame length – 32 ms.
  - Serial implementation in MATLAB R2012a on a node with INTEL(R) Xeon(R) CPU E5-2609 @ 2.40 GHz processor and 32 GB RAM.

### Execution Time

	MATLAB	MEX-C	MEX-CUDA
Size 256: Overall Appl.	2132	305	22.861
Size 256: Kernel	2116	294	9.861
Size 512: Overall Appl.	3260	603	26.558
Size 512: Kernel	3239.73	584.3	10.863



### CODE OPTIMIZATIONS

- Memory coalescing.
- Shared memory usage
- Efficient device memory usage by reuse instead of allocation every time across kernel calls

### APPLICATIONS

- ✓ Mobile Phones : A cloud-based framework offloading this high computational overhead to a cloud service employing the parallel implementation of CBSE.
- ✓ Teleconferencing Systems
- ✓ Speech Enhancement as a preprocessing to Speech recognition.
- ✓ Hearing Aids

### CONCLUSIONS & FUTURE WORK

- Proposed a parallel GPU Implementation of Codebook-based speech enhancement technique .
- Each 32ms noisy input frame is enhanced in **10ms** making the implementation usable for real-time applications.
- One possible future work could be to port the structured CBSE technique which tackles the high computational cost of the CBSE onto mobile GPUs like Tegra K1 providing a non-cloud-based solution in mobile communication.

### REFERENCES

1. Sriram Srinivasan, Jonas Samuelsson, and W Bastiaan Kleijn. Codebook-based Bayesian speech enhancement for nonstationary environments. Audio, Speech and Language Processing, IEEE Transactions on, 15(2):441–452, 2007.
2. Philipos C Loizou. Speech enhancement: theory and practice. CRC press, 2013.

### ACKNOWLEDGEMENT

❖ We dedicate this work to our Founder Chancellor Bhagawan Sri Sathya Sai Baba who has been the guiding force behind all our endeavors.