

REFERENCES

- [1] Håkan Ardö, Kalle Åström, and Rikard Berthilsson. 2007. Real-time Viterbi optimization of Hidden Markov Models for multi target tracking. In *IEEE Workshop on Motion and Video Computing (WMVC)*. 2–2.
- [2] Mohamed Benzeghiba, Renato de Mori, Olivier Deroo, Stéphane Dupont, Teodora Erbes, Denis Jouviet, Luciano Fissore, Pietro Laface, Alfred Mertins, Christophe Ris, Richard Rose, Vivek Tyagi, and Christian Wellekens. 2007. Automatic speech recognition and speech variability: A review. *Speech Communication* 49, 10-11 (2007), 763–786.
- [3] Paolo Boldi, Marco Rosa, and Sebastiano Vigna. 2011. HyperANF: approximating the neighbourhood function of very large graphs on a budget. In *Proceedings of the 20th International Conference on World Wide Web (WWW)*. 625–634.
- [4] Hugo Braun, Justin Luitjens, Ryan Leary, Tim Kaldewey, and Daniel Povey. 2020. GPU-accelerated Viterbi exact lattice decoder for batched online and offline speech recognition. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 7874–7878.
- [5] Diamantino Caseiro, Hugo Meinedo, António Serralheiro, Isabel Trancoso, and João Neto. 2002. Spoken book alignment using WFSTs. In *Proceedings of the 2nd International Conference on Human Language Technology Research (HLT)*. 194–196.
- [6] George E. Dahl, Dong Yu, Li Deng, and Alex Acero. 2012. Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition. *IEEE Transactions on Audio, Speech, and Language Processing* 20, 1 (2012), 30–42.
- [7] Jon Feldman, Ibrahim Abou-Faycal, and Matteo Frigo. 2002. A fast maximum-likelihood decoder for convolutional codes. In *Proceedings of the 56th IEEE Vehicular Technology Conference (VTC)*. 371–375.
- [8] Pedro F. Felzenszwalb, Daniel P. Huttenlocher, and Jon M. Kleinberg. 2003. Fast algorithms for large-state-space HMMs with applications to web usage analysis. In *Advances in Neural Information Processing Systems 16 (NIPS)*. 409–416.
- [9] Philippe Flajolet, Éric Fusy, Olivier Gandouet, and Frédéric Meunier. 2007. HyperLogLog: the analysis of a near-optimal cardinality estimation algorithm. In *Conference on Analysis of Algorithms (AoFA)*. 137–156.
- [10] Mark J. F. Gales and Steve J. Young. 2007. The application of Hidden Markov Models in speech recognition. *Foundations and Trends in Signal Processing* 1, 3 (2007), 195–304.
- [11] John S. Garofolo, Lori F. Lamel, William M. Fisher, Jonathan G. Fiscus, David S. Pallett, and Nancy L. Dahlgren. 1993. *DARPA TIMIT acoustic-phonetic continuous speech corpus CD-ROM. NIST speech disc 1-1.1*. National Institute of Standards and Technology.
- [12] Alexandru-Lucian Georgescu, Alessandro Pappalardo, Horia Cucu, and Michaela Blott. 2021. Performance vs. hardware requirements in state-of-the-art automatic speech recognition. *EURASIP J. Audio Speech Music. Process.* 2021, 1 (2021), 28.
- [13] Sudipto Guha. 2005. Space efficiency in synopsis construction algorithms. In *Proceedings of the 31st International Conference on Very Large Data Bases (VLDB)*. 409–420.
- [14] Joachim Hagenauer and Peter Hoeher. 1989. A Viterbi algorithm with soft-decision outputs and its applications. In *1989 IEEE Global Telecommunications Conference and Exhibition 'Communications Technology for the 1990s and Beyond'*. 1680–1686.
- [15] Panagiotis Karras and Nikos Mamoulis. 2008. Hierarchical synopses with optimal error guarantees. *ACM Transaction on Database Systems* 33, 3 (2008), 18:1–18:53.
- [16] Pedro J. Moreno and Christopher Alberti. 2009. A factor automaton approach for the forced alignment of long speech recordings. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. 4869–4872.
- [17] Douglas B. Paul and Janet M. Baker. 1992. The design for the Wall Street Journal-based CSR corpus. In *The 2nd International Conference on Spoken Language Processing (ICSLP)*.
- [18] Jennifer Pohle, Roland Langrock, Floris M. van Beest, and Niels Martin Schmidt. 2017. Selecting the number of states in hidden Markov models: pragmatic solutions illustrated using animal movement. *Journal of Agricultural, Biological, and Environmental Statistics* 22, 3 (2017), 270–293.
- [19] Daniel Povey, Arnab Ghoshal, Gilles Boulianne, Lukas Burget, Ondrej Glembek, Nagendra Goel, Mirko Hannemann, Petr Motlicek, Yanmin Qian, Petr Schwarz, Jan Silovsky, Georg Stemmer, and Karel Vesely. 2011. The Kaldi speech recognition toolkit. In *IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*.
- [20] Sajid M. Siddiqi and Andrew W. Moore. 2005. Fast inference and learning in large-state-space HMMs. In *Proceedings of the 22nd International Conference on Machine Learning (ICML)*. 800–807.
- [21] Christopher Tarnas and Richard Hughey. 1998. Reduced space hidden Markov model training. *Bioinformatics* 14, 5 (1998), 401–406.
- [22] Keith Vertanen. 2008. Combining open vocabulary recognition and word confusion networks. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. 4325–4328.
- [23] Andrew J. Viterbi. 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory* 13, 2 (1967), 260–269.
- [24] Andrew J. Viterbi. 2006. A personal history of the Viterbi algorithm. *IEEE Signal Processing Magazine* 23, 4 (2006), 120–142.
- [25] Qi Wang, Lei Wei, and Rodney A. Kennedy. 2002. Iterative Viterbi decoding, trellis shaping, and multilevel structure for high-rate parity-concatenated TCM. *IEEE Transactions on Communications* 50, 1 (2002), 48–55.
- [26] Reza Yazdani, Albert Segura, José-Maria Arnau, and Antonio González. 2017. Low-power automatic speech recognition through a mobile GPU and a Viterbi accelerator. *IEEE Micro* 37, 1 (2017), 22–29.
- [27] Steve J. Young, Gunnar Evermann, Mark J. F. Gales, Thomas Hain, Dan Kershaw, Xunying Liu, Gareth Moore, Julian Odell, Dave Ollason, Dan Povey, et al. 2002. *The HTK Book*. University of Cambridge, Department of Engineering.
- [28] Steve J. Young, N.H. Russell, and J.H.S. Thornton. 1989. *Token passing: a simple conceptual model for connected speech recognition systems*. University of Cambridge, Department of Engineering.