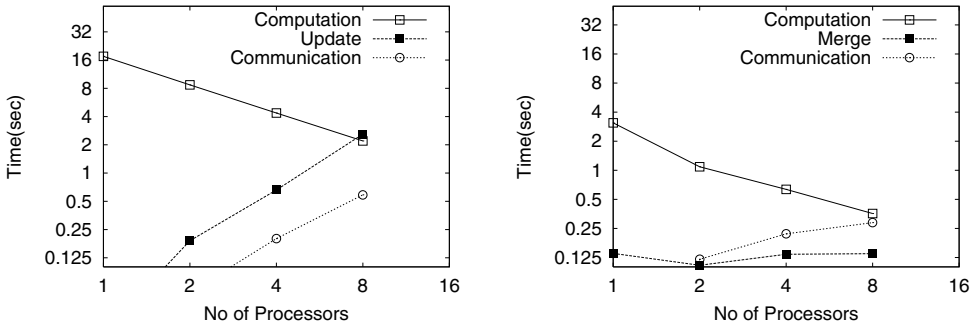
Parallelization of VQ Codebook generation by two algorithms: parallel LBG and aggressive PNN

Akiyoshi Wakatani(wakatani@konan-u.ac.jp)
Faculty of Science and Engineering, Konan University, Japan

We evaluate two parallel algorithms for the codebook generation of the VQ compression: parallel LBG and aggressive PNN. Parallel LBG is based on the LBG algorithm with the K-mean method, whose cost of both algorithms mainly consist of a) the computation part, b) the communication part and c) the update part. Aggressive PNN is a parallelized version of the PNN (pairwise nearest neighbor) algorithm, whose cost of both algorithms mainly consist of a) the computation part b) the communication part and c) the merge part.

We measured the speedups and elapsed times of both algorithms on a PC cluster system which consists of 8 CPUs (Celeron 1GHz) and LAN (100Mbps) under MPICH1.2.4 and Linux 2.4. The speedups of both algorithms are almost same except for an 8 CPU case. The speedups of the aggressive PNN with 3000 and 8000 training vectors on an 8 CPU cluster are 4.16 and 5.94, respectively. However, the speedups of the parallel LBG are degraded on an 8 CPU cluster, because the costs of the communication and update parts exceed the parallelization merit of the computation part, namely, the former are proportional to $P * \log P$ and P and the latter is inversely proportional to P. The costs of the update and communication part are nearly equal to or more than that of the computation part for all the cases on an 8 CPU cluster. So the scalability of the parallel LBG is not so good.

When the quality of images compressed by both algorithms are same, the number of training vectors required by the aggressive PNN is much less than that by the parallel LBG, and the aggressive PNN is superior in terms of the elapsed time. For example, the quality of images compressed by the parallel LBG with 8000 training vectors and the aggressive PNN with 3000 training vectors are almost same (32.46dB and 32.10dB), but the total elapsed times on an 8 CPU cluster are 5.4 sec and 0.78 sec for the parallel LBG and the aggressive PNN, respectively. It should be noted that the elapsed time of the parallel LBG can be improved by using the all reduce collective communication instead of the combination of the broadcast communication and the update part to a certain degree, but it will not outperform the aggressive PNN.



(a) Parallel LBG (T=8000, 32.46dB) (b) Aggressive PNN (T=3000, 32.10dB)

Figure 1: Elapsed time (lenna)

Proceedings of the 2005 Data Compression Conference (DCC'05)
1068-0314/05 $ 20.00 **IEEE**