A Fully-Pipelined FPGA Design for Tree-reweighted Message Passing Algorithm

Wenlai Zhao*†, Haohuan Fu*†, Guangwen Yang*†
*Department of Computer Science and Technology,
Tsinghua University, Beijing, 100084, P.R. China
†Ministry of Education Key Laboratory for Earth System Modeling, and Center for Earth System Science,
Tsinghua University, Beijing, 100084, P.R. China
Email: zhaowl-11@mails.tsinghua.edu.cn, {haohuan,ygw}@tsinghua.edu.cn

***Abstract*---A Markov random field (MRF) is a set of random variables demonstrating a Markov property in the form of an undirected graph. Maximum a posteriori probability (MAP) inference is a class of methods that seek solutions of problems modeled by MRF. MRF has been a very popular and powerful tool in computer vision problems such as stereo matching and image segmentation [1]. Finding the optimal solution of the MRF MAP problem is an NP-hard problem. Inference algorithms often involve a heavy computation load. Therefore, most related works have focused on improving the performance and efficiency of algorithms. Hardware-based acceleration is one of the most practical solutions.**

Sequential tree-reweighted message passing (TRW-S) [2] is one of the inference algorithms based on message passing scheme. TRW-S shows good quality in finding optimal solutions but its performance on CPU is poor due to the sequential scheme and high cost of computing resources. Some efforts have been made to implement more efficient TRW-S using FPGAs and have achieved some improvement in performance [3]. But the sequential message passing pattern of TRW-S still remains a significant challenge for parallel and pipelined hardware designs and constrains the performance to a certain level.

In our work, we propose a modified algorithm based on TRW-S, called parallel tree-reweighted message passing (TRW-P), which eliminates the sequential data dependencies and can be fully pipelined in hardware. According to the description of TRW-S in [2], TRW-P can be described in brief as follow:

- **Step 0. Initialize all messages $M^{(0)}$ to zero.**
- **Step 1.**
  For iteration i from 1 to $MAXITER$
  - For each node $s \in V$, compute and normalize the information θ of s as (1),

$$\hat{\theta}_s^{(i)} = \overline{\theta}_s + \Sigma_{(u,s)\in E} M_{us}^{(i-1)} \quad (1)$$

  - For each edge $(s,t) \in E$, update and normalize the message M on (s,t) as (2),

$$M_{st;k}^{(i)} := \min_j \{(\gamma_{st}\hat{\theta}_{s;j}^{(i)} - M_{ts;j}^{(i-1)}) + \overline{\theta}_{st;jk}\} \quad (2)$$

- **Step 2. Check the stop criterion.**

Based on TRW-P, we propose a fully pipelined design and map it into an FPGA. Figure 1 shows the design and dataflow of our implementation. Then we build a hybrid CPU/FPGA system to test the performance of the design using stereo matching application.

Experimental results show that TRW-P can find an approximate solution to the global optimum, which is nearly the same with TRW-S and better than other algorithms. In terms of performance, our design provides about 100 times speedup over a single-core software implementation of TRW-S and compared with existing multi-FPGA design in [3], we can achieve 2 times speedup on the overall performance and 8 times speedup for one FPGA card. Moreover, our design can work at a video rate in many cases (167 frame/sec for a standard stereo matching test case), which makes it a promising work for many real-time applications.

Keywords*---Markov Random Field, Energy minimization, Tree-Reweighted Message Passing, FPGA

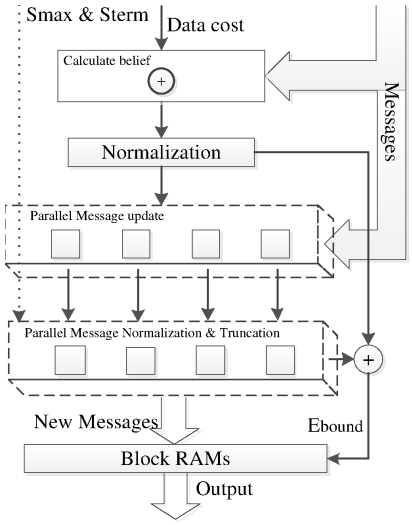

Fig. 1. Architecture and dataflow of our TRW-P kernel

References

[1] R. Szeliski, O. Veksler and M. Tappen, ``A Comparative Study of Energy Minimization Methods for Markov Random Fields with Smoothness Based Priors", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 30, no. 6, pp. 1068-1080, Jun. 2008.

[2] V. Kolmogorov, ``Convergent Tree-Reweighted Message Passing for Energy Minimization", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 10, pp. 1568-1583 Oct. 2006.

[3] J. Choi and R. Rutenbar, ``Hardware Implementation of MRF MAP Inference on an FPGA Platform", *Field Programmable Logic and Applications (FPL), 2012 22nd International Conference on.*, pp. 209-216, 2012.

978-1-4799-5111-6/14 $31.00 © 2014 IEEE
DOI 10.1109/FCCM.2014.59