Hardware Design of Dynamic Time Warping Algorithm based on FPGA in Verilog

Kriti Suneja, Dr. Malti Bansal
Department of Electronics and Communication Engineering,
Delhi Technological University
Delhi-110042, India

***Abstract* — The Dynamic Time Warping (DTW) algorithm is a software technique to quantify the diversity between two time varying sequences. In spite of its computational complexity, it is often used in speech recognition systems. In this work, hardware implementation of DTW modeled in VERILOG Hardware Description Language (HDL) has been done, with an aim to use it as an independent unit or as a co-processor in large embedded systems. The target device used for synthesis is xc3s400-4-pq208 in Xilinx. Simulations were performed in ModelSim. Hardware utilization analysis is provided to show its efficacy in warping a sequence in non linear fashion.**

Index Terms —dynamic time warping, verilog, time sequence data, Field Programmable Gate Array (FPGA), Look Up Tables (LUTs), Application Specific Integrated Circuits (ASICs)

I. INTRODUCTION

Dynamic Time Warping (DTW) algorithm finds the edit distance, which is the minimum number of editing operations required to convert one sequence into another, between two sequences by analysing three operations: substitution, insertion, and deletion [1]. Besides, it is a search algorithm which finds an optimal warping path dynamically for the minimal cost alignment of input on template.

Measuring the edit distance is best reflected as a 2-dimensional grid with the two sequences on the top and left side respectively. The central idea is to formulate optimal path to any intermediate point X in the grid in terms of optimal paths of all its immediate antecedents.
Let $\gamma(n,m)$ = Minimum path cost from origin to any point (n,m) in the grid.

$$\gamma(n,m) = d(n,m) + \min \{\gamma(n-1,m), \gamma(n,m-1), \gamma(n-1,m-1)\} \quad (1)$$

where d(n,m) is the square of the Euclidean distance between the two elements of the target sequences at X.
Hence, starting from the origin and computing minimum path cost for every grid point entry, proceeding from top left to bottom right corner provides the minimum edit distance as the last entry.

To determine the alignment, a back-pointer is maintained from X to its antecedent, which provides the minimum cost to it. At the end, back tracing gives the best alignment.

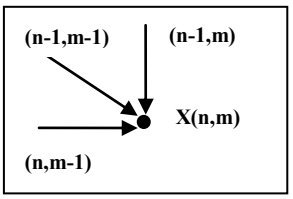

Fig.1 Point X and its antecedents to find minimum cost

DTW is very efficient in software computing. It has been used in isolated word speech recognition system for human computer interaction [2], hand gesture recognition [3], isolated musical pattern recognition [4], human action recognition [5]. Also, it provides excellent results in terms of accuracy. B.C. Bhokse and B.S. Thakare used DTW for *Devnagari* handwriting recognition system [6]. However, its application in real time processing is limited to literature because of space and time restrictions in its hardware implementation. In order to use specialized hardware for similarity search using DTW, we need to investigate the limitations of the target device and DTW computation. Ding et al. [7] stated the outperformance of DTW among all elastic distance measures. Sart et al. [8] investigated acceleration of DTW with GPUs and FPGAs using an open source C to VHDL compiler system, ROCCC. Tai, Li and Elmiligi [9] used VHDL for hardware space utilization and speed analysis of DTW which has its limitations as the data size grows.

In this paper, we propose an elementary DTW processing unit designed in VERILOG [10], in order to explicate its use in speech processing, such as, plagiarism detection in music by finding the distance between the song pairs after features extraction. Features' extraction can be done in MATLAB and then the scalar or vector data can be given as input to DTW hardware for comparison.

Though, implemented on SPARTAN FPGA, the goal of using VERILOG is to make DTW ASIC compatible and to exploit the use of VERILOG Programming Language Interface (PLI) and VERILOG Analog and

ISSN: 2278 – 909X

Mixed-Signal Extension (AMS) for large samples of speech.

The organization of this paper is as follows. In Section II, we describe the pseudo code and MATLAB implementation of DTW. We then describe how this can be incorporated in VERILOG. Next, we present results of experiments. Finally, we present our conclusions and suggestions for future work.

II. AN n-ELEMENT DTW PROCESSOR

Let us consider two sequences A and B of length n and m respectively. Pseudo code of DTW to find the edit distance between them is as follows:

```
for i := 1 to n
        for j := 1 to m
                cost :=  d(A[i] , B[j])
DTW[i,j] :=cost + minimum (DTW[i-1 , j],
                           DTW[i , j-1],
                           DTW[i-1,j-1])
        return DTW(n,m)
w=[]    //Optimal Path
k= n;
l=m;
                while (k+l ~=2)

[value , number]=
min([DTW(k-1,l),DTW(k,l-1),DTW(k-1,l-1)])
switch  number
case 1      k=k-1
case 2      l=l-1
case3       k=k-1;l=l-1
w = cat (1, w , [k,l])
```

where d(A[i],B[j]) is the square of Euclidean Distance and DTW(i,j) is the minimum edit distance between ith and jth elements of sequences A and B respectively.

For example, let A and B be :

A= [1 2 3 4 7 8] ; B= [1 1 2 2 3 3 3 4]

MATLAB simulations will give the results shown in Fig. 2 .

1	0	0	1	2	6	10	14	23
2	1	1	0	0	1	2	3	7
3	5	5	1	1	0	0	0	1
4	14	14	5	5	1	1	1	0
7	50	50	30	30	17	17	17	09
8	99	99	66	66	42	42	42	25

Fig.2 Matlab simulation results for the considered example

A. DTW in VERILOG with Xilinx ISE

In this paper, we propose the hardware design of DTW using VERILOG Hardware Description Language (HDL).

The system is implemented using Xilinx Integrated Software Environment (ISE) which is used to synthesize and process HDL algorithms on FPGA. Target device Spartan 3 with package pq208 has the following properties: 3584 Slices, 7168 Look Up Tables (LUTs) , 141 Input/ Output Blocks (IOBs). The input is taken as an array of binary numbers. Here, we have done it for example referred in Section II. If used for speech recognition, features of the samples, for example, Mel Frequency Cepstral Coefficients (MFCCs), tonalcentroids, zerocrossovers, etc., can be used as input.

First, the square of Euclidean distance is calculated for the first element of the matrix. Then, first row and first column elements, are calculated which require the previous element value and no comparisons. Finally, the rest of the entries in matrix are calculated with recursion. The minimum edit distance is found using ModelSim.

B. Experimental Results

The minimum edit distance in ModelSim for the example stated in section II is 25, as shown in Fig.3.

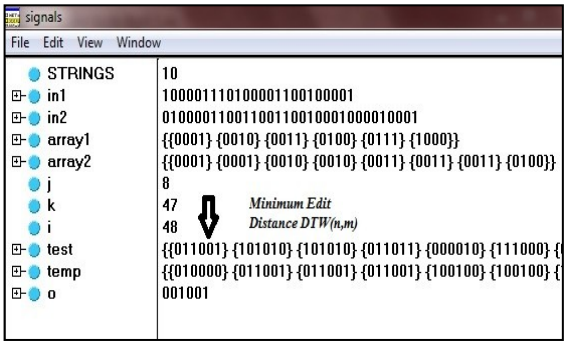

Fig.3 ModelSim simulation Results

MATLAB simulations give the same result along with the optimization path as shown in Fig.4.

The delay analysis shows that the delay increases linearly with the increase in number of elements in the sequences to be compared as shown in Fig.5.

When simulated in MATLAB, the execution time for 6 × 8 unit is 0.000138s and the delay obtained for the same unit in Xilinx is 95.49ns.

In the synthesis of DTW with input sequences of length 6 and 8 , 1024 out of 3584 Slices , 1982 out of 7168 4 input LUTs and 57 out of 141 IOBs are used which are only 28% , 27% and 40% respectively of the available resources.

However, number of slices and LUTs increase exponentially with the increase in input size; and IOBs increase linearly, as shown in Fig.6, Fig.7 and Fig.8 respectively.

ISSN: 2278 – 909X

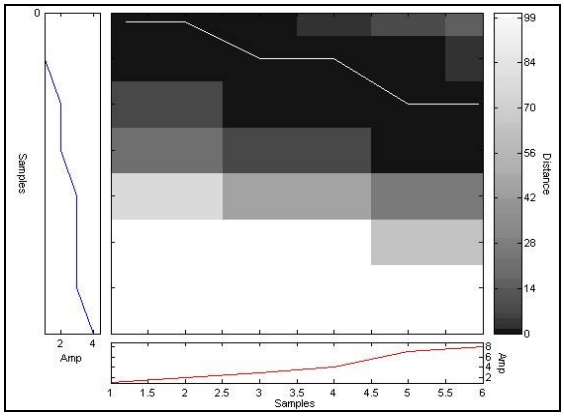

Fig.4 Accumulated Distance Matrix and Optimal path

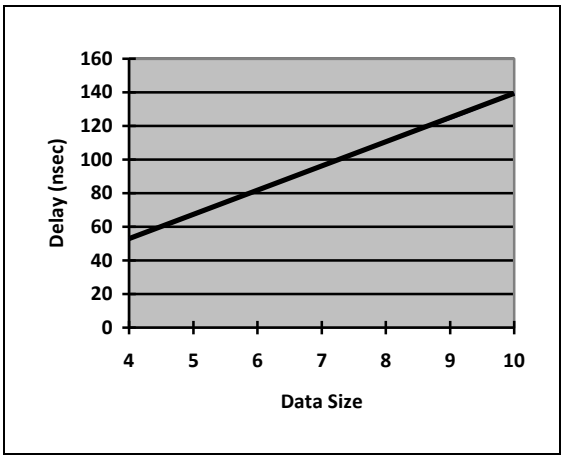

Fig.5 Delay Variation with number of elements

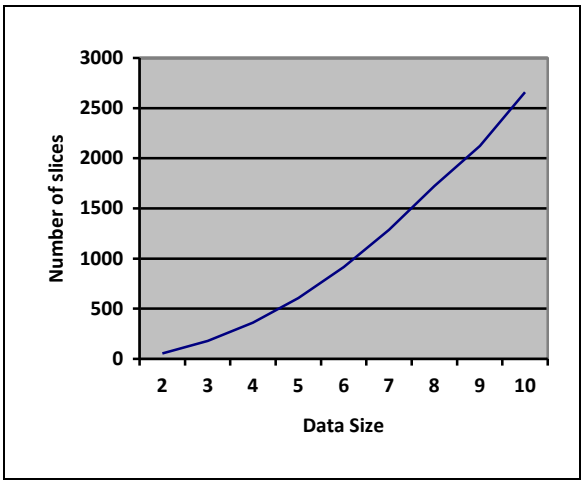

Fig.6 Number of Slices variation with data size

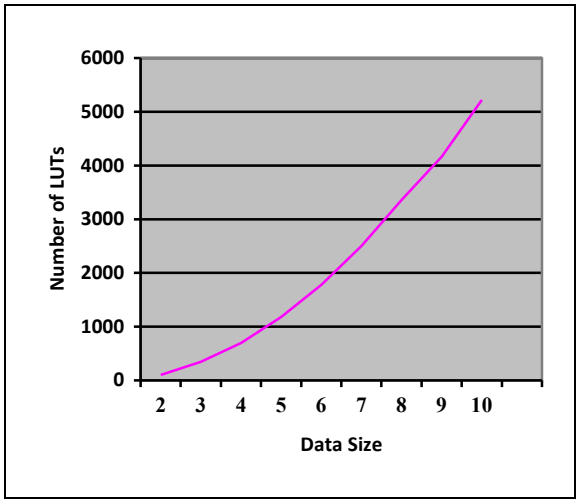

Fig. 7 Number of LUTs variation with Data size

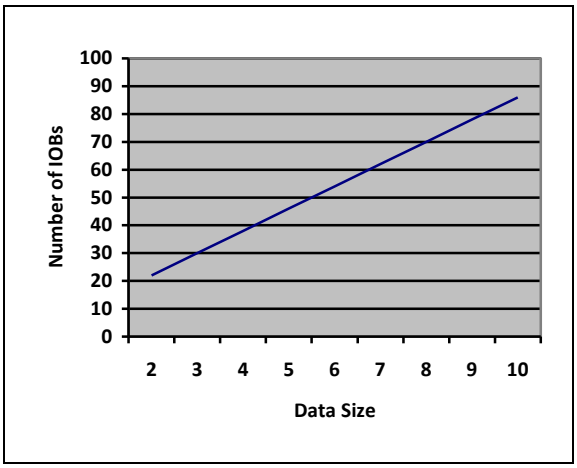

Fig.8 Number of IOBs variation with Data Size

III. CONCLUSION AND FUTURE WORK

DTW is an efficient software algorithm for comparing time sequences and is extensively used in speech and pattern recognition. However, in real time systems, speed and hardware utilization are critical parameters. Implementation of a basic DTW unit in this paper demonstrates that it can be used as an integral part of large processing units in FPGAs and ASICs. The device used in this paper has limited number of slices and LUTs available. However, if Virtex 7 is used, large data size can be used because it provides 408,000 programmable slices registers and 203,000 LUTs [9].

With the successful hardware implementation of DTW, it can be used as a built in coprocessor for speech processing embedded systems. Since, the algorithm requires a large number of multipliers proportional to the square of data size, there is a need of an efficient multiplication algorithm which can reduce the hardware utilization because the number of built in multipliers are limited.

ISSN: 2278 – 909X

REFERENCES

[1] M. Ravishankar, "Isolated word recognition using dynamic time warping," Design and implementation of ASR systems. http://asr.cs.cmu.edu/spring2010/class3/class3.25jan.dtw.pdf

[2] R. Makhijani and R. Gupta, "Isolated word speech recognition system using dynamic time warping," International Journal of Engineering Sciences & Emerging Technologies (IJESET), vol. 6, no. 3, pp.352-367, December 2013.

[3] K. Barczewska and A. Drozd, "Comparison of methods for hand gesture recognition based on dynamic time warping algorithm," IEEE Federation Conference on Computer Science and Information Systems (FedCSIS), pp. 207-210, September 2013.

[4] A. Pikrakis, S. Theodoridis and D. Kamarotos, "Recognition of isolated musical patterns using context dependent dynamic time warping," IEEE Transactions on Speech and Audio Processing, pp. 175-183, May 2003.

[5] S. Sempena, N.U. Maulidevi and P.R. Aryan, "Human action recognition using dynamic time warping," IEEE International Conference on Electrical Engineering and Informatics (ICEEI), pp. 1-5, July 2011.

[6] B.C. Bhokse and B.S. Thakare, "Devnagari handwriting recognition system using dynamic time warping algorithm," International Journal of Computer Applications(IJCA), vol. 52, no. 9, pp.7-13, August 2012.

[7] H. Ding, G. Trajcevski, P. Scheuermann, X. Wang, and E. Keogh, "Querying and mining of time series data: Experimental comparison of representations and distance measures," PVLDB, vol. 1, no. 2, pp.1542-1552, August 2008.

[8] D. Sart, A. Mueen, W. Najjar, E. Keogh and V. Niennattrakul, "Accelerating dynamic time warping subsequence search with GPUs and FPGAs," IEEE 10th International Conference on Data Mining, pp. 1001-1006, 2010.

[9] J.S. Tai, K.F. Li and H. Elmiligi, "Dynamic time warping algorithm: A hardware realization in VHDL," IEEE International Conference on IT Convergence and Security(ICITCS), pp. 1-4, December 2013.

[10] S. Palnitkar, "Verilog HDL- A guide to digital design and synthesis," SunSoft Press, 1996.

[11] B.K. Yi, H.V. Jagadish, C. Faloutsos, "Efficient retrieval of similar time sequences under time warping," IEEE 14th International Conference on Data Engineering, pp. 201-208, February 1998.

[12] Z. Qian, K. Takaya, "Dense stereo disparity maps by dynamic time warp with sparse features and FPGA acceleration" 24th Canadian Conference on Electrical and Computer Engineering, pp. 874-877, May 2011.

ISSN: 2278 – 909X