Analysis of DNA Signal Representation Applying Dynamic Time Warping (DTW) and Derivative Dynamic Time Warping (DDTW)

Ab. Kader Magdy, Ahmed Khadragi, Magdy Saeb, A. Baith Mohamed

Arab Academy of Science, Technology and Maritime Transport
College of Engineering and Technology, Computer Engineering Department,
Alexandria, EGYPT

Abstract: The Deoxyribonucleic Acid (DNA) is a double stranded helix of nucleotides consists of: Adenine (A), Cytosine (C), Guanine (G) and Thymine (T). In this work, the DNA code is converted into an equivalent digital signal quaternary code. The Dynamic Time Warping (DTW) and the Derivative Dynamic Time Warping (DDTW) algorithms are applied on two different species' DNA to calculate the distance between them. We apply these two techniques to compare human DNA to several species of different biological families by calculating the warping path values.

Keywords: DNA, Nucleotides, Digital Signal, Dynamic Time Warping, Derivative Dynamic Time Warping, DTW, DDTW.

1. Introduction

All living organisms are characterized by the same type of genetic blueprint that acts as the deciding factor of organism specification. This is the Deoxyribonucleic Acid (DNA) which is a double-stranded helix of nucleotides that carries the genetic information of a cell. DNA is a combination of 4 nucleotides: Adenine (A), Cytosine (C), Guanine (G) and Thymine (T) [1].

Various properties and characteristics are governed to exist based on the millions of combinations of these 4 nucleotides. In this work, we compare between multi species DNA by transferring the stored data from its biological form to digital form. The proposed approach takes advantage of the techniques of signal processing [2]. The Dynamic Time Warping and the Derivative Dynamic Time Warping (DDTW) algorithms are applied on the resulting DNA signal representations to extract information from this code.

We apply the two techniques to compare human DNA to several species of different families by calculating warping path values. The organisms compared are Mammals, Reptiles, Fish and Amphibians.

The human is considered a member of the Ape family. There are 193 living species of apes, 192 of them are covered with hair. The exception is a naked ape self-named Homo-sapien [3]. The recent publication of the complete chimp genome [4], marked by a celebratory issue of the journal "Nature" recounts that humans and chimps share 96 percent of the same genetic material. The number of genetic differences between humans and chimps is ten times smaller than that among mice and rats [5].

In the following sections, we discuss how to compare between different organisms using the DTW and DDTW algorithms. In section 2, we present the binary, quaternary representation of DNA sequences and the length of the DNA code. In section 3 and 4, the Dynamic Time Warping and the Derivative Dynamic Time Warping algorithms are applied on the two resulting signals to enable us to extract information from DNA codes. In section 5, we compare between Human DNA and different species of different families. In section 6, various organisms with the nearest distance from human DNA are illustrated.

2. DNA SIGNAL REPRESENTATION

In this section, the DNA code is converted into a signal form; it is a simple straight forward procedure, and applying the DTW algorithm on the two different signals to calculate warping path values between these two signals and the distance between them.

2.1. DNA Binary and Quaternary Representations

Mapping the DNA sequences to binary representation is a simple and a straightforward procedure. For most tasks, a flat encoding of 2 bits per nucleotide, assigned in an alphabetical order would be a sufficient starting point [6].

$$A = (00)_2 \text{ or } A = 0Q$$
$$C = (01)_2 \text{ or } C = 1Q$$
$$G = (10)_2 \text{ or } G = 2Q$$
$$T = (11)_2 \text{ or } T = 3Q$$

For example the DNA sequence:

ACTGGTTTAAACTC

Will be represented in binary format as:

$(00,01,11,10,10,11,11,11,00,00,00,01,11,01)_2$

It will be represented in quaternary format as:

$(0,1,3,2,2,3,3,3,0,0,0,1,3,1)_q$

2.2 DNA Genomic Length

The length of the DNA sequence may reach millions of bases. The DNA sequence of most known organisms can be downloaded from the gene bank [7]. When the length of the DNA sequence increases, the resulting resolution increases accordingly and vice versa. The data encoding representations are used to distinguish between different organisms.

The following examples compare the distance between Human and other species such as Cheetah, Shaping Frog and Eurasian Wolf in a small and large DNA sequence. Only the first 100 and 5000 DNA sequences of the species are chosen from the gene bank [7] to perform such a procedure.

2.2.1 Choosing 100 DNA sequences for each species

Applying DTW to calculate the distance between two various signals, the results will be as follows:

Human versus Cheetah = 24.4652777777777778567269706400111
Human versus Frog = 24.03496503496503322594435303472
Human versus Wolf = 24.2551724137931046243465971201666

2.2.2 Choose only 5000 DNA sequences for each species

Applying DTW to calculate the distance between two signals, the results will be as follows:

Human versus Cheetah = 1116.6109426681366585398791357875
Human versus Frog = 1090.3453296703296473424416035414
Human versus Wolf = 1103.6768081869727211596909910044

Examining the above results, it can be shown that when the length of the DNA signals decreases the distances between Human and other species signals are too close and the resultant resolution is quite clear. However, when the length of the DNA signal increases, the distance between Human and others species are more apart and the resultant resolution is clear.

The DNA of any organism contains millions of DNA sequences. In our experiments we use only the first 5000 base pairs of this DNA sequence to increase the resultant resolution. This number is employed to reduce the execution time [1].

3. Dynamic Time Warping

DTW algorithm has earned its popularity because of its extreme efficiency as a "time-series similarity measure". It minimizes the effects of shifting and distortion in time by allowing "elastic" transformation of time series in order to detect similar shapes with different phases [8]. Dynamic Time Warping algorithm (DTW) calculates an optimal warping path between two time series. The algorithm calculates both warping path values between the two series and the distance between them [9]. In the next few lines, we are summarizing the DTW procedure.

Suppose we have two signals:

$$A = (a_1, a_2, \ldots a_n) \text{ and } B = (b_1, b_2, \ldots b_m)$$

Where 'n' and 'm' represent signals A and B lengths respectively.

The lengths of both signals need not be the same. DTW calculates the difference between the elements in the two signals; the most popular method for distance calculation between values of signals is the Euclidian Distance measure. It results in a matrix of distances having n rows and m columns of the general term given by:

$$d_{ij} = |a_i - b_j|, \quad i = \overline{1,n}, j = \overline{1,m} \qquad (1)$$

Where i and j are the indices of the matrix rows and columns respectively. The minimal distance matrix between sequences is determined using a dynamic programming algorithm and the following optimization criterion:

$$a_{ij} = d_{ij} + \min(a_{i-1,j-1}, a_{i-1,j}, a_{i,j-1}) \qquad (2)$$

Where a_{ij} is the minimal distance between the subsequences $(a1, a2, ..., ai)$ and $(b1, b2, ..., bj)$ [9].

3.1. Finding an Optimal Path

A warping path is a path through the minimal distance matrix from a_{11} element to a_{nm} element consisting of those a_{ij} elements which have formed the a_{nm} distance [9]. It is not necessary to compute all possible paths P and the corresponding distances to find the optimal path. Out of the huge number of theoretically possible paths, only a fraction is reasonable for our purposes [10]. The alignment path built by DTW must satisfy the following criteria:

1. Boundary Condition:

$$p_1 = (1; 1) \text{ and } p_K = (N; M)$$

The starting and ending points of the warping path must be the first and the last points of the aligned sequences.

2. Monotonicity Condition:

$$n_1 \le n_2 \le \dots \le n_k \text{ and } m_1 \le m_2 \le \dots \le m_k$$

This condition preserves the time-ordering of the points.

3. Continuity Conditions

The path advances gradually, step by step where the indices i and j increase by maximum 1 unit on a step. The global warp cost (GC) of the two sequences is defined as shown below:

$$GC = \frac{1}{p} \sum_{i=1}^{p} w_i \qquad (3)$$

Where the terms w_i, refer to those elements that belong to the warping path with p being their number.

4. Derivative Dynamic Time Warping

Dynamic Time Warping (DTW) algorithm is useful to align two signals that are similar except for local accelerations and decelerations in the time axis. The algorithm has a problem when the two sequences are differing in the Y-axis such as different scaling (amplitude scaling) [11]. Another Algorithm is proposed to give more accurate results which is the Derivative Dynamic Time Warping algorithm it is a modified DTW by first Deriving the sequences of the both DNA input signals [11]. The algorithm details are summarized as follows:

Algorithm Details

The distance measured between two signals in Derivative Dynamic Time Warping is not Euclidean but rather the square of the difference of the estimated derivatives of the two signals, it is defined as shown below:

$$D_x[q] = \frac{(q_i - q_{i-1}) + ((q_{i+1} - q_{i-1})/2)}{2} \qquad (4)$$

Where q is the first DNA input signal, as the same on the second signal. DDTW time complexity is O(MN), which is the same as DTW [11].

5. METHODOLOGY

In this section, the DNA code of the human and other different organisms DNA codes are compared by applying the Dynamic Time Warping and the Derivative Dynamic Time Warping algorithms to calculate both warping path values and the distance between them.

5.1. Comparing Human DNA with the Mammalian Family

Mammals are warm-blooded vertebrates which, with the exception of a few notable species, nurse their young with milk produced by the female's mammary glands. They give birth to live young, and have bodies insulated by hair [1],

[12]. Table 1 compares between the distance results of Human and some of the different species of the Mammals family after applying the Dynamic Time Warping Algorithm (DTW) on the DNA signals.

Table 1: The Distance between Different Species of Mammals and Humans

#	Family	Species	Distance
1	Carnivorous	Spectacled Bear	1084.7442212438083970482693985105
2		Sloth Bear	1094.6056494045969884609803557396
3		Malayan Sun Bear	1088.0787737145999471977120265365
4		Asiatic Lion	1111.4275681630404051247751340271
5		Snow Leopard	1108.2985383342527256900211796165
6		Cheetah	1116.6109426681366585398791357875
7	Hoofed	Bactrian Camel	1081.7890085611709309887373819947
8		American Bison	1080.1953769950466721638804301620
9		Eurasian Elk	1103.7682808380741334985941648483
10		Horse	1089.4071922544951576128369197249
11	Apes	Chimpanzee	1045.1434301521437646442791447043
12		Gorilla	1074.1851392335263426502933725715
13		Western Lowland Gorilla	1076.4661612680909001937834545971
14	Monkeys	Grivet Monkey	1078.1153582682559317618142813444
15		Proboscis Monkey	1094.0886850152905935829039663076
16		Black Snub-Nosed Monkey	1091.2661122661122590216109529138

This table demonstrates, the distances between human DNA and some of the mammalian species. Each species is indicated by a family type. In the carnivorous part, all species of the same type are close to each other like bears and tigers family. In hoofed part, some of these results are very close to each other like Bactrian Camel and American Bison, but the other results are relatively close. In the Ape and monkey parts, some of these results are very close to each other, but the distance between Human DNA and Chimpanzee are very close than any other distances between Human and other species.

Table 2 compares between the distance results of Human and some of the different species of the Mammals family after applying Derivative Dynamic Time Warping Algorithm (DDTW) on the DNA signals.

Table 2: The Distance between Different Species of Mammals and Humans

#	Family	Species	Distance
1	Carnivorous	Spectacled Bear	1403.06551144549393939908716738224
2		Sloth Bear	1443.19738520408168369613122219501
3		Malayan Sun Bear	1425.72737465961881753173656676165
4		Asiatic Lion	1424.37419974391809773806016899219
5		Snow Leopard	1428.49796553374812901893164966253
6		Cheetah	1439.60782518707219423959003170109
7	Hoofed	Bactrian Camel	1422.48235578472440465702675282296
8		American Bison	1424.44661708253352117026224247324
9		Eurasian Elk	1478.52241323907446712837554551441
10		Horse	1422.43179251371407190163154155502
11	Apes	Chimpanzee	1353.87872985479498311178758774043
12		Gorilla	1394.79328310355890607752371579413
13		Western Lowland Gorilla	1396.55712344099765687133185565479
14	Monkeys	Grivet Monkey	1403.77336523125995881855487822349
15		Proboscis Monkey	1431.13579800899174188089091330773
16		Black Snub-Nosed Monkey	1431.78971364581661873671691866473

This table illustrates, the distances between human DNA and some of the mammalian species. Each specific species is indicated by a family type. The results of this table have increased than the results of Table 1, and the results are relatively close to each other but chimpanzee is still a shorter distance to human than any other species.

5.2. Comparing Human DNA with the Reptile Family

The Reptile family is a cold-blooded, scaly-skinned vertebrates. Most reptiles reproduce by laying leathery eggs. However, many lizards and snakes give birth to live young [1], [12]. Table 3 compares between the distance results of Human and some of the different species of the Reptile family after applying the Dynamic Time Warping Algorithm (DTW) on the DNA signals.

Table 3: The Distance between Different Species of Reptiles and Humans

#	Family	Species	Distance
1	Alligators and Crocodiles	Nile Crocodile	1072.57354352296101751562673598533
2		American Alligator	1072.73515667949413909809663891793
3		Chinese Alligator	1067.54456401766015005705412477253
4	Snakes	Ball Python	1083.45556774547458189772441983223
5		King Cobra	1093.51157662624041222443338483573
6	Tortoises and Turtles	Vietnamese Big-Headed Turtle	1088.48351951454969821497797966
7		Egyptian Tortoise	1074.78846683181950538710225373513
8		Annam Leaf Turtle	1064.66816939890713911154307425023

This table shows, the distances among Human DNA and some of reptile species. Each species is indicated by a family type. In the Alligator and Crocodile family, the results are close to each other. However, the results of the other families are relatively close to each other.

Table 4 compares between the distance results of Human and some of the different species of the Reptile family after applying Derivative Dynamic Time Warping Algorithm (DDTW) on the DNA signals.

Table 4: The Distance between Different Species of Reptiles and Humans

#	Family	Species	Distance
1	Alligators and Crocodiles	Nile Crocodile	1390.26585581245990397292189300006
2		American Alligator	1393.43103310410992889956105500046
3		Chinese Alligator	1374.53454415954411160782910888343
4	Snakes	Ball Python	1393.29978561219627408718224614863
5		King Cobra	1403.12832245742470149707514792683
6	Tortoises and Turtles	Vietnamese Big-Headed Turtle	1429.43251192368848023761529475453
7		Egyptian Tortoise	1424.20120501030601189995650202043
8		Annam Leaf Turtle	1399.68711093375895870849490165713

This table illustrates, the distances between human DNA and some of the Reptiles species. Each specific species is indicated by a family type. The results of this table have increased than the results of Table 3., and the results are relatively close to each other.

5.3. Comparing Human DNA with the Amphibian Family

Amphibians include frogs, toads, newts, salamanders and the curiously worm-like caecilians. Some of amphibians live permanently on land while others, such as the axolotl, never leave the water [1], [12]. Table 5 compares between the distance results of Humans and some of the different species of the Amphibian family after applying Dynamic Time Warping Algorithm (DTW) on the DNA signals.

Table 5: The Distance between Different Species of Amphibians and Humans

#	Family	Species	Distance
1	Frogs and Toads	Lake Victoria Clawed Frog	1092.80934605281345284311100084057
2		Shaping Frog	1090.3453296703296473424416035414
3		Chusan Island Toad	1061.5463320463320542330062081039
4	Salamanders and Newts	Ryukyu Spiny Newt	1079.4857966241252142935991287231
5		Hong Kong Warty Newt	1078.27799862919800943927839339838

This table demonstrates, the distance between Human DNA and some Amphibian species. Each specific species is separated by a family type. In the frogs and toads family, some of these results are very close like Shaping Frog and Lake Victoria Clawed Frog. The same is true with the Salamanders and Newts family.

Table 6 compares between the distance results of Humans and some of the different species of the Amphibian family after applying Derivative Dynamic Time Warping Algorithm (DDTW) on the DNA signals.

Table 6: The Distance between Different Species of Amphibians and Humans

#	Family	Species	Distance
1	Frogs and Toads	Lake Victoria Clawed Frog	1413.9930833200064625998493283987
2		Shaping Frog	1427.7448515325670541642466560006
3		Chusan Island Toad	1388.6673306772909199935384094715
4	Salamanders and Newts	Ryukyu Spiny Newt	1431.6704833597464130434673279524
5		Hong Kong Warty Newt	1430.8383306962025471875676885247

This table shows, the distances between human DNA and some of the Amphibian species. Each specific species is indicated by a family type. The results of this table have increased than the results of Table 5., and the results are relatively close to each other.

5.4. Comparing Human DNA with the Canis Family

Canis family includes dogs, wolves and foxes. Table 7 compares between the distance results of Human and some of the different Canis species family after applying Dynamic Time Warping Algorithm (DTW) on the DNA signals.
.

Table 7: The Distance between Different Species of Canis and Humans

#	Species	Distance
1	Domestic Dog	1109.97440863189922231099400669336
2	Eurasian Wolf	1103.676808186972721159690991044
3	Mongolian Wolf	1102.40805313408054644424192905426

This table illustrates, the distance between human DNA and some Canis species. The results of this family are relatively the same for different species.

Table 8 compares between the distance results of Humans and some of the different species of the Canis family after applying Derivative Dynamic Time Warping Algorithm (DDTW) on the DNA signals.

Table 8: The Distance between Different Species of Canis and Humans

#	Species	Distance
1	Domestic Dog	1447.7264964086193685943726450205
2	Eurasian Wolf	1445.2515612489992236078251153231
3	Mongolian Wolf	1452.3620290549170022131875157356

This table demonstrates, the distances between human DNA and some of the Canis species. Each species is indicated by a family type. The results of this table have increased than the results of Table 7., and the results are very close to each other.

5.5. Comparing Human DNA with the Felines Family

The feline family includes cats, lions, tigers, and cheetahs [1]. Table 9 compares between the distance results between Human and some of the different species of the Feline family after applying Dynamic Time Warping Algorithm (DTW) on the DNA signals.

Table 9: The Distance between Different Species of Felines and Humans

#	Species	Distance
1	Asiatic Lion	1111.4275681630404051247751340027
2	Leopard	1094.1465784286117468582233414054
3	Snow Leopard	1108.2985383342527256900211796165
4	Clouded Leopard	1117.7243704417228400416206568848
5	Cheetah	1116.6109426681366585398791357875

This table illustrates, the distance between human DNA and some Feline species. The results of this family are relatively the same for different species.

Table 10 compares between the distance results between Human and some of the different species of the Feline family after applying Derivative Dynamic Time Warping Algorithm (DDTW) on the DNA signals.

Table 10: The Distance between Different Species of Felines and Humans

#	Species	Distance
1	Asiatic Lion	1425.1498239718355307559249922633
2	Leopard	1384.6044231084852071944624185562
3	Snow Leopard	1429.4998006061573505576234310865
4	Clouded Leopard	1428.3566428003182409156579524279
5	Cheetah	1440.8438494589433958026347768188

This table shows, the distances between human DNA and some of the Felines species. Each specific species is indicated by a family type. The results of this table have increased than the results of Table 9., and the results are relatively close to each other.

5.6. Comparing Human DNA with the Fish Family

Fish were the earliest vertebrates to appear on Earth, having evolved more than 500 million years ago. Fish typically have fins and are covered in scales, are cold-blooded and breathe using gills [1], [12]. Table 11 compares between the distance results of Human and some of the different species of Fish family after applying Dynamic Time Warping Algorithm (DTW) on the DNA signals.

Table 11: The Distance between Different Species of Fish and Humans

#	Family	Species	Distance
1	Shark	Shark Mullet	1078.47053964757719768385868519 54
2	Shark	Elephant Shark	1083.42028583321780388359911739 83
3	Shark	Gummy Shark	1083.35347682119208911899477243 42
4	Whale	Blue Whale	1085.42184687802841835946310311 56
5	Whale	Striped Dolphin	1090.45582551076745403406675904 99
6	Other	American Angler	1052.4144907723855340000932238996
7	Other	Swordfish	1057.1645258267285498732235 28266

This table demonstrates, the distance between human DNA and some Fish species. Each species are separated by a family type. In the shark family, the results of elephant and gummy sharks are almost the same but the results of mullet shark are relatively close to the other shark results. In the Whale family, the results are relatively close to each other.

Table 12 compares between the distance results between Human and some of the different species of the Fish family after applying Derivative Dynamic Time Warping Algorithm (DDTW) on the DNA signals.

Table 12: The Distance between Different Species of Fish and Humans

#	Family	Species	Distance
1	Shark	Shark Mullet	1396.1869024856596297468058764935
2	Shark	Elephant Shark	1459.0834002569868061982560902834
3	Shark	Gummy Shark	1434.8754394375200718059204518795
4	Whale	Blue Whale	1429.1688659628919140232028439641
5	Whale	Striped Dolphin	1438.4635658294673703494481742382
6	Other	American Angler	1375.1492389408592771360417827964
7	Other	Swordfish	1376.8337736151136141415918245912

This table illustrates, the distances between human DNA and some of the Fish species. Each specific species is indicated by a family type. The results of this table have increased than the results of Table 11., and the results are relatively close to each other.

6. Verification

Table 13 demonstrates the closest results of Human and some of the species of different families after applying Dynamic Time Warping Algorithm (DTW) on the DNA signals.

Table 13: The Distance between Different Species of Different Families and Humans

#	Species	Distance
1	Pig	1042.6019257221457792184082791209
2	Chimpanzee	1045.1434301521437646442791447043
3	American Angler	1052.4144907723855344000932238996
4	Swordfish	1057.1645258267285498732235282266
5	Chusan Island Toad	1061.5463320463320542330620810390
6	Annam Leaf Turtle	1064.6681693989071391115430742502
7	Chinese Alligator	1067.5445640176601500570541247725
8	Nile Crocodile	1072.5735435229610175156267359853
9	American Alligator	1072.7351566794941390980966389179
10	Gorilla	1074.1851392335263426502933725715
11	Egyptian Tortoise	1074.7884668318195053871022537351
12	Western Gorilla	1076.4661612680909001937834545970

Table 14 shows, the closest results of Human and some of the species of different families after applying Derivative Dynamic Time Warping Algorithm (DDTW) on the DNA signals.

Table 14: The Distance between Different Species of Different Families and Humans

#	Species	Distance
1	Pig	1289.6494919021911300660576671362
2	Chimpanzee	1353.8787298547949831117875874043
3	Chinese Alligator	1374.5345441595441116078291088343
4	American Angler	1375.1492389408592771360417827964
5	Swordfish	1376.8337736151136141415918245912
6	Chusan Island Toad	1388.6673306772909199935384094715
7	Nile Crocodile	1390.2658558124599039729218930006
8	American Alligator	1393.4310331041099288995610550046
9	Gorilla	1394.7932831035589060775237157941
10	Western Gorilla	1396.5571234409976568713318556547
11	Annam Leaf Turtle	1399.6871109337589587084949016571
12	Egyptian Tortoise	1424.2012050103060118999565020204

These tables illustrate, the closest DTW and DDTW distance results between human and different species of different families. The pig is added as a new species. Amazingly, with 5000 base pairs-based computation, the distance between human and pig is much closer than the ratio between human and chimpanzee.

Dr. L. Schook and J. Beever at University of Illinois animal geneticists have created a side-by-side comparison of the human genome and the pig genome that reveals remarkable similarities. Dr. Schook said: "we took the human genome, cut it into 173 puzzle pieces and rearranged it to make a pig, everything matches up perfectly. The pig is genetically very close to humans" [1] [13]. The two appendices A and B, charts lay out the first 200 input DNA values of different species as mentioned in table 13 and 14 respectively.

7. Summary and Conclusion

Within the cells of any organism is a substance called Deoxyribonucleic Acid (DNA) acting as the genetic blueprint. DNA sequence consists of A, C, G and T. Applying Dynamic Time Warping (DTW) algorithm on the quaternary signal equivalent of humans and other selected species, we were able to compare different genetic codes quantitatively. In this work we have discussed the following

- Converting a DNA biological signal into a digital signal.
- The digital representation of DNA sequences enables us to apply the Dynamic Time Warping (DTW) to the developed signal.
- Dynamic Time Warping (DTW) and the Derivative Dynamic Time Warping (DDTW) algorithms are utilized to compare Human DNA and other selected species by calculating both the warping path values and the distance between them. The results are depicted in Table numbers 1 to 14.
- Tables 13 and 14 show the DTW and DDTW the closest distance results between human DNA and different species of different families signals.

Applying this methodology, different biological families can be quantitatively reconstructed.

References

[1] A. Magdy, M. Saeb, A. Baith, A. Khadragi, "The Haar Wavelet Transform of the DNA Signal Representation," International Conference on Computing and Information Technology," (WASET), January 2013.

[2] El-Zanaty, M. Saeb, A. Baith, S. K. Guirguis,"Haar Wavelet Transform of The Signal Representation of DNA Sequences," The International Journal of Computer Science and Communication Security (IJCSCS) Volume 1, July 2011.

[3] D. Morris, The Naked Ape: A Zoologist's Study of the Human Animal, Dell publishing, 1967, pp. 1.

[4] E. Keogh, S. Lonardi, V. B. Zordan, S. H. Lee, M. Jara, "Visualizing the Similarity of Human and Chimp DNA," University of California, Riverside, USA, 2005.

[5] Site: news.nationalgeographic.com   "Chimps, Humans 96 Percent the Same, Gene Study Finds," http://news.nationalgeographic.com/news/2005/08/0 831_050831_chimp_genes.html,   accessed   on 15/12/2012.

[6] M. El-Zanaty, M. Saeb, A. Baith, S. K. Guirguis, E. El-Abd,"Virus Classifications Based on the Haar Wavelet Transform of Signal Representation of DNA Sequences," The International Journal of Computer Science and Communication Security (IJCSCS) Volume 2, February 2012.

[7] Site: ncbi.nlm.nih.gov   Gene   Bank, http://www.ncbi.nlm.nih.gov/genbank, Accessed on 9/12/2012.

[8] P. Senin, "Dynamic Time Warping Algorithm Review," University of Hawaii at Manoa Honolulu, USA, December 2008.

[9] T. F. FURTUNĂ, "Dynamic Programming Algorithms in Speech Recognition," Revista Informatica Economică , 2008.

[10] P. Lama, M, Namburu, "Speech Recognition with Dynamic Time Warping using MATLAB," PROJECT REPORT, SPRING 2010.

[11] E. J. Keog[†], M. J. Pazzani, "Derivative Dynamic Time Warping", Department of Information and Computer Science University of California, Irvine, California 92697 USA.

[12] Site: arkive.org   , http://www.http://www.arkive.org, Accessed on 9/12/2012.

[13] Site: aces.uiuc.edu "Human to Pig Genome Comparison   Complete," http://www.aces.uiuc.edu/Discover/discover37.cfm, Accessed on 12/12/2012.

BIOGRAPHY

Abdel-kader Magdy received the BSc. In Computer Engineering, Alexandria Higher Institute of Engineering & Technology (AIET), in 2010. He is a Master degree student in Computer Engineering at Arab Academy for Science, Technology & Maritime Transport, Alexandria, Egypt; He is an instructor in Computer Engineering at Alexandria Higher Institute of Engineering & Technology (AIET).

Ahmed Khadragi received the BSc. And MSc. in Computer Engineering, College of Engineering and Technology, The Arab Academy for Science, Technology & Maritime Transport, Egypt, in 2000 and 2003 respectively. He received the PhD in Computer Science, School of Computing, Science & Engineering, College of Science & Technology, The University of Salford, UK, in 2011. In 2000 Ahmed started his academic career as a Graduate Teaching Assistant in the Department of Computer Engineering, The Arab Academy, Egypt. He was then promoted to an Assistant Lecturer in 2003. Currently he is a lecturer in the same department. Through the study of his PhD, he has acquired and built an advanced knowledge base in computer engineering, microcontroller based system design, data acquisition, modeling and animation. Dr. Khadragi has recently designed and implemented a system that interprets and teaches sign language that met a lot of success locally and internationally. He has constructed his own hardware and software lab through assembling all the different equipments and supporting materials from USA, Egypt, Italy, France and UK. Currently his main research interests include intelligent systems design, graphic modeling and animation, steganography and cryptography.

A. Baith MOHAMED, received the BSc. In Computer Science, Vienna University, MSc. And Ph.D. in Computer Science Vienna University, in 1992. He is a Professor at the Arab Academy for Science and Technology (AASTMT), Computer Engineering Department.   In addition, he holds the position of Vice Dean for Training and Community Services, College of Engineering and Technology. His research interests include computer and Network Security, Bioinformatics, Steganography, cryptography, and Genetic Algorithms. He was also a member of an International projects teams in Europe, for design and implementation and maintenance of subsystems in the environment of peripheral processor controls as part of a large Public Switched Systems (EWSD) in SIEMENS, AG. Austria. Also, he was a scientific researcher in the department of Information Engineering, Seibersdorf Research Institute (Atomic Energy Agency) in Austria, for design and implementation of security software system in the domain of railway automation project (VAX/VMS, DEC systems). He was also a member of software testing for distribution points in an international project in AEG, Vienna, Austria.   He is a senior member of IEEE Computer Society, USA since 2001.

Magdy Saeb received the BSEE, School of Engineering, Cairo University, in 1974, the MSEE, and Ph.D. degrees in Electrical & Computer Engineering, University of California, Irvine, in 1981 and 1985, respectively. He was with Kaiser Aerospace and Electronics, Irvine California, and The Atomic Energy Establishment, Anshas, Egypt. Currently, he is a professor and head of the Department of Computer Engineering, Arab Academy for Science, Technology & Maritime Transport, Alexandria, Egypt; He was on-leave working as a principal researcher in the Malaysian Institute of Microelectronic Systems (MIMOS). He holds five International Patents in Cryptography. His current research interests include Cryptography, FPGA Implementations of Cryptography and Steganography Data Security Techniques, Encryption Processors, Mobile Agent Security.www.magdysaeb.net

Appendix A

The following chart laying out the first 200 DNA values of some species without Derivation to be input to Dynamic Time Warping (DTW) algorithm. These patterns are to be used to compare DNA codes.
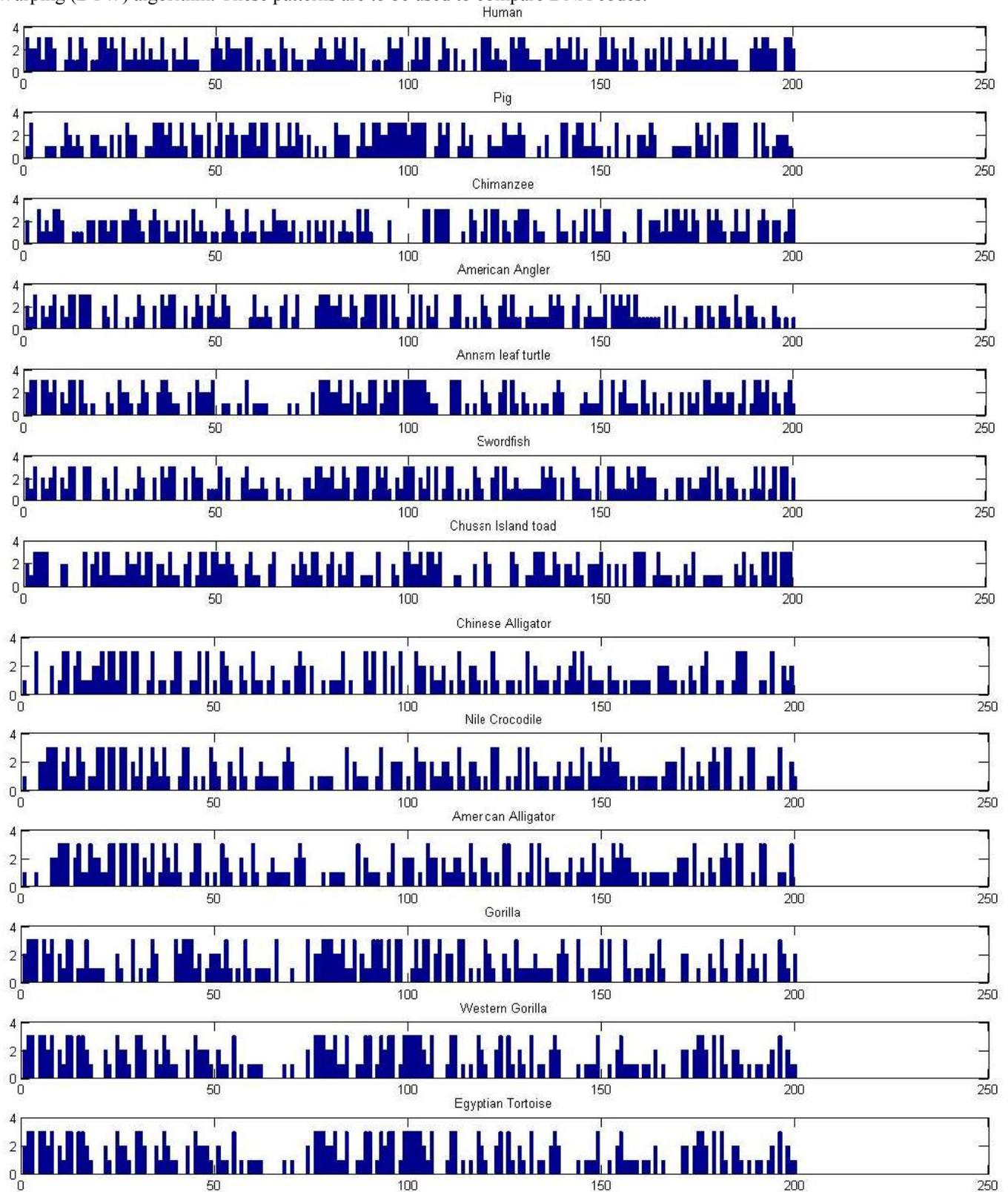
The DNA Input Signal of Dynamic Time Warping

Appendix B

The following chart laying out the first 200 DNA values of some species after the first derivative to be input to Derivative Dynamic Time Warping (DDTW) algorithm. These patterns are to be used to compare DNA codes.

The DNA Input Signal of Derivative Dynamic Time Warping