

Project Proposal - Digital Image Processing
JPEG-based image coding solution
for data storage on DNA
By Athul V
2021-2022

1.Introduction

I am a double major student in EEE and Biology and when I came across this project on DNA based storage, I thought it would be the ideal for long term storage needs. Today's harddisks have a storage life of no longer than 10-20 years. And recent studies have shown that DNA is a promising future for long term storage devices.

I would like to look into making a simulator and looking into the efficiency of JPEG based DNA storage of Data.

The project would involve applying conventional JPEG algorithms combined with DNA compression algorithms to achieve an efficient model.

2.Idea

Environment

I am currently using OpenCV with C++, I am happy to shift to Matlab too.

Prototype

My current Prototype involves converting the Image into grayscale then black and white, then converting each black and white pixel into 1 and 0 (binary) then assigning this binary data to DNA sequence using the following table.

Two bits	Nucleotides
00	A (Adenine)
10	G (Guanine)
01	C (Cytosine)
11	T (Thymine)

This is the image I am using (lena1.png)



Now this image with Threshold Black/White 50% Gray Threshold, I get



Now I assign each pixel , Black = 0 and White = 1. And from that I get the binary sequence of <https://pastebin.com/7wnQgyEd>

Here each line which is 199 (Width of the Image) is 0 and 1.

Now I convert this 0 and 1 into a DNA sequence of AGCT from the table.

<https://pastebin.com/vFNMJtVi>

As you can see even my simple prototype has reduced the width by half(100 characters).

Now this can again be converted back and retrieved back into our black and white image.



I clearly see room for a lot of improvement here. And the possibilities are endless. I need to sequencing and color images etc.

3. Conclusion

This study would be on state of the art technology and probably would be pioneering in nature. I would love to continue this project as my master thesis, if the project can be done.

4. References

<https://www.cs.utexas.edu/~bornholt/dnastorage-asplos16/>

<https://jpeg.org/jpegdna/>

<https://arxiv.org/abs/2103.09616>

https://www.researchgate.net/publication/350131512_A_JPEG-based_image_coding_solution_for_data_storage_on_DNA