Log

Eirik Lie Hegre

16. april 2020

Innhold

1 2020–03–24 **2**
1.1 Brief summary 2
1.2 Databending 2
1.3 Method and Experiments 2

2 2020–03–25 **3**

3 2020–03–26 **3**

4 2020–04–04 **3**

5 2020–04–05 **3**

6 2020–04–06 **4**

7 2020–04–07 **4**

8 2020–04–08 **4**

9 2020–04–09 **4**

10 2020–04–10 **4**

11 2020–04–13 **5**

1 2020–03–24

1.1 Brief summary

My goal is to try and figure out more in depth about the jpg file format and how I can databend it. Hopefully with some predictable results. If I open an image and start randomly changing bytes the results often come out somehow similar. Bright saturated colours, horizontal lines, compression artefacts, and repetitions of the image content amongst other things.

1.2 Databending

The kind of databending I want to do is by opening the images in a hex editor and edit the ASCII text. I want to see if there are ways that I can affect the image in a predictable way by finding out more about how the format works. To do that I will have to do some experiments.

1.3 Method and Experiments

My method will include finding four images, two that are different, but the same resolution, and two more that are in different sizes. To sum up, four images, three resolutions. These will all be scrubbed of data in a hex editor, down to the size that removal of another byte will cause the file to break. By only keeping the header of the original images I hope I will have a more common ground to start from. The source for my images will be NASA since their images are licensed under the Creative Commons.

So first of all I will scrub the data from the images. This turns the images into grey canvases for the new data to be reapplied. Having done that I will start writing in new data in the form of ASCII in a hex editor. With the text that I write I will see if I can find patterns in how the computer interprets the data I write into the image. An easy experiment will be to discover how many characters make a line, I expect this to be dependent on resolution. After that it will be to see if I can predict the colour as well. These things will be tried out on all four images at the same time so that I will be able to continuously compare the results of my changes.

2 2020–03–25

Made the four templates, copied them, and made test files for my experiments. Having scrubbed the templates down to the point where a single more deleted byte would break the file. This left me with four templates of the following sizes, template1 is 1041×694, template2 is 4053×3461, template3 and template4 are both 1280×720. These were all copied into four files, test1–4 for easy experimenting. What I wanted to try first was to see how many characters do I need make a line of pixels in the particular image I am editing. I expected the different sized images to behave differently, but I found out the same—sized also acted differently. Not only in colour, but what made one line test3, made several lines in test4.

3 2020–03–26

Did not do much today, but added a license, GPL v.3.0 and added text to the README

4 2020–04–04

Long Break in the work. Worked on an application for exchange in the fall, and did not get back in the rythm afterwards. But I am back in business now! I need to produce more content, and today I started writing about some of my thoughts regarding open source. Each image adds a new line of pixels to the existing image. In the end I hope to have a full image, and maybe make som animations with them.

5 2020–04–05

I renamed notes.txt to log and moved it to the main directory. In the different image directories I created new notes.txt files for each with the plan to write my ideas for each specific folder. I also made 29 images in the Alphabet directory.

6 2020–04–06

I have started updating the log and the notes.

7 2020–04–07

Made two new test directories. One named non_repeating_text and one named metadata_experiments. The plan for these directories would be to try out metadata manipulation and just feeding random key strokes into an image. I started out with the random key strokes and I ended up with a broken image. I will try it one more time, but it seems like just random typing might break an image.

8 2020–04–08

Today I worked with a program called exiv2 which is a metadata manipulation program. I had hoped by using it I could make two seemingly similar images act differently with the same sets of data put into them. That did not happen. Two identical images will still look the same even if I manipulate the metadata in one of them. I did not delve very deep into this though. So I might get some new results with some further exploration.

9 2020–04–09

I did not do any work today, had the need for a think about what to do next.

10 2020–04–10

I wanted to readress the non—repeating text experiments, and turn them into the direct opposite. I found out that it was time consuming putting in the non—repeating text and have decided on using repeating patterns of text instead. First out I will use a sonnet by Shakespeare and see where it will take me.

11 2020–04–13

Took a break over the weekend. Today I finished the sonnet experiments. But instead of only applying one template to the first ten sonnets, I used all four template instead.