Visual Memory: Storing “The Entire History of You”

Will Bozic
The Centre for Interdisciplinary Science, University of Leicester
23/04/2018

Abstract
This paper investigates the plausibility of the video recording achieved by the bionic grain depicted in season one, episode three of the TV series Black Mirror. The grain grants the wearer the ability to record all sight and sound experienced through their eyes and ears respectively. Initially the storage required to record all visual information as if captured uncompressed by the human eye, for the average lifespan of both a female and a male was calculated. This was determined to be $6.65 \times 10^9$ and $6.37 \times 10^9$ Terabytes respectively. The physical volume required to store this using the most efficient hard-drive currently available was then investigated and compared to the volume of the grain. It was determined that when uncompressed and compressed the volume required was $1.18 \times 10^{11}$ and $5.34 \times 10^9$ times larger than a grain for a female and $1.13 \times 10^{11}$ and $5.11 \times 10^9$ times larger for a male respectively; thus showing the impossibility of this design given the deficit in current storage technology.

Introduction
In season one, episode three of the British sci-fi show Black Mirror: “The Entire History of You”, a reality is presented in which most people have bionic ‘grains’ implanted within their bodies [1]. These grains grant the wearer the ability to record their every move, seemingly using their eyes and ears to record all sight and sound they experience. These recordings are then stored on the grain and can be played back at a later date, either in front of the individual’s own eyes, or projected onto a screen. This paper investigates the computer storage space required to record all visual events, and the plausibility of storing that quantity of data on a chip the size of a ‘grain’ as depicted in the episode.

Theory and Discussion
The method of recording in the episode is left ambiguous and the video compression technique is unknown. A number of scenes within the episode showcase the ability of recorded video to be significantly magnified without any apparent loss of quality. As such, recorded video is investigated as if it were recorded uncompressed by the human eye. The size of an uncompressed video can be calculated using equations 1, 2 and 3 [2]:

\[
\text{Frame Size in MB} = \frac{\text{Resolution} \times \text{Colour Data} \times \text{MetadataFactor}}{8 \times 1024 \times 1024} \tag{1}
\]

\[
\text{MB of Video Per Second} = \text{Frame Size in MB} \times \text{Frame Rate per Second} \times \text{Chroma Factor} \tag{2}
\]

\[
\text{Colour data} = \text{Colour Bit Depth per Channel} \times \text{Total Number of Channels} \tag{3}
\]

The resolution of the human eye was calculated by Clark [3] as being approximately 576 megapixels. The colour bit depth per channel is assumed to be 24-bit “true colour” [3]. This is because the human eye can distinguish ~10 million different colours; 24-bit is the closest colour depth precision to this, expressing 16,77,216 colour variations across its total colour gamut [4]. The video is assumed to be in the YUV format, which has three channels: Y’ for brightness and C_b and C_r for colour information [5]. This is used as YUV is a video format evolution of the RGB image format, which encode approximately the same colours as perceived by the human eye [6].

It is assumed that the chroma factor, based on the level of chroma subsampling, experiences no compression. It is therefore equal to a ratio of 4:4:4 for Y’, C_B and C_R respectively, where Y’ is the luma brightness component, C_B is the blue difference component and C_R is the red difference component [7, 5]. This results in this variable equalling 1 and hence becoming negligible for the purpose of this calculation.
The metadata factor is the information about the recorded material, such as the date of capture and specific camera settings used for capture [8]. This adds approximately 2-3% more data to the file on average [2]. In the episode the grain is shown to store facial recognition information and the time and date of recording, so an upper estimate of 3% additional data is used.

In their study Davis et al. [9] showed that the human eye can perceive flicker artefacts over 800 Hz, but that the average viewer’s threshold for noticing artefacts was approximately 500 Hz. As 1 Hz = 1 fps, it is assumed that the “average framerate” of the human eye is 500 frames per second.

Using the above parameters and equation 3, the colour data is determined to be:

\[24 \times 3 = 72.\]

Hence using equation 1 the size of each frame of video is:

\[
\frac{5.76 \times 10^8 \times 72 \times 1.03}{8 \times 1024 \times 1024} = 5092.163 \text{ MB.}
\]

Using this and equation 2 the MB of video per second is:

\[5092.163 \times 500 \times 1 = 2546082 \text{ MB s}^{-1}.\]

In the episode the grains are shown to constantly record; assuming they are implanted at birth this gives an average recording time of \(2.55 \times 10^6\) s for men and \(2.61 \times 10^9\) s for women [10]. Hence the amount of video data stored in a grain over an average female \((V_{DF})\) and male \((V_{DM})\) lifetime is:

\[V_{DF} = (2.55 \times 10^6)(2.61 \times 10^9) = 6.65 \times 10^{15} \text{ MB} \quad \therefore V_{DF} = 6.65 \times 10^9 \text{ Terabytes}\]

\[V_{DM} = (2.55 \times 10^6)(2.50 \times 10^9) = 6.37 \times 10^{15} \text{ MB} \quad \therefore V_{DM} = 6.37 \times 10^9 \text{ Terabytes}\]

Currently the hard drive with the best volume to storage ratio is the Seagate 60 TB SAS SSD [11]. This hard drive has 60 terabytes (TB) of storage in a 3.5 inch form factor [12]. In order to store \(V_{DF}\) or \(V_{DM}\), \(1.11 \times 10^9\) or \(1.06 \times 10^9\) of these hard drives would be required respectively. As this hard drive is specialist, details about its exact dimensions are unreleased, however, the 3.5 inch form factor gives approximate dimensions of 147.00 mm \(\times\) 101.60 mm \(\times\) 42.00 mm [13]. This gives a volume of \(0.000627 \text{ m}^3\) per hard drive. Therefore the volume required to store \(V_{DF}\) and \(V_{DM}\) is approximately \(6.95 \times 10^4 \text{ m}^3\) and \(6.6 \times 10^4 \text{ m}^3\) respectively. At the end of the episode, the main protagonist removes his grain and throws it in a sink (figure 1).

![Figure 1 – Screenshot of final scene where the main protagonist removes his grain [1].](image)

Video compression drastically reduces video size; an uncompressed 1080p video uses approximately 149 MB s\(^{-1}\), while a H.264 compressed video uses approximately 6.75 MB s\(^{-1}\) [16, 4]. This example shows compression reducing file size by 22.1 times. If this reduction was applied to \(V_{DF}\) and \(V_{DM}\) they would be equal to \(3.01 \times 10^8\) TB and \(2.89 \times 10^8\) TB. This reduces the volume to \(3.15 \times 10^3 \text{ m}^3\) and \(3.02 \times 10^3 \text{ m}^3\) respectively. While this is a significant decrease the volumes are still \(5.34 \times 10^9\) and \(5.11 \times 10^9\) times larger than the grain. This highlights the current implausibility of this design given current hard drive technology.

**Conclusion**

The total amount of uncompressed video data stored on a grain over the course of an average female and male lifetime was determined to be \(6.5 \times 10^9\) and \(6.37 \times 10^9\) TB respectively. Using the most size efficient hard drive available today this would take up \(6.95 \times 10^3 \text{ m}^3\) (female) and \(6.6 \times 10^4 \text{ m}^3\) (male). This is a volume \(1.18 \times 10^{11}\) and \(1.13 \times 10^{11}\) larger than a grain. Even when compressing video volumes \(5.34 \times 10^9\) and \(5.11 \times 10^9\) times larger than the grain are required, showing the current implausibility of the grain design.
References


