



NEW ZEALAND
ARCHAEOLOGICAL
ASSOCIATION

ARCHAEOLOGY IN NEW ZEALAND



This document is made available by The New Zealand
Archaeological Association under the Creative Commons
Attribution-NonCommercial-ShareAlike 4.0 International License.

To view a copy of this license, visit
<http://creativecommons.org/licenses/by-nc-sa/4.0/>.



SCANNING HISTORIC AERIAL PHOTOS: A COMPARATIVE EXPERIMENT

M.G. HUTCHINSON

A recent research project required the scanning and georeferencing of hundreds of aerial photographic frames from the initial government survey of the Waikato Valley in 1943, conducted by NZ Aerial Mapping Ltd (Gumbley & Hutchinson 2013).

The available funding did not stretch to the purchase of specialised scanning equipment, and acquiring image files scanned from negative plates was also beyond the financial reach of the researchers. Instead, paper prints of the aerial survey were accessed from the libraries of the University of Waikato and the Waikato Regional Council, and scanned on a portable consumer-level device. The frames were scanned at the maximum resolution available in the hopes that this would be sufficient to identify and locate visible archaeological features captured by the photography.

This paper relates the results of an attempt to assess the quality of the data yielded by the scanning process.

Methodology

For comparative purposes two scanners were used. These were a Hewlett Packard Deskjet 2050 (an A4 combination printer/scanner purchased new for about \$60) and a Brother MFC-J6510DW (an A3 combination device, worth about \$250). The Hewlett Packard scanner was the one used to capture data for the research project.

The software to drive the scanners and capture the resulting image files was SANE (Scanner Access Now Easy) running in a Debian GNU/Linux desktop environment. For consistency, both scanners were operated from the same computer, with the same neutral settings for gamma, brightness and contrast.

Files were captured into an uncompressed TIFF image format, and no additional editing was done to the captured images; files were viewed

and compared as they came out of the scanners. Also, all compression in the SANE configuration was turned off.

For the present experiment, frames were scanned from a small collection held in the NZAA Site Recording Scheme file (in Hamilton) and were essentially selected at random. For each frame, a series of scans were captured on each scanner, at increasing resolutions. Each scan was timed, and the resulting file sizes recorded.

A visual analysis of the resulting images was conducted. This consisted of choosing a target area near the centre of the frame in order to minimise the effects of parallax. The target required strong contrast and diagonal lines – the edges of buildings in strong sunlight – in order to determine the effects of pixelisation.

The images were viewed in the Eye of Gnome (EoG) image viewer. This application has the ability to dither, or blur, images viewed at magnifications higher than 100%. For the images presented here, the dithering function was disabled, so individual pixels could be seen.

A note on figures

Many of the figures in this paper are screenshots of the EoG viewer showing two image files side-by-side for comparison. In order to produce figures of suitable density for print publication, the two viewing windows were made as large as possible on a 1920 x 1080 px flatscreen monitor, and screenshots of the image area were captured into PNG files. The original files are 8-bit greyscale (256 levels) in uncompressed TIFF format, which have been converted into 24-bit colour images by the screen capture process. They have then been transformed into 1-bit screened paper prints for publication. Most of the original information will have been lost during these various transformations, and it may not be possible to see effects mentioned in the text in these printed copies.

Results

Table 1 lists the times taken to capture scans, and the resulting file sizes, for the two scanners used. Data from a single frame (SN266/707/1) are presented. File sizes are as reported by the unix DU (disk usage) utility.

The maximum resolution possible from the Hewlett Packard scanner is 1,200 x 1,200 px (Hewlett Packard n. d.). The Brother device is capable of generating images to 9,200 x 9,200 px, although anything above 2,400 px is interpolated rather than capturing actual data values for individual pixels (Brother n. d.).

Scanner	dpi	Time (s)	File size (B)	Pixel area	
Hewlett Packard	200	17	2,468	2.5M	1,336 x 1,887
Brother	200	8	2,388	2.4M	1,312 x 1,861
Hewlett Packard	300	21	5,548	5.5M	2,004 x 2,831
Brother	300	8	5,372	5.3M	1,968 x 2,791
Hewlett Packard	600	69	22,188	22M	4,008 x 5,662
Brother	600	33	21,572	22M	3,952 x 5,582
Hewlett Packard	1,200	228	88,740	87M	8,016 x 11,324
Brother	1,200	51	86,264	85M	7,904 x 11,164
Brother	2,400	132	344,868	337M	15,808 x 22,328
Brother	4,800	443	1,379,108	1.4G	31,616 x 44,656
Brother	9,600	1,762	5,515,024	5.3G	63,232 x 89,312

Table 1: Time and filesize data for frame SN266/707/1.

The first thing to note is that the more modern and expensive Brother scanner is much faster than the Hewlett Packard model, with times per scan approximately half that of the HP device. File sizes are predictably comparable between images of the same resolution produced on different devices, and (also predictably) doubling the resolution yields a file roughly four times larger.

Comparison of devices

Images created by the two devices at the same resolution were displayed side by side on the screen and compared visually.

Figure 1 shows the two 200 dpi images at full frame. This renders a magnification of 53%, given as the ratio of the number of image pixels per monitor pixel.

At this resolution and magnification, the two scans look almost identical. The Hewlett Packard-produced image (left) is slightly brighter, but otherwise no difference is discernible.

Figure 2 shows the same two images at 400% magnification.

At this magnification the individual pixels comprising the image can be seen. The Brother-produced image (right) is much smoother, whereas the HP image has higher contrast and appears more jagged. These effects become more pronounced at higher magnifications.

Figure 3 shows the two 200 dpi images at 800% magnification, cropped to the farmhouse and outbuilding complex near the centre of the frame. Here, not only are the individual pixels visible, but so too are clusters of pixels car-



Figure 1. 200 dpi scans. Left: HP. Right: Brother.

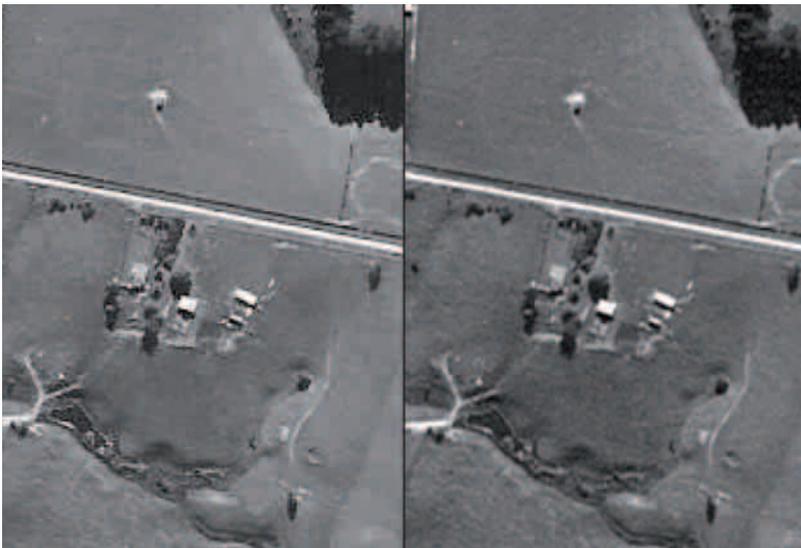


Figure 2. 200 dpi scans at 400%. Left: HP. Right: Brother.



Figure 3. 200 dpi scans at 800%. Left: HP. Right: Brother.

rying anomalous values in the HP-produced image on the left. Vertical lines spaced at intervals of eight pixels indicate these clusters.

Also apparent in the HP image are areas where adjacent pixels carry the same values, forming blocks of grey that are not reflected in the Brother image. These blocks are very similar to artefacts seen in images subjected to high compression using the lossy JPEG algorithm.

Figure 4 shows the same location near the centre of the frame, with the two 1,200 dpi images at 100% magnification.

Once again we see the smoother, more graduated grey values produced by the Brother device (right).

Figure 5 illustrates these differences more clearly. These are the 1,200 dpi scans presented at 1,000% magnification, centred on the farmhouse. The rectangular white area at centre is the highly-reflective gabled roof of the building, with dark areas below this representing shadows cast by the house. Linear features below this shadow area are fences.

The HP-produced image (left) shows the mottled appearance, jagged outlines and vertical lines similar to compression artefacts, whereas the Brother image (right) shows smooth graduations and higher definition in areas surrounding the house.

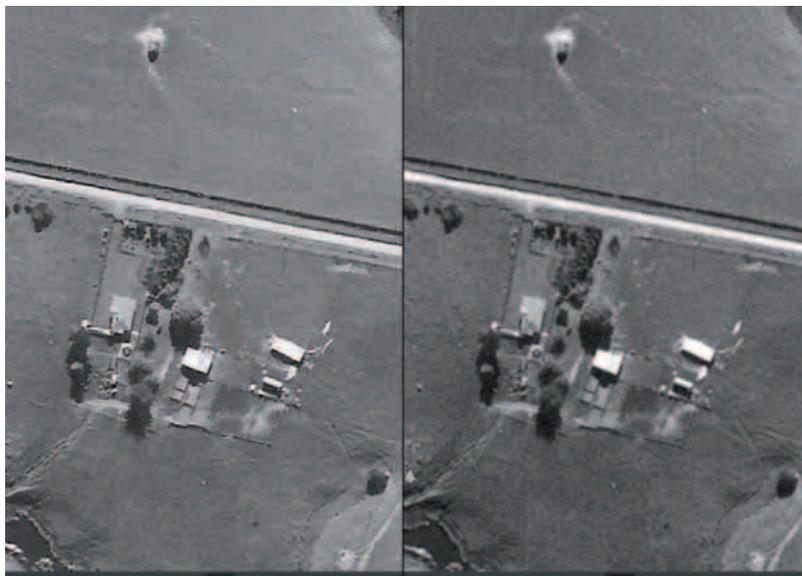


Figure 4. 1,200 dpi scans. Left: HP. Right: Brother.

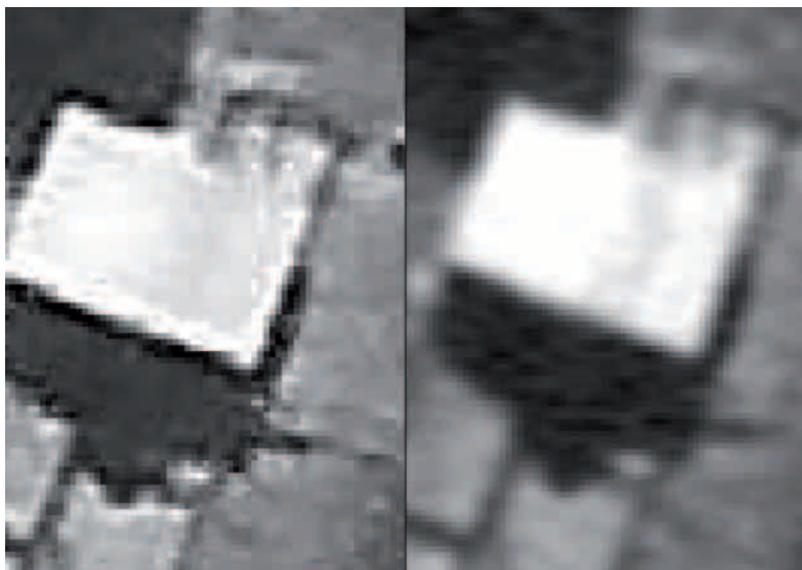


Figure 5. 1,200 dpi scans at 1,000%. Left: HP. Right: Brother.

Comparison of resolutions

Several image files at differing resolutions were compared for density of visual information. Given the jagged, artefactual nature of scans produced on the Hewlett Packard device, and the restriction to 1,200 dpi resolution, the following section displays images generated only on the Brother scanner.

Also, the files dealt with in this section are too large to be displayed in the simple viewer. For the purposes of comparing files above 2,400 dpi, the Gnu Image Manipulation Program (GIMP) was used.

Figure 6 shows the 4,800 dpi scan at 50% magnification. Of note is the bitty nature of the image, particularly visible in the diagonal lines representing the edges of the buildings. That such features are visible at less than full magnification indicates that the real resolution of the image is something less than that rated by the scanner and driver software.

The same image projected at 800% reveals that the pixelated nature is caused by sets of pixel clusters in which values are graduated evenly. This is presumed to be the results of the interpolation process whereby image data are “made up” by the scanner to create a denser image than the device is physically capable of.

Figure 7 shows the edge of one of the farm buildings. The pale area is the roof, and the dark area is the shadow cast by the building. Pixel clusters are clearly visible as squares with faded edges.

Figure 8 is provided by means of comparison. This is the 1,200 dpi image at 1,600% magnification, and the square shapes visible in it are individual pixels.

The specifications for the Brother printer claim that the device is capable of producing scanned images to an optical resolution of 2,400 dpi (Brother n. d.). Anything higher than that will be interpolated. We would predict, therefore, that scans at 2,400 dpi will be a similar but finer representation of the image from the 1,200 dpi scan depicted in Figure 8.

Figure 9 is a detail from the 2,400 dpi scan captured on the Brother scanner, presented at 100% magnification. This image shows the interpolation tiles in the pixelated area at the edge of the buildings.

Figure 10 is a comparison of the 1,200 dpi scan (left) with the 2,400 dpi scan, presented at magnifications to make the image sizes similar. The 1,200 dpi image is at 400% and the 2,400 dpi one at 200%.

The image on the right is clearly more pixelated than the lower-resolution image on the left. This implies a level of interpolation in the denser image, which is not indicated in the manufacturer’s technical data.



Figure 6. 4,800 dpi scan at 50% magnification.



Figure 7. 4,800 dpi scan at 800% magnification.



Figure 8. 1,200 dpi scan at 1,600% magnification.



Figure 9. 2,400 dpi scan at 100% magnification.

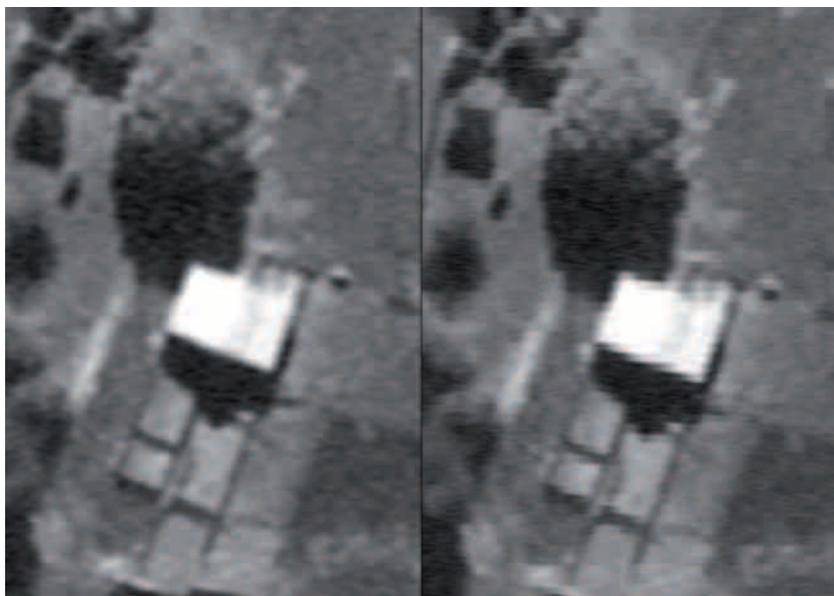


Figure 10. Comparison of 1,200 dpi (left) and 2,400 dpi scans.

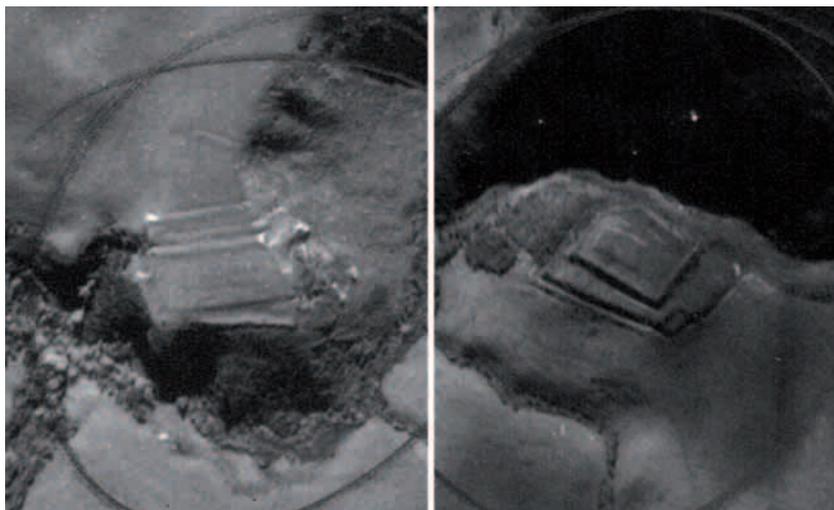


Figure 11. Pā sites S15/159 (left) and S15/139 on frame 707/1.

Discussion

We are able to see that image quality is to a certain extent dependent on the device used. Perhaps unsurprisingly, the more expensive of the two devices tested provides a better quality product.

What was surprising, however, was the level of software interpolation coming into the scanned products from both machines, especially the compression artefacts from the less expensive HP scanner. If this is the case, it is being done in the scanner, as all compression was turned off in the SANE software driver.

It would appear from this admittedly cursory experiment that the optimum resolution for scanning photographic prints is 1,200 dpi. This is the maximum resolution provided by both scanning devices, and despite claims by the manufacturer that the Brother scanner is capable of producing optical scans at twice this density, the results presented here call this into question.

The point

The reason why we do this is shown in Figure 11. This depicts detail from two areas on the frame SN266/707/1. Both images clearly show ditches, banks and other defensive features associated with two pā sites. Once this frame has been georeferenced, we will be able to make measurements of the visible features and map them, with high confidence.

References

- Brother (n. d.) MFC-1065DW Product specifications. Accessed 2014-11-15 from <http://www.brother-usa.com/MFC/ModelDetail/4/MFCJ6510DW/spec#.VGbMZxF6Rh>
- Gumbley, W and Hutchinson, M. G. (2013) Pre-European Maori Garden Sites in Waipa District: An assessment of the state of the resource. Unpublished report for the NZ Historic Places Trust.
- Hewlett Packard (n. d.) HP Deskjet 250 scanner specifications. Accessed 2014-11-15 from https://h10025.www1.hp.com/ewfrf/wc/document?docname=c02063036&tmp_task=prodinfoCategory&cc=us&d1c=en&lc=en&product=4027469#N1053.