The clue in the dot of the 'i': Experiments in quick methods for verifying identity via handwriting

Tom Benton Research Division

Summary

This article demonstrates some simple and quick techniques for comparing the style of handwriting between two examinations. This could potentially be a useful way of checking that the same person has taken all of the different components leading to a qualification, and form one part of the effort to ensure qualifications are only awarded to those candidates that have personally completed the necessary assessments. The advantage of this form of identity checking is that it is based upon data (in the form of images) that is already routinely stored as part of the process of on-screen marking. This article shows that some simple metrics can quickly identify candidates whose handwriting shows a suspicious degree of change between occasions. However, close scrutiny of some of these scripts provides some reasons for caution in assuming that all cases of changing handwriting represent the presence of imposters. Some cases of apparently different handwriting also include aspects that indicate that they may come from the same author. In other cases, the style of handwriting may change even within the same examination response.

Introduction

In order for assessments to be any use at all, it is crucial that they are taken by the same people to whom results will be issued. As such, we need measures to discourage any attempts at malpractice by one person completing an assessment on behalf of another. Reports of such forms of cheating are currently extremely rare in the UK; however, they are frequently reported in other countries and it is important that we should be prepared for the possibility of this type of cheating.

There are many possible ways of checking the identity of candidates. For example, for certain assessments internationally, candidates are required to take identification documents with them to the exam centre in order to be permitted to take the exam. As an alternative, for some tests produced by the exam board Cambridge English Language Assessment, test day photos are taken so that users of examination results are able to verify for themselves the identity of the person who actually took the assessment. However, in addition to such checks, there may be value in examining the handwriting used within assessments to verify that all of the different elements of a qualification are being taken by the same individual.

The advantage of using handwriting for identity verification is that it is a source of information that is already freely available to exam boards. The vast majority of Cambridge Assessment's examinations are taken using pen and paper and, furthermore, due to the rise of on-screen marking, scanned images of most candidates' scripts are already stored within our systems. Thus, it is theoretically possible for us to examine the handwriting used across all of the assessments taken by an individual to help reassure ourselves that the correct individual is receiving credit for his or her work.

Manually checking handwriting from different assessments against one another would be both laborious and expensive in terms of labour costs. For this reason, the aim of this research project was to begin to explore the extent to which such a process could be automated by computers.

Automatic handwriting recognition is a widely researched area (Dolega, Agam, & Argamon, 2008) with a wide variety of available algorithms. However, many of these algorithms are slow – requiring detailed tracing of the strokes used to form each of the words and letters on a page and could not be quickly applied to the thousands of digital images we hold. Instead, this project looks for whether there are any metrics of handwriting that are relatively quick to calculate which could provide a reasonable indicator of whether the author of two separate pieces of handwritten text was the same.

Source of images

The images for analysis were extracted from Cambridge Assessment's Digital Script Repository (DSR). Since this is the first time we have undertaken analysis of this kind, a relatively simple example was chosen. Two compulsory Higher Tier papers, taken two days apart as part of a General Certificate of Secondary Education (GCSE) in English Literature in June 2014, were chosen for use in the analysis. Throughout this article, the two papers will be referred to as 'Unit A' and 'Unit B'. Both examinations required candidates to provide essay-type responses written on lined paper. Excluding the front and back covers of the script, for the vast majority of candidates 6 scanned pages were available from Unit A and 14 from Unit B, although it was rare for candidates to actually write on all of the available pages.

The consistent format of responses between the two assessments simplified the process of analysis. The aim of the project was to develop some simple measures for the style of handwriting and explore the extent to which such metrics remain stable between different examination occasions. Metrics that are highly stable between occasions might be useful for verifying that the same person has taken each examination.

Software, methodology and metrics

All of the analysis within this article was undertaken using the free statistical software package R version 3.2.2 (R Core Team, 2015). The majority of the work of reading, segmenting and manipulating images was done using the package *EBImage* (Pau, Fuchs, Skylar, Boutros, & Huber, 2010).

Pre-processing of images

Before analysis of handwriting can begin, a few pre-processing steps are necessary. The key steps of this process are shown in Figure 1. For reasons of space, the images in Figure 1 are restricted to a portion of text at the top of one page of a candidate's response.

The top left-hand of this image gives an example of what (part of) a single page of a candidate's response might look like before any pre-processing has been applied. To begin with, the full-page image is read into R as a grayscale matrix. The data matrix has one value for each of the 2300 × 1620 pixels in the image and, as it is standard for data representing images, higher values are given to whiter sections of the image and lower values to the blackest sections – that is the sections where there is actual writing.

The first step of pre-processing is to attempt to distinguish shapes that represent actual handwriting from those that represent margins, the typed text of the question, or dotted lines. This task requires a number of steps.

To begin with, 5 per cent of the original image is removed on both the right and the left. This is done to remove dark black lines that may be created at the edge of the image as part of the scanning process. Next, the image is converted from grayscale to black and white using *Otsu's method* (https://en.wikipedia.org/wiki/Otsu's_method). Now each pixel in the image is either represented by a 0 (white) or a 1 (black). Next, we break the image into sections of joined up black pixels. Due to the medical context of the software (which treats white sections as indicating the presence of something and black sections as absence), it is necessary to take the negative of the image before doing this. The identified separate sections of joined up pixels are shown in different colours in the top right-hand corner of Figure 1.

The size of each segmented section in the top right-hand image can be used to identify two sets of items of interest: dots from the dotted lines and written words and letters. Some experimentation revealed that the dots printed within dotted lines on the exam paper tended to contain between 10 and 55 pixels. Using this rule of thumb, we could count the number of pixels within such dots in each row of the image. Rows where between 80 and 200 pixels were within these identified 'dots' were deemed likely to represent a dotted line[1]. Thus, we could identify the first and last such rows in the matrix as likely representing the top and bottom dotted lines, and restrict the matrix to writing between these two only. One downside of this approach was that if the candidate wrote on top of the first dotted line then this text was lost. In addition any text below the final dotted line is also lost.

Some experimentation showed that most sections that represented handwriting (that is, words or elements of words) tended to contain between 60 and 3,000 pixels. Any joined groups of pixels outside of this range were set to be white as they were unlikely to represent writing. However, this often meant that the dots of handwritten 'i's or 'j's were deleted.

Applying the steps above led to images of the type shown in the bottom left-hand corner of Figure 1. Some simple metrics of handwriting, to be described later, were calculated based purely upon this image.

However, one problem with using the image so far was that the metrics of handwriting may be affected by the thickness of the pen used by the candidate. This was addressed by using a crude form of 'thinning'. Thinning is the process of trying to find a skeleton form of any given shape that is only one pixel wide at any point but that preserves the essence of the shape. Many algorithms have been proposed for this procedure (see Lam, Lee, & Suen, 1992). However, the formal approaches

1. This obviously ignores the height of the 'dots'. We are just looking at the number of pixels within the dots, within each row of the matrix.

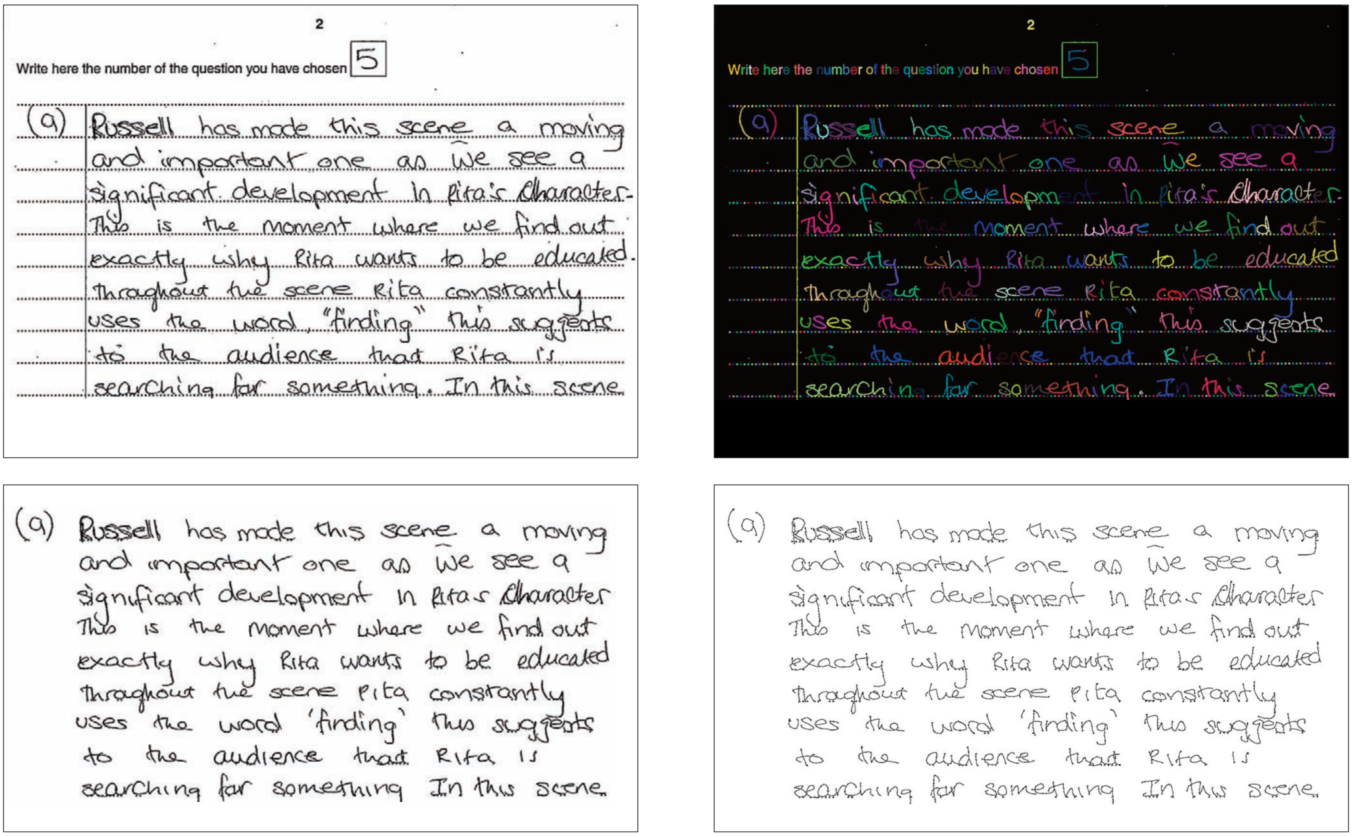

Figure 1: Steps in pre-processing images (Original image – top left; Segmented original image – top right; Cleaned image – bottom left; Thinned image – bottom right)

in the literature are fairly slow, requiring each pixel of an image to be considered in turn in relation to the surrounding pixels, and then either left alone or deleted as necessary. The decision for each pixel may then affect decisions for subsequent pixels. This means that the calculations need to be processed one at a time (at least within connected areas of an image).

As an alternative, an approximate but fairly fast approach was adopted. For the purposes of this method, the blurred density of each pixel was calculated by taking a weighted average of the pixels in the surrounding area (including the pixel itself) with more weight given to pixels that were nearby. Then, pixels of writing were only retained if the blurred density was greater than that of the pixels on either side in at least one direction. That is, the blurred density was either greater than both of:

- the pixel on the left and the pixel on the right, or
- the pixel below and the pixel above, or
- the pixel above and to the right and the pixel below and to the left, or
- the pixel above and to the left and the pixel below and to the right.

This method could be applied fairly quickly to each page and certainly helped address issues relating to the thickness of the pen. However, it could not be said to be a true 'thinning' method as the resulting image was often two or three pixels wide in certain areas rather than one. An example of why this occurred is given in Figure 2. This image represents a section of a letter 'p' from an image with the grey squares representing shaded pixels in the original image and the black squares those pixels that remain shaded after thinning. The numbers in the chart represent the blurred density at each point. As can be seen, in most places the image is reduced to be one pixel wide (in at least one direction). However, in the first two rows of the chart the image remains two pixels wide in all directions. This is because, whereas one pixel on the second row (with the value 0.634) has greater blurred density than that of the pixel on the left, this other pixel has greater density than the pixel below and to the right, as well as the pixel above and to the left. For this reason both pixels are retained and the shape is not completely thinned.

0.396	0.557	0.661	0.661	0.556	0.394	0.246	0.170
0.339	0.499	0.615	0.634	0.545	0.390	0.241	0.154
0.287	0.444	0.571	0.608	0.538	0.394	0.244	0.147
0.244	0.396	0.531	0.588	0.539	0.408	0.258	0.150
0.208	0.354	0.496	0.573	0.547	0.431	0.282	0.162
0.177	0.315	0.462	0.558	0.556	0.457	0.311	0.178
0.149	0.279	0.429	0.541	0.562	0.480	0.336	0.195
0.128	0.249	0.400	0.524	0.563	0.496	0.357	0.211

Figure 2: Example of the approximate thinning algorithm

Nonetheless, this algorithm can be applied fairly quickly, and, to a large extent, accounts for the thickness of the pen in any piece of writing. The bottom right-hand section of Figure 1 shows an example of a final image after thinning.

Having thinned the image, handwriting metrics are calculated based on each word. Ideally, we would rely on pupils having joined-up writing to identify words as any continuous sequence of marked pixels within the image. Sadly, as seen earlier, few candidates have completely joined-up writing. For example, in forming the letter 'f', most candidates will pause their writing to cross the letter rather than immediately joining on to the next one. Similar behaviour can also occasionally be found with letters such as 't', 'i' and 'j'. In addition, capital letters at the start of sentences are not usually joined to others in the same word. For this reason, a dilated version of the writing was created whereby, every time we find a marked pixel within the image, seven pixels to the left and seven to the right are also marked. For most candidates, applying this dilation step ensured that the majority of the letters within a word were joined together whereas separate words usually remained separated within the image. Note that, although this step is used to identify the location of words, the thinned version of the image produced in the previous step is also retained in order to calculate handwriting metrics. Figure 3 illustrates the thinning and word segmentation steps. The dilated image portions used to identify separate words are shown in different colours. Within each 'word' a blue line shows the thinned version of the handwriting.

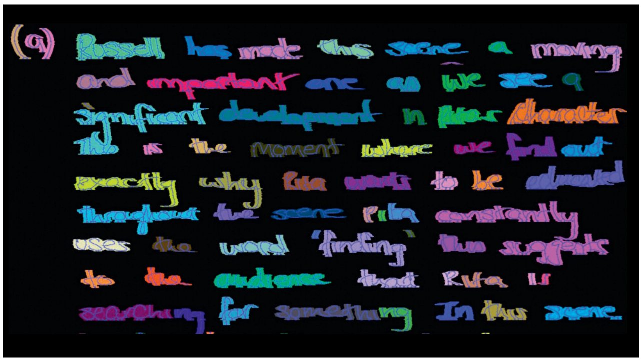

Figure 3: Thinned writing within post-dilation connected areas

Metrics

Once pre-processing was complete, a number of metrics were calculated for both Unit A and Unit B, for each page submitted by each candidate.

The first two metrics were computed prior to the thinning step of pre-processing as this saved considerable time in computation and it was of interest to discover whether we could get a reasonable indicator of candidate identity at this stage. Specifically we calculated:

1. **Median pixels per line (PPL) prior to thinning:** Specifically, we restricted analysis to the rows of the matrix (representing the image) where the variance of the values in the row was greater than the median. This was done to ensure that rows that were either almost completely blank or (possibly) almost completely black (such as might occur at the margins due to scanning) were removed. The columns of the matrix were restricted in exactly the same way. Then, the proportion of pixels in each row that were black was calculated and the median of this value across all relevant rows was taken.

2. **80th percentile of pixels per line (PPL) prior to thinning:** Similar to Metric 1 but with the 80th percentile of the row density stored rather than the median. The idea behind using this metric was to ensure that density was being calculated within parts of the text where a full line of writing had been completed thus excluding shortened lines that might occur at the beginning or end of a paragraph.

A further five metrics were calculated after all the pre-processing steps (including thinning) had been completed.

3. **Rough word count:** This simply counted the number of separate joined sections of writing identified after the dilation step. For example, this would count the number of separate sections identified in Figure 3 (but across a full page). Technically, this metric is not entirely related to the handwriting style. However, it was useful for identifying which pages contained sufficient writing for meaningful analysis as well as being an interesting variable for analysis in its own right.
4. **Writing density within words (sometimes labelled within this article as 'word density'):** This metric calculated the percentage of pixels within the segments identified in the image that contained the thinned version of the writing after all holes within the segment had been filled in. For example, within the pink area identifying the first word ('Russell') in Figure 3, this metric calculates the percentage of the pixels that are covered by the thinned version of the writing. This metric is designed to distinguish writing that is small and tightly packed for writing that is large and loopy. The median of this value was taken across all of the words on the page.
5. **Standard deviation of writing density within words:** Similar to Metric 4 but, rather than focussing on the median density across words, this metric calculates the extent to which the writing density varies across the words. This metric was intended to capture the consistency of handwriting within the page.
6. **Area of words:** This metric separately calculated the number of pixels covered by each of the sections of the image identified as words (including dilation and filled holes). The median of this metric was taken across all of the words within the page. This metric was designed to measure the size of a candidate's writing.
7. **Perimeter of words:** Similar to Metric 6, and again designed to measure the size of writing, but calculated via the perimeter of the identified sections rather than the area. Once again the median value of the perimeter was taken across all of the words on a page.

The above metrics were calculated for each page in a candidate's response. In order to compare metrics between Unit A and Unit B it was necessary to reduce the data to one observation per candidate (rather than per page). This was done by removing any pages where the rough word count was below 10 and then taking the median value of each metric across the remaining pages. The only exception to this procedure was for word count where it was of interest to take the total word count across all pages (including those with less than 10 words) rather than the median word count per page.

Due to the obviously close relationship between area and perimeter, it was decided at the end of the calculations to combine the two to create one final metric:

8. **Shape:** This was defined as (median) word perimeter squared divided by (median) word area. This metric is similar to *circularity* (also known as the *isoperimetric quotient*) which is an existing measure of the shape of an object. In theory, this metric will assign higher values to writing that is low and wide than to writing that is tall and square.

The effectiveness of these different metrics is evaluated in the next section. However, from the metric descriptions, it is immediately obvious that these were not the only set of metrics that could have been chosen. For example, why focus on the median metric across words or lines of an image? Why not calculate metrics relating directly to the height and width of words? Should the density of pixels within a word be calculated within a dilated version of this same text, or should it be calculated within a box defined by the top, bottom, leftmost and rightmost pixels? The decisions that were made in regard to these questions were somewhat arbitrary and fairly strongly influenced by the availability of existing functions within the *EBImage* package to perform each task. Further research could explore the effect of different choices. However, as we will see later, the metrics performed relatively well and give a reasonable idea of what can be achieved using simple metrics.

Computing speed

Despite the relative simplicity of the described metrics, processing each page from each script was still relatively slow – taking around 7.5 seconds. Thus processing 6 pages for Unit A and 14 pages for Unit B took around 2.5 minutes for each candidate. Given that more than 26,000 candidates took both exams, more than 1,100 hours of computing time were required to compute all of the metrics for all candidates. In real terms this was reduced considerably by using multiple machines and splitting the processing across multiple cores on each machine. Nonetheless, processing these images was slow, requiring an entire weekend to calculate all of the necessary metrics for all candidates on both examination papers.

Results from the trials

Before beginning the analysis, any candidates with highly unusual handwriting metrics were removed from the data. This was done as the aim of the analysis was to identify candidates where the style of handwriting changed between occasions – not to simply identify scripts with very unusual handwriting or features. For this reason, any scripts where any of the described metrics were more than four standard deviations above the mean on either Unit A or Unit B were excluded[2]. In particular, this process helped to remove atypical scripts where the response had been typed as well as other unusual cases, including one instance where the candidate had decided to draw a series of cartoon images (unrelated to the exam question) rather than write an essay. A total of 25,450 candidates were retained within the analysis.

Performance of metrics

The stability of each of the eight metrics between Unit A and Unit B is examined in Figure 4. As we can see, for each metric there is a clear positive correlation between the values calculated on each exam.

2. Since all of the metrics had a natural lower bound of zero, it was not necessary to exclude candidates with unusually low values.

The smallest correlation (0.46) is for the standard deviation of pixel density within words (Metric 5). However, all of the remaining seven metrics display a correlation greater than 0.8 between occasions, and four of them display correlations above 0.9. The highest correlation relates to median pixel density within words (Metric 4) which displays a correlation in excess of 0.95 between occasions. For comparison, the correlation between the marks awarded to candidates on Unit A and those achieved on Unit B was just 0.51. In other words, the metrics of handwriting developed in the previous section are far more stable between examinations than the performance of candidates. All three of the most successful metrics were calculated after thinning had been applied to the images. This suggests that this is a worthwhile step.

Correlations between the different measures indicated that, to a large extent, they provided separate pieces of information about candidates' writing style. As might be expected, given that both metrics relate to word size, perimeter and word area displayed a correlation in excess of 0.9. In addition, perhaps due to the way perimeter was used in its definition, perimeter and shape had a correlation of 0.76. Pixel density within words had a negative relationship with both word area (-0.66) and word perimeter (-0.56) – the slightly obvious point being that candidates with bigger writing will tend to leave more space within the words themselves. Aside from these obvious relationships, the correlations between the different metrics tended to be small, with the majority being below 0.2.

Initial analysis attempted to make use of all of the above metrics simultaneously in order to identify candidates with a large change in handwriting style[3]. However, manual inspection of script images from the 20 candidates showing the biggest overall change from Unit A to Unit B revealed some problems with this approach. Only eight of these scripts displayed clearly different handwriting between the two occasions. In some other cases the style of handwriting was inconsistent within examinations rather than between, and in other cases they appeared to have been identified as different for reasons other than a change of handwriting style. For example, in two cases the handwriting looked similar but it was likely that a major change in the type of pen used for writing led to a major change in values for Metrics 1 and 2 – underlining the importance of the thinning step. In another two cases, an extreme change in the length of the submission (i.e., the word count) appeared to be the main reason for the candidate being identified, rather than any obvious difference in the style of the handwriting.

As an alternative, a second, much simpler, approach was adopted. The best metric from Figure 1 (pixel density within words) was chosen. It should be noted that the mean absolute difference in this metric between occasions for any candidate was just 0.004. In contrast, the mean absolute difference between two randomly chosen candidates was

3. Linear discriminant analysis (https://en.wikipedia.org/wiki/Linear_discriminant_analysis) was used to combine the metrics.

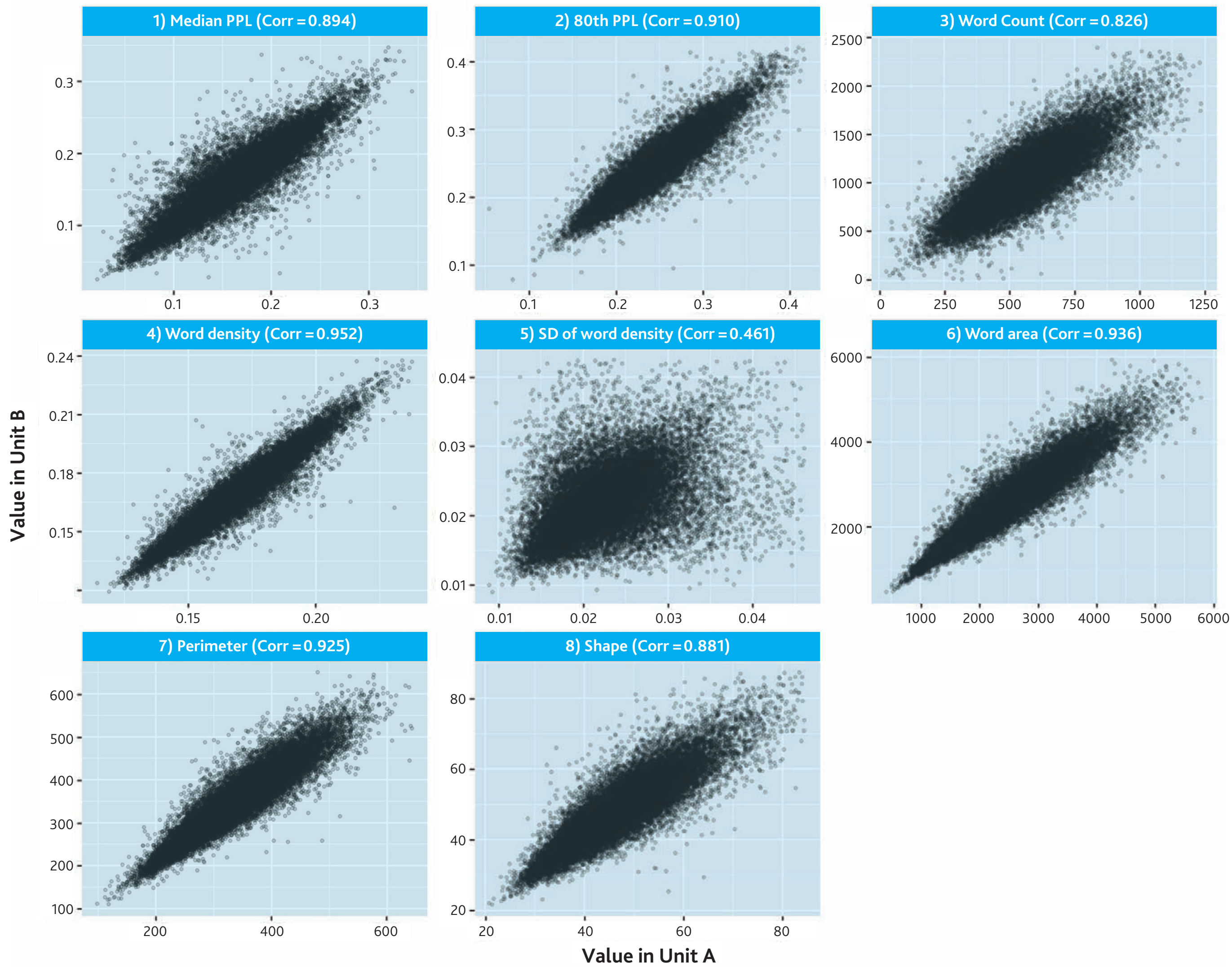

Figure 4: Relationship between each metric for Unit A and for Unit B

five times higher at 0.02. Only 213 candidates out of more than 25,000 with available data displayed such a large difference between occasions. This indicates that a focus upon this metric alone could yield interesting scripts for inspection.

Rather than examining all such instances, the 20 candidates' scripts showing the greatest change in this metric between Unit A and Unit B were inspected by eye. Fifteen of the scripts identified in this way had visibly different handwriting between the two occasions. Further details are given the next section.

Examples

Table 1 provides a list of the 20 candidates with the greatest changes in the chosen metric between occasions. The table begins with the candidate showing the greatest difference in pixel density within words between Unit A and Unit B and works through them in order, noting the qualitative impression of why the candidate has been identified, as well as the grades they achieved on each paper.

Table 1: Notes on 20 candidates with greatest difference in pixel density within words between Unit A and Unit B

Case No. (ranked starting from greatest difference)	*Possible reason why identified*	*Grade on Unit A*	*Grade on Unit B*
1	Visible difference in handwriting	C	B
2	Visible difference in handwriting	B	D
3	Visible difference in handwriting	A	D
4	Visible difference in handwriting	A	C
5	Visible difference in handwriting	B	D
6	Visible difference in handwriting	B	C
7	Visible difference in handwriting	B	C
8	Visible difference in handwriting	A	B
9	Visible difference in handwriting	A	A
10	Visible difference in handwriting	D	D
11	Visible difference in handwriting	B	E
12	Visible difference in handwriting	B	B
13	Visible difference in handwriting	U	E
14	Very little writing; Change of pen	D	U
15	Inconsistent handwriting	A*	A*
16	Inconsistent handwriting	B	D
17	Visible difference in handwriting	B	B
18	Visible difference in handwriting	B	A
19	Inconsistent handwriting	C	C
20	Not obvious why flagged	C	D

As noted above, in fifteen of the cases identified by this method the style of handwriting was visibly different between Unit A and Unit B. For example, Figure 5 compares part of the first page of writing on Unit A to part of the first page of writing on Unit B for the candidate with the largest change between occasions. As we can see, there is a marked difference in handwriting style. For Unit A, the handwriting is tidy, with curved characters and a uniform height. In contrasts in Unit B, the writing has a messy, uneven and angular style. The different styles shown in these small portions continued throughout the examination scripts. Nonetheless, having manually checked the names as well as the centre and candidate numbers entered on the front of both scripts, it is clear that both pieces of writing supposedly belong to the same candidate.

However, before leaping to the conclusion that one or other of these responses (or perhaps both) was provided by an imposter, there are

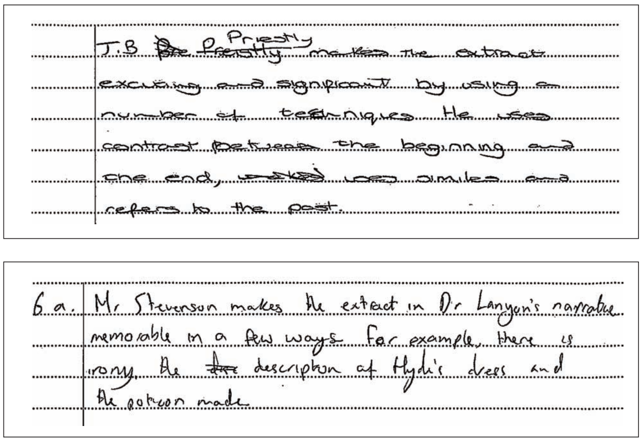

Figure 5: Portion of image of writing in Unit A (top) and Unit B (bottom) for the candidate with the greatest change in median pixel density per word

some other pieces of evidence to consider. To begin with it can be seen from Table 1 that, although the handwriting changed, the level of performance achieved was fairly similar on both examinations with a C grade awarded for Unit B and a B grade awarded for Unit A. This in itself suggests little motive for impersonation.

Secondly, examples of handwriting from other candidates examined as part of this research revealed cases potentially indicating that a single person might use very different handwriting styles in two examinations. An example (found in a separate analysis) is shown in Figure 6. Again, this example shows a marked change, from a large and looping style used in Unit A, to a small and neat style adopted in Unit B. However, the response to Unit B also displays another clear characteristic – the fairly large circles, almost like hollow umlauts, used to dot the 'i's. This same trait is also visible in Unit A. Given the unusual nature of this trait, it would appear at least possible that both sets of writing were produced by the same person. This suggests that we need to exercise some caution before concluding that a change in handwriting style indicates a change of author – an important fact when we consider Figure 5.

To emphasise this point further it is possible to find candidates where the style of handwriting changes *even within the same examination*.

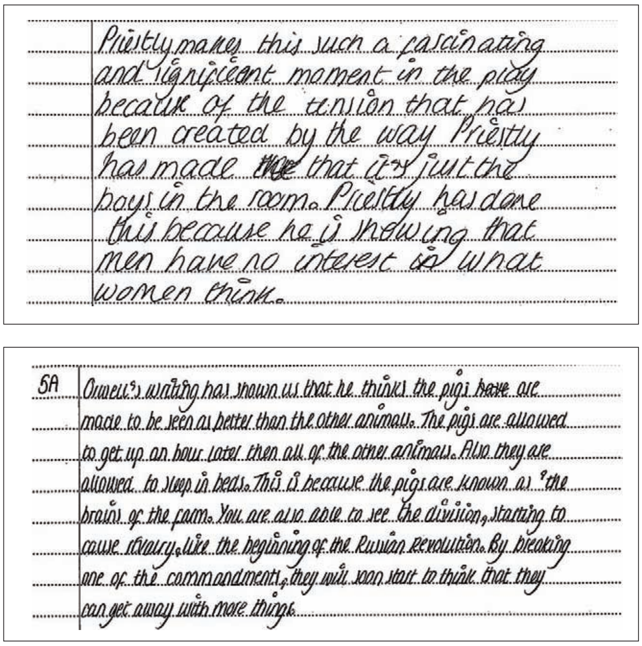

Figure 6: Example of a candidate with a consistent trait (the dots of the 'i's) but a different handwriting style (Unit A on top, Unit B on bottom)

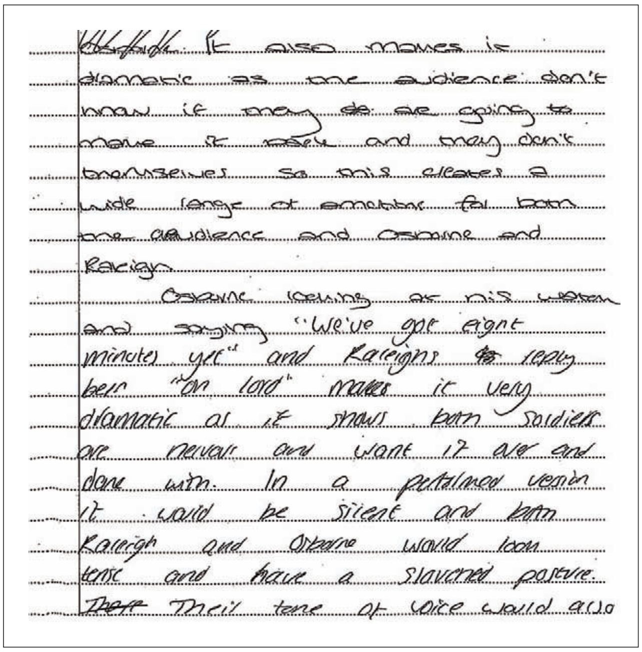

Figure 7: Example of candidate changing writing styles within the same page

This situation was evident in Cases 15, 16 and 19 in Table 1. Figure 7 shows a page from Case 19 that illustrates the issue most clearly. As we can see, the squat and curved lettering at the top of the figure gives way to taller and more angular writing at the bottom. We do not know what has caused the change although we might guess that time pressure or stress has led to a change of style. Indeed, one possibility that may be worth investigation is whether the writing at the end of this page was added by another author. However, this page does suggest a need for caution in interpreting the results. If handwriting can change even within a single page of writing, we cannot necessarily conclude that a change of handwriting style between Unit A and Unit B indicates any form of malpractice.

Returning to Table 1, it is notable that in nearly all cases the grade achieved on Unit A was similar to that awarded for Unit B. This fact, together with the examples discussed and the fact that this form of cheating is not widely reported in the UK, in any case suggest that something other than a change of the person taking the exam may explain most (and possibly all) of the cases identified in Table 1. It should be noted that other researchers in this area (see Dolega et al., 2008) have noted a "lack of stability of human handwriting" which fits with the results we see here.

Discussion

This article has proposed a number of metrics of handwriting style that are relatively easy to calculate. The majority of the suggested metrics displayed very high correlations between occasions suggesting that they may provide a reasonable indicator of whether the same candidate has indeed taken all of the relevant examinations leading to a qualification. The most effective metrics required thinning methods to be applied to the image as part of pre-processing indicating that, although computationally burdensome, this is a worthwhile step. The most effective metric (median pixel density within words) displayed a correlation in excess of 0.95 between separate examination occasions.

Out of more than 25,000 candidates taking both of the exams being studied, the metrics allowed us to quickly find a number of examples where a candidate's handwriting style changed between occasions. However, the fact that we were able to identify cases where the handwriting style had changed, but other aspects of the writing indicated the author may have been the same, suggests that a change of handwriting style in itself is not proof of malpractice. The same applies in cases where the style of handwriting changed within an individual exam.

Of course, in the UK school context of the scripts analysed in this article, the use of an imposter for one exam but not another is rarely reported as an issue. As such, any cases where handwriting is identified to have changed are perhaps more likely to be explained by other factors than by the presence of an imposter in one or more exams, and the automated methods of checking handwriting styles we propose here are unlikely to be useful. However, in other contexts, where we are more suspicious that an imposter may be used for one or more exams, the methods we suggest here may be helpful as they provide a relatively quick means by which candidates displaying inconsistent handwriting between exams can be identified. Thus, in contexts where we are more worried about this form of cheating, this may provide an efficient means of identifying the candidates worthy of further scrutiny.

On a more general level, this research has begun to develop our expertise in processing images from the DSR to procure useful information about candidates' responses. For example, one by-product of this research has been to calculate a rough word count for candidates' essays – a potentially interesting variable for further research. Further work could build upon this basis to explore further automated methods of collating information from candidates' script images for use in research.

References

Dolega, B., Agam, G., & Argamon, S. (2008). *Stroke frequency descriptors for handwriting-based writer identification. Proc. SPIE* 6815, Document Recognition and Retrieval XV, 68150I (January 28, 2008). doi:10.1117/12.767227

Lam, L., Lee, S. W., & Suen, C. Y. (1992). Thinning methodologies – A comprehensive survey. *IEEE Transactions on pattern analysis and machine intelligence*, *14*(9), 869–885. doi: 10.1109/34.161346

Linear discriminant analysis. (n.d.). In *Wikipedia*. Retrieved January 12, 2017, from https://en.wikipedia.org/wiki/Linear_discriminant_analysis

Otsu's method. (n.d.). In *Wikipedia*. Retrieved January 12, 2017, from https://en.wikipedia.org/wiki/Otsu's_method

Pau, G., Fuchs, F., Sklyar, O., Boutros, M., & Huber, W. (2010): EBImage – an R package for image processing with applications to cellular phenotypes. *Bioinformatics*, *26*(7), pp.979–981. doi:10.1093/bioinformatics/btq046

R Core Team. (2015). R: *A language and environment for statistical computing*. The R Foundation for Statistical Computing, Vienna, Austria. Available online at: https://www.R-project.org/.