THE PIXEL DIFFERENCE METHOD FOR COMPARING THE 2D SHAPE OF HANDAXES

Hannah Cutler

ABSTRACT

Palaeolithic handaxe shape variability is an important avenue of investigation into hominin cognition and behaviour. There is, however, a problem in that metrical methods often rely on multiple arbitrary dimensions or subjective visual typology to define shape similarity. With this in mind an entirely new method for the morphological analysis of handaxes is outlined. Comparing and grading handaxes based on the difference of 2D silhouettes’ pixel counts allows shape to be compared as a whole, without pre-judging the specific characteristics of 2D shape which make shapes similar or different to one another. This method allows a numerical value to be obtained that describes the relative difference in planform of two handaxes. This can be used in settings where numerically defined similarity or difference is preferable to subjective typology, notably inter- or intra-handaxe assemblage description and comparison. Also, it offers the ability to create definable parameters for the typological inclusion of individual handaxes into a sample, in the absence of dating context. This methodology would not have been easy a decade ago, but combinations of new software, much of which is freely available, has allowed this inventive technique to be devised.


Keywords: Palaeolithic, archaeology, handaxe, morphometrics, biface, image analysis.

INTRODUCTION

The shape of handaxes and bifacial tools has been central to discussion of hominin technology, time investment, mental templates, aesthetics and culture (Iovita & McPherron, 2011). Their persistence throughout long periods of the Pleistocene and over a wide region of the Old World means that they represent a key source of information about Palaeolithic hominins. That they were made by more than one species of hominin (Homo erectus and Neanderthals) also means that they have the potential to show technological and evolutionary change over time. As tools made by façonnage they are generally believed to represent deliberate choices by Palaeolithic humans in their production. From this, there has been much discussion as to the role of symmetry, reduction and raw material on the shape of handaxes (Noble & Davidson 1996; Ashton & McNabb 1994; McPherron 1995, 2000 & 2003; White 1998; Wynn 2002).

To better understand the variation in form of handaxes many methods have been devised and adapted over the years. Many of these involve the descriptive metrical measurements of pre-determined dimensions (Bordes, 1961; Roe 1964, 1967 & 1968; Crompton and Gowlett 1993; White 1996). These are contrasted by methods which use radial measurements from a central point (Wynn & Tierson 1990) and Elliptical Fourier Analysis (Iovita 2009; Iovita & McPherron 2011). Visual comparative methods (Wymer 1968 & 1985; McNabb et al. 2004) have also been used, but could be considered problematic due to the potential for variation between their applications by different researchers. More recently computing is allowing for innovation which is resulting in new and interesting methods designed specifically for the study of handaxe morphology; most notably, The FlipTest, a program by Hardaker & Dunn (2005) looking at bilateral symmetry. The use of new methods within pre-existing software is also of great benefit to the archaeological community making such methods widely accessible (Foulds 2013; McNabb 2013). This paper presents a new method for approaching the comparison of 2D handaxe shape. It aims to provide a single variable to describe shape similarity/difference as a whole without resorting to subjective typology or further assumptions about the importance of particular handaxe features.

THE PIXEL DIFFERENCE METHOD

The Pixel Difference Method (henceforth PDM) was initially devised for the purpose of studying British Late Middle Palaeolithic (Neanderthal) handaxes (c. MIS3...
To assign an undated handaxe to the British Late Middle Palaeolithic (henceforth BLMP) a suite of morphometric measurements was hoped to provide constraints derived from available BLMP-dated handaxes. However, this was considered unsatisfactory as a result of the Texas Sharpshooter fallacy. This is whereby, given enough variables, some will likely show a correlation and lead to the false conclusion that they are the specifics of biface form that are important.

A single method of defining similarity/difference in handaxe shape (purely in the sense of how alike is their shape) is required that can be expressed numerically to offer a measure of objectivity to counteract the difficulty in definition and implementation of typology of BLMP handaxes and the so-called bout coupé (for discussion of this issue see Roe 1981; Coulson, 1986; Tyldesley 1987; White & Jacobi 2002; Wragg Sykes 2010; Cutler 2013). This method was devised with the BLMP in mind, due to the high temporal resolution of BLMP handaxes (by Palaeolithic standards) and the high shape conformity between them. The LMP in general is a period where very distinctive and highly similar handaxe forms are present (Richter 2000 & 2006; Ruebens 2007; Soressi 2004; Ruebens 2006; Cutler 2013).

The PDM was influenced by the Flitpest (Hardaker & Dunn 2005). The Flitpest was designed to show numerically how symmetrical is a 2D image of a single handaxe; the lower the output number the more symmetrical the image. Similar to the Flitpest (though without the associated equation used within the software) the initial pixel difference output is obtained by calculating how many pixels two regions/shapes do not share within a raster image. Unlike the Flitpest, the PDM is designed to initially show similarity/difference between two shapes rather than comparing two halves to describe the symmetry of one. The method is as follows (see also Figure 1).

1. Using Inkscape (version 0.48, Windows 7) images of handaxe plans are isolated from their background and vector outlines are created.
The resulting silhouette of the handaxe is coloured black with an arbitrary medium transparency setting; in this case 116, out of 255.

The image is scaled to a predetermined maximum length in Inkscape and all other dimensions by the same proportion automatically. The author suggests 100 mm as an acceptable maximum length for this purpose. The primary concerns here are to keep file sizes small to avoid overly long processing times and to make sure the images are large enough so that pixels do not simplify the outline. Use of a single transformation within the imaging software also expedites the process.

This is repeated for the comparison handaxe.

The two images are overlaid onto one another using a central guide which divides the maximum width by half, not the maximum length bisector (this still informs the orientation of the handaxe and which face is used, after Wynn & Tierson (1990). Manual image creation avoids problems that could be caused by handaxes (and thus their source images) with different widths.

The transparency causes three colours to be present: the white background, a dark grey area where the handaxes overlap and a light grey area made from the areas of both handaxes that do not overlay one another.

The composite image is exported as a PNG file (using the default dpi of 90); this produces an image 354 pixels in height with width depending on the width of the image. To preserve the colour transparency this image is converted to a JPEG using batch PNG to JPG (version 1.0, Windows 7; available at Design-Lib.com). It produces an image at 96 dpi with no loss at the 100 (%) JPEG quality. A higher image quality could have been used, but to avoid human error and to maintain small file size, defaults are used. Also, at this level one pixel is as wide as 0.28% of the resized maximum length, beyond this is the scale at which there is colour variation caused by compression at the interface of two colours. Therefore any appreciable difference in shape overlay will be registered.

The image is opened in Fiji (ImageJ version 1.42l, Windows 7) (Schindelin et al. 2012), and the histogram function is performed, visualised as a list or a graph.

There are three peaks representing the three colours present. Using the transparency of 116 used here the dark grey is given the value 76, the light grey 139, and the white 255.
• The pixel count of the light grey (139) is the count of the number of pixels that the two overlaid handaxes differ from one another. There are small counts of other shades caused by JPEG compression artefacts, but they are negligible. This process can be sped up using a Macro, which is available from the author on request.

USES FOR PIXEL DIFFERENCE VALUES

Using the pixel difference value produced by the methodology outlined above, an objective measure of the difference between the planform of overlaid handaxes is possible. Following the method for the comparison of 2D shape, each handaxe will have a numerical measure that describes how many pixels they do not share when overlaid on one another. The raw results can then be presented in a pixel difference matrix (see Figure 2).

Three main uses for the pixel difference value are proposed:
• Means of minimum pixel difference.
• Agglomerative hierarchical cluster analysis.
• Quantifiable typological dating limits

The first two of these options are simple processes using the pixel difference value of a sample/assemblage; while the third is a more involved process. With all observations of patterning within handaxe assemblages one must consider the finished artefact fallacy as defined by Davidson & Noble (1993), notably the pitfalls associated with assuming causal relationships between visible patterns and prior intent. Having said this, certainly in the BLMP, pixel difference values between handaxes can be so low as to make this highly unlikely (Cutler 2013).

Means of minimum pixel difference

The arithmetic mean of the minimum pixel difference from each column of the similarity matrix suggests the average amount of the lowest difference between handaxes in the sample. Using the minimum values rather than the entire matrix of pixel differences allows similarity rather than difference to be analysed. The mean of the minimum pixel differences could be used as a measure of assemblage shape homogeneity in comparison with other assemblages. This could be purely descriptive in nature or involve tests of statistical significance. Assemblage shape homogeneity, whilst not offering indication of the planform type(s) within assemblages, has great potential to describe how assemblages vary in time and place. For example minimum pixel difference mean might indicate the degree of imposition of standardised form(s) (or at least standardised processes) in the knapping method; standard deviation would suggest the flexibility with which standardisation might have been applied and overall minimums (or their frequency in a pre-determined upper percentile) could show the extent of the highest level of knapping skill/imposition of form present.

Agglomerative hierarchical cluster analysis

The Pixel difference matrix can be input into MultiDendrograms (version 2.1, Windows 7) (see Fernández & Gómez (2008) for input formats and file types). The Unweighted Average or UPGMA (Unweighted Pair-Group Method using Averages) clustering algorithm is suggested as most applicable though others are available. This method forms clusters formed from pairs of previously formed clusters, based on the arithmetic mean distances between the individual members. Vertical distance in the resulting dendrogram shows the similarity between the clusters directly below. Use of the dendrogram in this manner is an effective method of visualising shape relationships, however it may be more applicable to handaxe assemblages, rather than to samples composed of individual isolated finds. The clustering software can be applied to multiple handaxes to identify or describe proportions of shape groups in the more traditional sense (ovate, cordate etc...). However such groups, if they exist, tend to create themselves. An output example from this technique is shown in Figure 3. Identifying handaxe shape proportions might offer insight into cultural associations, site function or other traditional question. Importantly it offers a method of visualising if there exists continuum of forms or more defined discrete forms. Also assessing the relationships of less clustered examples (or those further up the dendrogram) could offer insight into whether they are isolated examples, thus potentially describing
conceptual leeway of the imposition of pre-conceived forms, or skill of the knappers.

Quantifiable typological dating limits

Using pixel difference values to formulate acceptable 2D shape limits to ascribe an undated isolated handaxe to a particular period is more complex. This of course rests on 2D shape being considered an important enough, or possibly the most important feature of a handaxe, over technological concerns as the sole sample refinement method. This is an acceptable compromise over using multiple morphometric variables.

Which handaxes to compare?

The starting point must be dated handaxes. Using dated handaxes which have a high degree of relative 2D shape conformity allows patterns to be tested against other items in a numerically gradable fashion. Before this can be tested the dated sample itself needs refining in several ways. Each handaxe needs to:

- Be dated to the correct period as defined by the investigation.
- Not be made on a nodular blank or on a flake with a majority of cortical edge. This is an unknown factor that may or may not have influenced shape in individual cases. This means is that the remaining sample can be said to consist of items that are more likely, though to an unknown degree to result only from a maker’s intentions before, or during production, be they related to a particular form or the processes that lead to it. It ignores the use, or accommodation of existing form, simply because it is too difficult to quantify. Therefore any identified consistencies (see below) will be less likely to be as a result of chance.
- Without shape altering breaks, noticeably on the tip which would alter resizing. This is most obvious with difference in staining or patination of broken sections. Sometimes silhouettes can be digitally repaired; i.e. have part of its circumference extrapolated at the point of a break in vector graphic software, but this would be an individual choice on a case-by-case basis. It would also have to be explicitly stated in each case, to make the effects of the interpretation clear.

Comparison of 2D images of Dated Handaxes

With pixel difference values for each dated handaxe, limits must be imposed to apply the pixel test to undated individual artefacts. What requires defining are:

- What are the acceptable limits of similarity that one can say an undated handaxe is most likely to be of correct age?
- Which shape or shapes should be used for such a test/comparison?

While similarity is an abstract concept, limits applied to any undated sample must be based in reality. What pixel difference level is sensible as a threshold can be worked out from the mean minimum pixel difference. The upper 95% limit of the population mean is suggested as the most useful broad similarity threshold, while any value below is likely to be equal to or less than the average minimum pixel difference. An upper limit of the mean is preferable to a median value, which would exclude moderately higher values and negate the only broadly quantifiable concept of similarity.

It is desirable to pick a small number of different dated shapes to test against (referred to here as testing shapes), rather than test all...
dated handaxes against all undated forms. This is because the aim is to quantify similarity or difference from a typological norm/extreme form, to decide on an items inclusion in a sample and possible variation in space (such as when plotting find spots using GIS). This is better served by having few comparable values. It is necessary to know whether an undated handaxe is likely to be of correct age, not whether it is within the broadest limits of shape variation. Broader limits, if required, can be better controlled through the flexible application of the pixel difference threshold, not in testing against multiple shapes. Therefore, for any example to be used for typological dating, it would have to be shown that it was not a freak individual. It is entirely plausible that shape similarity (defined with reference to well dated examples) exists for one or many reasons. These can include the same maker, cultural group, and intentional shaping or use/reduction pathways. However these concepts are not relevant at this stage. What needs to be demonstrated is that, for whatever reason, a particular shape appeared more than once.

A sample of dated examples can be reduced to only include individuals with at least one pixel difference of ≤ the upper 95%, but this limit will not provide enough of a refinement to observe the most extreme examples (see Wragg Sykes (2010) for a discussion of hyper shapes in the context of the Late Middle Palaeolithic). The problem remains of defining similarity, this time at the highest level. Deciding on testing shapes is to some extent an intuitive process relying on not one, but several methods to determine consistent shapes within a dated handaxe sample. Using a clustering dendrogram, an ideal testing shape will be:

- One of a pair at the lowest level (highest similarity) of a hierarchical arrangement of values of pixel differences.
- Within a tight cluster (relative to other clusters) below the pixel difference threshold.

Other shapes can be chosen due to the frequency with which they are the most similar shape (minimum pixel difference) to other dated shapes. Deciding on which member of a pair is used can depend on this or using the lowest symmetry as defined by the Fliptest as a deciding factor. The number of shapes chosen to test against will likely depend on how many consistent shapes present themselves or are desired and other such concerns. Once testing shapes are decided upon they can be compared to undated examples using the pixel difference test. Ideally, those below the pixel difference threshold to at least one testing shape would be accepted as being above the average handaxe similarity of the dated examples, and thus acceptable for inclusion into a sample for further analysis.

The PDM provides no comment on the applicability of typological dating as a general concept within the Palaeolithic, but simply how it may be undertaken in a quantifiable manner using pixel difference values. It maybe be suggested at certain times and places but not in others. For example it has been shown to be very effective in the case of the BLMP. Very high handaxe similarity and small temporal windows for occupation prove that handaxe shape is acceptable for typological dating of handaxes at least once in the Palaeolithic (Cutler 2013). The existence of dated consistent forms at the lowest pixel difference, if identified, legitimises the extrapolation out to undated examples by pixel difference comparison with testing shapes. It will perhaps be overly strict in that it will only accept handaxes at a broad level to the testing shape(s), so that even if a handaxe does not fall below the pixel difference threshold to one testing shape, it may still be of correct age, but those that do will be more likely to be of correct age within limits set. Having said this, depending on the nature of the material it may be acceptable to date items/sites/occupations by association, but this would be a choice of a particular implementation rather than a definite application of the method. Using the PDM, typological dating, based on 2D form as a binary question, is replaced with a quantifiable description of typological assumption. The age of Palaeolithic material will always make interpretation more difficult than archaeology from recent periods. Therefore an approach which is flexible but explicit in its compromises is preferable where the only alternative would be the exclusion of all but the most unproblematic data sources.
ISSUES

One major problem with the PDM is that it relies on input handaxes being scaled to an equal height. This means that it will not identify similar artefacts where one area (namely the tip) has been worked or is different but elsewhere they are identical; this is also true of asymmetrically reduced handaxes, unless asymmetry itself common. This is preferable to using equal widths or only part of an item as neither could offer across-sample comparable pixel difference values, as they would be restricted to being interpreted one item at a time. Having said this, equal width may be applicable in a non-quantifiable fashion when looking at tip retouch where size may be a factor in handaxe shape. If shape is important enough to show consistencies within a particular period, resizing also attempts to account for the lack of accurate measurement-use by Palaeolithic hominins. Indeed, knowing where to superimpose artefacts would be a problem for pixel-differences tests as shown in Figure 4. If portions of differently sized (real-life sized) objects appear to fit together when overlaid, knowing whether it is a function of wishful thinking is impossible. Therefore a consistent method for use over many items is also impossible as one would be reduced to aligning comparisons by eye which is exactly what this method is trying to avoid by quantifying 2D similarity. Looking at technological concerns may give insight, but would have to be done on a case by case basis, undermining the strengths of this procedure which are that it is quantifiable and its relative speed. These points cannot be understood with enough accuracy to influence methodology, introducing unknown error rather than uniform methodological caveats. With this in mind, handaxe shapes must be treated as they are, rather than what they might have been, prior to any reduction which may have occurred.

CONCLUSION

This paper has presented an entirely new approach to the study of handaxe morphology. It is hoped that the PDM will provide a valuable addition to the suite of techniques available for the study of handaxe morphology. Considering an items shape as a whole, avoids the subjectivity of qualitative descriptions of shape or portions of shape, or the potential for false correlations of multiple quantitative variables. By conceiving of shape purely in terms of entire similarity or difference to another object, the problematic issues of arbitrary typological separation of forms can be avoided. If 2D shape as a whole is an important feature of handaxe morphology, it can be studied purely in relative terms without having to specifically define what shape is, as we cannot say of how it was conceived in the Palaeolithic. While the author’s focus has been the Late Middle Palaeolithic, the PDM has applications for a wider range of bifacial lithic artefacts, and potentially to any artefact type where 2D shape is, or maybe important, perhaps even outside the Palaeolithic. Using the pixel difference values within clustering dendrogram is an excellent objective method for typological grouping of objects and visualisation of variation. The outlined method for placing handaxes within quantifiable typological dating limits, at best offers a way of extrapolating dating information from items
from dated contexts to those where this information is lacking. At the very least it will allow more detailed future criticism and testing of conclusions with new discoveries or better dated material. Finally it must be mentioned that Inkscape, FIJI and MultiDendrograms are available as open source software. This greatly increases the worth of this method opening it up to a wider research community.

ACKNOWLEDGEMENTS
This research was undertaken as part of a PhD at the University of Cambridge funded by a Domestic Research Scholarship and by Corpus Christi College. The author would like to thank Paul Mellars, Frederick Foulds, Claire Ross and Matthew Bailey for their feedback on earlier drafts of this paper and to three anonymous reviewers who suggested improvements to this paper.

REFERENCES

Bordes, F. 1961. Typologie du Paléolithique Ancien et Moyen, 1, Bordeaux, Mémoires de l'Institut Préhistoriques de l'Université de Bordeaux.


