Schriften des Instituts für Dokumentologie und Editorik — Band 2

Kodikologie und Paläographie im digitalen Zeitalter

Codicology and Palaeography in the Digital Age

herausgegeben von | edited by

Malte Rehbein, Patrick Sahle, Torsten Schaßan

unter Mitarbeit von | in collaboration with

Bernhard Assmann, Franz Fischer, Christiane Fritze

2009

BoD, Norderstedt

Bibliografische Information der Deutschen Nationalbibliothek:
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über **http://dnb.d-nb.de/** abrufbar.

Leicht veränderte Fassung für die digitale Publikation (siehe Vorwort).

Slightly modified version to be published digitally (see preface).

Publication réalisée avec le soutien d'Apices
Association Paléographique Internationale
Culture – Écriture – Société
Dotation J.M.M. Hermans.
http://www.palaeographia.org/apices/

Herstellung und Verlag: Books on Demand GmbH, Norderstedt
ISBN: 978-3-8370-9842-6
Einbandgestaltung: Katharina Weber
Satz: XeTeX und Bernhard Assmann

Computer-Aided Palaeography, Present and Future*

Peter A. Stokes

Abstract

The field of digital palaeography has received increasing attention in recent years, partly because palaeographers often seem subjective in their views and do not or cannot articulate their reasoning, thereby creating a field of authorities whose opinions are closed to debate. One response to this is to make palaeographical arguments more quantitative, although this approach is by no means accepted by the wider humanities community, with some arguing that handwriting is inherently unquantifiable. This paper therefore asks how palaeographical method might be made more objective and therefore more widely accepted by non-palaeographers while still answering critics within the field. Previous suggestions for objective methods before computing are considered first, and some of their shortcomings are discussed. Similar discussion in forensic document analysis is then introduced and is found relevant to palaeography, though with some reservations. New techniques of "digital" palaeography are then introduced; these have proven successful in forensic analysis and are becoming increasingly accepted there, but they have not yet found acceptance in the humanities communities. The reasons why are discussed, and some suggestions are made for how the software might be designed differently to achieve greater acceptance. Finally, a prototype framework is introduced which is designed to provide a common basis for experiments in "digital" palaeography, ideally enabling scholars to exchange quantitative data about scribal hands, exchange processes for generating this data, articulate both the results themselves and the processes used to produce them, and therefore to ground their arguments more firmly and perhaps find greater acceptance.

Zusammenfassung

Das Forschungsfeld der »Digitalen Paläographie« hat in den letzten Jahren verstärkte Aufmerksamkeit erfahren; zum Teil weil die Paläographen in ihren Urteilen subjektiv zu sein scheinen oder weil sie ihre Argumentation nicht offen legen (können), so dass eine Gruppe von Autoritäten entstanden ist, deren Meinungen außerhalb der Diskussion stehen. Eine Antwort auf diese Situation ist der Versuch, die paläographischen Argumente quantitativer zu machen. Dieser Zugang wird jedoch durch die Mehrheit der

* I wish to thank the Leverhulme Trust and the Isaac Newton Trust for their financial support, without which this research would not have been possible.

Kodikologie und Paläographie im Digitalen Zeitalter – Codicology and Palaeography in the Digital Age. Hrsg. Malte Rehbein, Patrick Sahle und Torsten Schaßan, unter Mitarbeit von Bernhard Assmann, Franz Fischer und Christiane Fritze. Schriften des Instituts für Dokumentologie und Editorik 2. Norderstedt: Books on Demand, 2009. 309–338.

Fachgemeinschaft nicht akzeptiert, die unter anderem argumentiert, dass Handschrift per se nicht quantifizierbar sei. Der Beitrag untersucht deshalb, wie man paläographische Methoden objektiver und so somit auch von Nicht-Paläographen leichter akzeptierbar gestalten kann, ohne dabei eine fachliche Diskussion auszuschließen. Zunächst werden ältere Vorschläge für objektive Methoden aus der Zeit vor dem Computer und einige ihrer Defizite diskutiert. Im Anschluss daran wird die forensische Schriftanalyse diskutiert, wobei der Beitrag zu dem Ergebnis kommt, dass ihre Methoden (mit Einschränkungen) auch für paläographische Forschung nutzbar sind. Schließlich werden die neuen Techniken einer »Digitalen Paläographie« vorgestellt. Sie haben sich in der forensischen Schriftanalyse als erfolgreich erwiesen und setzen sich dort immer mehr durch, während sie in den Geisteswissenschaften noch nicht die gleiche Akzeptanz gefunden haben. Der Beitrag diskutiert die Gründe hierfür und macht einige Vorschläge, wie die Software verändert werden müsste, um auch hier größere Akzeptanz zu finden. Abschließend wird ein Prototyp eines Frameworks vorgestellt, der eine gemeinsame Basis für Experimente in »Digitaler Paläographie« bereitstellen soll; idealerweise indem er Forschern die Möglichkeit gibt, quantitative Daten über Schreiberhände auszutauschen, Austauschprozesse für die Erzeugung dieser Daten organisiert, Ergebnisse und die ihnen zu Grunde liegenden Verfahren darstellt und so die paläographischen Urteile besser nachvollziehbar macht und ihnen vielleicht eine größere Akzeptanz verschafft.

1 The Problem: How Many Scribes?

Three questions are commonly asked of palaeographers about medieval handwriting: "when was this written", "where was this written", and "were these different things written by the same person." All three questions are extremely important when working with manuscripts, and all three require slightly different approaches and methods to answer, but the focus of this discussion will be on the third. Many research projects depend more or less explicitly on scribal identification. For example, three major projects are running now on glossed manuscripts from Anglo-Saxon England or early medieval Ireland: the Boethius Commentaries project led by Malcolm Godden, the Irish Glosses project lead by Paul Russell, and the Manchester database of scribes and spellings and its successor, lead by Don Scragg, and all of these depend to some extent on identifying glossing hands (Scragg et al.; Russell et al.; Godden et al.). However, glosses appear in a wide range of scripts and hands, and recognising which glosses were by the same scribe is important but extremely difficult. Even more challenging is identifying where main text and glosses were written by the same scribe or scribes, since glosses so often show different letter-forms, proportions, and aspect because of the different circumstances in which they were written (Stokes 2004; Stokes 2005 Ch. 5). Another example, this one of fundamental importance to late Anglo-Saxon history, is Cotton Tiberius B.iv, the "D" version of the Anglo-Saxon Chronicle. This manuscript contains yearly entries of great historical interest and was written some time between the mid-eleventh century and the early twelfth, probably by a number of scribes. Our understanding of the manuscript and the accuracy of the text depends crucially on whether scribes wrote up the annals year-by-year as they happened or in one block some time after the fact, and to decide this depends utterly on our knowing precisely when the scribes changed. However, at least six different opinions on this subject can be found in print, and there is little if any agreement which is correct (Plummer; Keller; Howorth; Ker; Cubbin; Guimon). These examples are drawn from late Anglo-Saxon England, but the same questions arise again and again in other contexts, and much of medieval studies depends on these questions, whether in the context of history, literature, language, and so on. However, it is difficult at best to answer these questions, and the answer is very often uncertain, something that is not always admitted. A survey of the literature often reveals widely varying opinions even among palaeographers, and that is if the palaeographers even give a firm opinion: in many cases they produce frustratingly vague statements instead, particularly with difficult cases like the D-chronicle mentioned above, or glossed manuscripts, regarding one of which I myself once wrote that "[a]lthough ten or more scribes may have written vernacular glosses in this manuscript, most of the glosses can probably be attributed to just two or three" (Stokes 2005 2:323; compare also Ker). Bernhard Bischoff wrote that "[t]he definite establishment of the identity of medieval scripts in more than one manuscript, or even the establishment of several different hands within a single codex, can [...] be made rather difficult" (1990 44–5). Colleen Sirat took this further when she wrote that "[i]t is obvious that one cannot prove that two texts were penned by the same hand. The only way to persuade other people that this is so is to show them, to give them the feeling that it is the same hand" (2006 493). If Sirat is right then we as medievalists have a problem, since this implies that much of our work depends on nothing more certain than our feelings, and one may well ask how such a discipline can claim to be academic. But it remains to be asked if she really is right, and, if so, what can (or should) be done about it.

2 The Need for Quantitative Palaeography

The issue has been discussed actively since the 70s, if not before, the problem being that palaeographers tend to express qualitative opinions rather than objective arguments and to issue pronouncements that cannot be debated or engaged with meaningfully.[1] To some extent this is necessary, as very little can be argued with certainty in the hu-

[1] At the risk of unfairly highlighting only a few of the very many examples, see the commentaries of the Oxford Palaeographical Handbooks (Bishop 1961; Bishop 1971; Wright) and others such as Bischoff.

manities and particularly in early medieval studies. Furthermore, no argument is ever truly objective insofar as it must depend on assumptions about the data and the interpretation of evidence (Sculley and Pasanek). Nevertheless, palaeography has perhaps been accused of vagueness and subjectivity more often than most other disciplines, and the accusations often seem justified. This is not to doubt the opinions of highly skilled and experienced experts, but rather, as Albert Derolez has noted, that "the method applied hitherto in palaeographical handbooks has produced an authoritarian discipline, the pertinence of which depends on the authority of the author and the faith of the reader" (9). Although in a different context, David Ganz has similarly argued that "the evidence for dating manuscripts must be explained, so that we can learn how a problem may be defined and resolved" (18). However, there remains a methodological problem of how to articulate palaeographical arguments in objective ways; in Derolez's words again, "[h]ow is it possible to proceed in such a way that the description of a specimen of handwriting is as clear and convincing to its reader as it is to its author?" (7) This is not trivial, part of the problem being that subjective impressions are inherently difficult to communicate and cannot be engaged with effectively. But Derolez has proposed a more specific method, namely that "by replacing qualitative data by quantitative ones [...] there is very much to be said in favour of a quantitative approach to a matter so difficult to treat adequately with other techniques" (7–8). Derolez was not the first to suggest this and there had already been significant debate before he wrote, with some scholars protesting that handwriting is inherently fluid and "human" and therefore cannot ever be quantified (Costamagna et al.; Gumbert 1998; Pratesi). Although this possibility must be acknowledged, it has not yet been demonstrated, and indeed those studies which have taken place suggest on the contrary that modern handwriting can indeed be quantified and measured with some significant success (Section 3.3). It therefore seems reasonable to ask whether the same applies to medieval handwriting and, if so, then how digital tools can help.

3 How to Go About It? Theoretical Overview

What is needed, then, is an objective, quantitative method of representing and describing handwriting, of analysing the similarities and differences between scribal hands, and of judging (or at least arguing) whether or not two or more stints of handwriting were written by the same scribe. But is this even possible? As noted in the previous section, scholars have been arguing about this for years, and various methodologies have been proposed. Before turning to the possibilities that computers bring, it is worth first surveying some of the methods that have already been tried, to see how these lessons might be brought into the so-called "digital age."

3.1 Objective Criteria in Palaeography

A useful starting-point for this is a dissertation on the "scribal fingerprint" of Cristine de Pizan which includes detailed discussion of several scholarly criteria for the identification of scribal hands (Aussems). One of the earliest listed there is that of Jean Mallon who proposed seven points for analysis (22–3; Aussems 53):[2]

1. Form, "the morphology of the letters."
2. Pen angle (*l'angle d'écriture*) "in relation to the base line."
3. Ductus, "the sequence and direction of a letter's different traces."
4. Modulus, the proportions of the letters.
5. Weight, "the difference in thickness between the hair lines and the shadow lines."
6. Writing support.
7. Internal characteristics, "the nature of the text."

Other lists since then are similar but tend to specify more criteria and to demand increasing levels of detail. One example is that proposed by Jan Burgers (1:501; Aussems 56–67):

1. Slant.
2. Writing angle.
3. Weight.
4. Modulus.
5. Format.
6. Width of the margins.
7. Ruling and irregularities of the base line.
8. Flourishes and other decoration.
9. "Text structure", punctuation and use of majuscules and capitals.
10. Abbreviations.
11. Cursiveness between letters.
12. Cursiveness within letters.
13. Characteristic letter forms.

Aussems himself essentially followed this list, although he omitted numbers 5, 6, and 9 as irrelevant for his case, and also number 10 for reasons that will be discussed shortly (70–78). A similar list has been proposed by Michelle Brown in her study of the Book of Cerne (25–26):

1. Aspect, "the overall appearance of the script."
2. Ductus.

---

[2] Aussems has also referred to the criteria established by Lothar Michel but these are for modern handwriting and require measurements such as the speed and pressure of the pen which are not applicable to medieval manuscripts. See Aussems 54–6, citing Michel 237–61; see also Aussems and Brink in this volume.

3. Pen angle.
4. Weight.
5. Letter-forms.
6. Mannerisms.
7. Orthography.
8. Abbreviations.
9. Punctuation.
10. Textual Apparatus, namely “devices [...] to assist layout and facilitate reading.”

Alexander Rumble has also printed guidelines for distinguishing scribal hands, “the key to [which ...] is the accurate description of the hands involved” (1994 13). His suggested features for description are (13–15):

1. Treatment of ascenders (including proportions).
2. Treatment of descenders (including proportions).
3. Minims (including weight and shading).
4. Letter-forms.
5. Ligatures.
6. Abbreviations.
7. Punctuation.
8. Orthography.

Finally, a relatively early list was provided by a palaeographer but in the context of detecting modern forgeries, but much of it applies equally well to medieval writing (Brown 1993 259–60):

1. Aspect, including shakiness, layout on the page, beauty, clarity, and tidiness.
2. Spacing.
3. Writing angle.
4. Treatment of loops.
5. Modulus.
6. Punctuation.
7. Formation of common words and syllables.
8. Common groups of letters which may differ according to their position in the word.
9. Ligatures between letters and between words.
10. Dotting of **i** and crossing of **t**.
11. Figures, capital letters, and “odd” letters such as **k**, **g**, **x** and **z**.
12. Pairs of similar letters, such as **n** and **u**.
13. Letters which often have more than one form, such as **d** and **e**.
14. Inconsistencies.

This survey of criteria is far from complete, but already several common elements can be seen in all of them. However, even these apparently scientific criteria do not produce a purely objective analysis, not least because they still include qualitative terms such as "aspect", but also because even the more quantitative criteria are often imprecisely defined and impractical to measure. The criteria involve sitting in libraries and taking very many measurements from potentially hundreds of scribes and thousands of letters; they are therefore impractical, and the best-known study to use such an approach was also methodologically flawed (Gilissen; Derolez 8). But more fundamental problems remain. How does one accurately measure pen-angle, for example, particularly as some scribes deliberately altered the angle of the pen when writing? What does "writing angle" really mean, and which strokes should be measured when determining it, particularly if a scribe wrote a very round or inconsistent hand (see further Maarse)? Some scholars have focused on just a few forms or abbreviations which they considered distinctive, an approach which derives ultimately from that pioneered by Ludwig Traube (1907; Brown 1959 363).[3] However, as Michelle Brown and others have observed, there are difficulties with interpreting even these results. How much does a scribe reproduce the punctuation or abbreviations of his or her exemplar? What about the orthography? Or even the distinctive letter-forms? As Bishop noted in 1961, "the more distinctive [the feature], the more easily imitated" (7–8), and letter-forms, mannerisms and abbreviations are the easiest of the imitable. Scribes certainly imitated script other than their own, such as twelfth or thirteenth-century English scribes consciously imitating (and forging) Anglo-Saxon script, and fifteenth-century scribes in England imitating twelfth-century script (Bishop 1961 7–9; Crick; Parkes 2008 142–4). These imitations are usually obvious, but the adoption of just one or two forms is much harder to detect.

Apart from obvious imitation, this problem of how much scribes were incidentally influenced by their exemplars has been raised many times but answered very rarely. One important response is a series of articles by Angus McIntosh, Michael Benskin and Meg Laing who collectively produced a typology of scribal copying with respect to Middle English dialect. They outlined some ways in which scribal interventions in a text can be identified and suggested ways to undo these interventions and recover the forms in the exemplar (Benskin and Laing; McIntosh et al. 1986 1:12–23; McIntosh 1989a; McIntosh 1989b). They have proposed three categories of scribal copying, "*litteratim*" in which the precise orthography of the exemplar is preserved, "translation" where the orthography is altered to match the scribe's own practice, and a mixture of the two. There are also different sub-categories, such as a scribe who began copying *litteratim* but then lapsed into "translation", or "constrained" scribes who generally follow the exemplar but sometimes give their own orthography instead. Perhaps most

[3] Examples include Gumbert 1976; Muir; Davis 1998; Beneš; and McGillivray; compare also the lists of features presented by Brown 1996 52–60; Brown 1993 259–60; and Rumble 1994 14. A similar philosophy has also been taken by Scragg et al., for discussion of which see especially Rumble 2005 221–5, and Rumble 2006 14–16.

significantly, though, they have found that the vast majority of scribes in the Middle English period fall into the category of "translation"; that is, they tend to adapt the spelling of their exemplars to match their own practices.

An approach like this seems very relevant to palaeographical analysis as well, and some possible approaches have already been suggested, particularly the distinction between a scribe's "graphic" and "linguistic" profile, namely, the handwriting on one hand and the orthography and punctuation on the other (Parkes 1994, esp. p. 24, citing McIntosh 1974 and McIntosh 1975). Specifically, we may ask whether scribal practices in script, punctuation and capitalisation follow the same patterns as in orthography and dialect, and, perhaps more interestingly, whether the methods proposed to recover the dialect-forms of an exemplar can be used to recover the letter-forms as well. Certainly we can find examples of both "*litteratim*" and "translation" in letter-forms. The latter is the norm, insofar as most scribes copied with their own natural script, but we have already seen examples where this was not so. There are also examples of scribes starting with one script and lapsing into another, such as Cambridge, University Library MS Ff.1.23, the so-called "Winchcombe" or "Canterbury" Psalter which was written in the first quarter of the eleventh century, probably at Canterbury (Dumville 1991 40–41; Stokes 2005 1:41–42). This scribe began writing the Latin text of the psalms in a careful Anglo-Caroline script but introduced more and more vernacular letter-forms, apparently by mistake, and after about a page or so gave up entirely and wrote the Latin and Old English in the same English Vernacular minuscule (Stokes 2005 1:68, reproduced by Robinson 2:pl. 17). This example probably says less about the exemplar and more instead about the new requirement to differentiate between languages by script (for which see Bishop 1971, Dumville 1993, and Stokes 2005), and this is one point where orthography and script diverge.[4] Nevertheless script was certainly taught, and there seems often to have been a strong sense that certain manuscripts or even texts should be written in certain scripts (Brown 1993 201–2; Lieftinck 1964 1: xiii–xvii). Furthermore, even vernacular orthography may have been standardised in some places even in the early Middle Ages, such as in England from the late-tenth through to the early-twelfth centuries where it may have been more consistent than the vernacular script was (Gneuss esp. 70; Gretsch 69–83). Similarly, the apparent readiness of scribes to write very different scripts alongside each other in the early eleventh century in England, as demonstrated by many vernacular manuscripts such as the so-called "Beowulf manuscript", also suggest that the pressure to write a particular style of script may not have been as strong as we might like to think (Stokes 2005; reproduced by Zupitza and Kiernan). Of course to address this question properly requires very much more research, and this is precisely something that databases of scripts and spellings and its

[4] Compare Bishop: "The difference in aspect between the Latin and the vernacular script need hardly be considered [and was at] no time so marked as to disguise a scribe's identity" (1961 4).

successors should help us answer (Scragg et al.), but the point remains that even the apparently objective criteria given above still require a good deal of interpretation.

There is a further difficulty with the different approaches listed here: although similar, there are important differences between them, and yet there is no clear way of testing them or deciding which should be used, or indeed having any meaningful way of knowing which, if any, lead to valid results. However much one might argue that "all knowledge is situated and contingent" (refuted by Drout §§10–12 and Shippey §26) the fact remains that a surviving manuscript was once written by one or more people at one or more place(s) and time(s) in history, and in this sense handwriting identification has a "correct" answer that palaeographers seek, whether or not it can be attained in practice. However much one might discuss individual strengths and weaknesses in each of the methods discussed above, it is ultimately difficult or impossible to know which is most "correct". Furthermore, the very question of scribal identity depends in turn on the assumption that the handwriting of no two people is the same, and yet this assumption is not normally questioned by palaeographers. To some extent this uncertainty is inevitable, and we should neither demand nor expect that palaeographical results will always be certain. Nevertheless, it is not ideal to have an entire discipline the validity of which has been assumed but not firmly demonstrated.

3.2 Objective Methods in Forensic Document Analysis

Fortunately most of these difficulties have already been raised in forensic document analysis.[5] Forensic document analysts have been critisized for some of the same shortcomings as palaeographers, including the inability to verbalise their methods and the variation in their results, and these uncertainties have even reached national headlines in the United States (Kam Wetstein and Conn 6–7 and 12; Liptak). When forensic document analysts do articulate their methods, furthermore, they seem to follow principles much like those of palaeographers, referring to features such as aspect, slant, writing-angle, shading, cursiveness between and within letters, characteristic letter-forms, and particular idiosyncracies (Kam Wetstein and Conn 12). However, document analysts must withstand cross-examination in court, and so they are forced to provide clear objective arguments. Furthermore, the United States Supreme Court has a recent guiding principle that judges must evaluate expert testimony for factors such as whether or not the method has been tested, what the potential rate of error is, and how reliable (and therefore reproducible) the results are (Srihari et al. 2002 856–7). This guiding principle has twice brought to bear on recent cases involving document analysis, as a result of which the Supreme Court was asked to rule on the objectivity and validity of handwriting identification (US v. Prime; US v. Thornton). New York state justices

[5] Forensic document analysts (FDAs) are also known as questioned document analysts (QDAs), or forensic or questioned document examiners (FDEs or QDEs).

also commissioned a study to determine if the identification of handwriting is objective and if more objective methods are possible, and it is significant for our discussion that "objective" was understood to mean "automatic" and determined entirely by computer (Srihari 2001). As a result, the State University of New York has a very active centre for the study of individuality in modern handwriting, and they have been studying questions of importance to forensic document analysts and palaeographers alike, questions such as whether handwriting is indeed individual (Srihari et al. 2002), how accurately trained and untrained people can identify samples of handwriting written by the same person (US v. Prime 11–12), whether computers can identify writers automatically (Srihari 2001), and even whether and to what degree handwriting varies between twins who have the same education and (presumably) much the same biological mechanisms (Srihari Huang and Srinivasan). Fortunately they found that handwriting is indeed discriminable and that trained experts can correctly identify passages written by the same person with a fairly consistent level of accuracy (US v. Prime 11–12; Kam Wetstein and Conn; Kam Fielding and Conn). They also found that experts are significantly better than untrained people at identifying which samples were written by the same people even without the benefit of any laboratory equipment; more interestingly, they found that "nonprofessionals" tend to produce many more "false positives", that is, untrained people were found generally to underestimate the degree of similarity in different people's handwriting (US v. Prime 11–12; Kam Wetstein and Conn; Kam Fielding and Conn).[6] They also found, again unsurprisingly, that the "handwriting of twins is less discriminable than that of non-twins", and that "error rates with identical twins were higher than with fraternal twins", but that the handwriting of twins can still be identified with an error-rate of about 13%. Perhaps more importantly for our purposes, they developed fully automatic systems which could correctly identify the writer of a given sample 95% of the time or more, and that the success-rate in almost all of these tests was about the same for human specialists and purely automatic systems (Srihari et al. 2002 871; Srihari Huang and Srinivasan 2008).

These are all precisely the questions that have been raised about palaeography, and indeed the relationship between the two fields is being noticed more and more recently (Davis 2007; Stokes 2007/8). However, although modern forensic processes have a lot to teach students of medieval handwriting, there are also important differences between the two. One is that forensic document analysts often can (and ideally must) obtain large samples of the suspect's handwriting, preferably written at different times and

6 The study in question found that untrained people incorrectly attributed two similar samples to the same person 38% of the time compared to 6.5% of the time for experts, and that they correctly matched documents written by the same person with about the same average accuracy as experts (Kam Fielding and Conn). It is an interesting question how expert palaeographers would fare in controlled tests of (modern) handwriting identification as described by Kam Fielding and Conn and US v. Prime 10, or those provided by CEDAR.

in different circumstances, to build up a full picture of the individual and his or her variation (Davis 2007 255; US v. Prime 11–12). However, the palaeographer rarely has this luxury, and even if a substantial corpus has been attributed to a scribe, it is rare (at least for the early medieval period) for those attributions to be certain. Forensic analysts also rely on features that are not generally available to the palaeographer. For example, the pressure exerted on the pen has been cited as important for forensic analysis, particularly for detecting forgeries, and this can be measured with an electrostatic detection apparatus which detects indentations in paper (Kam Wetstein and Conn 12). However, medieval quills do not require pressure to write in the way that modern pens and pencils do, and parchment does not hold indentations, particularly not for centuries, so electrostatic devices would not yield any information.[7] One also suspects that modern writers are much less practiced than medieval scribes, and therefore that modern handwriting varies much more than medieval does, at least for samples from the same region and period.[8] Indeed, the methods employed by forensic analysts, and particularly the automatic handwriting-identification systems they use, often explicitly exclude skilled forgeries or even imitations, suggesting that they may be inappropriate for medieval script (Srihari et al. 2002 857; Kam Wetstein and Conn 7). All these issues suggest that forensic techniques cannot be applied uncritically to medieval script, and although some techniques of forensic analysis have been successfully trialled on medieval documents, the results require further improvement (Davis 2007; Bulacu and Schomaker 2007; Stokes 2007/8).

3.3 "Digital" Palaeography

Returning to the issue of objective methods in palaeography, the question remains what methods can and should be used, and what we can learn from other related fields such as forensic document analysis. One striking aspect of recent research in forensics is the use of computing in the attempts to quantify the field and particularly to develop objective methods. Such an approach has also been hinted at, although not stated explicitly, by Derolez when he suggested replacing qualitative measurements with quantitative ones (7–8), since quantitative methods now normally imply digital ones. Indeed, most of the approaches discussed in Section 3.1 benefit from the use of computers. Features such as the angle and width of strokes can be measured much more easily with high-quality images than they can from manuscripts; images can easily be magnified if the resolution

[7] One calligrapher I have spoken to, Michael Gullick, has described writing with a pen as pushing ink across the parchment, a process which exerts almost no pressure on the page. Pressure on the quill can sometimes be detected by the strokes that remain, but this is much more difficult and less objective than electrostatic devices.

[8] For some examples of variation see the samples provided by Srihari et al. 2002 857 or CEDAR, and contrast those with plates of medieval script such as those by Watson, Robinson or Lieftinck and Gumbert.

is good enough; examples of letter-forms can be cut out and stored in databases for comparison, and so on (Stokes, Palaeography and the 'Virtual Library'). However, these methods are not new and do not depend on computing, they have just become easier and therefore widespread with the advent of the so-called "digital age". In contrast to this are some entirely new approaches which have emerged just in the last five years or so and which have only become possible with the combination of powerful computers and high-quality digital images in what is starting to be called "digital palaeography" (Ciula; but for a very different use of the term see Hirtle). In essence this is a logical extension of the older methods: it is again taking statistical measures of handwriting and then using these measures to make inferences and quantify similarities and differences between hands. The crucial point is that the earlier methods use statistics which were developed by a person sitting down and doing all the counting and measuring. In contrast, this new approach is fundamentally different: we now take images of handwriting and feed these into a computer, then ask the computer to make comparisons and find out which hands are closest to others (Bulacu and Schomaker 2007; Ciula; Stokes 2007/8; see also Hofmeister et al. in this volume).

The underlying principle is to use techniques in computer science, especially in image-processing and data-mining, to generate statistical measures of handwriting and to use these to compare the handwriting in ways that could never have been done previously. These approaches therefore constitute two stages, the first is sometimes known as "feature extraction" and involves generating the numerical measurements, and the second, "data mining", constitutes finding similarities and classifying handwriting based on these measurements (Stokes 2007/8). One example of feature extraction is to take many examples of a given letter (or ligature) written by a single scribe and generate a composite "average" letter from all of them (Ciula). Alternatively, one might break down every stroke in a sample into thousands or millions of tiny line segments and measure the angle of every such segment, or indeed the angle between adjacent sections. We can also do less obvious things like overlay a sample of writing on top of itself and then slide the top sample across and see how much the writing overlaps; the more regular the hand, the greater the overlap (Bulacu and Schomaker 2007; Stokes 2007/8). One or more of these sets of data, these quantitative measure of features, can then be used to generate a statistical profile of each sample of handwriting, and these profiles are then used to compare different samples and to measure the mathematical distance between them. Neither feature-extraction nor measuring distances needs to be especially complex: software can be written quite easily which does this at a basic level, and forensic document analysts have already tested these methods and have developed systems which can correctly identify the writer in 90–95% of cases or more (Srihari et al. 2002 871; Bulacu and Schomaker 2006 285). Other possible methods are much more complex, however, often rely on postgraduate-level mathematics, and their potential is far from fully exploited (for one possibility see Kingsbury).

However promising this may seem, "digital palaeography" seems to have received almost no acceptance and very little interest from so-called "traditional" palaeographers. This is partly because the technology is not yet mature, but this is not a complete explanation. It may also be because most work in this field to date has involved small groups working for relatively short periods, rather than the large, interdisciplinary groups with extended funding that digital humanities often requires. The System for Palaeographic Inspections, for example, was developed by postgraduate students in computer science and one doctoral student in digital humanities, and was never completed (Ciula §§13–14). Software developed for modern forensic analysts has been applied to medieval writing but apparently without the directly involvement of scholars in the humanities (Bulacu and Schomaker 2007). A project to identify medieval scribal handwriting led by a computer scientist and a palaeographer was announced in 2004, and the UK Arts and Humanities Research Council awarded funding to the principle investigator in 2006 but the results have not yet appeared publicly to my knowledge (Intute; AHDS). Finally, the software to quantify differences in scribal hands described in Section 4 has been developed by one person, the author of this paper, working as both computer scientist and palaeographer, as part of a two-year project funded by the Leverhulme Trust. Although none of these is trivial, and although other projects are now emerging, including some described elsewhere in this volume (see contributions by Hofmeister *et al.*; Aussems and Brink; and Ciula), we have not yet had the large groups with experts in a range of fields, computing and humanities, palaeography, image-processing, data mining, but also interface design, database design, developing XML schemas, and so on, and these interdisciplinary groups are now normally required for work in the digital humanities (Pierazzo).

These difficulties, the relative immaturity and the lack of sustained interdisciplinary research, can both be resolved relatively easily given time and resources. However, another problem is perhaps less obvious but still significant, namely that of understanding and engagement. In some cases of software designed for palaeographical analysis, as indeed for other applications of digital humanities, it is not clear exactly what the computer is doing, either because the particular technique requires a lot of *ad hoc* human intervention which is not properly documented, or because the software is proprietary rather than open-source. One example of this is the image enhancement performed by Fotoscientifica, a company which recovers text from damaged manuscripts using multispectral photography and image enhancement (Fotoscientifica). Their results are spectacularly successful and yet their services are sometimes not used because of concern about the degree to which they enhance and the lack of openness about what they have enhanced and how they have done it (Craig-McFeely 2007/8 §§62–3).[9] Even if the

[9] I myself encountered such concerns among colleagues when working for the British Library on the *Rinascimento Virtuale* project to recover Greek palimpsests in 2003–4.

methods are fully documented and reproducible, however, and even if they are communicated clearly using recognised standards and terminology, scholars still require a good understanding of many complex fields to fully appreciate and engage with the results. Indeed, this concern has already been raised explicitly by Tom Davis in a footnote on computing in palaeography. He noted that "these methods are unlikely to replace, though they may supplement, the work of the document analyst, because, however powerful computers will (surely) become, it will probably not be possible to cross-examine them" (266 n. 27). Similarly, researchers in forensic document analysis have argued that "black box" answers rather than verbal reports have contributed to juries tending not to accept automatic methods (Schomaker 2007 §6), and much the same has recently been said for data-mining in literary criticism (Sculley and Pasanek 421). This is by no means to underestimate the ability of medievalists to move between disciplines and to grasp very complex concepts outside their main field. But it does seem fair that we as medievalists in general and palaeographers in particular cannot be expected to understand the intricacies of postgraduate-level mathematics and computer science, and if we cannot understand them then we cannot evaluate them properly or debate their results. We therefore have the same authoritarian discipline as before, with final pronouncements that must be either accepted or rejected wholesale. The difference is that the authority is now a machine.

On the other hand, the results of "digital" palaeography look very promising and should not be discarded lightly; as noted in Section 3.2, experiments with modern handwriting have given successful identifications in 95% or even 98% of the time. Indeed, as discussed above, projects today in digital humanities routinely involve large groups of experts who cannot fully understand each other's fields and who have to trust their validity to some extent. One might even cite the precedent of digital methods such as genetic algorithms which have been used to design new and very effective electronic circuits even though the engineers are sometimes surprised by the results and cannot always explain how they work (Rahmat-Samii and Michielssen 1999, especially 245–6 and 272–7). Nevertheless the computer is a tool to aid us, and like any tool it must be understood before it can be used correctly. In this respect palaeography is not like electronic engineering, or indeed like some branches of digital humanities, insofar as the engineers (and some digital humanists) can test the results of their algorithms, and as long as the results are valid then the details of how they were obtained are not important. However, as has been stated several times already, palaeographers cannot easily test their results, computer-generated or otherwise, and even if they can then those results do not necessarily hold when applied to different scripts or different types of manuscripts, since the methods depend on assumptions that may or may not still be valid. If we do not understand the algorithms, though, then we cannot know on what assumptions our algorithms depend, and therefore we cannot know if they still hold in the new situation. This uncertainty means that our tests are no longer useful.

With this in mind it follows that we *must* be able to "cross-examine" the computer, to use Davis' phrase; that is, we must program the computer in such a way that it is cross-examinable. In other words, even if we as palaeographers or medievalists cannot readily understand the method we should still be able to interpret the results. Rather than having a computer announce that Hand A and Hand B are by the same scribe, it seems much more useful for it to state that Hand A and Hand B both have an average inclination of X°, and an average proportion of width to height of Y, and ascenders of relative length Z, and so on. This sort of meaningful information is perhaps more likely to be trusted than vast quantities of meaningless data or electronic pronouncements of scribal identity; as noted above, this point has also been suggested for forensic document analysts and data-mining in literary criticism. It has recently been argued that systems for forensic document analysis should present information in verbal reports, including margins of error in their results, if their results are to be accepted by juries (Schomaker 2007 §6), and we may reasonably argue the same for medieval handwriting and palaeographers as well. One may well argue further that the computers should not even try to judge scribal identity, but instead that they should present data for experts to interpret; either way, though, it seems useful for that data to be intelligible and ideally to give new insights into ways of seeing and comparing handwriting. This is perhaps a more beneficial way for computers to be used in palaeography. This also suggests renewed scope for the old style of quantitative methods now that we can use computers to organise and process our data. The difference now is that we can handle a much large volume of data than before by using databases and spreadsheets in what could be called "statistical" or perhaps "computer-aided" rather than purely "digital" palaeography.

3.4 Suggestions for a Successful System

Now that the background and previous attempts have been considered, it is worth asking what criteria are necessary for a computer-based system of handwriting identification to be as successful as possible while still being acceptable to palaeographers and medievalists. The first criterion to emerge from this discussion is that whatever is done should be reproducible; this is a basic criterion for acceptability in the scientific world, and reproducing the results of lab experiments in the sciences is considered valid (and necessary) research in its own right. Such reproduction is not considered valid research in the humanities, but it has been recommended for data-mining in literary criticism and for digital humanities more generally (Sculley and Pasanek 423). Similarly, a judge in the US Supreme Court has criticised one set of studies on handwriting identification precisely because the data was not released and so the work could not be verified (US v. Prime 12). Even though it is unlikely that anyone will reproduce a long and detailed study of handwriting *in toto*, it still should be possible in principle to reproduce the experiment and verify its accuracy, otherwise the authority of the study will again be dependent entirely on the person who produced it.

As well as being reproducible, the process should also be debatable; that is, it should be possible to understand and evaluate the assumptions which underlie the analysis and each of the stages used in getting to a result. This also suggests that the results should not be final but should themselves allow for (human) interpretation, or at least understanding, and should also indicate the mathematical level of confidence in the result. Sculley and Pasanek have recently demonstrated that interpretability is not a sufficient criterion for evaluating results in data mining (421), but I would suggest that it is a necessary one. This is another aspect of "communicability" but applying to the outcome rather than the process. The computer should generate evidence which can be "cross-examined" and interpreted by us, the scholars, rather than producing the impenetrable answer of a final authority.

In practice, these criteria imply that the process should be documented and made open in a way that can be communicated effectively and understood by those in the field (compare Pitti 482). This is rarely the case in studies to date, and is particularly rare in computer-based work in manuscripts. To take a slightly different example, some scholars (including myself) use Adobe Photoshop or the GIMP to enhance images of damaged manuscripts and thereby recover lost readings (Craig-McFeely and Lock; Stokes, Recovering Anglo-Saxon Erasures). However, it is very difficult indeed to record precisely what enhancements are done on a particular image, and this in turn makes it difficult for anyone else to verify the results. Furthermore, proprietary software is by definition opaque, and any documentation is usually platform-dependent and unsustainable in the longer term. Thus, to use the example of Photoshop once again, even if one carefully notes every minute enhancement that one makes, this information is still only useful if someone else has exactly the same version of Photoshop and will almost certainly be useless in a few years time when the software has changed and the old version is no longer available. Furthermore, it is not at all clear exactly what software like Photoshop does in particular cases, and so it is difficult or impossible for anyone else to evaluate. On the other hand, if one uses an open standard such as METS to record the precise details of all the algorithms employed then this can be interpreted by anyone else with the required skills, it can be repeated in future, potentially with other software, and it is not tied to a single version of a single application. In this way the criteria to be reproducible and debatable imply a further one, namely to be communicable: a system should allow the entire process to be documented, preferably automatically and without the user having to intervene, and employing open standards for information interchange (Stokes, Recovering Anglo-Saxon Erasures).

In addition to these so-called "digital" aspects, there are also some "humanities" requirements. Certainly any system must allow for a lot of scribal variation. Manuscripts are written by people, not machines, and people change according to many different factors, so this must be taken into account by any palaeographical method, digital or otherwise (Bishop 1961 4–9; Costamagna et al.; Gullick 23). Nevertheless, the entire field of palaeography (and forensic document analysis) is based on the assumption that everyone's handwriting has some deep, innate and inherently individual characteristics which do not vary, or at least vary slowly and can be documented.

So our hypothetical method can (and surely must) assume some underlying commonality in handwriting, although it is by no means clear what kinds and degree of variation this should entail. A successful method must therefore not be too rigid and must accommodate this flexibility, ideally allowing the user to program how much and what sorts of variation she or he has in mind. Indeed flexibility seems to be the key, not least at this relatively early stage when there is still so much uncertainty about what methods will work best and how the software should be developed.

As well as flexibility in allowing variation of handwriting, a system such as this should also allow for flexibility in the methods used for generating the measurements. As discussed in Section 3.3, there are quite a number of competing algorithms which are already being used by forensic document analysts, but it is by no means clear which one, or which combination, will be most successful for medieval documents. Indeed, it seems entirely possible that different types of document will respond better to different combinations of algorithms, and Sculley and Pasanek have advocated always using different methods for data-mining in literary studies (423). Furthermore, many methods have been developed already which have not yet been applied to medieval handwriting but which will probably be useful in this context, and new methods will continue to develop (for one promising example see Kingsbury). Therefore, any software designed to analyse handwriting must allow users to easily add new functionality, otherwise it will quickly become obsolete. Rather than providing a fixed process for analysing handwriting, it seems much more useful at this early stage to provide a common framework which allows researchers to test different methods in a consistent way, allowing one to compare the different results. Just as it is difficult to assess the efficacy of methods in "traditional" palaeography, so also is it hard to comparing digital methods. For two different algorithms to be usefully compared, they must be run in the same circumstances with exactly the same images, the same "correct" classification, and so on, and this assumes that the required outcome is already known. These conditions can potentially be achieved but only if all researchers release the data-sets that they used to test their systems, and the more that this requirement is built into the system the better.

4 A Practical Suggestion: The Hand Analyser

Now that these criteria have been presented, it remains to ask how they might be put into practice. To this end software has been developed to implement a framework for the analysis of scribal hands. As discussed in Section 3.3, this has been developed entirely by the author of this paper as part of a two-year project to investigate objective methods in palaeography. Being a "lone scholar" working on all aspects of the topic, theoretical and practical, "digital" and "humanities", has necessarily limited the amount that can reasonably be achieved, and the framework is certainly not considered to be final or even necessarily usable by palaeographers in its present state. Instead, it is designed to be a platform for testing the methods and principles that would ultimately form part of such a tool for palaeographers and scholars in the humanities more generally.

4.1 Design Principles

Several design decisions follow from the principles outlined in Section 3.4. The requirement to provide a common framework for disparate scholars to share information and test each other's results, along with the basic requirements for sustainability, make Java the obvious choice of programming language. Software written in Java is multi-platform by nature and (in principle, at least) should work on future computers without needing to be rewritten or recompiled. Furthermore, Java now has a very large number of standard libraries of pre-existing code which can be incorporated into any new software, and it has such a wide user-base that these libraries are unlikely to be discontinued for some time, if at all. In particular, the Java Advanced Imaging library (JAI) provides a lot of useful functionality for image-processing. This is a standard library which can be distributed freely and which ships with most installations of Java; it is therefore already installed on most computers and can be freely downloaded from the Java website if not. The source code is under licence to Sun Microsystems but is "open" and may be modified for research use.[10]

Perhaps the most fundamental is the requirement for extensibility and the easy inclusion of different modules. To this end, the system has been designed as a fully modular framework in which the processing is done almost entirely by plugins, where each plugin runs a process to generate a single set of measurements from a single set of features, probably by implementing one or more algorithms in image-processing.[11] For

[10] The source code is available under the terms of the Java Research Licence for non-commercial use and the Java Distribution Licence for commercial use (JAI-Core). Note that this use of Java is a change from previous work discussed by Stokes 2007/8, for which C++ was used and an imaging library from Delft University. Not only was this older software platform-dependent but the imaging library was discontinued between the first development of the software in 2004 and the beginning of the Leverhulme fellowship in 2007. A "second generation" has now continued support and released a version of the library for MacOS X but this still cannot be distributed freely and has a much smaller user-base than the JAI and so is in much greater danger of being discontinued again.

[11] This structure was inspired by John Bradley's discussions of Pliny which in turn draws on that of the Eclipse workbench for software development (Bradley; Birsan).

example, the directions of the edges of strokes has been used as one way of measuring handwriting (Bulacu et al.), and so a plugin can then be written to implement this process. Other plugins can implement different processes, and the users can hence choose which processes they want to test on their particular samples and thereby determine the optimum combination for their particular cases, as well as writing and testing their own new processes.

To allow the exchange of information, as well as accommodating practical issues such as the long time that is often necessary to process high-quality images, the system treats "hands" as distinct objects, where each "hand" contains a complete set of information about a given scribal hand. It therefore includes the URL and other relevant metadata of the image, a full record of the processes that have been carried out on that image, and the full set of measurements generated by each process. The system therefore comprises one or more processes that are run in turn on one or more hands, the results of which are then stored by a "hand" object and can be used to measure the statistical distances between the various samples. A "hand" file combined with the necessary plugins and image file is therefore sufficient to reproduce the process of analysis, but the "hand" file alone contains sufficient information to compare it with other scribal hands. This allows the exchange of information and also means that users need not rerun all the plugins on each hand every time they wish to access the data, an important benefit for such an intensive process.[12] Since each "hand" includes a record of which process has been run on it, and the data generated by that process, it follows that any two hands can check which processes they have in common and use the data generated by the common processes to measure the distance between them. For example, Hand A might have had three processes run on it, say Horizontal Runs, Vertical Runs, and Autocorrelation, in which case it will include three sets of measurements, one for each process.[13] Hand B might also have had the Horizontal Runs and Vertical Runs, but then had Edge Directions and Hinge Directions, and therefore contains four sets of data. However, each hand "knows" what has been done to it and "knows" that it is not meaningful to compare data generated by different plugins. A comparison of Hands A and B will therefore use the results of the Horizontal Runs and Vertical Runs and ignore the other sets of data. This then allows different scholars to run processes on different sample images and pool the resulting "hands" and in this way a very large database of scribal hands could be built up by many different scholars contributing their data from around the world.

The framework itself is further divided into two packages, one containing the core modules which drive the system, store the data, and coordinate the plugins, and the

---

[12] The five processes described by Stokes 2007/8 on a single image of 1370×490 pixels can take approximately 90 seconds on average when running on a MacBook Pro (2.4 GHz Intel Core 2 Duo processor with 2 GB RAM running the Java 2 VRM version 1.5 under MacOS 10.5).

[13] For these terms and those that follow see Stokes 2007/8.

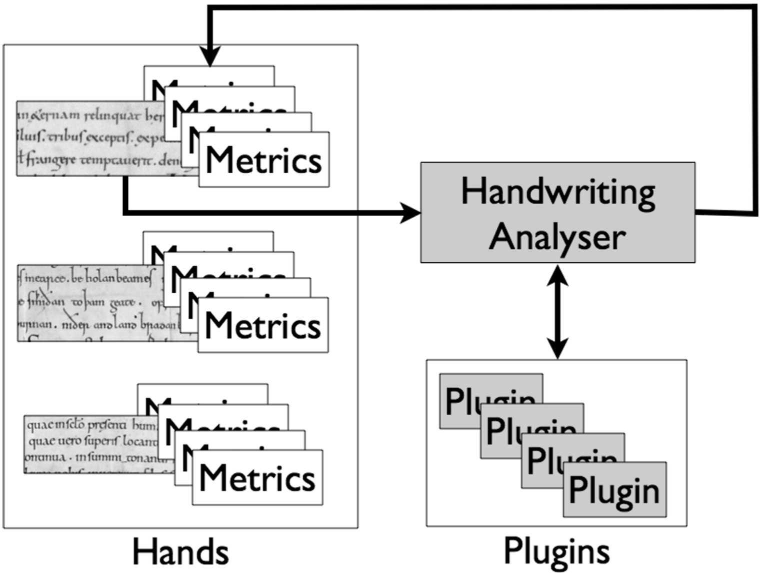

Figure 1. Structure of the Image Analyser.

second providing a graphical user interface (GUI). This allows the GUI to be adapted and expanded independently of the processing and, if desired, to be entirely rewritten, thereby allowing different users to have different interfaces depending on their needs such as, for example, having a desktop application and a web-based service.[14]

At a more basic level, each plugin must automatically log every step taken by the user and allow this log to be exported in MIX/METS format.[15] This is relatively easy to achieve, except of course that one has no way of knowing that all future plugins will conform to this requirement. Instead, everything possible should be done to encourage those writing plugins to conform to this requirement, and specifically the capability for

[14] This is similar to, but different from, the principle that content and presentation should be separated to allow long-term survival of digital resources (O'Donnell 2004 §2; TEI v1). Here, the data can indeed be exported (as is discussed shortly), and the use of Java also makes it more sustainable, as Java is itself separated into different layers and so is inherently platform-independent, and this combined with its very widespread use makes it comparatively sustainable (Gosling and McGilton 1.2.3; compare also O'Donnell 2007 65–6 on the relative longevity of the Java produced by Kiernan).

[15] MIX is an XML implementation of ANSI/NISO Z39.87–2006, §10 of which applies to image processing, and MIX itself is often used to extend the METS standard for technical and administrative metadata (NISO; MIX; METS).

generating such logs is an inherent part of the system's design and should be as easy as possible for authors of plugins to use.

The requirements of reproducibility and sharing of information create a practical difficulty. On the one hand, this requirement implies that images should be stored in the system and circulated as a crucial part of the data; after all, no process can possibly be reproduced if the original images are not available. However, circulating images in any way is not "private use" and therefore violates the terms of use for many digital repositories and libraries (Padfield 165; Case and Green; Stokes, Palaeography and the 'Virtual Library'). For this reason, the system should ideally store URLs to images which are freely available instead of storing the images themselves. As long as the entire process is recorded in suitable detail then users should still be able to reproduce it by using the online images. Unfortunately this has several drawbacks (Stokes, Palaeography and the 'Virtual Library'), not least that the images which are made freely available are often not of sufficiently high quality and are almost always in JPEG format, and this lossy format is not suitable to very demanding applications (Craig-McFeely and Lock; Craig-McFeely).

This requirement of working from publicly available images also has implications for the way in which plugins should be designed. Specifically, they should not assume any preprocessing but should instead begin with the raw image as found at the relevant URL and should log every single step required to produce the output. Indeed the plugins that have been developed to date are all designed to operate on the image as a whole with minimal human intervention. This is in contrast to some other systems in which the user must select portions of the image and often classify those portions, such as selecting all examples of a particular letter, for example. The system being discussed here certainly allows this alternative approach, the slight difficulty being that the sections of images being processed must be carefully defined and recorded in order to comply with repeatability requirement. However, there are now standards for recording this (TEI §11.1) and indeed tools for selecting portions of images and even developing databases of letterforms from images (Holmes; IDP), and in principle these could be turned into plugins for this system.

Another concern relates to the information generated by each plugin and stored in each hand. As discussed in Section 3.4, many of the algorithms used for handwriting identification are different from the principles used by human palaeographers or forensic document analysts, and the algorithms can generate hundreds or even thousands of numbers which are not readily intelligible to human users. It is therefore desirable that plugins present data which can be understood by human users. This need not preclude very large data-sets, as one possibility would be to allow access to the raw data but also provide the mechanism for displaying that data in a human-readable way such as graphically or even interactively.[16] This cannot be enforced, but the structure of the plugins allows the possibility, and indeed the use of Java means that even interactive interfaces can be designed relatively easily and in a way that is platform independent and relatively sustainable (see p. 326).

One example of such an interactive plugin is the Image Enhancement module. This has been designed to address the need to "clean up" images by removing as much of the background as possible and presenting only writing to the computer for analysis, since otherwise the computer can be mislead into interpreting the parchment, folds, stains and so on as ink. Although this can be done easily enough with software such as Adobe Photoshop or the GNU Image Manipulation Program (GIMP), this software has limited value in a scholarly context (Stokes, Recovering Anglo-Saxon Erasures), not least because it does not allow users to easily record the steps taken in the way discussed above, and even if it did this information cannot be readily incorporated into the system for handwriting analysis. For this reason a module was designed which lets the user process the image in a controlled way to remove as much background noise as possible, the objective being to have simply black ink on a white background, and the result of this is documented as part of the overall process of analysis and can be exported, reproduced, and so on. Indeed, the need for a controlled system for image enhancement with automatic documentation is not limited only to handwriting identification but is a desideratum for manuscript studies more generally (Stokes, Recovering Anglo-Saxon Erasures; Craig-McFeely and Lock; Craig-McFeely), and so the module has been designed in such a way that it can also be used as a standalone application.

5 Conclusions

Most of the principles discussed above have been implemented in the prototype at the time of writing, and they are successful insofar as they provide a framework for experimentation, although the framework is nowhere near as sophisticated as that described by Birsan. However, I make no claim that the software is ready to be used by palaeographers or anyone else to establish scribal identity, and it is partly for this reason that no results are presented here (but see Stokes 2007/8). Indeed, as discussed throughout this paper, this system (and probably every other) should not be taken as a "black box" and used uncritically, and its success can only be judged when it is made widely available and tested by the community at large.

On another methodological note, the processes tested in the framework so far all rely on analysing the image as a whole, without the computer having any knowledge of letter-forms *per se*. This approach has the advantage that it can be documented and reproduced easily and is not subject to *ad hoc* human intervention and interpretation. It

[16] For some possible models see Schomaker *Java Demos*.

therefore stands in contrast to other approaches which require the user to segment the image into letters or to build up large databases of graphemes (Ciula §§26–30; Bulacu and Schomaker 2007 282). This approach has also been used with some success on both modern and medieval material (Srihari et al. 2002; Stokes 2007/8; but compare Bulacu and Schomaker 2006 and Bulacu and Schomaker 2007 283–4), and indeed it has been proposed that such an approach is better than considering morphology when comparing scribal hands. "Even more than in ductus and sense of form and proportion the idiosyncratic is to be found in the production of single strokes, in the behaviour of the pen as it turns a curve or a corner, in features defying verbal analysis but offering a limit beyond which sustained imitation, with any appearance of spontaneity, becomes exceedingly difficult" (Bishop 1961 9). One might reasonably ask if this is where "digital" palaeography (as opposed to computer-aided) may be at its best: not by examining letter-forms which can be imitated and which vary anyway depending on the script being written, but instead by looking at the minutiae of strokes over a relatively large sample and which might reflect the individual writer whatever the script.

This raises a further question: even if these digital methods can be used to identify writers, what about the two other questions asked of palaeographers, namely where and when the sample was written? The techniques have been tested for this sort of use (Ciula), but they can only associate similar samples, and it is then up to the palaeographer to interpret the results. Indeed, it has been suggested that morphology, the forms of letters, is a more appropriate criterion for establishing chronological and geographical styles (Derolez 6–7; compare also Stokes 2005 but note Ciula §11), and it may well be that this sort of study requires both evidence that is less amenable to computerisation but much more to human interpretation than is the case for handwriting identification. Whether this is so requires investigation, and even if it is then one might hope that some of the methodological lessons learned here can be applied to these other problems as well.

Even the most objective method still necessarily involves interpretation, and this holds as much for the hard sciences as for the humanities. Palaeography, like every other field, therefore cannot ever be purely objective. However, the more we can articulate our methods and our results, the more we can debate our different interpretations, the more we can aid communication and interpretation and analysis, and the more quantitative and new evidence we can bring to the discussion, the stronger our conclusions will be.

Bibliography

Arts and Humanities Data Service (AHDS). AHRC Award Winners with Technical Components. 2006. <http://ahds.ac.uk/collections/ahrc-research-grant-winners.htm>.

Aussems, Mark. "*Christine de Pizan and the Scribal Fingerprint: A Quantitative Approach to Manuscript Studies.*" MA Thesis. Utrecht University, 2006. <http://igitur-archive.library.uu.nl/student-theses/2006-0908-200407/UUindex.html>.

Beneš, Carrie E. "The Appearance and Spread of the E-Cedilla." *Manuscripta* 43/44 (1999–2000): 1–43.

Benskin, Michael and Margaret Laing. "Translations and Mischsprachen in Middle English manuscripts." *So Meny People Longages and Tonges: Philological Essays in Scots and Medieval English Presented to Angus McIntosh.* Ed. Michael Benskin and Michael L. Samuels. Edinburgh: Middle English Dialect Project, 1981. 55–106.

Birsan, Dorian. "On Plug-ins and Extensible Architectures." *Queue (ACM)* 3:2 (2005): 40–6. <http://queue.acm.org/detail.cfm?id=1053345>.

Bischoff, Bernhard. *Latin Palaeography: Antiquity and the Middle Ages.* Trans. D. O Cróinín and D. Ganz. Cambridge: Cambridge University Press, 1990.

Bishop, T. Alan. *Scriptores Regis.* Oxford: Clarendon Press, 1961.

Bishop, T. Alan. *English Caroline Minuscule.* Oxford: Clarendon Press, 1971.

Bradley, John. "Collaborative Tool-Building with Pliny: A Progress Report." *Digital Humanities 2008: Book of Abstracts.* Ed. Lisa Lena Opas-Hänninen, Mikko Jokelainen, Ilkka Juuso and Tapio Seppänen. Oulu: University of Oulu, 2008. 65–7.

Brown, T. Julian. "Latin Palaeography Since Traube." *Transactions of the Cambridge Bibliographical Society* 3 (1959–63): 361–81.

Brown, T. Julian. *A Palaeographer's View: Selected Writings.* Ed. J. M. Bately, M. P. Brown and J. Roberts. London: Harvey Miller Publishers, 1993.

Brown, Michelle P. *The Book of Cerne: Prayer, Patronage and Power in Ninth-Century England.* London: British Library, 1996.

Bulacu, Marius, Lambert Schomaker, and Louis Vuurpijl. "Writer-Identification using Edge-based Directional Features." *Proceedings of the Seventh International Conference on Document Analysis and Recognition* 2 (2003): 937–41.

Bulacu, Marius and Lambert Schomaker. "Combining Multiple Features for Text-Independent Writer Identification and Verification." *Proceedings of the 10th International Workshop on Frontiers in Handwriting Recognition (IWFHR)* (2006): 281–6.

Bulacu, Marius and Lambert Schomaker. "Automatic Handwriting Identification on Medieval Documents." *Proceedings of the 14th International Conference on Image Analysis and Processing (ICIAP 2007).* Ed. Rita Cucchiara. Los Alamos, CA: IEEE Computer Society, 2007. 279–84.

Burgers, Jan W. J. *De paleografie van de documentaire bronnen in Holland en Zeeland in de dertiende eeuw.* Louvain: Peeters, 1995.

Case, Mary and David Green. "Rights and Permissions in an Electronic Edition." *Electronic Textual Editing.* Ed. Lou Burnard, Katherine O'Brien O'Keeffe, and John Unsworth. New York: Modern Language Association of America, 2006. 565–88.

Center of Excellence for Document Analysis and Recognition (CEDAR). *Individuality in Handwriting: Demo Programs.* CEDAR, State University of NY, 2008. <http://www.cedar.buffalo.edu/NIJ/programs.html>.

Ciula, Arianna. "Digital Palaeography: Using the Digital Representation of Medieval Script to Support Palaeographic Analysis." *Digital Medievalist* 1 (2005). <http://www.digitalmedievalist.org/journal/1.1/ciula/>.

Costamagna, Giorgio, et al. "Commentare Bischoff." *Scrittura e Civiltà* 19 (1995) 325–48 and 20 (1996) 401–7.

Craig-McFeely, Julia and Alan Lock. *Digital Image Archive of Medieval Music: Digital Restoration Workbook.* Oxford: Oxford Select Specialist Catalogue Publications, 2006.

Craig-McFeely, Julia. "Digital Image Archive of Medieval Music: The Evolution of a Digital Resource." *Digital Medievalist* 3 (2007/8). <http://www.digitalmedievalist.org/journal/3/mcfeely/>.

Crick, Julia. "St Albans, Westminster, and Some Twelfth-Century Views of the Anglo-Saxon Past." *Anglo-Norman Studies* 25 (2002): 65–84.

Cubbin, Geoffrey P., ed. *MS D: A Semi-Diplomatic Edition. The Anglo-Saxon Chronicle: A Collaborative Edition*, vol. 6. Cambridge: D. S. Brewer, 1996.

Davis, Lisa F. "Towards an Automated System of Script Classification." *Manuscripta* 42 (1998): 193–201.

Davis, Tom. "The Practice of Handwriting Identification." *The Library (7th series)* 8 (2007): 251–76.

Derolez, Albert. *The Palaeography of Gothic Manuscript Books from the Twelfth to the Early Sixteenth Century.* Cambridge: Cambridge University Press, 2003.

Drout, Michael D. C. "Anglo-Saxon Studies: The State of the Field?" *The Heroic Age* 11 (2008). <http://www.mun.ca/mst/heroicage/issues/11/foruma.php>.

Dumville, David N. "On the Dating of Some Late Anglo-Saxon Liturgical Manuscripts." *Transactions of the Cambridge Bibliographical Society* 10 (1991–95): 40–57.

Dumville, David N. *English Caroline Script and Monastic History: Studies in Benedictinism, A.D. 950–1030.* Woodbridge: The Boydell Press, 1993.

Fotoscientifica. <http://www.fotoscientificarecord.com>.

Ganz, David. "'Editorial Palaeography': One Teacher's Suggestions." *Gazette du Livre Médiéval* 16 (1990): 17–20.

Gilissen, Léon. *L'expertise des écritures médiévales. Recherche d'une méthode avec application à un manuscrit du XIe siècle: le Lectionnaire de Lobbes, codex Bruxellensis 18018.* Ghent: Éditions scientifiques E. Story-Scientia, 1973.

Gneuss, Helmut. "The Origin of Standard Old English and Æthelwold's School at Winchester." *Anglo-Saxon England* 1 (1972): 63–83.

Godden, Malcolm, Rohini Jayatilaka and Rosalind Love, eds. *Boethius in Early Medieval Europe: Commentary on The Consolation of Philosophy from the 9th to the 11th* *centuries.* Oxford: Faculty of English, 2007. <http://www.english.ox.ac.uk/boethius/>.

Gosling, James and Henry McGilton. *The Java Language Environment: A White Paper.* Sun Microsystems, 1996. <http://java.sun.com/docs/white/langenv/>.

Gretsch, Mechthild. "Winchester Vocabulary and Standard Old English: The Vernacular in late Anglo-Saxon England." *Bulletin of the John Rylands Library* 83 (2001): 41–88.

Guimon, Timofey V. "The Writing of Annals in Eleventh-Century England: Palaeography and Textual History." *Writing and Texts in Anglo-Saxon England.* Ed. Alexander R. Rumble. Cambridge: D. S. Brewer, 2006. 137–145.

Gullick, Michael. "The Hand of Symeon of Durham: Further Observations on the Durham Martyrology Scribe." *Symeon of Durham: Historian of Durham and the North.* Ed. David Rollason. Stamford: Shaun Tyas, 1998. 14–31.

Gumbert, J. Peter. "A Proposal for a Cartesian Nomenclature." *Essays presented to G. I. Lieftinck, IV, Miniatures, Scripts, Collections.* Ed. J. Peter Gumbert and Max J. M. de Haan. Amsterdam: Van Gendt, 1976. 45–52.

Gumbert, J. Peter. "Commentare 'Commentare Bischoff'." *Scrittura e Civiltà* 22 (1998): 397–404.

Hirtle, Peter. "Editorial." *D-Lib Magazine* 6.4 (2000). <http://www.dlib.org/dlib/april00/04editorial.html>.

Holmes, Martin. *The UVic Image Markup Tool Project.* <http://tapor.uvic.ca/~mholmes/image_markup/>.

Howorth, Henry H. "The Anglo-Saxon Chronicle, its Origin and History." *Archaeological Journal* 69 (1912): 312–370.

International Dunhuang Project (IDP): Technical Resources. London: British Library. <http://idp.bl.uk/pages/technical_resources.a4d>.

Intute. *Identification of Medieval Scribal Handwriting.* 2004. <http://www.intute.ac.uk/artsandhumanities/cgi-bin/fullrecord.pl?handle=humbul12875>.

JAI-Core Project Home Page. <https://jai-core.dev.java.net>.

Kam, Moshe, Joseph Wetstein, and Robert Conn. "Proficiency of Professional Document Examiners in Writer Identification." *Journal of Forensic Sciences* 39 (1994): 5–14.

Kam, Moshe, Gabriel Fielding, and Robert Conn. "Writer Identification by Professional Document Examiners." *Journal of Forensic Sciences* 42 (1997): 778–86.

Keller, Wolfgang. *Die litterarischen bestrebungen von Worcester in angelsächsischer zeit.* Strasbourg: K. J. Trübner, 1900.

Ker, Neil R. *Catalogue of Manuscripts containing Anglo-Saxon.* Oxford: Clarendon Press, 1957.

Kiernan, Kevin S. *Electronic Beowulf.* CD-ROM. London: The British Library, 1999.

Kingsbury, Nick. "The Dual-Tree Complex Wavelet Transform: A New Efficient Tool for Image Restoration and Enhancement." *Proc. European Signal Processing Conference, EUSIPCO 98, Rhodes* (1998): 319–22.

Lieftinck, Gerard I. "Pour une nomenclature de l'écriture livresque de la période dite gothique." *Nomenclature des écritures livresques du IXe au XVIe siècle: Premier colloque international de paléographie latine, Paris, 28–30 avril 1953.* Ed. Bernhard Bischoff, Gerard I. Lieftinck, and Giulio Battelli. Paris: Centre National de la Recherche Scientifique, 1954. 15–34.

Lieftinck, Gerard I., ed. *Manuscrits datés conservés dans les Pays-Bas: catalogue paléographique des manuscrits en écriture latine portant des indications de date. Tom. 1, Les manuscrits d'origine étrangère: 816–c.1550.* Amsterdam: North-Holland Publishing Co., 1964.

Lieftinck, Gerard I. and J. Peter Gumbert, eds. *Manuscrits datés conservés dans les Pays-Bas: Catalogue paléographique des manuscrits en écriture latine portant des indications de date, 2: Les manuscrits d'origine néerlandaise (XIVe-XVIe siècles) et supplément au tome premier.* Leiden: Brill, 1988.

Liptak, Adam. "Prosecutors Hope New Study of Handwriting Analysis will Silence Skeptics." *New York Times.* 26 May 2002: A14.

Maarse, Frans J. and Arnold J. W . M. Thomassen. "Produced and Perceived Writing Slant: Difference between Up and Down Strokes." *Acta Psychologica* 54 (1983): 131–47.

Mallon, Jean. *Paléographie romaine.* Madrid: Consejo Superior de Investigaciones Científicas, Instituto Antonio de Nebrija, de Filología, 1952.

McGillivray, Murray. "Statistical Analysis of Digital Paleographic Data: What Can it Tell Us?" *Computing in the Humanities Working Papers* A.33 (2005). <http://www.chass.utoronto.ca/epc/chwp/Casta02/McGillivray_casta02.htm>.

McIntosh, Angus. "Towards an Inventory of Middle English Scribes." *Neuphilologische Mitteilungen* 75 (1974): 602–24.

McIntosh, Angus. "Scribal Profiles from Middle English Texts." *Neuphilologische Mitteilungen* 76 (1975): 218–35.

McIntosh, Angus, Michael L. Samuels, and Michael Benskin, eds. *A Linguistic Atlas of Late Mediaeval English.* Aberdeen: Aberdeen University Press, 1986.

McIntosh, Angus. "A New Approach to Middle English Dialectology." *Middle English Dialectology.* Ed. Margaret Laing. Aberdeen: Aberdeen University Press, 1989a. 22–31.

McIntosh, Angus. "Word Geography in the Lexicography of Middle English." *Middle English Dialectology.* Ed. Margaret Laing. Aberdeen: Aberdeen University Press, 1989b. 86–97.

MIX: NISO Metadata for Images in XML Schema. 2008. <http://www.loc.gov/standards/mix/>.

METS: Metadata Encoding and Transmission Standard Official Website. 2009. <http://www.loc.gov/standards/mets/>.

Michel, Lothar. *Gerichtliche Schriftvergleichung: eine Einführung in Grundlagen, Methoden und Praxis.* Berlin: De Gruyter, 1982.

Muir, Bernard J. "A Preliminary Report on a New Edition of the Exeter Book." *Scriptorium* 43 (1989): 273–88.

National Information Standards Organization (NISO). *ANSI/NISO Z39.87 Data Dictionary: Technical Metadata for Digital Still Images.* Bethesda, MD: NISO Press, 2006. <http://www.niso.org/kst/reports/standards/>.

O'Donnell, Daniel P. "The Doomsday Machine, or, 'If you Build it, Will they Still Come Ten Years from Now?': What Medievalists Working in Digital Media Can Do to Ensure the Longevity of their Research." *The Heroic Age* 7 (2004). <http://www.mun.ca/mst/heroicage/issues/7/ecolumn.html>.

O'Donnell, Daniel P. "Disciplinary Impact and Technological Obsolescence in Digital Medieval Studies." *A Companion to Digital Literary Studies.* Ed. Ray Siemens and Susan Schreibman. Oxford: Blackwell, 2007. 65–81.

Padfield, Tim. *Copyright for Archivists and Records Managers.* London: Facet Publishing, 2007.

Parkes, Malcolm B. "Latin Autograph Manuscripts: Orthography and Punctuation." *Gli autografi medievali: problemi paleografici e filologici : atti del convegno di studio della Fondazione Ezio Franceschini, Erice, 25 settembre-2 ottobre 1990.* Ed. Paolo Chiesa and Lucia Pinelli. Spoleto: Centro italiano di studi sull'alto Medioevo, 1994. 23–36.

Parkes, Malcolm B. *Their Hands Before Our Eyes: A Closer Look at Scribes.* Aldershot: Ashgate, 2008.

Pierazzo, Elena. "Editorial Teamwork in a Digital Environment: The Edition of the Correspondence of Giacomo Puccini." *Jahrbuch für Computerphilologie* 10 (2008). <http://computerphilologie.tu-darmstadt.de/jg08/pierazzo.html>.

Pitti, Daniel V. "Designing Sustainable Projects and Publications." *A Companion to Digital Humanities.* Ed. Susan Schreibman, Ray Siemens and John Unsworth. Oxford: Blackwell, 2004. 471–87.

Plummer, C., ed. *Two of the Saxon Chronicles Parallel, with Supplementary.* Extracts from the Other Edition, 2 vols. Clarendon Press, Oxford, 1892–99. Revised, on the basis of an edition by John Earle.

Pratesi, Alessandro. "Commentare Bischoff: un secondo intervento." *Scrittura e Civiltà* 22 (1998): 405–8.

Rahmat-Samii, Yahya and Eric Michielssen, eds. *Electromagnetic Optimization by Genetic Algorithms.* New York: J. Wiley, 1999.

Rinascimento Virtuale. Digitale Palimpsestforschung. Rediscovering written records of a hidden European cultural heritage. <http://www1.uni-hamburg.de/RV/>.

Robinson, Pamela R. *Catalogue of Dated and Datable Manuscripts, c. 737–1600, in Cambridge Libraries.* Cambridge: D. S. Brewer, 1988.

Rumble, Alexander R. "Using Anglo-Saxon Manuscripts." *Anglo-Saxon Manuscripts: Basic Readings.* Ed. Mary P. Richards. New York: Garland, 1994. 3–24.

Rumble, Alexander R. "Palaeography, Scribal Identification and the Study of Manuscript Characteristics." *Care and Conservation of Manuscripts 8: Proceedings of the 8th International Seminar.* Ed. Gillian Fellows-Jensen and Peter Springborg. Copenhagen: Museum Tusculanum Press, 2005. 217–28.

Rumble, Alexander R. "The Study of Anglo-Saxon Manuscripts, Collections and Scribes: in the Footsteps of Wanley and Ker." *Writing and Texts in Anglo-Saxon England.* Ed. Alexander R. Rumble. Cambridge: D. S. Brewer, 2006. 1–17.

Russell, Paul, Sharan Arbuthnot, and Pádraic Moran. *Early Irish Glossaries Project.* Cambridge: Dept. ASNC, 2006. <http://www.asnc.cam.ac.uk/irishglossaries/>.

Schomaker, Lambert. "Advances in Writer Identification and Verification." *Proceedings of the 9th International Conference on Document Analysis and Recognition (ICDAR 2007), 26 September, Curitiba, Brazil* 2 (2007): 769–73.

Schomaker, Lambert. *Java Demos for Handwriting Recognizers.* University of Groningen. <http://www.ai.rug.nl/~lambert/recog/java-demos.html>.

Scragg, D. G., Alexander R. Rumble, Kathryn Powell, Susan D. Thompson, and Joana Soliva. *MANCASS C11 Database Project.* Manchester: Manchester Centre for Anglo-Saxon Studies, 2005. <http://www.arts.manchester.ac.uk/mancass/C11database/>.

Sculley, D. and Bradley M. Pasanek. "Meaning and Mining: The Impact of Implicit Assumptions in Data Mining for the Humanities." *Literary and Linguistic Computing* 23 (2008): 409–24.

Shippey, Tom. "Response to Three Papers on 'Philology: Whence and Whither?' Given by Drs Utz, Macgillivray, and Zolkowski, at Kalamazoo, 4th May 2002." *The Heroic Age* 11 (2008). <http://www.mun.ca/mst/heroicage/issues/11/foruma.php>.

Sirat, Colette. *Writing as Handwork: A History of Handwriting in Mediterranean and Western Culture.* Ed. Lenn Schramm. Turnhout: Brepols, 2006.

Srihari, Sargur N. *Handwriting Identification: Research to Study Validity of Individuality of Handwriting and Develop Computer-assisted Procedures for Comparing Handwriting.* New York: Center of Excellence for Document Analysis and Recognition, 2001.

Srihari, Sargur N., Sung-Hyuk Cha, Hina Arora, and Sangjik Lee. "Individuality of Handwriting." *Journal of Forensic Sciences* 47 (2002): 856–72.

Srihari, Sargur N., Chen Huang, and Harish Srinivasan. "On the Discriminability of the Handwriting of Twins." *Journal of Forensic Sciences* 53 (2008): 430–46.

Stokes, Peter A. "Shoots and Vines: Some Models for the Ascenders and Descenders of English Vernacular Minuscule." *Quaestio Insularis: Selected Proceedings of the Cambridge Colloquium in Anglo-Saxon, Norse, and Celtic* 5 (2004): 98–109.

Stokes, Peter A. "*English Vernacular Script, ca 990 – ca 1035*." PhD Thesis. Cambridge University, 2005.

Stokes, Peter A. "Palaeography and Image Processing: Some Solutions and Problems." *Digital Medievalist* 3 (2007/8). <http://www.digitalmedievalist.org/journal/3/stokes/>.

Stokes, Peter A. "Recovering Anglo-Saxon Erasures: Some Questions, Tools and Techniques." *Palimpsests and the Literary Imagination of Medieval England*. Ed. Raeleen Chai-Elsholz, Tatjana Silec and Leo Carruthers. New York: Palgrave, forthcoming 2009.

Stokes, Peter A. "Palaeography and the 'Virtual Library'." *Digitizing Medieval and Early Modern Material Culture*. Ed. Brent Nelson and Melissa Terras. Forthcoming.

TEI Consortium, The. *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Ed. Lou Burnard and Syd Bauman. Oxford: Humanities Computing Unit, University of Oxford, 2009.

Traube, Ludwig. *Nomina sacra: Versuch einer Geschichte der christlichen Kürzung*. Munich: C. H. Beck, 1907.

United States v. Prime: Order Regarding Defendant's Motion In Limine. US Supreme Court, 9th Circuit, 2002. 220 F. Supp. 2d 1203 (W. D. Wash. 2002).

United States v. Janet L. Thornton. 2003. Case No. 02-M-9150-01, decided January 24, 2003.

Watson, Andrew G. *Catalogue of Dated and Datable Manuscripts, c. 700–1600, in the Department of Manuscripts, The British Library*. London: British Library, 1979.

Wright, Cyril E. *English Vernacular Hands from the Twelfth to the Fifteenth Centuries*. Oxford: Clarendon Press, 1960.

Zupitza, Julius. *Beowulf: Reproduced in Facsimile from the Unique Manuscript British Museum MS. Cotton Vitellius A.xv*. London: Oxford University Press, 1959.