Rendering Digital Type: A Historical and Economic View of Technology

---

R. A. MORRIS*

Department of Mathematics and Computer Science, The University of Massachusetts at Boston, Boston, MA 02125, USA†

The design and rendering of digital type poses new problems to the type and software designers who must bring the type to fruition. But the stages through which this process is travelling reflect those of previous transitions to new technologies which have accompanied letter rendering since the beginning of written language.

Received 20 June 1989, revised 12 August 1989

1. INTRODUCTION

The intent of this survey is to show what questions arise in the quest for highly legible text produced by computers, along with some of the present answers. The appoach is to describe the interaction – often noted by typography scholars and repeated through the history of written language – of economic, technical and visual factors. Some, but not all, of what is discussed applies to graphics in documents and to other visual facets of complex documents.

Written languages appear to have developed in several societies by about 5000 years ago, but alphabetic writing probably originated in the Middle East less than 3500 years ago. The earliest known indication of the present alphabetical *order* is found on a tablet dated 1500 B.C.E. from the city of Ugarit in what is now Syria. The idea of an alphabet seems to have arisen in the probably earlier writing systems of Semitic languages, and the Ugaritic tablet organised its phonemes, which were the same as those of the North Semitic alphabet but represented by cuneiform, in the order which has reached us today through Hebrew, Greek and Latin.

Alphabetic writing offers artistic freedom not available in logographic languages. Unlike the shapes of logograms, the shapes of letters are free of requirements to look like something from nature. Nevertheless, these shapes are not fully under the control of the letter designer. For the moment, let us ignore the important requirement that letter shapes should not depart radically from what a literate reader has learned to recognise. Even then, letterforms are not free of all natural constraints. Among the most important ones are the physical and economic constraints arising from a given rendering technology (for example, it is quite difficult to make uniform curved shapes in clay) and requirements of the human visual system (for example, an adult with normal vision cannot resolve two lines which are closer than 1 minute of visual angle, nor detect a displacement of less than about 12 seconds of visual angle).

The extent to which a typographic enterprise – starting from the letter design, through the layout and to the rendering – takes account of these constraints has a large impact on both the legibility (the reader's ability to *see* and *decode* the text on the page) and the readability (the reader's ability to *understand* the text on the page).‡

In recent years, these factors have become even more important because technology has given authors capabilities which formerly were in the control of skilled experts. Amateurs can (and do) determine letter, word and line spacing, type size, margins and sometimes even which pixels are black in a laser printer bitmap of a particular character.

2. RENDERING

In Europe by 1500 a letter-rendering technology had emerged which was to remain unchanged in fundamental ways until the mid-twentieth century (although certain detailed technological advances – or regressions! – specifically affected letter shapes, as sketched below). The punch cutter, or typeface implementer, cut steel punches in a shape articulated by the type designer. (The punch cutter and type designer were often the same person.) These punches were used to impress a negative intaglio image of the letterform usually in brass matrices. Molten metal was poured into these matrices to comprise type which was then the shape of the original punches. The movable type was assembled into lines, inked and then pressed with paper. An approximation of the original letterforms was thereby recovered. Late in the 19th century, machines with keyboards were invented which permitted lines of the matrices to be rapidly composed for molten lead to be poured into them to make an entire line of type at one time. (The first of these was trade-marked the 'Linotype' machine, a designation which had become generic by the time of the machines' obsolescence 75 years later).

In technology historically intermediate between line-casting machines and digital rendering, a photographic reproduction was made of text set by machine. The reproduction was eventually transformed into plates of the same 'polarity' as the lead type. The printing area of the plate was distinguished from the non-printing area either by raising it (*letterpress*), lowering it (*gravure*) or by chemically treating the plate to reject ink where it was not desired, and accept it where it was desired (*offset*

---

* Supported in part by NSF grant IRI-87-15960 at the University.

† At the time of submission the author was also a member of the technical staff at Interleaf Inc., 10 Canal Park, Cambridge, MA 02141, USA.

‡ There is not complete terminological agreement. This corresponds to the usage of Tracy, chapter 4[44] and probably most typographers. Rubinstein uses visibility and legibility (the latter as readability). Vision scientists would charge the typographers with failing to separate the visual task into detection and discrimination, and they wouldn't be interested in readability at all! (Ref. 34, p. 174).

lithography). These are still the fundamental mass-reproduction technologies in use today, although the photographic master is made in other ways, which we describe below. In high-volume applications the plates are wrapped on cylinders in the press and ink is transferred from the plates as the cylinder rotates. It is curious that the first cylindrical printing technology appeared near the dawn of western writing (Ref. 32, p. 242).

> Equally noteworthy [to a series of stamps with interchangeable sign units rather like movable type] was the practice of the scribes in Elamite Susa, who resorted to the use of cylinders on which were engraved curses that could be transferred to the soft clay surface by rolling the cylinder over it in order to save themselves the bother of writing these tiresome formulas by hand.[32]

Several important consequences of the metal type scheme disappeared with the emergence of electronic technologies. The first is that the brass matrices, and hence ultimately the metal type, have a fixed base which cannot be narrowed arbitrarily. Specific pairs of letters, for example, 'A' and 'V' will appear to have too much space between them if separated by the normal width of the matrices. Indeed, a portion of the 'A' must lie under a portion of the 'V'. Thus, in order to *kern* specific pairs of letters, i.e. set them closer together than their matrices would allow, it is necessary to have a number of specially formed variants of the matrices. Another typographic consequence of static matrices is that it is sometimes cumbersome mechanically to arrange symbols two-dimensionally, as is typical, for example, in mathematics typesetting, which also had unsatisfiable needs for some symbols, e.g. radical signs, of rather arbitrary size. (Indeed, mathematics is called *penalty copy* in the printing trade.) Finally, the great time and expense of producing the matrices led to the availability of a relatively small number of sizes and faces (although one may argue that this was an advantage, not a disadvantage, since the work was often commissioned only of skilled artists, who in turn produced designs which lasted hundreds of years).

The first fast (analogue) phototypsetters appeared in 1950. These machines carry the letters on negative masters through which photographic paper is exposed one character at a time. The photographic paper is then reproduced using a photosensitive plate (originally metal, now plastic), which is used for printing in ways similar to those described above for pages photographically derived from hot metal type.

Phototypesetters permit the optical scaling of type through a wide – sometimes continuous – range of sizes, just as is possible with the current digital technologies – such as those of Adobe Systems, Bitstream and Folio Corporation – used in PostScript devices and their emulations. Photocomposing machines commonly were operated with only one photographic master per face; however, typographers are well aware that maximum legibility of type demands that small letterform shapes be somewhat wider and heavier than those of larger sizes, and on some machines it was possible to have two or three photographic masters for each face.

With phototypesetting, several of the economic constraints of punch and matrix technology disappeared. The original master could be made photographically from the production artwork derived from the designer's drawings, eliminating time-consuming punchcutting. Most of the physical constraints of the intermediate metal form were gone; for example, kerning became a simple matter of moving optics – the same problem that must be solved to set letters anyway. The quality of the resulting photographic output was limited by the quality and accuracy of the optics and transport mechanisms, as well as the rate at which composition took place. (A longer exposure, hence slower typesetting, permits smaller aperture. As is familiar to all photographers a smaller aperture results in a less optical blur.)

How final copy is printed is a very important factor in rendering technology, which has changed as much over time as has the master-making technology. This is especially consequential when the final copy is produced by a laser printer, but even in conventional printing with liquid inks, the physics of ink dispersal is understood to have important mediating effects that typeface designers cannot ignore. The range of these effects can be enormous. For example, high-speed newspaper presses demand inexpensive inks and paper which will not tolerate typefaces with sharp corners (so-called 'ink traps'), because these corners fill with ink and distort the letterform.

In the development of letterforms for phototypesetting one sees a phenomenon noted by many observers: when a new rendering technology emerges it takes some time for the type designs to catch up with the new technology, initially resulting in worse typographic quality than with the old technology. In a sense, it is not that this lesson remains unlearned so much as that the technology must be explored before new designs can be contemplated. It happened again in the next stage, as remarked by the eminent type designer Hermann Zapf (Ref. 48, pp. 28–29).

> The first years of phototypesetting machines offered designs which in most instances were simple copies of alphabets originally drawn for metal. Yes, some firms took the same master enlargements they used before cutting the metal type punches. One of the excuses given was that the customer wanted familiar designs and, in addition, some salesmen of the new phototypesetting systems told their clients the foolish story that they could mix metal and phototype. If the new machine broke down they could easily put a fire under the lead pot of their good old Linotype machine, and be back in business.
>
> As was done at the beginning of phototypesetting, it looks now as though we are making some fundamental mistakes in the digitizing of some old alphabets. This even includes companies with electronic printing systems, the so called 'printing on demand' installations...Today we have to find new solutions.
>
> The additive effect of letterpress printing, which was taken into account by all the skilled punchcutters and craftsmen, was completely ignored in the designs or, let us better say, in the early adaptations of existing alphabets for phototypesetting. But it was not always realized that using negatives to shoot characters onto film resulted in a subtractive effect which was completely different from metal type. The text produced by a phototypesetting machine was used for offset printing, in which there is no inkspread as in letterpress, and the result was that the majority of those alphabets looked oversharpened, weak, and vigorless.

Classical phototypesetting had such a short life because it was quickly realised that by using a cathode ray tube, more than a single character could be imaged in a short time without moving optics. Based on television technology, this became digital phototypesetting. The CRT is illuminated at fixed points on the raster. The problem is now one of choosing which *pixels* to illuminate to shape the character.

With digital type, ink must be deposited at fixed locations on a discrete grid – an entirely new constraint for the type designer. There is no longer a continuum of sizes of features with which to form characters. The problem becomes more severe at low resolutions, since at small sizes there may be a very few pixels available to form the character, making some distinctions difficult to achieve.

The most immediate impact of digital rendering is the requirement to be conscious of resolution. Digital phototypesetter vendors found by experimentation that resolution of more than 900–1200 dots/inch (dpi) does not result in improved text quality even on the photographic film from which the printing masters are made. (This does not include requirements for halftone screening, the process by which continuous-tone images are converted to varying size clumps of dots to give the illusion of multiple levels of grey from black-and-white. For such a process, resolutions of 1600–2000 dpi and higher are needed.) Media are not the principal limiting factor since photographic paper generally exceeds the resolution demanded by even the highest-resolution typesetters. Many factors affect the appearance of type images on the film, and even more mediate the final result on paper. The highest resolution needed for the best possible results is somewhat controversial among typographers. Walter Tracy writes (p. 46):

> Whether we can actually see the effect of the [digitisation] process in print depends entirely on the resolution at which the face was digitised and the typesetting machine operates. To quote Charles Bigelow: 'Trained typographers claim that resolutions below 600 lines per inch seem crude or coarse; resolutions between 600 and 1200 lpi seem acceptable or adequate; resolutions above 1200 lpi seem as good as traditional analogue type – that is, type in which the edges of the letters are natural and unchanged. Clearly the quality of the subsequent plate-making, paper and presswork will be factors in the matter. Typesetting at a resolution of 1200 lpi may show no sign of its digital genesis in printing of ordinary quality, but may look not quite sharp when the presswork and paper are of the superior quality demanded for some sorts of publicity work; hence the very high resolution of 2000 lpi obtainable from some modern systems.[44]

Resolution is really a measure of addressability; but a number of photographic and optical factors can affect the shape of the spot on the film and so the final effect. These include non-linearities in exposure, scattering of light in the emulsion, intensity distribution of the illuminating spot and diffraction in the optics. An anonymous referee of the first version of this paper observed that smoothing effects due to some of these factors allowed CRT phototypesetters to produce acceptable results at relatively low resolutions.

In more recent devices, laser beams illuminate the film, and resolution is determined largely by the rate at which the beam can be turned on and off as it sweeps across the paper. Generally, the sweep is accomplished by reflecting the beam off a multi-faceted rotating mirror. There are two ways the resolution can be increased with such a scheme. Either the beam can be turned off and on more rapidly or the mirror can be rotated more slowly. The first increases the complexity and hence the cost of the device; the second decreases the throughput. Thus, the quality/cost/throughput tradeoffs are similar to those for optical phototypesetters and, indeed, for all rendering technology throughout history.

Why, then, is there no advantage to increasing resolution beyond 1200 dpi? Most of the limitations are in the human visual system, not the intermediate master rendering. They arise principally from two factors which, wondrously, are perfectly matched. First, as with any lens, the human lens has a resolution limit imposed by its optics. In adults with normal vision, this is about one minute of visual angle. This visual capability is called *grating acuity* by vision scientists, who prefer to describe it in terms of spatial frequency, which is roughly the alternation rate of black and white with respect to visual angle. Visual angle, not linear measure, is clearly the most appropriate variable for such descriptions, since it makes the viewing distance irrelevant. Resolving lines one minute apart is equivalent to a grating acuity of 60 cycles per degree (cpd) of visual angle.

In addition to the optical limits, the human retina is itself digital. The images passing through the optics are sampled at the retina by virtue of the fact that there are discrete photoreceptors in the eye. The receptors responsible for high illumination vision, the cones, have a density at the fovea (the centre of vision) of about 120 per degree. A famous theorem of signal processing, the Nyquist–Shannon sampling theorem, asserts that all the information in an image with a maximum spatial frequency is contained in any set of samples that are taken at no less than twice the image's maximum frequency. Thus the cone mosiac is an ideal sampler; no more cones are needed at the fovea. (Peripheral vision is less acute, but reading always takes place by moving the eyes to bring text into the fovea.) Although it is unclear if this has any implications for type, the retinal design is even more amazing: the foveal cone mosaic is close-packed, and, because the cones have circular cross-section and are approximately the same size, they are hexagonally packed. The optics of the eye impose radially symmetric acuity limits (although the visual system as a whole has highest acuity at vertical and horizontal orientation, lowest at 45 degrees). For such a 'circularly bandlimited' system, it is not difficult to show that among all the optimal sampling mosaics (i.e. those sampling at exactly twice the maximum frequency), hexagonal sampling requires the fewest number of samples per unit area.

As it happens, *vernier acuity*, which is the ability to visually distinguish an *offset* of one line from another, is normally about five times greater than the grating acuity described above, i.e. about 12 seconds of arc. The explanation of this and other so-called 'hyperacuities' is beyond the scope of this paper, but the interested reader can find some discussion and pointers to the literature in Ref. 10. Unlike grating acuity, vernier acuity can, with practice, be doubled. This means that 'jaggies' are undetectable if they result from displacements of less than 12 seconds (or, with practice, about 6 seconds) of visual angle. Simple trigonometry will show that this corresponds to a displacement of about 1/900 inch at normal reading distances (18 inches). At these distances, a displacement of 1/1200 inch subtends about 9 seconds of visual angle, which might conceivably be normal acuity for an artistically trained viewer. Even with practice at the vernier resolution task itself, achieving the 6-second acuity limit would demand a resolution of no more than 1800 dpi. Thus, human vernier acuity suffices to account for the empirical findings of the typesetter manufacturers.

Besides vernier acuity, it is likely that other hyperacuities affect our perception of type. Rather recent psychophysical results[46] offer a somewhat surprising view of curvature discrimination. There is evidence that discrimination of both high- *and* low-curvature shapes is mediated by high-spatial-frequency visual mechanisms. This is counter-intuitive, because the latter are changing more slowly. Any claims that skilled observers can distinguish inaccuracies in curve rendering at 1800 dpi could find confirmation or rejection in these results. In particular, they do suggest that lowering resolution will affect the appearance of gentle curves as well as sharp ones.

3. LOW-RESOLUTION DEVICES

By the early 1970s, researchers at the Xerox Corporation had already realised that the electrostatic principles successful in photocopy machines could be coupled with the digital imaging techniques then beginning to be used in CRT typesetters. Instead of exposing photographic paper, they illuminate a photosensitive drum. The commercial variants of this now in wide use generally illuminate the drum by modulating a laser beam or controlling a row of light-emitting diodes. With so-called 'write-black' devices, the drum accepts electric charge where the light hits it, and charged particles of carbon are attracted to the drum. Next the particles are transferred to charged plain paper and fused with heat, and the paper is discharged. The toner thus can make a pattern of spots at almost arbitrary locations on the paper. Alternative arrangements permit the removal of the charge at the point of illumination, in essence causing the toner to be cut out much as a cookie cutter would do to dough. Such 'write-white' devices are also in use, and require some letterform design considerations described below. Xerox loaned three prototypes of the eXperimental Graphics Printer, XGP, to the computer science departments at Stanford, Carnegie Mellon and MIT, where there was a lot of experimentation with the bit-mapped fonts necessary to drive the printers. The Stanford machine was first operational in January 1973. By the time it was retired in 1985, it had 527 fonts available, counting various sizes and styles.[14] A substantial amount of medium-resolution experience was gained using these printers and in research at Xerox.

The XGP was illuminated by a single line of video on a CRT. It was originally designed as the receiver for a fascimile system. The first commercial laser printer was probably the IBM 3800, sold first in 1975. It had no graphics capability, so was used as a sophisticated line printer. Two years later Xerox announced the 9700, a high-speed printer which cost $500,000. Still in wide use, it was only given graphics capability in recent years by addition of Xerox's Interpress page description language. Not until late 1981 was there any commercially available laser printing system under $100,000. The first of these was offered by Imagen Corp., now a subsidiary of QMS. The Imagen printer was based on the Canon LBP 10 marking engine, and was soon replaced by the Canon CX engine. That low-priced marking engine and its successor, the Canon SX, are found in the widely available printers from Hewlett-Packard, Apple and others now commonly for sale at prices as low as $1200.

The size and resolution of the toner spots and the throughput of a laser printer are once again determined by physics, engineering and economics. At the time of writing, common office laser printers operate at speeds of eight pages/minute and above, with a resolution of 300 dpi. The ability of the photosensitive media to keep small charges close to one another and the physical properties of the toner (both before and after fusing) are among the things which limit resolution and spot size. For most laser printers (and laser phototypesetters), the resolution perpendicular to the paper motion direction is determined by the rate at which the beam can be modulated. Generally, this limitation is not as severe as the limits on resolution in the direction of paper motion imposed by the rotating mass of the photosensitive drum and by the accuracy of paper motion. As always, increasing the electronic or mechanical complexity may be expected to increase cost also. Laser printer systems with 600 dpi resolution presently sell for 10–20 times the price of lower-resolution (and usually lower-throughput) machines.

In selecting which pixels to blacken, the type designer (or, increasingly more commonly, software) has a number of factors to consider. For example, the 'cookie cutter' behaviour of write-white devices subtracts from the stems, rather than adding to them as do write-black devices. This results in thinner stems, which must be compensated for in the outline design or choice of bitmaps. Also, at low resolutions (600 dpi or less), quantisation error – which, as mentioned earlier, may only be relevant below the limits imposed by the vernier acuity of the viewer – militates against shapes with many curves. Sometimes this makes it easier to render sans-serif faces than serif faces on low-resolution devices, but even the fundamental shapes need to take account of the resolution. In what is probably the first face whose shapes were designed with low-resolution rendering in mind, Bigelow and Holmes' Lucida, the serifs are polygonal instead of curved, arches are joined to stems relatively deeply, and other considerations based on the final effect are made.[4] More recently other faces, for example, ITC Stone,[41] have been designed based on experiences with rendering on low-resolution laser printers.

4. OUTLINES AND THEIR MANIPULATION

Type designers generally design continuous letter shapes, even if the ultimate target is digital rendering. Although it has recently become economically important to digitise these outlines completely by algorithm, the history of mathematical description of letter outlines is approximately as long as that of movable type itself.[22] The earliest rigorous attempts to put such descriptions in formal terms suitable for digitisation are probably Coueignoux's 1975 MIT Ph.D. thesis[8] and the lesser-known system of Mergler and Vargo,[27] in use by 1967 and described in Ref. 35 (p. 5). At about the same time, various vector drawing schemes were also being investigated; the most widely known is that of Hershey.[17]

The Mergler–Vargo and Coueignoux systems appear to have been the first to parameterise the shapes of character pieces and to provide ways of amalgamating the pieces to form characters with consistent features.

In the early 1970s there also emerged what soon became the principal digitisation tool in the phototypesetter industry, the Ikarus system by Peter Karow.[19] Still in wide use today, Ikarus addresses a number of the issues required to digitize type, including the now common technique of producing a spline representation from outlines provided by the designer. Ikarus permits the selection of specific control points on the letter, input with a digitising tablet. It then automatically computes a spline representation suitable for further manipulation and rasterisation.

One the most influential typesetting programs in use today is Knuth's TeX.[24] Its paradigm of boxes and glue underlies many contemporary composition systems, especially those driving laser printers and modern typesetters. Besides developing this typesetting program, Knuth sought an even more mathematical description of type outlines. More precisely, he wanted the description to be the foundation of the letterform design. This led to the Metafont system, an actual programming language which describes letter shapes by equations, rather than deducing those equations by numerical methods as does Ikarus. Metafont's most important contribution is that it permits the extreme use of parameterisation, making it easy to design very large collections of highly related fonts and typefaces.[21–23] In this regard, Metafont is more general than the automatic scan conversion software described below for PostScript and related printers. Those schemes only produce the font, in this case the bitmap, and require each variant to be regarded as a separate face with no particular connection to other members of its family.

Unfortunately, the metaphors of linear algebra and differential geometry upon which Metafont rests do not seem to correspond to the skills acquired by design artists, who often find Metafont difficult to use. Only a few complete families of type have been designed with Metafont by experienced designers. The Computer Modern family, designed by Knuth with much critical and design assistance from several established type designers (Ref. 20, pp. vi–vii), is not in use outside TeX applications. The American Mathematical Society commissioned Hermann Zapf to create Euler, for mathematics[38] (although Euler used Metafont more to enhance its TeX utility than to take advantage of Metafont principles). A few efforts of smaller scope have been aimed at special purposes, such as designing a face for library catalogues[43] or for research in the tool itself.[5, 39]

Just as producing hot metal type required two independent skills – the art of the type designer and the craft of the punchcutter – so it appears Metafont requires both the art of the designer and the craft of a programmer to produce a production-quality typeface. And as with metal types, expert skill at each appears to be needed. Whether with Metafont or some other production method, there is no reason to expect the kind of necessary synergy to emerge any more rapidly than the several centuries it took to reach the highest state of the corresponding art with metal type production. For the forseeable future, digital 'punchcutters' will probably continue to use tools such as Ikarus, plus a few more modern ones, described below in the context of on-the-fly scan conversion. These tools are not dedicated to new ways to describe letters but rather to making traditional ways more amenable to software implementations.

5. RASTERISATION ON THE FLY

Unlike the case of graphic arts, many text-processing applications have little typographic need for characters at orientations other than horizontal and vertical, or for a huge range of type sizes. But even in such applications it is often not feasible to store the pixel descriptions of all the needed characters in a low-cost digital printing system. A 10-point lower-case character at 300 dpi is typically about 18–20 pixels high and 20–24 pixels wide, requiring, say, $18 \times 24/8 = 54$ bytes to store. Upper-case characters might require 2–3 times as much storage, and an entire font, together with enough metric information to be useful, might require about 15 kB to store at 10 point, and storage increases quadratically with size. Thus a single size in roman, italic and bold styles might consume 45–50 kB at 10 point and each face might require upwards of 0.5 mB. Even a modest collection of faces, sizes and styles will require auxiliary storage, raising the price of the printing system. Using precomputed bitmaps also restricts the selection to specific sizes and orientations. For example, rotating the bitmap itself is satisfactory for multiples of 90°, but this is rarely the case for other angles. Bitmap compression schemes not fundamentally different from those in use in other image-processing applications can alleviate the storage problem somewhat. But these probably yield results principally of interest in typesetter applications, where the bitmap storage requirement at 1200 dpi is 16 times that of a 300 dpi laser printer. Other encoding schemes based on letterform features can yield substantially greater compression than the one order of magnitude typically achieved by two-dimensional run-length encoding (Ref. 33, Sec. 22.3.1). Such schemes have been successfully used to give compression ratios of 500:1 for Chinese characters in applications requiring 7000 characters in 25 styles in 20 sizes.[47] But such representations cannot be decoded at high enough speeds to satisfy current laser printers without special hardware.

These factors, and the graphic arts needs of the users, motivate laser printer manufacturers to include software which computes all sizes and all orientations from one or more outlines of the type. This computation is called 'rasterization on the fly'.

The first widely available printer with this capability was the Apple Laser-Writer, introduced in 1985. It uses the PostScript page-description language, which has since become widely accepted as an industry standard. It is available on several dozen printing systems with software written either by its originators, Adobe Systems, or by a number of PostScript clone manufacturers. Although it is generally regarded as typographically inappropriate,[2] PostScript makes a linear transformation of a single outline for each character in a given typeface and computes, on demand, the correct pixels to darken from the scaled outline.

Linear scaling is certainly inappropriate on purely visual grounds because human visual response is not a linear function – in fact, not even a monotonic function – of spatial frequency (Ref. 37, p. 155). This may imply that linearly scaled type is, under some circumstances, less legible. In particular, since visual response drops off at lower as well as higher spatial frequencies, psychophysical evidence would predict that less information is actually presented to the visual cortex by larger type and smaller type than the type around 10 point which typographers have found optimal for text faces. In particuar, legibility should be increased for small sizes by widening them to decrease the spatial frequencies present. This is in accord with typographic wisdom.

Although a rigorous relationship between legibility and readability has not been established, some recent evidence in this direction is discussed below.

Five centuries of printing history notwithstanding, its wide commercial acceptance suggests that linear scaling on an inexpensive printer represents an acceptable economic tradeoff, at least among the consumers of laser-printed documents. Note that this class of documents may itself be of a restricted nature, since laser printers are used largely in offices. PostScript typesetters have not made nearly the impact that PostScript laser printers have, although it is too soon to tell whether this is for economic or for design reasons. On the other hand, Adobe Systems' PostScript outlines apparently are optimized for 10 to 12-point rendering. Those are historically the most common text sizes, probably for visual reasons described above. Thus, it is possible that the deficiencies are not blatant enough at those sizes to bar acceptance of linear scaling as a way to size type. Traditionally, the smaller a typeface is in height, the more the designer expands its width and darkens it for greater legibility. This tradition is threatened by the current designs of on-the-fly rasterisation systems, but could well be revived as the skill of the software punchcutters improves. The author considers the current disregard for this typographic tradition as an anomaly ascribable to the youth of the discipline.

Automatic scan conversion involves several important problems besides how to size and shape the letters. These problems might be thought of with the distinction made by Southall[40] between *shape* and *appearance*, the former often being more under control of the letterform designer than the latter (indeed, Southall concludes that present typeface description languages cannot adequately support control of appearance). One such critical issue is 'grid fitting', which roughly describes the choices that must be made when the width of a character's stroke is not an integral number of pixels. Clearly, the problem is exacerbated at low resolutions. The principal non-proprietary work on this issue is that of Hersch,[16] which describes and solves some of the problems of maintaining fixed relationships among character parts as the outline is moved around the grid. Hersch's scheme also addresses curve-generation errors, especially those induced by scaling outlines down. Grid fitting can also uniformise the inter-stroke spacing, which might vary significantly with single-pixel errors at small sizes In the spatial frequency domain, this gives it an advantage similar to that gained by proportional faces over fixed-width faces. Due to their non-uniform inter-stroke spacing, fixed-width faces tend to have spatial frequency energy smeared over the spectrum, which may make them harder to detect. Rubinstein illustrates this (Ref. 34, pp. 46–47) and the author discusses similar issues in Ref. 28.

Proprietary systems, most notably those of Adobe Systems, Bitstream, Inc. and Folio, Inc. (recently acquired by Sun Microsystems), add so-called 'hints'—proprietary additions to the data structures which describe the splines which in turn represent the character outlines. Some aspects of the Bitstream system are described in Ref. 1. The hints in such systems are used by the rasterisation software to improve the pixel choices and increase the uniformity of appearance of character pieces both within the character and across various size manipulations of the typeface. Hints, and the rendering algorithms using them, form the major source of revenue protection for typeface vendors and the printer vendors reselling their type. That is because typeface designs cannot at present be protected by patent or copyright in North America, although in Europe they can be copyrighted. Very recently, controversial software patents have been issued in the United States, and it is conceivable that vendors could argue in the future that, as with Metafont, their outlines are programs and thus patentable.

6. VIDEO DISPLAYS

By and large, the current technology of automatic scan conversion has not produced widely acceptable fonts at screen resolutions, which are typically 75–100 dpi. This is a real problem for document production systems which attempt visual fidelity between the screen and the printed page ('What you see is what you get, or WYSIWYG). Because a typical screen has $\frac{1}{4}$–$\frac{1}{3}$ as many pixels on each axis as a laser printer has, the problems mentioned above become even more severe.

There is the additional problem of matching the setwidths of the screen font to those of the final output device. The composition software must calculate line breaks using the widths which the final output device uses. If the screen widths are too much wider than the printer widths, the letters will crash against one another; if they are too much narrower, excessive inter-character space will result. At least two approaches to the setwidth matching problem have appeared in print,[26, 13] but it appears to be largely ignored, probably because screens of adequate resolution to make good representations of a wide variety of letters have been largely in the domain of engineering workstations until recently. Dramatic decreases in cost of such displays and of computers fast enough to deal with this kind of software have occurred in the last year or so, which may lead to more interest in the setwidth-matching issue.

7. GREYSCALE DISPLAYS

One inexpensive way to improve the apparent resolution of screens is to take advantage of the human visual system's willingness to trade intensity resolution for spatial resolution. Roughly speaking, when the edge of a character is 50% grey, its appearance is of an edge one-half pixel wide. Choosing the appropriate grey level for an edge pixel is a matter of taking the weighted averages of some of the surrounding pixels. This process can be expressed in terms of blurring, or digital low-pass filtering of the character image. This idea was first suggested by Warnock,[45] who found that the precise shape of the digital filter had little effect, although its width (that is, the extent of adjacent pixels over which the averaging takes place) did. He also argued that about four bits (16 grey levels) seemed to provide all the discriminability possible for text. Both of Warnock's results remain uncontested by later experimenters.

Image-processing theory can account for the apparent unimportance of filter shape. Further, this can be done in ways consistent with our knowledge of human vision. The basic issue is that none of the useful filters can affect that portion of the information which accounts for position.[12] As to more than four bits of resolution being irrelevant, one should observe that this is not the case in black-and-white vision in general, which has a contrast discrimination as low as 0.1–0.4%, i.e. requires about 9–10 bits.[30] Most of the current psychophysical research is in agreement with this, but it is usually carried out on images with spatial frequencies well below those which comprise serifs and other small features of text, and this may be a factor. Also, it is possible, though unlikely, that text has different properties from general images in the world. More pyschophysically rigorous experiments with greyscale fonts might settle whether increasing the contrast resolution beyond four bits has any advantage. (For example, the greyscale font literature does not appear to contain any reports of experimentation that controls carefully for the now well-understood non-linearity of spatial frequency response).

8. REPRISE OF HUMAN VISION, LEGIBILITY AND READABILITY

Among psychologists, the classic study of text readability is Tinker.[42] This work took form over many decades, and was published 15 years before the appearance of currently accepted models of 'early human vision' (i.e. the first stages of vision, not the vision of young humans), described in Ref. 37 or Ref. 10. The spatial frequency analysis of these models remains largely ignored in the literature of perception researchers who study visual cues in reading. But a series of rigorous psychophysical studies involving text on screens is reported in a collection of papers by Legge *et al.*[25] and its referenced predecessors. Experimental results in those studies (e.g. Ref. 25, p. 247) give reading-rate curves as a function of spatial frequency and give a shape which is generally consistent with psychophysical contrast sensitivity curves. This suggests that legibility and readability may, indeed, be highly correlated, and that psychophysical measures of legibility should be good predictors of readability.

Legge's data are taken with text on CRTs, which have a wide variety of visual differences from paper, including font quality, light emission instead of reflection, flicker, lower contrast and many others. No one of these factors accounts for the 20–30% reduction measured in reading rates for CRTs compared to paper (Ref. 34, pp. 189 ff.). Psychophysical experiments are done on CRTs because they give the investigator a greater measure of control, but nothing in contemporary vision models inherently makes them inapplicable to paper. Legge's results correspond to traditional typographic wisdom, which asserts that type around 10 point is the most legible for adult readers with normal vision. The fact that larger type is harder, not easier, to read is conventionally explained by arguments about requiring more eye motion,[29] but the Legge data, which are taken with moving text, suggest that more fundamental mechanisms may be relevant, as we have proposed above in the discussion of non-linear scaling. Some of the other ways in which traditional typographic wisdom can be seen to be consistent with current psychophysical theory can be found in Refs 6, 28 and 12.

9. SUMMARY

Type rendering has recently entered a new technological era. As seems to have happened in every previous change of technology, the quality of the resulting images is initially driven by economics and lack of experience in the new medium. The designs which evolved for the old technology are often ill adapted to the new. There is no reason to believe that the many orders of magnitude of computational enhancement given humans by computers will translate to a corresponding increase in the rate at which humans exercise the new technology to achieve its highest possible quality. Progress in this basically intellectual problem is limited by the rate at which designers and programmer-artisans progress toward understanding the relationship of the new technology to the fundamental impact of human vision and cognition on the task of reading. One should not expect this understanding to accelerate due to the acceleration of technology, except possibly for the decreasing real costs, which make the tools more rapidly available to a wider collection of practitioners. Perhaps this will make magnificent digital type available a mere 10 years hence, instead of the 100 or so from the invention of movable type to the introduction by Claude Garamond of type for which 'three generations of punchcutters laboured to create additional sizes',[2] and which even today is rendered, for better or worse, on millions of office machines. If this dramatic speed-up takes place, we will soon be rewarded. If the more tradtional pace obtains, we may see fine laser printer type by 2075, the approximate 100th anniversary of the availability of the Xerox XGP to a few explorers in digital typography.

10. ABOUT THE REFERENCES

In addition to the works already cited, the works below are of interest to students of the subjects discussed here.

The book *Digital Typography*, by Richard Rubinstein,[34] is a lucid and comprehensive study of digital type and composition systems. Its substantial annotated bibliography is of great value. The *Scientific American* article 'Digital Typography', by Charles Bigelow and Donald Day[3] has broader scope than the present paper, and it remains, as characterized in Rubinstein's annotated bibliography, 'the best introduction to the issues of digital typography in an article that can be read in one sitting'.

The journal *Visible Language* has several special species issues of interest, especially volume 15, number 2 (*Visual Cues in Word Recognition and Reading* and volume 19, number 1 (*The Computer and the Hand in Type Design, Proceedings of the Fifth ATypI Working Seminar, Part I*). The new journal *Electronic Publishing*, published by John Wiley & Sons, is devoted to all aspects of electronic publishing.

Avi Naiman's master's thesis[31] contains more detail about issues arising in the production of greyscale fonts,and his impending Ph.D. thesis will report on some careful experimentation on legibility of greyscale fonts.

The introductory text, *Perception*, by Robert Sekuler and Randolph Blake[37] is somewhat more current about spatial vision than the classic and still useful book of Cornsweet.[7] *Seeing*, by John Frisby[15] is also noteworthy. The recent book, *Spatial Vision*, by Russell and Karen DeValois[10] is slightly more mathematically sophisticated that either of the above, and is a useful reference when reading the psychophysical literature.

Writing Systems, by Geoffrey Sampson[36] is a study of written languages using linguistic principles usually applied only to spoken languages. Inquiries about the history of writing might well start here, before proceeding to the classical detailed works of Diringer[11] and Jensen.[18]

Production for the Graphic Designer, by James Craig,[9] has excellent surveys of the various stages in the production of typeset material.

Letters of Credit: A View of Type Design, by Walter Tracy,[44] is a handsome and inexpensive hardcover book about the art, craft and history of type design.

Acknowledgements and notes

Celia Morris, Karl Berry and K. Hargreaves suggested substantial improvements in the exposition. Katherine Nitchie and Chuck Bigelow corrected several factual errors and obscurities. Thanks to Les Earnest for digging out the Stanford XGP history and telling me the history of the early electrographic printers. Chuck Bigelow's writing, and many personal and electronic communications with him, have been particularly informative and influential (Bigelow's writing for the type community was probably the earliest to suggest that visual psychophysics could throw light on type design). One anonymous referee made several helpful criticisms about some factual errors and oversimplifications in the original manuscript, and pointed out that photographic halation effects work somewhat in favour of acceptable results at medium typesetter resolutions.

The following are trademarks: PostScript (Adobe Systems), Lucida (Bigelow and Holmes), ITC Stone (International Typeface Corporation), TeX (the American Mathematical Society), Metafont (Addison-Wesley Publishing Company), Apple LaserWriter (Apple Computer) Interpress (Xerox Corporation).

REFERENCES

1. Phillip G. Apley, Automatic generation of typographic images from outline masters. In *Proceedings of the Third Annual Artificial Intelligence and Advanced Computer Technology Conference*, 219–229 (1987).
2. Charles Bigelow, Form, pattern and texture in the typographic image. *Fine Print* **15** (2), 75–82 (1989).
3. Charles Bigelow and Donald Day, Digital typography. *Scientific American* **249** (2), 75–82 (1983).
4. Charles Bigelow and Kris Holmes, The design of Lucida: an integrated family of types for electronic literacy. In *Text Processing and Document Manipulation*, Cambridge University Press, Cambridge (1986). (Proceedings of the EP-86 Conference, Nottingham, 14–16 April 1986.)
5. Neenie Billawala, Pandora: an experience with Metafont. *Preprint* (1988).
6. Michael Brady, Toward a computational theory of early visual processing in reading. *Visible Language* **15** (2), 183–215 (1981).
7. Tom N. Cornsweet, *Visual Perception*. Academic Press, New York (1970).
8. Phillipe J. M. Coueignoux, Generation of Roman Printed Fonts. *Ph.D. thesis*, Massachusetts Institute of Technology (June 1975).
9. James Craig, *Production for the Graphic Designer*. Watson-Guptill Publications, New York (1974). (Published in the U.K. by Pitman).
10. Russell L. DeValois and Karen K. DeValois, *Spatial Vision*. Oxford University Press, New York (1988).
11. David Diringer, *The Alphabet*, 3rd edn, 2 vols, p. 151. Hutchinson London (1968).
12. Guozhen Duan and Robert A. Morris, The importance of phase in the spectra of text. *Electronic Publishing* **2** (1), 47–59 (1989).
13. Les Earnest, Making WYSIWYG characters shape up. In *Proceedings of ProText IV*, edited J. J. H. Miller, pp. 86–93. Boole Press, P.O. Box 5, Dun Laoghaire, Co. Dublin, Ireland (1987).
14. Les Earnest, The XGP at Stanford (Personal communication, 1989).
15. John P. Frisby, *Seeing: Illusion, Brain and Mind*. Oxford University Press, Oxford (1980).
16. Roger D. Hersch, Character generation under grid constraints. *Computer Graphics* **21** (4), 243–252 (1987). (Proceedings of SIGGRAPH-87.)
17. Allen V. Hershey, A computer system for scientific typography. *Computer Graphics and Image Processing* **4**, 373–385, (1972).
18. Hans Jensen, *Sign, Symbol and Scripts*, 3rd edn. Translated from the German by George Unwin. Allen and Unwin, London (1970).
19. Peter Karow, *Digital Formats for Typefaces*. URW, Kereienkoppel 55, 2000 Hamburg 65, (1987).
20. Donald E. Knuth, *Computers and Typesetting*, vol. E, *Computer Modern Typefaces*. Addison-Wesley, Reading, Massachusetts (1986).
21. Donald E. Knuth, Lessons learned from Metafont. *Visible Language* **19** (1), 35–54 (1985).
22. Donald E. Knuth, Mathematical typography. *Bulletin of the American Mathematical Society*, **1** (2), n.s. 337–372 (1979). (1978 Josiah Willard Gibbs Lecture.)
23. Donald E. Knuth, *Computers and Typesetting*, vol. C, *The Metafont Book*, Addison-Wesley, Reading, Massachusetts (1986).
24. Donald E. Knuth, *The TeXbook*. Addison-Wesley, Reading, Massachusetts (1984).
25. Gordon E. Legge, Gary S. Rubin and Andrew Luebker, Psychophysics of reading. V. The role of contrast in normal vision. *Vision* **27** (7), 1165–1177 (1987).
26. David Mellinger and Katherine Nitchie, Font matching with Flexifonts. In *Proceedings of ProText IV*, edited J. J. H. Miller, pp. 111–117. Boole Press, P.O. Box 5, Dun Laoghaire, Co. Dublin, Ireland, (1987).
27. H. W. Mergler and P. M. Vargo, One approach to computer assisted letter design. *Journal of Typographic Research* **2** (4), 299–322, (1968).
28. Robert A. Morris, Image processing aspects of type. In *Document Manipulation and Typography*, edited J. C. van Vliet. Cambridge University Press, Cambridge (1988).

(Proceedings of the International Conference on Electronic Publishing).

29. Robert E. Morrison and Albrecht-Werner Inhoff, Visual factors and eye movements in reading. *Visible Language* **15** (2), 129–146 (1981).
30. Jacob Nachmias and Richard V. Sansbury, Grating contrast: discrimination may be better than detection. *Vision Research* **14**, 1039–1042 (1974).
31. Avi Naiman, High Quality Text for Raster Displays. *Master's thesis*, Department of Computer Science, University of Toronto (1985).
32. A. Leo Oppenheim, *Ancient Mesopotamia, Portrait of a Dead Civilization* rev. edn (completed by Erica Reiner), p. 242. University of Chicago Press, Chicago, (1977).
33. William Pratt, *Digital Image Processing*. Wiley Interscience, New York (1978).
34. Richard Rubinstein, *Digital Typography, an Introduction to Type and Composition for Computer System Design*. Addison-Wesley, Reading, Massachusetts (1988).
35. Lynn Ruggles, *Letterform Design Systems*. Technical Report STAN-CS-83-971, Department of Computer Science, Stanford University (April 1983).
36. Geoffrey Sampson, *Writing Systems*. Stanford University Press, Stanford, Calif. (1985).
37. Robert Sekuler and Randolph Blake, *Perception*. Alfred A. Knopf, New York (1985).
38. David R. Siegel, *The Euler Project at Stanford*. Department of Computer Science, Stanford University, Stanford, Calif. (1985).
39. Richard Southall, Designing new typefaces with Metafont. In *TeX for Scientific Documentation*, edited J. Désarménien. Springer-Verlag, New York (1986).
40. Richard Southall, Shape and appearance in typeface design. In *ProText III*, edited J. J. H. Miller. Boole Press, P.O. Box 5, Dun Laoghaire, Co. Dublin, Ireland (1986).
41. Sumner Stone, The Stone family of typefaces: new voices for the electronic age. *Fine Print* **14** (3) (1988).
42. Miles Tinker, *Legibility of Print*. Iowa State University Press, Ames, Iowa (1963).
43. Virginia K. M. Tobin, The OCLC roman family of fonts. *TUGboat, The TeX Users Group Newsletter* **5** (1), 36–46 (1984).
44. Walter Tracy, *Letters of Credit: a View of Type Design*. David R. Godine, Boston (1986). (U.K. edition published by Gordon Fraser Gallery, London).
45. J. E. Warnock, The display of characters using gray level sample arrays. *Computer Graphics* **14** (3), 302–307 (1980). (Proc. SIGGRAPH 81).
46. Hugh R. Wilson and Whitman A. Richards, Mechanisms of contour curvature discrimination. *Journal of the Optical Society of America*, A **6**, 106–115 (1989).
47. Wang Xuan, Key problems in developing an advanced Chinese text processing and typesetting system. In *Proceedings of ProText IV*, edited J. J. H. Miller, pp. 65–76. Boole Press, P.O. Box 5, Dun Laoghaire, Co. Dublin, Ireland (1987).
48. Hermann Zapf, Future tendencies in type design: the scientific approach to letterforms. *Visible Language* **19** (1), 23–34, esp. pp. 28–29 (1985).

Book Review

J. André, R. Furuta and V. Quint (editors)
Structured Documents

Cambridge University Press (Cambridge Series on Electronic Publishing, ed. P. Hammersley, 2), 1989. 220pp. ISBN 0 521 36554 6. £35.

This book is based on a series of lectures given at an INRIA workshop in 1987. The first half of the book consists of three major papers that introduce the concepts of structured documents, look at document models and standards, and provide a survey of systems for manipulating structured documents. The second half is made up of six shorter papers considering the relationship between structured documents and areas such as document design, semantics, linguistics, and optical recognition of document structure.

Most of the authors have undertaken significant work in this field and are well known for their contributions. The book thus provides a sound introduction to the importance of structure in documents as well as a competent state-of-the-art guide to current systems. Anyone seriously interested in document-handling systems should read the introductory and survey papers in this book.

As the reviewer, I had two complaints about the book. The first is that it suffers from the usual problems associated with any book that is effectively a collection of separate papers. Some introductory material is repeated, material on standards and existing document-manipulation systems is scattered through the different papers in a manner that is occasionally confusing and irritating, and the index provided at the back of the book is woefully incomplete. The second complaint is simply that there is no consideration of the importance of document structures for hypertext. This is an unfortunate omission in view of the increasing importance of hypertext and the strong similarities between the needs of interactive document editors and hypertext systems. In spite of these complaints, however, I recommend the book as a worthwhile investment.

Heather Brown
Kent