

УДК 681

Долгова Т. А., доцент

## УЧЕТ ВЕРТИКАЛЬНОЙ ЕМКОСТИ ПРИ ОПРЕДЕЛЕНИИ ЭКОНОМИЧНОСТИ КОМПЬЮТЕРНЫХ ШРИФТОВ

Alongside the number of characters in the line the optimal number of lines in the image area is introduced as an important criterion for font efficiency estimation. Methods of calculating the mathematical expectation of the character width and principles of selecting the typeface leading are developed.

Современные настольные издательские системы обладают большим набором компьютерных гарнитур для всевозможных вариантов шрифтового оформления печатных изданий.

Модификация стандартных полиграфических шрифтов и появление все новых шрифтовых гарнитур привели к тому, что старая справочная информация безнадежно устарела. Свой вклад в изменение метрических характеристик шрифтов внес и переход к англо-американской полиграфической системе измерений, принятой в компьютерных технологиях.

Вместе с тем выбор варианта оформления будущего издания должен базироваться на соблюдении как эстетических, так и экономических требований. Выбор оптимальных параметров верстки и достаточно точные расчеты объема многостраничных изданий на донпечатной стадии подготовки позволяют снизить себестоимость полиграфической продукции [1]. Наличие точной метрической информации о компьютерных гарнитурах будет способствовать созданию профессионально подготовленных изданий.

Известно, что предварительный объем издания рассчитывается исходя из числа строк  $C$  полосы набора и среднего количества символов в строке для конкретного шрифта  $Z$ . Емкость полосы заданного формата определяется по классической формуле

$$E_{\text{п.н.}} = C \cdot Z, \quad (1)$$

где число строк  $C$  рассчитывается делением высоты полосы на интерлиньяж (в одних и тех же единицах измерения: в пт. или мм); емкость строки  $Z$  получается делением длины строки на среднюю ширину одного знака.

Современные технологии компьютерного набора повлекли за собой пересмотр многих устоявшихся понятий. Что можно отнести и к обсуждаемому понятию емкости шрифта. Эта величина должна в полной мере учитывать емкость всей полосы набора (1), а не только отдельной строки. Раньше число строк  $C$  полосы определенного формата зависело только от кегля, а интерлиньяж не был жестко связан с конкретной гарнитурой. Для компьютерных шрифтов эта связь существенна, и интерлиньяж, равный кеглю, уже не правило, а исключение. Большинство гарнитур для нормального восприятия текста требует увеличенного интерлиньяжа. Использование значений, заданных по умолчанию в верстальных пакетах (120% или +2 пт. от кегля), часто не позволяет получить оптимальный результат. Определение интерлиньяжа «на глаз» усложняет предварительный расчет объема издания на этапе выбора шрифтового оформления.

Конечно, интерлиньяж, обеспечивающий удобочитаемость, во многом связан с рисунком конкретных шрифтов, с его графическим характером. Рекомендации по его выбору должны учитывать контрастность, насыщенность, пропорциональность и другие характеристики шрифта, немалую роль играет формат наборной строки. Но при-

менительно к экономичности следует рассмотреть некоторое минимальное межстрочное расстояние, которое обеспечит нормальное восприятие текста. Эта величина и будет характеризовать вертикальную емкость конкретной гарнитуры.

Таким образом, актуальной является задача комплексного исследования емкости строки и вертикальной емкости компьютерных шрифтов. Рассмотрим сначала первый из сомножителей формулы (1).

Как известно, для одного и того же кегля в разных гарнитурах ширина одинаковых знаков довольно заметно отличается. Математически ожидаемая ширина знака  $E$  (по-другому — среднеуточненная ширина) учитывает ширину отдельной буквы  $e_i$  и вероятность ее появления в тексте  $p_i$  (удельную частоту встречаемости) в соответствии с формулой

$$E = \frac{e_1 \times p_1 + e_2 \times p_2 + \dots + e_n \times p_n}{p_1 + p_2 + \dots + p_n}, \quad (2)$$

где  $e_1, e_2, \dots, e_n$  — ширина знаков;  $p_1, p_2, \dots, p_n$  — удельная частота встречаемости соответствующих знаков русского языка [3].

От точности значения  $E$  для конкретной гарнитуры и кегля зависит степень соответствия расчетов реальным объемам, которые получаются после верстки издания.

Для определения ширины отдельных букв и математически ожидаемой ширины знака по описанной в [2] методике строка определенного формата целиком набирается одной из строчных букв русского алфавита. В этой строке рассчитывается количество знаков и определяется ширина одного знака  $e$ . После вычисления ширины всех букв алфавита  $E$  находится по формуле (2). В этой формуле знаменатель, представляющий собой сумму вероятностей всех знаков текста, должен быть равен единице. Но исключение из рассмотрения прописных букв, относительно редко встречающихся, значительно сокращает объем работы, практически не влияя на результат. Поэтому при учете только строчных букв сумма вероятностей несколько меньше единицы.

Определению среднеуточненной ширины знака компьютерных шрифтов посвящена и статья [1], авторы которой рассмотрели зависимость ширины знака от размера кегля. Но в данной заметке рассмотрены всего несколько гарнитур.

Расчет метрических характеристик большого числа компьютерных гарнитур требует уточнения методики и автоматизации расчетов. Для решения этой задачи может быть написана программа на каком-либо языке программирования. Однако использование стандартного программного обеспечения компьютерных издательских систем является более предпочтительным. Это позволит оператору НИС, или любому другому пользователю компьютера, самостоятельно и быстро определить необходимые параметры шрифтов.

При реализации такого подхода используется редактор для набора и верстки текста (PageMaker, Word, CorelDRAW или др.) и любая программа, имеющая возможность работы с электронными таблицами (можно даже использовать вычисляемые таблицы Word).

В текстовом редакторе создается полоса набора определенной ширины  $H$  (в мм); ширина должна быть достаточно большой и удобной для расчетов. Например, все приведенные ниже значения получены с использованием полосы набора редактора Word с шириной  $H = 200$  мм. Каждой строчной буквой русского алфавита набирается отдельный абзац, содержащий примерно полторы строки знаков. Когда курсор находится перед последним знаком полной строки, его местоположение (номер позиции по строке состояния) равно искомому количеству знаков  $n_i$ . Перемещая курсор вниз, остается только записывать длины полных строк. Если нужно изменить гарнитуру или кегль, то форматируется этот же файл. При этом «лишние» знаки в неполных строках позволяют не добирать буквы при увеличении числа символов в строке. Когда же число символов в строке

уменьшается, то неполная строка отдельного абзаца, куда переходят непоместившиеся буквы, не влияет на строку со следующей по алфавиту буквой.

Параллельно с текстовым файлом алфавита создается электронная таблица, которая может иметь примерно следующую структуру.

Первый столбец — буквы русского алфавита; второй — вероятность их появления  $p_i$ , последняя ячейка содержит автоматически вычисляемую сумму  $\sum p_i$ , т. е. знаменатель формулы (2); в третий столбец заносится количество  $n_i$  знаков в строке шириной  $H$ .

Табл. 1 демонстрирует пример электронной таблицы Excel описанной структуры (столбцы в этой программе обозначаются буквами А, В, С, ...). Полу жирным начертанием в примере выделены заголовки пользователя, т. е. собственно заполнение таблицы начинается с третьей строки.

Следующие столбцы предполагают расчет по формулам. Ширина знака  $e_i$  получится в четвертом столбце делением длины строки  $H$  на значение  $n_i$  из соответствующей ячейки предыдущего (третьего) столбца. Пятый столбец предназначен для перемножения значений второго и четвертого столбцов, т. е.  $p_i$  на  $e_i$ . Сумма содержимого всех ячеек пятого столбца  $\sum p_i e_i$  — числитель формулы (2), это значение может быть помещено в конце столбца. Затем в любой удобной ячейке рассчитывается искомая ширина  $E$  делением числителя формулы (2) на знаменатель (последние ячейки 2-го и 5-го столбцов).

В вычисляемых ячейках табл. 1 представлены соответствующие формулы (в столбцах D, E и двух последних строках).

Во время работы текстовый редактор с файлом алфавита и электронную таблицу удобно использовать одновременно, разместив их на экране рядом друг с другом (каскадом). Тогда число символов в строке  $n_i$  можно сразу заносить в таблицу.

Для оценки точности получаемых результатов были проведены эксперименты, позволяющие найти погрешность определения ширины каждой буквы отдельно и математического ожидания ширины знака для заданных гарнитур и кегля.

Для определения точного значения  $e_i$  использовался специальный редактор шрифтов FontLab. Здесь при описании символа шрифта используется координатная сетка  $1000 \times 1000$  единиц, 1000 единиц координатного пространства соответствуют 1 пункту. Ширина символа считалась от левой до правой границы символа (в этом случае кроме ширины собственно буквы учитываются и боковые поля символа). Полученное значение домножалось на кегль  $K$ .

В исследовании использовались шрифты 10, 12 и 14-го кеглей для гарнитур Lite-*turnaya* и Times New Roman Cyr (метрические параметры символов этой гарнитуры несколько отличаются от Times New Roman).

Таблица 1

Таблица Excel для расчета математического ожидания ширины символа

1	А	В	С	Д	Е
2	буква	$p_i$	$n_i$	$e_i = H/n_i$	$p_i \times e_i$
3	а	0,0748	118	200/D3	B3*D3
4	б	0,0147	106	200/D4	B4*D4
5	в	0,0415	118	200/D5	"
			...		
33	я	0,0079	121		"
34	сумма:	СУММ(B3:B33)			СУММ(B3:B33)
35	ширина:	E34/B34			

Для каждого  $i$ -го символа по описанной методике определялась ширина  $e_i$ , которая сравнивалась с точным значением  $e_{FLi}$ , для определения относительной ошибки использовалась формула

$$\Delta e_i = \frac{e_i - e_{FLi}}{e_{FLi}} \cdot 100\% . \quad (3)$$

Среднеарифметическая величина  $\Delta e$  по всем  $\Delta e_i$  представлена в табл. 2.

Точная ширина символов использована для определения математически ожидаемой ширины  $E_{FL}$ , относительно которой рассчитана ошибка  $\Delta E$ .

Погрешности  $\Delta e$  и  $\Delta E$ , как и следовало ожидать, имеют один порядок, однако разная удельная частота встречаемости букв приводит к небольшому отличию в точности определения среднеарифметической и среднеуточненной ширины.

Использованные гарнитуры имеют существенно отличающиеся друг от друга абсолютные значения  $E$ . Но в обоих случаях и при разных кеглях ошибка определения математического ожидания не превышает 1%.

Как известно, шрифты контурных форматов (Type 1, TrueType) хранятся в закодированном виде, а для воспроизведения символа нужного кегля специальные программы производят масштабирование контура символа и его заполнение (растрирование) в соответствии с разрешением устройства, на котором отображается текст. Таким образом, символы разного кегля одной и той же гарнитуры должны иметь одинаковые относительные размеры, т. е. для любого кегля  $K$  должно выполняться  $E/K = const$ .

Однако пересчет в абсолютные координаты точек символа происходит с округлением нецелых результатов. При этом могут проявляться некоторые специфические, особенно для малых кеглей, ошибки: нарушение пропорции символа, симметричности и т. д. Эти проблемы решаются с помощью специальных встроенных алгоритмов, которые в общем-то могут привести к небольшому отличию относительных размеров одинаковых символов разных кеглей [4].

Последняя строка табл. 2 демонстрирует очень близкие значения относительной ширины символа  $E/K$  при 10, 12 и 14-м кегле для двух гарнитур. Похожие цифры приведены в [1] для трех гарнитур. Для более детального анализа были проведены исследования 12-ти гарнитур для кеглей: 6, 8, 9, 10, 11, 12, 13, 14, 16, 18, 20 пт. Для всех шрифтов было найдено отношение  $E/K$ , которое сравнивалось со среднеарифметическим значением соответствующей гарнитуры. Средний по всем кеглям разброс значений  $\Delta E/K$  для каждой гарнитуры представлен в табл. 3, где номерами 1, 2, 3, ..., 12 обозначены соответственно гарнитуры Journal, School, Peterburg, Kudrashov, Arial Cyr, Europe, Pragmatika, Text Book, Encyclopedia, Bodoni, Bruskovaya, Bukvarnaya.

Как видно из табл. 3, отклонения в относительной ширине знака не превышают несколько процентов. Таким образом, зная математическое ожидание ширины знака некоторого кегля  $K^*$ , можно найти величину  $E$  для нужного кегля  $K$  той же гарнитуры, используя коэффициент  $K/K^*$ .

Таблица 2

**Математическое ожидание ширины символа для разных кеглей**

$K$ , пт.	Times New Roman Cyr			Literaturnaya		
	10	12	14	10	12	14
$E$ , мм	1,875	2,252	2,627	1,966	2,366	2,754
$E_{FL}$ , мм	1,866	2,239	2,612	1,956	2,347	2,738
$\Delta e$ , %	0,43	0,54	0,59	0,47	0,68	0,56
$\Delta E$ , %	0,45	0,54	0,55	0,48	0,78	0,56
$E/K$	0,1875	0,1877	0,1876	0,1966	0,1972	0,1967

Разброс значений относительной ширины знака

Гарнитура	1	2	3	4	5	6	7	8	9	10	11	12
Среднее $\Delta E/K$ , %	1,97	0,84	1,39	1,51	2,59	1,63	1,82	2,03	1,90	1,59	2,33	2,34

Сравнительный анализ ширины букв одного и того же кегля для различных гарнитур позволяет судить об экономичности шрифта по величине, обратно пропорциональной  $E$ . Эту величину  $q = 1 / E$  принято называть емкостью шрифта.

Так, при 10-м кегле для гарнитуры Times New Roman Cyr  $q = 0,53$ , Times New Roman — 0,57, Literaturnaya — 0,51, Journal — 0,51, Bodoni — 0,56, Antiqua — 0,54. Небольшая разница значений  $q$  на самом деле влечет значительную разницу в емкости строки набора. Например, при формате строки 6 кв. ее расчетная емкость может быть от 55 до 62 букв (для приведенных гарнитур), что при 50 строках полосы набора может привести к разнице в 350 символов (соответствует примерно 6 строкам).

Рассмотрим теперь, какие параметры шрифтовой гарнитуры влияют на вертикальную емкость.

Как известно, кегль включает в себя высоту буквы плюс некоторое расстояние для образования межстрочного просвета. Это расстояние при одинаковом кегле для разных компьютерных гарнитур существенно различно. Линия верхних выносных элементов может быть выше или ниже линии высоты прописных букв, надбуквенные значки (например, у буквы «й») тоже имеют разное расположение, существенно отличается высота очка строчной буквы. Межстрочное расстояние должно, по крайней мере, обеспечить нормальное отображение всех элементов шрифта, без «наползания» соседних строк друг на друга.

При описании статистической информации о конкретном цифровом шрифте в формате True 1 используется понятие минимального охватывающего прямоугольника шрифта — прямоугольника минимального размера, в который целиком помещаются все символы шрифта. Среди них имеются и редко используемые прописные буквы с надбуквенными элементами («Й») и специальные символы. Найти охватывающую высоту (т. е. высоту этого прямоугольника) для TrueType шрифтов можно с помощью какой-либо сервисной программы работы со шрифтами, а для True 1 шрифтов посмотреть непосредственно в заголовке шрифта (в AFM-файле).

Следует учитывать, что шрифты разных форматов и разных фирм-производителей с одинаковым названием гарнитуры могут иметь несколько отличающиеся метрические параметры. Например, для самой распространенной гарнитуры Times New Roman минимальный прямоугольник примерно в 1,12 раза больше кегля (1117 единиц в координатной сетке 1000 × 1000 формата True 1), а для Courier — всего в 1,05 раза.

В таких случаях интерлиньяж, равный кеглю, приводит к тому, что засечки и наплывы нижних выносных элементов практически касаются прописных букв следующей строки, а в Word такие буквы вообще не прорисовываются полностью. В настоящем издании, где также используется Times New Roman, при кегле 12 пт. интерлиньяж установлен 14 пт. (12 пт. × 1,12 = 13,44 пт., что при округлении в большую сторону и дает 14 пт.).

Определить относительную охватывающую высоту шрифта можно и с помощью графического редактора (CoreDraw или др.). Необходимо набрать в нем несколько характерных символов исследуемой гарнитуры, установить максимально большой кегль (например, 400 пт.) и определить необходимые размеры с помощью специальных средств редактора (в качестве единиц измерения желательно использовать пункты). Полученные значения следует разделить на величину кегля.

Именно это значение охватывающей высоты  $h$  шрифта определенного кегля и гарнитуры использует в качестве интерлиньяжа редактор Word при выборе межстрочного интервала — «Одинарный». Поэтому самый простой, но не самый точный, способ определения  $h$  состоит в использовании этой опции Word. Сначала определяется число строк полосы набора для конкретной гарнитуры, а затем вычисляется значение интерлиньяжа делением высоты полосы (в пунктах) на количество строк [5].

Интерлиньяж, равный  $h$ , гарантирует достаточный междустрочный просвет при применении любых символов. В этих же границах обычно находятся верхние и нижние индексы (чуть ниже выступают только индексы, имеющие нижние выносные элементы). В то же время это значение обычно меньше, чем интерлиньяж, устанавливаемый по умолчанию в верстальных пакетах.

Аналогично рассмотренной выше емкости строки  $q$  можно ввести понятие вертикальной емкости шрифта, которая будет рассчитываться по формуле

$$Q = \frac{1}{h}. \quad (3)$$

Поскольку интерлиньяж задается в пунктах, то в этих же единицах удобнее измерять охватывающую высоту  $h$ . Число строк полосы набора можно рассчитать по формуле  $C = QH$ . Так, при кегле 10 пт. для Times New Roman  $Q = 0,089$ , для Courier  $Q = 0,095$ . Тогда на полосе набора высотой 480 пт. поместится соответственно 42 и 45 строк. Если же использовать интерлиньяж, равный 120% кегля, — всего 39 строк.

Анализируя одновременно емкость строки и вертикальную емкость для компьютерных гарнитур, можно подобрать оптимальные параметры верстки и провести расчеты объема будущего издания.

## ЛИТЕРАТУРА

1. Волкова Л. А., Ревякова О. Н. Исследование параметров компьютерных шрифтов // Полиграфия. — 2001. — № 1. — С. 42—43.
2. Волкова Л. А., Решетникова Е. Р. Технология обработки текстовой информации. — Ч. 1. — М.: Изд-во МГУП, 1999. — 307 с.
3. Справочник технолога-полиграфиста: В 3-х ч. Ч. 1. Наборные процессы / М. В. Шульмейстер, Г. А. Таль. — М.: Книга, 1981. — 255 с.
4. Ярмола Ю. А. Компьютерные шрифты. — СПб.: ВHV — Санкт-Петербург, 1994. — 208 с.
5. Долгова Т. А., Гончарова З. В. Исследование емкости полосы набора для компьютерной технологии допечатной подготовки издания // Издательско-полиграфический комплекс на пороге третьего тысячелетия: Материалы Междунар. науч.-техн. конф. / БГТУ. — Мн., 2001. — С. 55—60.