СЕКЦІЯ
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ І АВТОМАТИЗАЦІЇ
ПОЛІГРАФІЧНОГО ВИРОБНИЦТВА

УДК 655.28: 681.3

ОБГРУНТУВАННЯ ОСНОВНИХ ПАРАМЕТРІВ КОМП'ЮТЕРНОГО ФОРМУВАННЯ ТЕКСТУ

І.В. Андріїв, В.М. Сеньківський

Аналізуються основні вихідні параметри процесів комп'ютерного опрацювання текстової інформації. Наводяться технологічні вимоги та пропонуються математичні моделі, що обгрунтовують способи і методи їх використання.

Анализируются основные входные параметры процессов компьютерной обработки текстовой информации. Приводятся технологические требования и предлагаются математические модели, обосновывающие способы и методы их использования.

За останнє десятиріччя книговидавнича діяльність в Україні зазнала відчутних змін, які пов'язані передусім з активним запровадженням прогресивних інформаційних технологій, що базуються на сучасній комп'ютерній техніці та програмних пакетах опрацювання текстової і графічної інформації й охоплюють як додрукарські (складальні, репродукційні), так і, в меншій мірі, друкарські процеси. Додрукарська підготовка видань майже повністю комп'ютеризована і здійснюється в основному засобами комп'ютерних видавничих систем, розроблених іноземними фірмами. Практично необмежені можливості програмних пакетів цих систем дозволяють готувати до друку видання довільної складності. Кардинально змінюються функції та розподіл обов'язків між підрозділами, що беруть участь у технологічному процесі підготовки і випуску друкованої продукції. Мінімальним стає коректурний обмін між видавництвом і друкарнею. Сучасне комп'ютеризоване (електронне) видавництво здатне самостійно підготувати остаточний варіант оригінал-макета — точну копію майбутнього видання. В залежності від потреб та наявного обладнання цей макет може бути переданий на електронному чи іншому носії для виготовлення друкарської форми або відтиражований без проміжних операцій, пов'язаних із формними процесами. Останній варіант набув поширення у зв'язку з появою цифрових друкарських машин. Досить ефективним може бути його використання в технологічних друкарських процесах оперативної поліграфії, де критичним є час виготовлення малотиражної продукції.

Незважаючи на значний поступ і доступність комп'ютерних технологій випуску друкованої продукції, глобальна комп'ютеризація в цій галузі не дала очікуваних результатів. Однією з основних причин такого стану є недостатній професійний рівень персоналу, що працює з комп'ютерними видавничими системами (КВС). Частими є порушення технологічних вимог щодо правил формування тексту та верстання сторінок: з одного боку, через їх незнання, з другого, через відсутність таких вимог стосовно сучасних технологій підготовки та випуску видань із застосуванням КВС. Задовільне знання основ комп'ютерної техніки та особливостей функціонування КВС недостатнє для досягнення необхідної якості видань, що готуються з використанням програмних засобів. В останні роки у зв'язку із значним поширенням видавничих організацій різного профілю і спрямування відчувається брак фахівців, які одночасно з навиками роботи в комп'ютерному середовищі володіли б достатньою видавничо-поліграфічною підготовкою. Усе це в деякій мірі призвело до невідповідності між можливостями комп'ютерних систем і потребами видавничо-поліграфічного комплексу.

Аналіз якості книжкової продукції, випущеної з використанням КВС, свідчить про наявність значних порушень. Вони мають різну природу і різні причини виникнення. Найбільш поширеним є недотримання вимог щодо розмірів пробілів між словами, що, в свою чергу, призводить до появи так званих коридорів в абзацах, тобто суміщення пробілів у декількох

сусідніх рядках. Так само неоднаковими бувають пробіли і в сусідніх рядках. Пов'язано це, перш за все, з відсутністю в деяких системах функцій розділення слів української мови переносом. У такому разі мінімальною неподільною частиною тексту стає слово, а не його склад, і величина пробілів тепер залежатиме від середньої довжини слів опрацьовуваного тексту. Перенесення слова, що не помістилося, у наступний рядок автоматично веде до збільшення сумарної величини пробілів на довжину цього слова. В результаті остаточне (реальне) значення пробілу

$$t_r = t_p + \frac{l_k}{k-1}, \qquad (1)$$

де t_p – початкове значення пробілу; l_k – довжина k-го слова в рядку; $k-1$ – кількість пробілів у рядку, що безпосередньо пов'язана з кількістю слів у ньому.

Часто на суміжних сторінках, розміщених на розгортках, вертикальні проміжки між рядками різні. Подібне спостерігається з проміжками, що розмежовують у виданні текст та ілюстрації, відділяють заголовки від тексту і т.п. Обтікання ілюстрацій текстом не завжди відповідає прийнятим вимогам. Не витримуються відповідності між посиланнями на ілюстрацію і фактичним місцем її розміщення. Порушення правил верстання сторінок частіше спостерігається в складних виданнях, які містять математичні та хімічні формули, заголовки, рисунки, виноски, таблиці, різні типи і форми виділення тексту, колонтитули, норми, сигнатури. Оскільки верстання сторінок невіддільне від їх художнього оформлення і в значній мірі залежить від професійного рівня, кваліфікації, нарешті, естетичного смаку виконавця, тобто є процесом творчим, повна автоматизація (читай комп'ютеризація) його практично неможлива. Це не означає заперечення фактичного стану справ. Автоматичне (програмне) верстання реалізоване для порівняно простих видань, автоматизоване (діалогове) — для складного тексту. У зв'язку з тим якість видання суттєво пов'язана із суб'єктивними факторами і не завжди вписується у формалізовані рамки комп'ютерної системи.

Отже, маємо різні фактори та причини, що прямо або опосередковано впливають на якість видань, одержаних засобами комп'ютерних технологій. Рангувати їх за важливістю чи мірою впливу на процес формування видання немає потреби, маємо реальну ситуацію та розуміння її недосконалості. Важливіше дослідити її та запропонувати методи і засоби, реалізація яких привела б до поліпшення становища в сучасному книгодрукуванні.

Звернемо увагу на особливості використання загальносистемних вихідних параметрів, які суттєво впливають на процес функціонування алгоритмів опрацювання текстової інформації та результат комп'ютерної підготовки видання.

Комп'ютерні видавничі системи, а точніше пакети прикладних програм, що складають їх основу, містять початкову інформаційну та метричну бази, призначені для формування структури і смислового наповнення майбутнього видання. Їх знання та вміле використання (кожного зокрема і в сукупності) суттєво впливають на технологічні особливості підготовки і випуску поліграфічної продукції, її якість та художню цінність.

Вихідні дані поділимо на основні та похідні. До перших віднесемо горизонтальний і вертикальний формати тексту, формати паперу, міжрядкові інтервали, абзацні відступи, гарнітури та розміри шрифтів, способи форматування (у тому числі вирівнювання) тексту. Похідними або такими, що певним чином залежать від перерахованих, вважатимемо пробіли між словами, міжсимвольні інтервали. Проаналізуємо способи задання і дії цих параметрів у класичних технологічних процесах складального виробництва та особливості їх використання в комп'ютерних видавничих системах.

Формат видання вибирають, зважаючи на такі фактори [1], як зручність і швидкість читання, характер тексту з врахуванням ступеня його складності, особливості користування виданням, його об'єм і тираж. Найбільша швидкість читання досягається, коли рядок містить 40–60 знаків. Це відповідає приблизно форматам від 3 до 5 квадратів для кегля 8, від 3,5 до 5,75 квадрата для кегля 9 і від 4 до 6,5 квадрата для кегля 10. Найбільш економічно вигідними вважаються співвідношення горизонтального і вертикального форматів сторінки 1:1,6; 1:1,8. Це відповідає форматам 4,5 × 8; 5,5 × 9; 7 × 12 квадратів. Формати паперу визначаються

характером видання. Для книжкових видань найуживанішими форматами є А4 (210 мм × 297 мм) та А3 (297 мм × 420 мм).

У комп'ютерних видавничих системах параметри сторінки найчастіше визначаються розмірами полів, тобто відстанню тексту від країв аркуша. Вони розраховуються залежно від встановлених форматів тексту та паперу. Задаються також відстані від краю тексту до верхнього чи нижнього колонтитула. При цьому слід враховувати, що корінцеве поле повинно бути найменшим, зовнішнє (бокове) більшим від верхнього, а нижнє – від зовнішнього. З огляду на це сторінка, тобто площа, заповнена текстом, розміщується на так званій оптичній середині.

Встановлення міжрядкового інтервалу в КВС звичайно здійснюється автоматично. Найчастіше його величина на 20% більша від кегля використовуваного шрифту і прямо залежить від довжини рядків. При заданні власного інтервалу слід пам'ятати, що надмірна щільність тексту погіршує його естетичний вигляд та ускладнює процес читання. Тут доцільно експериментально перевірити, як впливає на вигляд тексту кожне нове значення міжрядкового інтервалу.

Подібним чином поступають з абзацним відступом, який для кожної КВС може мати наперед задане або заздалегідь введене користувачем значення. Вимоги, котрі треба враховувати, стосуються не тільки естетичного вигляду тесту та процесу його конструювання, але й потреби перекриття абзацного відступу кінцевим рядком абзацу не менше ніж у півтора раза. Забезпечення цього може вплинути на величину пробілів між словами. У даній ситуації маємо взаємопов'язану сукупність вимог, виконання яких повинно бути узгодженим.

Одним з основних факторів, що суттєво впливає на формування видання та його зовнішній вигляд, є шрифтове оформлення. Графічна основа знаків визначає шрифти кириличний, латинський, грецький, готичний та інші. Стандартні шрифти розрізняють за рисунком, накресленням і розміром. Шрифти, що мають однаковий характер рисунка, об'єднуються в гарнітуру. В залежності від основних графічних ознак шрифти поділяються на шість основних та одну додаткову групи [1]. Це рублені, з малими засічками, медіювальні (різноширинні), звичайні (одноширинні), брускові, малоконтрастні. За нахилом основних штрихів розрізняють шрифти прямого, похилого та курсивного накреслень.

Відношення товщини основного штриха до висоти очка (від 0,23 до 0,34 і більше) дає насичення шрифту від світлого до жирного. В комп'ютерних шрифтах градації жирності набагато ширші і для деяких гарнітур сягають декількох десятків. Зміна товщини штрихів викликає зміну загальної щільності тексту, що може призвести до погіршення якості видання.

Щільність очка, тобто відношення між його висотою і шириною, визначає шрифти вузькі, нормальні, широкі і т. д. Так, наприклад, для букви "н" у різних гарнітурах при нормальній ширині ця величина складає 0,6–0,85. У широких шрифтах аналогічно одержимо 0,86–1,05.

Розмір або кегль, тобто висота шрифту, що містить висоту літери (очка) і частину вільного простору над і під очком (між верхньою і нижньою шрифтовими лініями), традиційно вимірюється в пунктах (1 пункт = 0,376 мм). У комп'ютерних шрифтах висота цієї умовної прямокутної рамки вимірюється в міліметрах і не завжди відповідає її типографському еквіваленту в пунктах.

В сучасних КВС шрифти звичайно векторні і масштабовані, їх називають ще "True Type-шрифти". Шрифти цього типу виглядають на екрані практично так само, як і на принтері, з врахуванням відмінностей, пов'язаних з різною роздільною здатністю цих пристроїв. Якщо треба забезпечити відповідність зображення на екрані його копії на папері, кожному принтерному шрифтові повинен відповідати свій екранний шрифт. Кожний кегль масштабованих шрифтів утворюється програмним способом на основі еталонного математичного опису форми літер, що містить комбінації прямих і кривих ліній. Такий підхід дозволяє економити пам'ять комп'ютера, оскільки для використання шрифтів певного кегля не треба заздалегідь формувати та зберігати окремі файли.

Унікальною особливістю комп'ютерних видавничих систем, з огляду на можливості шрифтового оформлення, є те, що введення (кодування) тексту можна здійснити одним шрифтом, вибравши певну гарнітуру, накреслення і кегль, потім у процесі роботи над виданням

у будь-який момент змінити шрифт в усьому виданні чи в окремих його частинах. Звідси випливає, що користувач системи (у даному випадку художній редактор) може експериментальним шляхом за короткий проміжок часу вибрати найкращий варіант шрифтового оформлення видання.

Форматування тексту згідно з існуючими вимогами або доведення його до заданого формату по горизонталі (формування рядків) і вертикалі (утворення сторінок) у КВС має свої особливості. Перш ніж перейти до викладення їх суті, розглянемо вихідні умови та вимоги, що стосуються формування рядків. Аналіз цих факторів дозволив поділити їх на три групи. До першої віднесено такі характеристики тексту, як горизонтальний формат, гарнітура і кегль шрифту, величина пробілу між словами та межі її зміни. Друга група стосується рядків, що містять підставний ініціал, абзацний відступ, різні способи форматування; рядків рубрик і виносок; правил організації колонтитулів, колонцифр, норм, сигнатур. Третю групу складають вимоги, виконання яких здійснюється після попереднього аналізу певних ділянок тексту. До них належать спеціальні правила переносів деяких словосполучень, наприклад, ініціалів і прізвищ, скорочених найменувань мір і чисел, до яких вони відносяться; правил кодування тире, номера, параграфа, багаторозрядних чисел.

Теоретичні викладки і технологічні обгрунтування, що стосуються правил форматування тексту та обмежень, яких треба при цьому дотримуватися, полягають у наступному. Текстовий рядок накопичується за словами. Визначається довжина кожного слова, що додається до поточної довжини рядка. Ця величина після чергового підсумовування порівнюється із заданим горизонтальним форматом

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i} a_{ij} + (n-1)t_{\min} \le F, \qquad (2)$$

де n, m_i – відповідно, кількість слів і символів у кожному з них; a_{ij} – ширина j -го символу в i -му слові; $t_{\min}$ – мінімальне значення пробілу між словами; F – формат рядка.

Доки у формулі (2) зберігається нерівність, накопичення слів продовжується. При досягненні рівності аналізується останнє слово. Якщо воно не є однобуквеним прийменником або скороченням типу "с.", "м.", "о." і т.п., то формування рядка завершується. Інакше останнє слово переходить у наступний рядок, а поточний вважається сформованим, якщо його кінець попадає в так звану зону форматування, і при цьому виконується така умова:

$$F - \left(\sum_{i=1}^{N+1}\sum_{j=1}^{m_i} a_{ij} + Nt_{\min} \right) \le N\delta t, \qquad (3)$$

де $\delta t = t_{\max} - t_{\min}$; $N = n - 2$. Вираз $N\delta t$ характеризує область закінчення рядка або зону форматування. Як бачимо, її величина залежить від кількості слів у рядку та параметрів пробілу між словами. Якщо умова (3) не виконується, тобто рядок не досягає зони форматування, то доведення його до потрібної довжини здійснюється за рахунок збільшення пробілів між словами.

У випадку, коли після додавання чергового слова довжина рядка перевищує формат, тобто

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i} a_{ij} + \sum_{p=1}^{r} a_p + nt_{\min} > F, \qquad (4)$$

де a_p – ширина p -го символу слова, що не помістилося в рядку, останнє слово ділиться на склади за спеціальним алгоритмом. Після цього один або декілька складів можуть залишитися в даному рядку, якщо його стан після завершення форматування відповідатиме нерівності

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i} a_{ij} + \sum_{k=1}^{s} a_k + nt_{\min} + t_{def} \le F, \qquad (5)$$

де a_k – ширина k-го символу частини слова $(s < r)$; t_{def} – ширина дефісу. Якщо перший склад не вміщається в рядку, то останнє слово повністю переноситься в наступний рядок.

Подібною є логіка верстання сторінок. Однак, оскільки цей процес важче піддається формалізації, можливі два варіанти формування сторінок — автоматизований для складних видань та автоматичний для простих. Якість автоматизованого верстання залежить в основному від кваліфікації персоналу та в деякій мірі від можливостей комп'ютерної системи. Автоматичний варіант передбачає комп'ютерне формування сторінок, що не виключає, однак, можливості модифікації чи удосконалення його результатів.

Верстання сторінок пов'язане із забезпеченням виконання як окремих вимог, так і їх сукупності. Одночасне виконання декількох вимог інколи викликає протиріччя, зняття яких вимагає складного аналізу всіх можливих варіантів закінчення сторінки. Автоматизація процесу верстання вимагає врахування не тільки суто поліграфічних показників, таких, як величина пробілів між словами, міжрядкові інтервали та рівномірність їх розподілу в сусідніх рядках, якість переносів, кількість і характер порушень при верстанні різних елементів складності, але й часових параметрів функціонування програм, об'ємів пам'яті для баз даних і т.п.

Один із варіантів алгоритму автоматичного верстання передбачає покрокове формування сторінок, поки деякий варіант не забезпечить одержання моделі з потрібними характеристиками. Для ефективнішого верстання попередньо формуються текстові та нетекстові блоки. Текстовий блок містить окремі рядки або їх сукупність. Нетекстовий блок складають, наприклад, ілюстрація або рубрика з рядками прикриття, спуск, формула і т.п.

Основою верстання є накопичення вертикального формату сторінки шляхом підсумовування вертикальних розмірів блоків. Процес продовжується до тих пір, поки сумарна величина блоків не перевищить заданий формат сторінки, тобто

$$\sum_{i=1}^{k} B_i + \sum_{j=1}^{k-1} P_j \geq H, \qquad (6)$$

де B_i – висота i-го блока; k – кількість блоків; P_j – розміри проміжків між блоками; H – вертикальний формат сторінки. Рівність в (6) означає завершення процедури верстання, тобто створена модель об'єкта формально відповідає заданим параметрам. У випадку строгої нерівності розв'язується задача втягнення або витягнення останнього блока. Вибір одного з цих варіантів залежить від співвідношення частин, які умовно утворюються на межі сторінки, та деяких її характеристик, а саме: щільності тексту, наявності змінних вертикальних проміжків чи спусків, довжини кінцевих рядків в абзацах та загальної кількості рядків у них.

Вище описано один з можливих підходів до побудови алгоритму верстання сторінок книжкових видань без врахування варіантів розміщення на них різних елементів складності. Відомості про ймовірні моделі подібних алгоритмів систем іноземних фірм, на жаль, не публікуються. Користувач оцінює їх ефективність тільки за результатами комп'ютерного опрацювання видання. Для ефективного верстання в переважній більшості КВС існує множина спеціальних стилів, тобто наперед визначених та описаних можливих макетів сторінок, що пропонуються кожною із систем. До стилю, звичайно, входить остаточне оформлення абзацу, інколи і всієї сторінки, а саме: шрифти, способи вирівнювання тексту, відстані тексту від лівого чи правого країв аркуша, міжрядкові інтервали, дозволені чи заборонені перенесення початкових або кінцевих рядків абзаца, способи оформлення тексту, розміри сторінки і т.д. Стилі зберігаються в спеціальних стилевих файлах. Стилі можна змінювати або утворювати нові. Для контролю якості комп'ютерного верстання можна вивести на екран структурну схему окремої сторінки або деякої частини видання в довільному масштабі.

Оскільки розглянуті варіанти моделей форматування тексту враховують технологічні вимоги до цих процесів, а їх реалізація закладена у відповідних алгоритмах і програмних пакетах, то користувач комп'ютерної видавничої системи тільки візуально бачить результат опрацювання видання і реально вплинути на нього може тільки через зміну вихідних параметрів текстової інформації. При цьому одержання видання з потрібними якісними та кількісними характеристиками може бути забезпечене розв'язанням оптимізаційної задачі. Її

формалізоване трактування полягає в реалізації ітераційного процесу знаходження таких вихідних параметрів (аргументів) $x_1, x_2, ..., x_n$ видання, при яких його результуючі параметри в процесі опрацювання в комп'ютерній видавничій системі набудуть найкращих у певному сенсі значень. Це означає, що деяка функція якості видання при заданих аргументах матиме оптимальне значення згідно з певними критеріями. Математична постановка та розв'язання такої задачі виходить за рамки даної статті і може бути темою окремого дослідження.

Таким чином, аналіз основних вихідних параметрів і технологічних вимог форматування тексту, математичних моделей формалізації цього процесу та відповідних їм комп'ютерних варіантів дозволив більш обгрунтовано підійти до розв'язання подібних проблем у технологічних процесах підготовки і випуску видань з використанням комп'ютерних видавничих систем.

1. Воскресенский М.И., Колосов А.И. Наборные процессы и переработка текстовой информации. М., 1989. 2. Левин А. Самоучитель работы на компьютере. М., 1998. 3. Молчанов А.А. Моделирование и проектирование сложных систем. К., 1988. 4. Сеньківський В.М., Андріїв І.В. Систематизація та використання засобів автоматизованого опрацювання текстової інформації // Квалілогія книги: Зб. наук. пр., Львів, 2000. Вип. 3. С. 28–32.

УДК 655.255

ВПЛИВ ПАРАМЕТРІВ СУБТРАКТИВНИХ КООРДИНАТ НА КОЛЬОРОПОДІЛЬНІ ХАРАКТЕРИСТИКИ КОМП'ЮТЕРНИХ ЗОБРАЖЕНЬ

М.М. Ференц

Проаналізовано вибір параметрів субтрактивних координат при визначенні кольороподільних характеристик комп'ютерних зображень.

Проанализирован выбор параметров субтрактивных координат при определении цветоделительных характеристик компьютерных изображений.

Графічний метод визначення кольороподільних характеристик, який полягає у побудові графіків ліній ефективних щільностей, дозволяє оцінити якість кольороподілу, проведеного в конкретних умовах, у тому числі за допомогою комп'ютерних видавничих систем. Лінії ефективних щільностей, розміщені на трьох координатних площинах, характеризують функціональну залежність між величиною отриманого ефекту (оптична щільність, кількість світлових градацій та ін.) на кожному з трьох кольороподільних приймачів, вираженого ефективною щільністю, і величиною вхідного параметра, що характеризує кількість жовтого, пурпурного чи голубого барвника або фарби, якими синтезується кольорове зображення оригіналу, тобто його субтрактивні координати.

Для побудови трьох графіків кольороподільних характеристик по осях абсцис відкладають величини, що характеризують кількісні характеристики кожної з однофарбових складових, а по осях ординат – ефективні щільності кольорових шкал. Тангенси кутів нахилу ліній ефективних щільностей, апроксимованих прямими, виражають кількісні характеристики кольороподілу. На практиці при кольороподільних випробуваннях матеріалів чи систем для побудови графіків по осях абсцис відкладають кількість кожної фарби, виражену поверхневою концентрацією барвників синтезу, або їх монохроматичні чи зональні щільності, або візуально- чи фотографічно-еквіваленті щільності [1].

Проведені автором дослідження кольороподільних характеристик комп'ютерних зображень показали, що на характер ліній ефективних щільностей та їх розміщення на координатній площині суттєво впливає вибір абсциси, тобто субтрактивної координати кожної кольорової складової. Для дослідження брали модельний оригінал, який складається з жовтої, пурпурної, голубої, червоної, зеленої і синьо-фіолетової ступеневих шкал, віддрукованих тріадою поліграфічних фарб, та ахроматичної шкали, виготовленої на фотопапері. Просканувавши даний оригінал у системі RGB, отримане зображення виводили на дисплей