

В. А. Вуль

# ЭЛЕКТРОННЫЕ ИЗДАНИЯ

- Структура и элементы гипертекстовых документов
- Мультимедийные издания
- Средства подготовки электронных изданий
- Информационная структура издательства



УЧЕБНОЕ ПОСОБИЕ



# Владимир Абрамович Вуль

## Электронные издания

*Текст предоставлен правообладателем*

*[http://www.litres.ru/pages/biblio\\_book/?art=6647742](http://www.litres.ru/pages/biblio_book/?art=6647742)*

*Электронные издания. / Вуль В. А.: БХВ-Петербург; Санкт-Петербург; 2003*

*ISBN 5-94157-047-3*

### Аннотация

Книга посвящена рассмотрению всей совокупности проблем, связанных с созданием, воспроизведением на компьютере и практическим использованием электронных и мультимедийных изданий. В частности, показано место этих изданий в современной медиа-индустрии. Рассмотрено использование различных HTML-редакторов и специализированных программных средств создания электронных и мультимедийных изданий (MS FrontPage, Macromedia Director, DreamWeaver и Flash, а также 3ds max). В книгу также вошли разделы, связанные с публикацией и хранением электронных изданий с помощью баз данных, а также семантическим анализом, индексацией и организацией поиска и долговременного хранения электронных документов. Описаны электронные библиотеки как одно из сетевых средств распространения изданий, в том числе учебной электронной книги. Отдельная глава посвящена распространению электронной книги. Каждая глава книги содержит перечень вопросов для самоконтроля. В ряде глав приведено описание и методика выполнения самостоятельных работ, с помощью которых читатели могут закрепить полученные знания на практике. Книга рассчитана на широкий круг читателей, начиная со студентов технических вузов и школьников 10—11 классов, интересующихся современными компьютерными технологиями, и заканчивая специалистами и профессорско-преподавательским составом вузов.

Для широкого круга пользователей РС

# Содержание

Предисловие	6
Глава 1. Общая характеристика электронных изданий	9
1.1. Что такое "электронное издание"	10
1.2. Составные элементы электронного издания	13
1.2.1. Тексты	13
1.2.2. Иллюстративный материал	13
1.2.3. Звуковое сопровождение	14
1.2.4. Анимация и видео	14
1.3. Форматы электронных изданий	15
1.4. Классификация электронных изданий	17
1.5. Современное состояние и перспективы электронного книгоиздания	20
1.6. Электронные СМИ и их роль в информационном пространстве	24
1.7. Конвергенция СМИ	30
Контрольные вопросы	38
Глава 2	39
2.1. Общая характеристика и структура HTML-документа	40
2.2. Тело документа и оформление его основных фрагментов	44
2.2.1. Тело документа	44
2.2.2. Тэги логического форматирования текста	44
2.2.3. Тэги физического форматирования текста	46
2.3. Оформление списков терминов и определений	51
2.3.1. Маркированные списки	51
2.3.2. Создание нумерованных списков	51
2.3.3. Списки определений	52
2.3.4. Многоуровневые списковые структуры	53
2.4. Использование графики при оформлении электронного издания	55
2.5. Включение в издание ссылок и комментариев	59
2.6. Оформление таблиц	62
2.7. Фреймовая структура электронного издания	67
2.8. HTML-формы	74
2.9. Понятие о динамическом HTML и его компонентах	78
2.10. Расширяемый язык разметки XML	81
2.10.1. Общая характеристика XML	81
2.10.2. Надстройки для XML	82
2.10.3. Перспективы развития XML	83
Контрольные вопросы	84
Самостоятельные (лабораторные) работы	86
Работа №2.1. Преобразование текстового документа txt-формата в HTML-документ	86
Порядок выполнения работы	86
Работа №2.2. Преобразование набора тематически связанных текстовых документов в сетевой HTML-документ	91

Порядок выполнения работы	92
Глава 3	93
3.1. Мультимедиа и ее роль в современных информационных технологиях	94
3.2. Форматы представления текстовых блоков электронного издания	97
3.3. Форматы представления графической информации	101
3.3.1. Общая характеристика графических форматов	101
3.3.2. Сжатие графических данных	103
3.3.3. Описание наиболее распространенных графических форматов	105
3.4. Форматы представления аудиофайлов	110
3.5. Форматы представления анимации и цифрового видео	113
3.5.1. Анимационные GIF-файлы	113
Конец ознакомительного фрагмента.	115

# Владимир Вуль

## Электронные издания

*Рекомендовано УМО вузов по университетскому политехническому образованию в качестве учебного пособия для студентов высших учебных заведений, обучающихся по направлению 654600 – "Информатика и вычислительная техника", специальности 220200 – "Автоматизированные системы обработки информации и управления"*

**Группа подготовки издания:**

Главный редактор *Екатерина Кондукова*

Зав. редакцией *Григорий Добин*

Редактор *Юрий Рожко*

Компьютерная верстка *Константина Ушакова*

Корректор *Евгений Камский*

Дизайн обложки *Игоря Цырульникова*

Зав. производством *Николай Тверских*

## Предисловие



В настоящее время производство печатной продукции все больше интегрируется в единое пространство медиа-индустрии. Составной частью этого процесса является интеграция печатных и электронных изданий. Причин такой интеграции несколько. Прежде всего, практически все информационные технологии перешли на цифровую форму, оснастились компьютерной техникой, программным обеспечением профессионального уровня, техническими средствами обработки изображений. С другой стороны, тот же процесс происходит и с распространением любых видов информационной продукции: появились компакт- и DVD-диски с электронными изданиями, а также диски содержащие различную музыкальную и видеoinформацию. Издательства внедрили новейшие сетевые технологии, связанные с распространением медиа-информации по интра- и экстрасетям, в глобальной сети Интернет.

Электронная книга значительно дешевле печатной и ее изготовление не связано с расходом трудновозобновляемых ресурсов (леса) и загрязнением окружающей среды. Во многих случаях электронные издания оказываются даже более функциональными, чем печатные. Так, справочное или учебное электронное издание позволяет более динамично построить процесс изучения материала и усилить его мотивацию, что, в конечном счете, позволяет ускорить процесс восприятия и запоминания информации. Электронные карты и путеводители помогают туристам и экскурсантам ориентироваться в незнакомых городах и впервые посещаемых музеях.

Художественная литература преимущественно распространяется в привычной для нас форме типографских изданий. Однако детские электронные издания уже существенно потеснили книги, так как последние не обладают многими возможностями электронных компьютерных технологий. Постепенно, но неуклонно продолжается наступление электронных изданий в учебной сфере, начиная со школьного обучения и до высшего образования.

Важнейшим преимуществом электронных изданий по сравнению с печатными является возможность их интерактивного оформления. Это относится не только к детским игровым, но и к любым изданиям сценарного типа. Такое издание может быть учебным, научно-популярным и даже художественным произведением, в котором потребителю отводится не пассивная роль читателя, а активная роль участника. Все фрагменты электронных изданий, моделирующие процессы, могут быть построены по этому типу, т. е. читатель сам становится участником событий и в определенных рамках может влиять на их исход, что сближает процесс работы с таким изданием с деловыми играми. Естественно, что такой эффект невозможно достигнуть при традиционной книжной форме представления информации.

Еще одним проявлением интерактивности является возможность моделирования и визуализации достаточно сложных физических явлений, таких, например, как ток и напря-

жение в электронных цепях, квантовое состояние атомов, их ядер и электронов, прохождение световых лучей в различных средах, включая анизотропные и т. п. Работая с соответствующими фрагментами электронного издания и самостоятельно определяя, например, величины входных и выходных сигналов или углы падения и преломления световых лучей, читатель чувствует себя полноправным участником компьютерного эксперимента, исследователем происходящих процессов и явлений. В учебном процессе такие операции вполне могут быть названы виртуальными лабораторными работами.

В последние годы непрерывно возрастает роль электронных изданий учебного и словарно-справочного характера вследствие возможности их быстрой модификации в соответствии с изменением достигнутого уровня знаний, т. е. достигается чрезвычайно высокий уровень оперативности электронного издания. Особенно это заметно для изданий, распространяемых по сетям. Если прибавить к сказанному возможность дистанционного общения преподавателя с одним или несколькими обучаемыми, дистанционного тестирования и оценки уровня знаний, то можно прийти к выводу, что электронный учебник преобразуется в постоянно развивающуюся обучающую, справочную и контролирующую уровень знаний специализированную информационную среду.

Электронные издания не только публикуются в базах данных, но и могут применяться в форме баз данных – реферативных и библиографических. Эти два вида изданий объединяет то обстоятельство, что они предназначены для квалифицированных пользователей, среди которых присутствуют как библиографы – работники крупных библиотек, так и научные работники, scrupulously следящие за изданиями в своей предметной области. Неудивительно, что первые электронные издания были связаны именно с этой предметной областью.

База данных, в которой хранятся (или публикуются) электронные издания, поддерживает каталогизацию и другие виды библиотечной деятельности, в частности, одну из важнейших – доступ читателей в публичный каталог, причем никаких специальных знаний помимо своей предметной области от читателей не требуется. База данных позволяет преодолеть основное ограничение в обслуживании читателей в библиотеках – она дает возможность многим пользователям работать с одним и тем же материалом одновременно.

Наконец, здесь проявилось и другое очень существенное преимущество электронного издания в сравнении с любым изданием на бумаге – качество хранимого материала (текста, иллюстраций и пр.) никак не зависит от интенсивности его использования: этот материал не изнашивается и не стирается. Еще один фактор – электронное издание занимает значительно меньшую площадь и объем, что также является немаловажным экономическим фактором.

Роль электронных изданий в составе издательской продукции постоянно растет. По прогнозам к 2010 году мировое производство печатной продукции в стоимостном отношении снизится с 60—70% до 35—50% от общего производства продукции медиа-индустрии, а производство электронных изданий в стоимостном выражении достигнет 25—30%. Естественно, что и в нашей стране следует учесть эти тенденции развития издательского дела.

Хотя в периодических изданиях множество материалов посвящены производству и распространению электронных изданий, в том числе и с мультимедийными компонентами, пока отсутствуют как учебные пособия, так и монографии, в которых были бы рассмотрены соответствующие проблемы. В результате в настоящее время возникла настоятельная потребность в издании книги по теме "Электронные и мультимедиа книжные издания". Автором накоплен определенный опыт в чтении этого курса или его фрагментов студентам различных специальностей. Опыт преподавательской, научно-исследовательской и методической работы, многочисленные доклады и публикации в этой области позволили автору освоить все аспекты данной темы, а в некоторых случаях – внести свой определенный вклад. Именно это послужило информационной основой предлагаемой книги. Кроме того, автор

счел целесообразным включить в книгу ряд разделов по смежной и близкой тематике, а именно:

- # распространение электронных изданий;
- # HTML-формы (Hyper Text Markup Language);
- # CGI-технологии (Common Gateway Interface) и их применение;
- # организация баз данных для хранения электронной литературы;
- # метаинформация и ее использование при поиске электронных изданий в локальных и сетевых базах данных;
- # семантический анализ электронных изданий;
- # современные поисковые стратегии и их применение при работе с электронной книгой;
- # электронная учебная книга.

Автор надеется, что книга будет полезна не только специалистам в области современных издательских технологий, но и студентам и учащимся старших классов в процессе изучения различных дисциплин, связанных с современными информационными и сетевыми технологиями. В частности, включение отдельных разделов книги поможет сделать гораздо более интересным и наглядным обычный курс "Информатики" не только в высшей, но и в средней школе. Книга окажется полезной всем интересующимся гипертекстовыми документами и системами, Web-дизайном, мультимедийными технологиями и базами данных, стратегиями индексации электронных документов и их поиском в локальных и распределенных базах данных. Все перечисленные направления активно внедряются в повседневную практику работы различных фирм и предприятий.



## **Глава 1. Общая характеристика электронных изданий**



В первой главе рассмотрены особенности электронных изданий и их место в современной медиа-индустрии. Обсуждаются составные элементы и форматы электронных изданий. Предложена классификация электронных изданий и показаны пути и перспективы их дальнейшего совершенствования и развития. Отдельно рассмотрено состояние и перспективы развития электронных средств массовой информации.

## 1.1. Что такое "электронное издание"

Книгопечатание появилось в XV веке благодаря изобретению Иоганна Гутенберга. Несколько веков печатные издания, т. е. книги, газеты и журналы были основным средством распространения визуальной информации. Большую часть этого периода основой изготовления печатной продукции служил оригинальный металлический набор, а металлическая матрица (в заключительной части периода – стереотип) являлся информационной основой для тиражирования.

Во второй половине XX века появилось новое понятие – "машинный носитель", под которым понималось любое электронное средство длительного хранения информации, на котором записано издание в целом или его часть. Чаще всего информация хранилась на магнитном носителе: магнитной ленте или магнитном диске. Однако такие носители оставались несовершенными: надежность хранения и плотность записи информации были низки. Поэтому такие средства в издательской практике использовались достаточно редко. В большинстве отечественных издательств процесс подготовки изданий основывался на использовании "бумажных" технологий. Автор сдавал рукопись, напечатанную на пишущей машинке. В издательстве она редактировалась и в окончательном виде также перепечатывалась машинистками издательства и этот издательский машинописный оригинал использовался в качестве основы для типографского набора.

Последние десятилетия XX века характеризуются быстрым совершенствованием и развитием электроники и компьютерных информационных технологий. Именно в этот период практически все издательства перешли на компьютерный набор и верстку газет, журналов и книг. В этом случае издание хранится в памяти компьютера во время набора и верстки книги, т. е. остается в электронной (машинной) форме в течение всего процесса подготовки, вплоть до вывода на принтер так называемого постраничного оригинал-макета. Полностью сверстанное и подготовленное к печати издание, хранимое в памяти компьютера (на жестком магнитном диске) или в специальном запоминающем устройстве долговременного типа можно назвать "электронным изданием".

Однако для того, чтобы электронная книга, журнал или газета действительно могли соперничать со своими печатными аналогами, необходимы средства их распространения, доведения до читателя. В конце XX века эти средства фактически стали массовыми, т. е. получили самое широкое распространение. Периодические электронные издания стали распространяться преимущественно по сетям, в частности, по глобальной сети Интернет. Информационной средой для распространения книг стали и остаются последние 15 лет компакт-диски.

Итак, вначале электронные издания существовали как аналог печатных, но на машинном носителе. Естественно, что для чтения электронных изданий, распространяемых по сетям, требовался персональный компьютер. Если же электронное издание подготовлено на компакт-диске, то дополнительно был еще нужен дисковод для чтения компакт-диска. Таким образом, издание на машинном носителе или электронное издание не может быть прочитано непосредственно, то есть требуется специальное дополнительное оборудование, чтобы сделать такое издание видимым для человеческого зрения или обеспечить его визуализацию.

Далее понятие "электронные издания" стало включать в себя тексты книг, журналов и газет, распространяемые в любом текстовом или ином формате, например, в гипертекстовом (HTML – Hypertext Markup Language) или одном из сжатых форматов (ZIP, ARJ, RAR, WINZIP и т. п.). В последнее десятилетие XX века в составе электронных изданий стали применяться еще и мультимедийные компоненты, под которыми подразумеваются цифровые звуковые или видеофрагменты, а также анимационные вставки в основную часть изда-

ния. В результате электронные издания стали средством комплексного информационного воздействия на человека, сравнимого с радио, кино и телевидением, а в чем-то даже превосходящем эти важные средства массовой коммуникации. Принципиальным отличием печатных изданий от электронных является возможность интерактивной реализации последних, при которой пользователь (читатель) может не только перемещаться по встроенным в текст гиперссылкам, но и активно вмешиваться в ход событий, моделировать процессы, в том числе производственные.

В последние несколько лет появились сетевые электронные издания. Большинство молодых людей еще в школьные годы знакомятся с глобальной сетью Интернет и встречаются с понятиями Web-сайт и Web-страница. Последняя определяется как электронная страница в соответствующем формате, а Web-сайт – это набор связанных друг с другом Web-страниц, объединенных общими признаками или назначением. Другое определение: Web-сайт – многоуровневая структурированная комбинация электронных страниц, с внутренними связями и ссылками, позволяющая перемещаться с одной страницы на любую другую, а также на другие сайты. В настоящее время в сети Интернет существует значительное количество сайтов, представляющих собой, по существу, электронные издания, т. е. аналоги книг или брошюр, но только в электронном сетевом представлении. В качестве примера можно сослаться на сайт "Электронные каталоги Российских библиотек", верхняя часть главной страницы которого представлена на рис. 1.1.

Размер этой страницы достаточно велик, так как содержит перечень из нескольких десятков библиотек с указанием тематики изданий, представленных в данной библиотеке, и особенностей поиска и извлечения информации из ее каталога. Название библиотеки, с точки зрения пользователя, подобно заголовку многотомного справочника, в котором описаны все книги, хранимые в этой библиотеке. Если подвести указатель мыши к названию библиотеки, то он приобретет форму указателя (кисть руки с вытянутым указательным пальцем). Таким образом, название служит указателем перехода к Web-странице, на которой организован поиск издания в соответствующей библиотеке. Для осуществления перехода следует щелкнуть левой кнопкой мыши по этому указателю перехода. Как создаются такие указатели и осуществляются переходы, мы рассмотрим во второй главе книги.

В описании каталога Государственной публичной научно-технической библиотеки есть ссылка на Сводный каталог по научно-технической литературе, который также служит указателем перехода. Щелкнув мышью по соответствующему фрагменту текста, мы переместимся на страницы этого сводного каталога. Таким образом, с помощью электронных страниц и гиперссылок в сети Интернет создан мощный электронный справочник, заменяющий много толстых книг, содержащих соответствующие библиотечные каталоги нескольких крупных библиотек. Причем поиск в этом справочнике вместо нас производит специальная поисковая система, которая работает во много тысяч раз быстрее и эффективней.

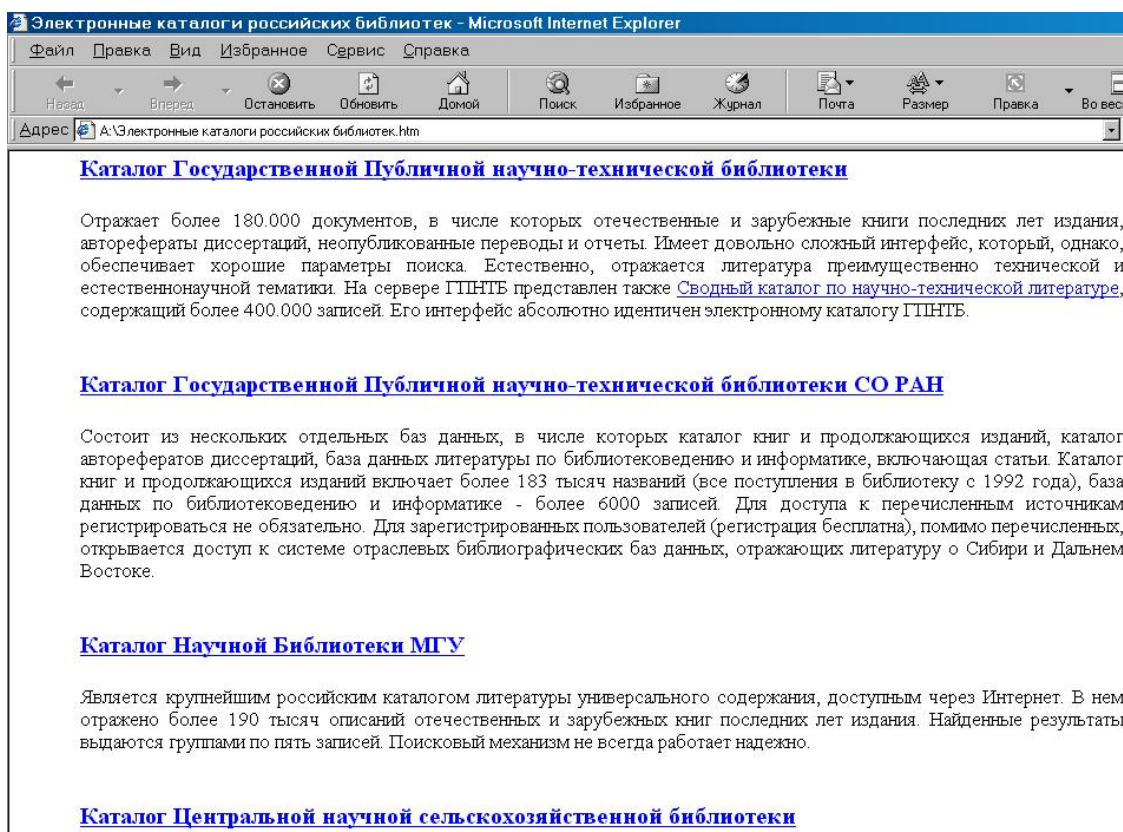


Рис. 1.1. Фрагмент лицевой страницы сайта **Электронные каталоги Российских библиотек**

Назначение Web-сайтов достаточно разнообразно. Большей частью они используются как представительства в сети различных фирм, производственных предприятий, учебных институтов и университетов. Есть свои сайты и у большинства учебных заведений: университетов, академий, институтов. Достоинства Web-сайта:

- # конкретность представленных данных, возможность использования самых различных форматов, простота и удобство навигации между страницами сайта;
- # неограниченное тиражирование информации с сайта, вследствие чего финансовые затраты на единицу информации оказываются на несколько порядков ниже по сравнению с печатной продукцией, радио и телевидением;
- # удобство восприятия и доходчивость представленной информации;
- # высокая оперативность и низкие затраты на обновления информации, в частности, на обновление учебных планов, программ и учебных пособий по отдельным курсам;
- # постоянная обратная связь с посетителями сайта на основе специальных технологий, включающих обработку форм (см. разд. 2.8 и 9.5) и электронную почту, что, в частности, позволяет студентам дистанционно выполнять практические и лабораторные работы, а преподавателям дистанционно проверять уровень знаний как студентов, так и школьников;
- # дистанционная связь и конфиденциальность информации для сотрудников, работающих вне главного офиса, и для партнеров на основе специальных интра- и экстрасетей.

## **1.2. Составные элементы электронного издания**

### **1.2.1. Тексты**

Основной частью большинства электронных изданий, так же как и книг, выпущенных типографским способом, являются текстовые фрагменты. Некоторые издания вообще целиком текстовые. Если такое издание набрано на компьютере в любом текстовом редакторе, то его можно назвать электронным изданием. Разновидностью такого издания являются так называемые гипертекстовые документы – издания в специальном HTML-формате. Язык HTML и методика создания HTML-документов будет рассмотрена в главе 2, поэтому здесь мы не останавливаемся на этом подробно.

Если заглянуть во Всемирную сеть Интернет, то там можно обнаружить некоммерческие электронные библиотеки [4]. Большинство изданий в этих библиотеках представлены в текстовых или упакованных (архивированных) текстовых файлах. Последние получены в результате предварительного сжатия исходного файла специальными программами – архиваторами. Такие издания характеризуются минимальным объемом и могут быть быстро загружены по сети на компьютер пользователя.

Простейший текстовый формат использует кодировку отдельных символов текста в ASCII-стандарте, в соответствии с которым каждый символ кодируется одним байтом. ASCII – это аббревиатура от American Standard Code Information Interchange или – в переводе – “Американский стандартный код для информационного взаимодействия”. Чем сложнее текстовый редактор, в котором набирается текстовые фрагменты, тем большее количество управляющих символов используется для форматирования текста, в результате чего объем текстового фрагмента заметно увеличивается. В графической оболочке Windows для кодирования каждого символа используется не менее 2 байт. В частности, в текстовом редакторе Word 97, в котором набирался этот учебник, на каждый кодируемый символ текста в среднем приходится около 5 байт, причем в состав редактора включены средства внутреннего сжатия информации в файлах документов. Тот же текст, набранный в редакторе Word 6 или Word 7 (Word 95) под Windows, занимает еще существенно большее информационное пространство.

В некоторых случаях текст кодируется в графическом формате (например, формат PDF (Portable Document Format), который будет рассмотрен далее). Это обеспечивает качественный визуальный интерфейс при просмотре издания, его независимость от шрифтовых гарнитур, установленных на компьютере пользователя, но требует применения специальных программных средств для его просмотра.

### **1.2.2. Иллюстративный материал**

Известно, что издательства и полиграфисты очень осторожны при включении иллюстраций, особенно цветных, в тиражируемые книги. Это связано с тем обстоятельством, что иллюстрации требуют специальной обработки, в том числе – растривания, а цветные – еще и цветоделения, в результате чего усложняется и удорожается производство книг.

В электронных изданиях этой проблемы не существует, так как абсолютное большинство компьютеров снабжены цветными мониторами и программными средствами для воспроизведения иллюстративного материала. Поэтому в электронных изданиях следует использовать такое количество иллюстраций, которое требуется для наилучшего восприятия и понимания материала, причем эта величина всегда больше, чем в изданиях, тиражируемых печатным способом. Иллюстративный материал содержит на несколько поряд-

ков больше информации, чем текст, занимающий то же самое пространство на странице, и гораздо эффективней воздействует на человека.

Скорость восприятия иллюстративной информации также многократно выше, чем текста. Это связано с особенностями визуального восприятия информации человеком. Зрительные образы в виде графических объектов воспринимаются целиком и непосредственно обрабатываются и сохраняются в долговременной памяти человека, без промежуточного преобразования в понятия, как это происходит с текстом.

Форматы графических файлов, особенности работы с цветом, количество используемых в изображении цветовых оттенков и многие другие вопросы будут рассмотрены в последующих главах.

### **1.2.3. Звуковое сопровождение**

Неотъемлемой частью многих изданий является звуковое сопровождение. Звуковое сопровождение может представлять собой авторский текст или ремарки, шумовые эффекты, иллюстрирующие происходящие события и делающие их описание более реалистичным. Звук может синхронно сопровождать включенные в издание видеокadres или анимацию, что будет подробно изложено в последующих разделах книги.

Скорость восприятия человеком звуковой информации имеет тот же порядок величин, что и для текста. Однако, одновременная работа с текстовой и звуковой информацией не только увеличивает общую скорость восприятия, но и способствует лучшему ее долговременному запоминанию, вероятно в результате образования определенных ассоциативных связей.

### **1.2.4. Анимация и видео**

Все мы смотрим телевидение, однако телевизионное вещание на всех отечественных каналах, кроме коммерческого канала НТВ+ и некоторых каналов кабельного телевидения, ведется в аналоговом формате. Цифровой формат, используемый в компьютерном видео, отличается несравненно более высоким качеством воспроизведения и разрешающей способностью, но требует большого объема памяти для хранения видеoinформации и высокой пропускной способности для ее воспроизведения с необходимой частотой кадров.

Поэтому гораздо чаще, чем видеофрагменты, в электронных изданиях используется компьютерная мультипликация или простые анимационные файлы. Они же очень часто встречаются на различных Web-страницах, причем, во многих случаях, применяются в качестве рекламы. Однако, даже такая простая мультипликация может использоваться для иллюстрации последовательности выполнения технологических операций, работы полиграфических машин и комплексов и даже последовательности выполнения операций в компьютерных программных пакетах. Во всех перечисленных случаях анимация играет вспомогательную роль, способствуя наглядности описания соответствующих процессов и лучшему пониманию и запоминанию этого описания. Широкое использование анимации связано с тем, что информационный объем анимационных файлов и требуемые для их воспроизведения информационные ресурсы сравнительно невелики, в то же время анимация оживляет страницы электронного издания, делает их более динамичными, способствует лучшему их восприятию. Цифровое компьютерное видео может успешно использоваться лишь на компьютерах с процессором и видеокартой высокой производительности, обеспечивающих скорости следования информационных потоков, достигающих, а иногда и превышающих 1 Гбайт/с. Его воспроизведение требует значительных аппаратных и программных ресурсов, причем эти требования пропорциональны площади, на которой воспроизводится видео.

### 1.3. Форматы электронных изданий

В предыдущем параграфе уже упоминались форматы электронных изданий. Здесь мы поговорим об этом подробнее, впрочем, и позднее мы будем еще не раз обращаться к этому вопросу. В электронном издании формат описывает то, каким образом в файле представлена информация. В настоящее время для современных и качественных электронных изданий используются два основных формата, а именно:

# PDF (Portable Document Format), разработанный фирмой Adobe и представляющий развитие и совершенствование известного издательского языка описания страниц документов (Page Description Language, PDL) – PostScript;

# HTML (HyperText Markup Language) – гипертекстовый язык разметки страниц, с помощью которого создано большое количество электронных документов, в том числе – абсолютное большинство изданий, циркулирующих в сети Интернет.

Главное достоинство обоих форматов состоит в возможности размещения в тексте издания гиперссылок, по которым осуществляются быстрые переходы как внутри данного издания, так и во внешние, по отношению к данному, документы. В частности, с помощью гиперссылки можно связаться с автором или издателем по электронной почте, просмотреть литературные источники, на которые ссылаются в данном издании, вызвать иллюстрацию, которая поясняет смысл некоторого текстового фрагмента и многое другое.

Большинство электронных материалов, циркулирующих в сети Интернет, оформлены в HTML-формате. Формат достаточно компактен и, наряду с текстом, позволяет включать в издание иллюстрации и мультимедийные фрагменты. Основу HTML-документов составляют обычные текстовые файлы, отдельные символы в которых представлены в ASCII-кодировке. Эти файлы доступны для просмотра и редактирования в любом редакторе текстов. Отличием HTML-издания от обычного текста является то, что в них присутствуют специальные команды-тэги, которые указывают правила форматирования документа при выводе информации на экран монитора. Язык HTML и правила составления HTML-документов и работы с ними, будут детально описаны в следующей главе. Недостатком этого формата является чувствительность к используемым в издании шрифтам: отсутствие нужного шрифта на пользовательском компьютере вызывает затруднения при просмотре документа, связанные с необходимостью замены отсутствующей шрифтовой гарнитуры.

Другим основным форматом для электронных документов является формат PDF (Portable Document Format), разработанный в 1993 году фирмой Adobe. Как говорит само название, он делает документы "portable" (переносными), то есть документы могут просматриваться независимо от программ, в которых они подготовлены и независимо от используемых шрифтов и компьютеров. PDF-формат позволяет преодолеть барьер, связанный с зависимостью от установленных на компьютерах шрифтов. Он впервые предоставил возможность распространять на машинных носителях (дискетах и компактдисках) и через Интернет выполненные в графическом режиме всевозможные документы, в частности оригинал-макеты изданий. При этом гарантируется их полная идентичность с оригиналом. Это делает PDF-формат чрезвычайно полезным не только для художественных и детских электронных изданий, но и для коммерческих и рекламных документов. В США PDF-формат стал основным для распространения правительственных материалов.

Изначально формат PDF был предложен как альтернатива печатным документам, средство для организации "безбумажного офиса". Просмотр таких документов должен быть организован с помощью универсальных средств, независимо от создавшего их приложения и установленных на данном компьютере шрифтов. Формат PDF может рассматриваться как компактный формат электронной документации. Практически вся документация разнобо-

разных программных пакетов распространяется на компакт-дисках именно в этом формате. Наконец, этот формат сохраняет всю информацию для выводных устройств, т. е. может служить для хранения оригинала и распечатки документа по запросу. По сравнению с форматом описания документов PostScript формат PDF имеет то преимущество, что он является странично-ориентированным, т. е. описывает каждую страницу в отдельности. Это резко упрощает вывод отдельных страниц на печать.

PDF поддерживает различные виды компрессии изображений, текста и графики, направленных на уменьшение размера конкретных фрагментов файла с помощью наиболее подходящих для этого алгоритмов, одновременно позволяя использовать гипертекстовые связи. Фирма Adobe разработала и комплект программного обеспечения для создания, просмотра, редактирования и распечатки PDF-файлов. Эти программные средства будут рассмотрены в разд. 4.3.



## 1.4. Классификация электронных изданий

Классификация электронных изданий может быть проведена по нескольким признакам. Наиболее существенными из них представляются следующие:

- # периодичность издания;
- # круг потребителей продукции;
- # вид издания;
- # способ распространения;
- # формат издания.

Сказанное иллюстрируется классификационной диаграммой, представленной на рис.

1.2.



Рис. 1.2. Классификация электронных изданий

Круг потребителей продукции или круг пользователей электронных изданий достаточно широк и разнообразен. К ним можно отнести учащихся, использующих электронные учебники, туристов и посетителей музеев, которые с помощью электронных справочников и путеводителей могут совершать не только реальные, но и виртуальные путешествия, и осмотр памятников культуры, специалистов, использующих справочные и иные пособия по своей специальности, детей, "посещающих" мультимедийные виртуальные спектакли и так

далее. Эти издания распространяются преимущественно на компакт-дисках и рассчитаны на мультимедийный компьютер (см. главу 5).

Научные и технические издания, преимущественно периодические, рассчитаны на широкий круг научных работников, специалистов и преподавателей и становятся все более популярными. Это, главным образом, объясняется опережающей по сравнению с печатными изданиями возможностью ознакомления с ними, большей простоте и дешевизне получения информации, легкости извлечения ее фрагментов и возможности контекстного просмотра авторских ссылок и сопутствующих материалов. В последнее время основной средой распространения таких изданий стала глобальная сеть Интернет. Многие из читателей, вероятно, не раз слышали высказывания, что в сети Интернет можно найти все. Это действительно так. Надо только знать, где и как найти нужную информацию. Соответствующие вопросы будут рассмотрены в главе 8.

Сопутствующими для всех рассмотренных выше разновидностей электронных публикаций являются рекламные материалы. В большинство электронных изданий, как впрочем и в печатные, вкрапливается реклама в самых различных формах. Самая простая и естественная из этих видов – реклама фирмы, производящей данное электронное издание, включающая в себя как минимум, информацию о других электронных изданиях данного профиля, выпущенных или подготавливаемых к выпуску этой фирмой. В периодических электронных изданиях реклама автоматически переносится со сверстанных страниц оригинала в электронное издание. В изданиях, распространяемых по глобальным сетям, само оформление сайтов, страниц и интерфейсов уже содержит рекламные элементы, в частности, анимационные – так называемые баннеры (banners).

Виды изданий, разнообразие которых уже затронуто в классификации по предыдущему признаку, тесно связаны с кругом пользователей. Здесь можно ограничиться признаками периодичности и тематической области, к которой издание относится. Ежедневные и еженедельные издания распространяются практически только в сетевых средах, причем они могут, в частности, распространяться путем рассылки, т. е. полное издание, а чаще – отдельные его тематические рубрики регулярно посылаются тем пользователям, которые на них подписались. По тематической области круг электронных изданий достаточно широк. Так, на машинных носителях распространяются:

1. Учебные пособия (обучение иностранным языкам, конспекты лекций, методические указания по выполнению практических и лабораторных работ, тексты рефератов и докладов и пр.).

2. Научные журналы (как правило, подборка из нескольких выпусков журналов вместе с программными пакетами и сопутствующими рекламными материалами).

3. Детские издания в форме мультимедийных сценариев сказочного, приключенческого, литературно-художественного и обучающего типа.

4. Художественная и научно-популярная литература (обычно в PDF-формате либо в архивированной форме). Эти издания представляют интерес для всех групп читателей, поэтому на рис. 1.2 этой группе соответствует позиция "Общие".

5. Энциклопедические и справочные издания, иногда многотомные, в том числе аналоги ранее изданных энциклопедий, например, Grolier, Laruss, и специфически компьютерные – интерактивный атлас мира, географические справочники, атласы автомобильных дорог и пр. В электронном виде создаются сейчас и другие справочно-картографические системы. Для многих крупных городов они уже созданы и позволяют не только найти нужное здание по известному адресу, но и оптимизировать маршрут до места назначения. Есть такие справочники, например, по Петербургу и его окрестностям.

В сетевых средах, в частности в сети Интернет, распространяются электронные издания преимущественно периодические, к которым можно отнести:

# не только отдельные учебные пособия, но и целые учебные циклы для дистанционного обучения и самостоятельного изучения;

# исключительно широкий круг научных, научно-популярных и технических журналов, начиная с компьютерных (Computer World, Computer Week/Moscow, CompuLog и других) и сетевых (Internet Journal, CrazyWeb, LAN Magazine) и кончая музыкальными и игровыми;

# общественно-политические журналы, например, широко известный "Огонек";

# литературно-художественные издания, в том числе журналы ("Новый мир", "Иностранная литература", "Октябрь", "Арт-Петербург", причем последний представляет собой культурный альманах, существующий только в Интернете), "Литературная газета" и пр.;

# библиографические указатели и рефераты типа Book Review и электронной библиотеки службы "ИНФОМАГ";

# газеты (например, "Аргументы и факты", "Аномалия", "Вести", "Натали", "Известия", "Учительская газета");

# развлекательные издания типа: "Знакомства", "Вечерний клуб", "Пятое колесо", "Вера, Надежда, Любовь";

# специализированные рекламные издания; в том числе "Центр-Плюс", "Реклама-Шанс".

По способу распространения все электронные издания можно разделить на 2 большие группы, а именно:

# распространяемые на физических носителях, преимущественно компактдисках;

# распространяемые в сетевых средах, как локальных (например, сетевая электронная библиотека учебного заведения), так и в глобальной сети Интернет.

Так же как и печатные, электронные издания могут быть классифицированы по формату. Однако, если в печатном издании формат характеризует физические размеры издания, то в электронном издании формат описывает то, каким образом в файле представлена информация, содержащаяся в данном издании. Начиная с восьмидесятых годов электронные публикации распространялись в текстовом формате, вначале под операционную систему MSDOS (txt), а затем под Windows (txt, doc) и другие платформы. В настоящее время используются, как уже отмечалось выше два основных гипертекстовых формата, а именно: HTML и PDF, причем последний хранит всю информацию в графической форме. Если в издании содержится цифровая анимация, а тем более цифровые аудио- и видеофрагменты, то такие электронные издания называют мультимедийными. Хотя это и не название формата, но важная характеристика того, какие именно цифровые форматы могут содержаться в издании (см. главу 3).

## 1.5. Современное состояние и перспективы электронного книгоиздания

Электронные издания относятся к динамически развивающемуся классу продукции. Их количество увеличивается быстрыми темпами, а качество непрерывно улучшается.

Интеграция полиграфической продукции с электронными документами приносит чисто практические выгоды. Так, переход на цифровую форму позволяет обеспечить сохранность многих уникальных видов продукции, таких как древние рукописи; даже обычные фотографии и картины теряют со временем свои качества. Хранение их электронных копий позволяет донести до последующих поколений уникальные культурные шедевры. Наконец, хранение документов и изданий в электронной форме позволяет организовывать электронные базы данных, четкая структура и развитые средства поиска и навигации, в которых облегчают процесс обнаружения нужных материалов и их фрагментов. В качестве примера можно привести американское специализированное издательство медицинской литературы Mosby-Year Book, в котором применение цифровой графической библиотеки позволило за считанные секунды находить нужные иллюстрации и помещать их в книги, в результате чего была достигнута значительная экономия времени и материальных средств.

Для ответа на вопрос об относительной ценности электронных изданий и перспективах расширения их производства следует провести исследования в следующих основных направлениях (иначе, получить ответы на перечисленные вопросы):

- # увеличивается ли ценность книги в электронном варианте, а если да, то в каких случаях и с чем это связано;
- # какие факторы влияют на качественные характеристики электронного издания с точки зрения читателя;
- # достаточен ли уровень развития технических и программных средств создания электронной книги;
- # насколько широко распространены индивидуальные и групповые средства для чтения электронных изданий;
- # насколько сформировался рынок распространения электронных книг;
- # какие нужны изменения в традиционном издательском бизнесе для успешного маркетинга и менеджмента электронных книг.

Несомненна более высокая потребительская ценность детских и учебных электронных изданий, так как повышается интерес к изучаемому материалу и обеспечивается возможность применения новых, более эффективных технологий не только самого обучения, но и постепенного вовлечения детей в процесс обучения в игровой форме. Эти методики уже опробованы во многих странах и на практике доказали свою эффективность. Справочные и научные издания предоставляют возможность более простыми средствами и в более короткие сроки получать необходимую информацию (или краткие сведения, позволяющие оценить ее необходимость пользователю и заказать ее, в том числе и в печатном виде). Естественно, что ценность таких электронных изданий по сравнению с печатными существенно возрастает. Степень увеличения ценности и роль определенных факторов может быть установлена в процессе проведения соответствующих исследований.

С точки зрения факторов, влияющих на качественные характеристики электронных документов, наибольший интерес, по мнению автора, представляет оптимизация пользовательских интерфейсов, причем не только графических в широком смысле, включая анимацию и цифровое видео, но так же и цифровое аудио. Все это, разумеется, аналогично требует исследований и сбора статистических данных.

Постепенно, но неуклонно продолжается наступление электронных изданий в учебной сфере, начиная со школьного обучения, далее в сфере среднего и высшего образования. Во многих случаях оправдан частичный или полный переход к электронным учебникам и компьютерным образовательным технологиям. Эффективность такой замены тем выше, чем больше разнообразие учебных пособий и ниже их тиражи. Ответ на вопрос об изменениях в традиционном издательском и печатном бизнесе, необходимых для адаптации к новым информационным технологиям, также требует проведения достаточно широких и глубоких научных исследований.

Так, у нашего северного соседа Финляндии с 1996 по 2000 год проводилась рассчитанная на 5 лет исследовательская программа по электронным издательским и печатным процессам с финансированием около 60 млн. финских марок. В США же в рамках поддерживаемой правительством программы по интерактивным электронным журналам около 10 лет работает специальная служба **OCLC Service** (Online Computer Library Center Service – Интерактивная служба компьютерного библиотечного центра), которая имеет огромную базу данных и не только экономически поддерживает университетские издательства, но и сама выступает как издатель ряда научных изданий – как электронных, так и печатных. Отделения этой службы имеются на всех континентах и в основных странах. Служба имеет мощный собственный сервер ([www.oclc.org](http://www.oclc.org)), который связан с библиотечной базой данных и ежедневно обслуживает большое число запросов.

Ответ на вопрос, достаточен ли современный уровень развития технических и программных средств создания электронной книги, автор надеется дать в главах 4 и 6 этой книги. В частности, в главе 4 будет проведен анализ программных и технических средств подготовки электронных изданий, а в главе 6 – освещены возможности реализации мультимедийных изданий в специализированных программных пакетах, таких как FrontPage фирмы Microsoft, Director, DreamWeaver и Flash фирмы Macromedia и некоторых других.

Насколько широко распространены индивидуальные и групповые технические средства для работы с электронными изданиями уже говорилось в самом начале главы. Аппаратному и программному обеспечению рабочего места читателя электронных изданий посвящена глава 5, где анализируются современные средства, возможности их дальнейшего совершенствования и развития и перспективы их распространения в нашей стране. Отметим, что уже в конце 2000 года в Санкт-Петербурге насчитывалось свыше 500 тысяч учтенных персональных компьютеров, более 40% которых были снабжены дисководом для компакт-дисков и звуковой картой, т. е. таких, которые могут быть использованы для работы с мультимедийными изданиями. В настоящее время, количество подобных компьютеров существенно превышает эту цифру, так как сейчас активизировался вторичный компьютерный рынок, да и сам процесс их сборки настолько упростился, что стал вполне доступен даже неспециалистам.

Насколько сформировался рынок распространения электронных изданий, и какие нужны изменения в традиционном издательском бизнесе для успешного маркетинга и менеджмента электронных книг детально исследуется в главе 9, предпоследней в данной книге. Здесь лишь укажем, что отечественные издательства в большинстве своем успешно адаптировались к современным медиа-технологиям. Самыми передовыми из них оказались те, которые занимаются выпуском литературы по компьютерным и сетевым технологиям. Практически все они используют глобальную сеть для маркетинга и распространения литературы. Наряду с печатными изданиями, выпускают издания и на компакт-дисках, причем нередко случаи выхода интегрированных изданий, когда в печатную книгу вкладывается сопутствующий компакт-диск, содержащий программный и иллюстрационный материал, дополняющий и развивающий те положения, которые содержатся в той части, которая тиражируется печатным способом. В качестве одного из примеров подобного издания можно

сослаться на книгу [29], выпущенную издательством BNV и ставшую в настоящее время редкостью. Именно приложенный к ней компакт-диск придает книге дополнительную ценность.

В дальнейшем широкое распространение мультимедиа-технологий и неудержимое развитие электронной коммерции наложит жесткие ограничения на конкурентоспособность, и даже "выживаемость" предприятий отрасли печати, ориентированных на широкий спрос. Преимущества в реализации даже самой высококачественной продукции получают те, кто быстрее и эффективнее освоил электронные способы коммерции и обслуживания. Фирмы, не имеющих собственной аппаратной и программной базы мультимедиа-технологий будут испытывать трудности с получением заказов, и их продукция будет подвергаться опасности устареть раньше, чем дойдет до потребителя. Причем это невозможно компенсировать в рамках собственно процесса подготовки и выпуска продукции. Таким образом, сетевые электронные издания имеют несомненные преимущества в части распространения перед любыми печатными изданиями.

Производство же – как электронных изданий, так и печатной продукции – все более интегрируется в единое пространство медиа-индустрии. Одновременно, в пределах той же медиа-индустрии происходит процесс интеграции печатных и электронных изданий. Состав коммерческих средств медиа-индустрии и приблизительное распределение доходов между ними представлено в табл. 1.1 [5]. В этой же таблице приведены данные прогноза на 2005 и 2010 годы долевого участия различных средств медиа-индустрии. Прогноз предусматривает два крайних варианта развития: консервативный и прогрессивный, которые определяют диапазон разброса значений при различных сценариях развития мирового производства.

Таблица 1.1. Прогноз развития средств медиа-индустрии в мире

Средства медиа-индустрии	2000 год	2005 год		2010 год	
		Консервативный	Прогрессивный	Консервативный	Прогрессивный
Печатные	65	62	54	54	37
ТВ и радио	15	16	18	18	22
Кино, видео, музыка	10	10	10	10	11
Электронные издания (на дисках)	6	7	10	9	14
Электронные издания (сетевые)	4	5	8	9	16

Прогрессивный сценарий отражает в таблице результаты, соответствующие существующим в развитых странах темпам развития составляющих медиа-индустрии. Консервативный сценарий подразумевает замедленное развитие перспективных средств медиа-индустрии, характерное для слаборазвитых и развивающихся стран. В целом, следует ожидать ускоренных темпов развития производства электронных изданий, в особенности мультимедийных и тех, для которых используются сетевые технологии распространения. Кстати, заметное снижение вклада печатных средств не означает абсолютного снижения производства печатной продукции. Вероятнее всего рост ее производства будет продолжаться (осо-

бенно это касается производства этикеток и упаковки), однако темпы этого роста будут постепенно снижаться.

## 1.6. Электронные СМИ и их роль в информационном пространстве

Сеть Интернет – самый емкий информационный ресурс. Мы уже говорили о том, что первыми в него пришли электронные журналы, связанные с компьютерными технологиями. Лишь на полгода отстали электронные газеты, причем первыми были опять-таки компьютерные (PC Magazine Daily и Computer World). Трудности публикации газет в сети связаны, прежде всего, с тем, что электронная газета серьезно отличается от печатной. Прежде всего, это связано с ограниченным размером экрана монитора. С другой стороны, на компьютере имеется возможность одновременно открыть несколько экранов, куда можно поместить различные страницы одной или нескольких газет. Из других отличий следует отметить возможность включения в электронную газету звуковых фрагментов, например, интервью, подачу материалов в виде набора графических фрагментов (включая диаграммы, гistogramмы и пр.), при обращении к которым появляется текст во всплывающих окнах. Наконец, в электронной газете материал может обновляться практически непрерывно, причем в любое время суток.

Естественно, что все эти различия порождают совершенно новый тип информационного издания. Электронная газета – это не аналог, а скорее дайджест печатного издания. Любой пользователь компьютера, имеющий доступ к сети Интернет, может подписаться на саму газету или ее электронную версию. В последнем случае пользователь может войти в базу данных этого издания и просматривать еще не опубликованные новости и сообщения.

В первой половине 1996 года число различных газет и журналов в глобальной сети превышало тысячу. В их числе были крупнейшие мировые издания, такие как "New York Times", "Washington Post", "Boston Globe" – одни из пионеров компьютеризации, "Le Monde", "The Guardian", наши "Известия" и "Независимая Газета" и множество других. Инициаторы создания сетевых версий газет не пытались переманивать читателей печатной версии на электронную, они просто осознали важность освоения глобальной сети. Примерно в это же время появились сетевые издания, для которых печатные версии являлись вторичными, а иногда и просто отсутствовали. Учредителями виртуальных информационных проектов выступали крупные концерны и корпорации, работающие на стыке компьютерной, издательской (СМИ – средств массовой информации) и индустрии новостей (телевидение, радио, информационные агентства).

В чем же состоят преимущества электронных СМИ? Для СМИ США – прежде всего, новое направление (сетевые СМИ), в котором для каждого уважаемого издания необходимо быть представленным. Далее – это автоматическая доставка и отсутствие процедуры тиражирования. Из других факторов отметим возможность прочтения без всякой задержки удаленных газет, возможность автоматического перевода с иностранного языка и легкость просмотра "подшивок" и подготовки информационных подборок. Далее – возможность и простота избирательного просмотра материалов, которые представляют реальный интерес для читателя, в том числе – объявлений и прайслистов, что значительно труднее выполнить в печатном издании.

Сайт [washingtonpost.com](http://washingtonpost.com) был создан в середине 1996 года. На нем ежедневно публикуется содержимое 20-страничного номера этой газеты. Экстренные новости и сенсации публикуются практически сразу же после создания, не ожидая выпуска следующего номера газеты. Помимо текста и графики посетитель сайта может просмотреть видео и прослушать аудио для интересующих его событий. На сайте ежедневно происходит около 10 виртуальных встреч с журналистами, известными персонами и даже спортивными обозревателями и психологами. Внешний вид лицевой страницы сайта показан на рис. 1.3.



Сайт вполне типичен для электронных СМИ. Справа вверху на лицевой странице находится предложение о подписке на газету. В верхней части размещаются основные рубрики, из которых по умолчанию выбирается **News** (Новости). Слева расположены основные темы для поиска материалов в газете: национальное, весь мир, метро, спорт, бизнес и пр. Имеется даже раздел, посвященный книгам. На сайте предусмотрены средства поиска информации. Кроме того, имеется также возможность поиска и во всей сети Интернет. Справа размещена ссылка на музыкальный фрагмент в формате MP3 (см. главы 3 и 5).

Посещаемость сайта [washingtonpost.com](http://washingtonpost.com) газеты "Вашингтон Пост" достигает 3 миллионов посетителей в день. Это огромное количество. И этот сайт входит в число трех самых популярных в США (а может быть, и во всем мире). Естественно, что штатных сотрудников газеты не хватает для информационного и технического обслуживания данного сайта. Этим занимаются около 220 дополнительных работников. Виртуальная газета арендует специальные площади для размещения сотрудников и оборудования. Убытки, связанные с содержанием сайта, составляют около 70 миллионов долларов в год. Почему же это процветающее газетное издание осознано идет на значительные убытки? Почему руководству газеты так важно с минимальной задержкой по времени сообщать о любом важном событии? Ведь радио и телевидение все равно это делают быстрее. Ответ на этот вопрос для СМИ США очевиден – нельзя отставать от прогресса, всегда следует идти в ногу со временем! Конечно, существует риск потерять деньги в том случае, если новшество окажется бесполезным. Но гораздо опасней пропустить реальный технологический прорыв, так как отставание потом уже не наверстать.



Рис. 1.3. Лицевая страница сайта **washingtonpost.com** газеты "Вашингтон Пост"

В сентябре 1999 г. "Вашингтон Пост" начал издавать сетевую электронную газету "PM Extra" ("Экстренно в полдень"). Она появлялась в сети ровно в 12 часов дня, когда средний американец поедает бутерброды на своем рабочем месте и использует возможность подключения к сети Интернет по своей служебной (бесплатной для него) линии. Экстренный online-выпуск газеты состоит из 10—12 кратких сообщений, каждое из которых содержит 3—4 абзаца и напоминает сообщения телеграфного агентства новостей. Этот выпуск значительно повысил популярность газеты, способствовал повышению ее рейтинга.

В нашей стране мощный газетный холдинг сформировался на базе еженедельника "Аргументы и факты" (АиФ). В него входят газеты в формате таблоида, выходящие с различной периодичностью, в частности: "Интерфакс-АиФ", "АиФ Здоровье", "Дочки-Матери", "Я молодой", "АиФ-Любовь", "АиФ На даче", "АиФ Семейный совет", "Разбор", а также детские издания "Детская энциклопедия", "Колобок и два жирафа" и т. п. Под "таблоидом" подразумевается сравнительно небольшой формат газеты, не превышающий А3, т. е. 297×420 мм. В крупных городах выходит 35 региональных приложений к АиФ, которые обеспечивают стабильный спрос на еженедельник в провинции. Это единственное массовое издание в России, которое отличается высоким качеством, достоверностью информации и независимостью редакционной политики. Это издание является абсолютным лидером по охвату аудитории среди еженедельных и ежедневных изданий, т. е. лидером общественно-политической

прессы. Неотъемлемой принадлежностью холдинга стал его сайт, который представлен на рис. 1.4. С его помощью можно ознакомиться с содержанием любой из издаваемых газет, достаточно подвести указатель мыши к нужному изданию в правой части лицевой страницы сайта и щелкнуть левой кнопкой мыши.

Как видно из рис. 1.4, информация на сайте четко структурирована. Наиболее интересные новости приводятся в рубриках "Пресс-центр АиФ", "Интервью", "Новости", "КОНКУРС АиФ", "ВИТРИНА AIF.RU-PRIKUP.RU". С каждой из этих рубрик можно перейти на соответствующую страницу, где тема раскрывается подробно. Здесь же предусмотрены рекламные элементы и сведения о подписке на "АиФ". Читатель может познакомиться с интересующей его информацией с помощью предусмотренной на сайте поисковой системы в рубрике "ПОИСК ПО САЙТУ". Лицевая страница сайта имеет значительную протяженность, поэтому на рис. 1.4 показана лишь ее небольшая часть. По сравнению с сайтом газеты "Вашингтон Пост" можно отметить лишь более слабый дизайн с точки зрения цветовой гаммы и отсутствие ссылки на аудиофайл в любом формате.

Наличие электронной версии газеты в последние годы стало и элементом престижа издания, поддержания его имиджа. Самое ценное в любой газете – ее содержимое, информация. Она должна предоставляться пользователю удобным для него способом, в том числе и посредством сети Интернет. В то же время большинство специализированных сетевых газет (online) являются убыточными. Это касается даже такой известной международной газеты как "Washington Post". Убыточность сетевых периодических изданий связана, прежде всего, с особенностями современной сетевой рекламы, которая рассмотрена в разд. 9.3. Тем не менее, электронные сетевые СМИ будут продолжать существовать, невзирая на убыточность.

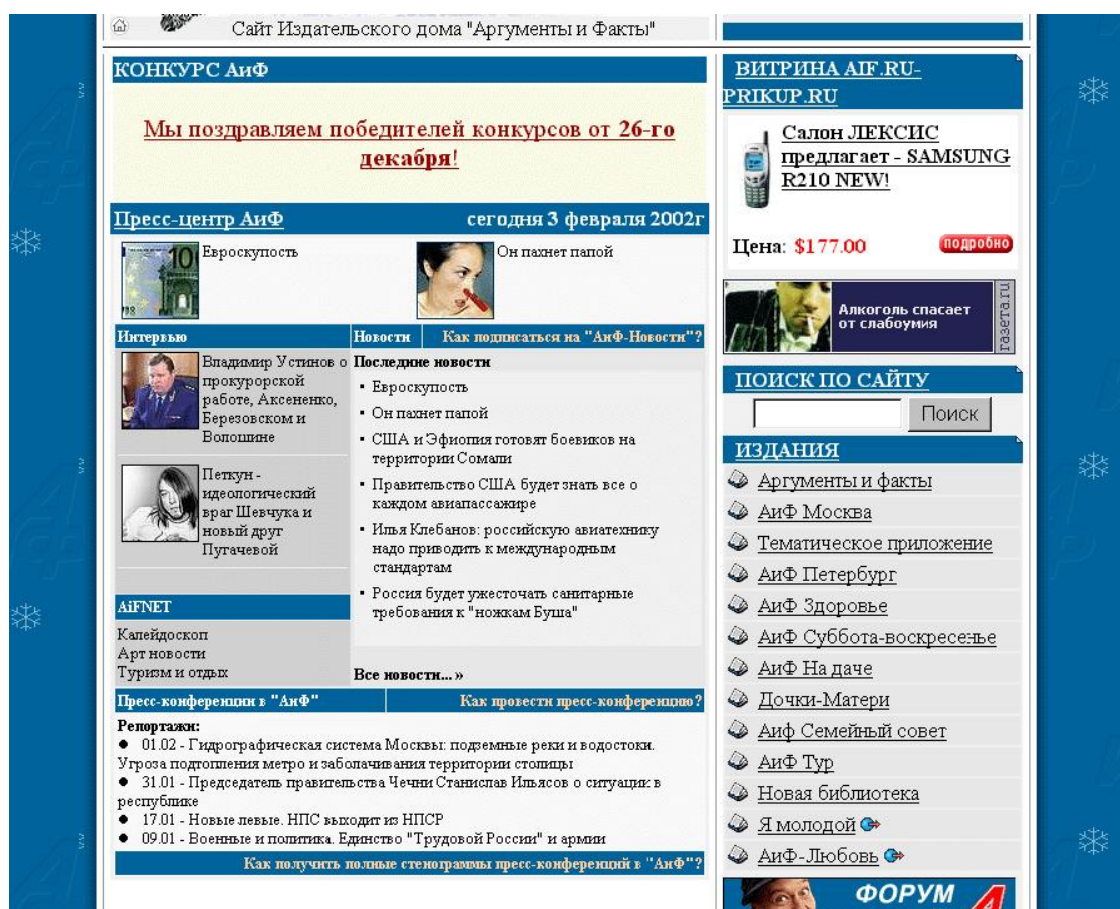


Рис. 1.4. Фрагмент лицевой страницы сайта издательского дома "Аргументы и Факты"

Уже и в нашей стране появляются не только аналоги и представительства печатных изданий в сети, но и специализированные сетевые издания, сумевшие привлечь внимание и завоевать популярность у читателей. Чем можно объяснить успех таких виртуальных СМИ? Скорее всего, умелым использованием специфики сетевой информационной среды и особенностей восприятия информации с экрана монитора. Главная особенность – это "нелинейность" материала, связанная с наличием гиперссылок. Читатель имеет дело не с линейным текстом, а с расширяющейся информационной средой. Другая особенность – это различные проявления интерактивности, которые дают возможность читателю не только высказать свое мнение, но и принять участие в форумах, конференциях, оставить свою запись в гостевой книге. В пределах сетевая информационная среда "размывает" границы между автором и читателем, читатель сам становится участником производства информационного продукта. Роль этого фактора, по нашему мнению, будет непрерывно возрастать. Наконец, электронная пресса во многих случаях оказывается оперативнее не только печатных СМИ, но и радио и телевидения, так как последние связаны жесткой сеткой вещания, где новости передаются в определенные часы.

Примером, подтверждающим высказанные положения является политическая сетевая газета "Полит.Ру", зарегистрированная Министерством Печати РФ. Политику этой газеты определяет фраза: "Мы рассказываем новости такими, какими сами их видим, а Ваш выбор – принять или отвергнуть это видение". Газета эта издается на 2 языках: русском и английском. Ее основные рубрики – новости, сюжеты, статьи, интервью, монитор (обзоры СМИ), дискуссии и несколько совместных проектов. Сайт газеты "Полит.Ру" ежедневно посещают около 20 тысяч человек, что, по нашим меркам, выглядит достаточно масштабно. На поисковом сервере Rambler газета занимает первое место в разделе "Политика" и стабильно входит в первую десятку в общем рейтинге.

Сама убыточность электронных газет, скорее всего, является болезнью роста. Ведь все сетевые информационные ресурсы очень молоды. Технологии их производства и распространения находятся в состоянии становления. То же самое касается и сетевой рекламы, неэффективность которой определяет убыточность сетевых периодических изданий.

Действительно, в последние 2—3 года пришел новый инструмент для работы с информационными ресурсами сети – информационно-аналитический портал или просто портал. Портал интегрирует возможности СМИ, архивов, аналитических агентств, средств коммуникаций, оперативных баз данных. Портал открывает совершенно новые возможности для проведения адресных рекламных компаний, изучения спроса и покупательской способности, для осуществления прямого маркетинга посредством сети Интернет.

С точки зрения пользователя возможности портала перекрывают все ранее существовавшие комплексные информационные услуги. Однако, несмотря на всю привлекательность такого проекта пока в мире существуют лишь небольшое число порталов. Причина такого явления состоит в том, что для вывода проекта хотя бы на уровень рентабельности потребуются огромные инвестиции, не менее нескольких сотен тысяч долларов. Они будут расходоваться на разработку специализированного программного обеспечения, обучение и подготовку обслуживающего персонала, способного обеспечить регулярное обновление информационного содержания и оперативное управление системой.

На сайтах некоторых электронных СМИ уже реализованы платные подсистемы распространения информации. Это способствует росту популярности данного СМИ. Коммерческий успех обычно свидетельствует о более высоком уровне, достигнутом данным электронным СМИ. В то же время, на каждом информационном уровне существует определенный предел достижениям. Поэтому, достигнутый успех служит признаком целесообразности перевода на более высокий информационный и коммерческий уровень деятельности. Верх-

ний предел отдельного СМИ гораздо ниже, чем его же, но в составе холдинга. Именно по этой причине преимущества электронных газет раскрываются гораздо полнее при их объединении в порталы.

В рамках портала СМИ объединены таким образом, чтобы пользователям (не только любителям, но и профессионалам) было проще и удобнее находить, извлекать и анализировать информацию. В рамках портала электронные версии различных газет связаны системой гиперссылок, позволяющих пользователю легко извлекать контекстно связанную информацию, оставаясь на портале и используя все возможности встроенного в него поискового и аналитического аппарата. Это же обстоятельство резко увеличивает информационный потенциал портала и его потенциал как носителя рекламы. Поэтому в следующем разделе мы отдельно остановимся на проблеме интеграции всех средств массовой информации и путях для достижения этой цели.

Футурологи предрекают, что бумажные газеты постепенно будут отмирать, а журналисты и репортеры будут отправляться на редакционное задание не с авторучкой и блокнотом, а с переносным компьютером и цифровой фото– или видеокамерой. Репортажи таких "электронных" журналистов могут быть немедленно опубликованы в глобальной сети. Однако пока никто не знает ответа на вопрос, каким образом можно превратить сетевое издание в доходный бизнес. И, вероятнее всего, печатные газеты отомрут совсем не скоро, иначе зачем бы тот же "Вашингтон Пост" вложил сотни миллионов долларов в строительство новых типографий в окрестностях Вашингтона. Ясно уже сейчас, что бизнес-модель газетной индустрии необратимо меняется – объявления с газетных страниц перемещаются в сетевую среду. А объявления уже сотни лет служили основным источником дохода.

## 1.7. Конвергенция СМИ

В XX веке началось объединение средств массовой информации в тресты, холдинги и прочие экономические монополии, зачастую принадлежащие одному лицу – медиа-магнату. Тенденции эти плавно перешли в XXI век. Известны такие медиа-магнаты в США, Англии и даже в нашей стране. Нередко эти люди владеют СМИ в различных регионах одной страны и даже во многих странах.

Первые такие объединения появились в газетной индустрии, затем стали объединяться газеты и радиостанции. С середины прошлого века в корпорации стали входить телевизионные каналы. На рубеже веков в состав монополий СМИ начали включать системы кабельного телевидения, а также Web-серверы и Web-сайты, распространяющие разнообразную информацию в глобальной сети Интернет. Хозяйственное объединение различных СМИ, единое руководство ими и единая информационная политика породили процессы информационной интеграции, т. е. попытки совмещения процессов подготовки информации для публикации в газетах, передач по телевидению и радио и демонстрации на информационных сайтах в сети Интернет. Единое информационное пространство СМИ, образовавшееся в результате такой интеграции, назвали английским словом "Newsroom", т. е. "комната новостей" в буквальном переводе. Соответствующие информационные процессы получили название конвергенции СМИ.

Процесс конвергенции нашел отражение на первой в новом тысячелетии конференции СМИ (IFRA) в Барселоне в мае 2001 г. Конференция была организована IFRA – международной организацией информационных и консалтинговых услуг, предоставляемых организациям, работающим в сфере издательского (в частности, газетного) бизнеса. IFRA издает свой журнал, переводимый в числе прочих и на русский язык, а также имеет международный сервер [www.ifra.com](http://www.ifra.com) и его представительство в русской части сети Интернет – [www.ifra.ru](http://www.ifra.ru). На конференции был поднят вопрос, станет ли конвергенция через 10–20 лет повсеместной, на который отвечал директор американской фирмы интерактивных и консультативных услуг Advanced Interactive Media Group С. Клейн. По его мнению, конвергенция неизбежна, так как никакое отдельное (не интегрированное) СМИ не сможет соперничать по своему воздействию на человека с объединенным эффектом, определяемым суммарным действием различных СМИ. Иначе говоря, чтобы остаться конкурентоспособным, неизбежно придется принять и активно применять на практике стратегию конвергенции.

Директор Центра передовых газетных проектов IFRA NewsOps К. Нортэп отметил, что совмещение информационных процессов в различных видах СМИ – это не технологический мутант или досужая выдумка журналистов, а фактор современной информационной экономики, определяемый требованием объективной необходимости для людей все большего количества информации, представленной в самых различных формах.

Уже сейчас большинство крупных газет и телекомпаний пытается использовать Web-сайты как вспомогательные источники информации, средство интерактивного взаимодействия с читателями и даже средство улучшения имиджа данного СМИ. Например, сайт газеты "Спорт-Экспресс" весьма популярен среди читателей – пользователей сети Интернет, так как он позволяет читателям оперативно (с задержкой в 1–2 минуты) получить самую оперативную спортивную информацию (например, результаты всех футбольных матчей очередного тура). Наличие такого сайта газеты значительно повышает ее рейтинг среди читателей, способствует росту ее популярности.

Условно все сетевые СМИ можно разбить на 3 группы. Первая группа – это сетевые СМИ, которые служат дополнением к печатной версии (представительство в сети газет "КоммерсантЪ", "Деловой Петербург", "Смена", "Час Пик", "Спорт-экспресс" и пр.). Такие



сетевые представительства обычно являются дотационными и используются для повышения имиджа основного издания и самого издательства.

Ко второй группе относятся чисто сетевые информационные СМИ. Здесь представлены информационные агентства "Интерфакс", "Итар-Тасс", "Прайм-Тасс", предоставляющие оперативные новости, и сетевые средства, которые, помимо новостей, предоставляют и аналитические материалы. К последним относится Северо-Западное представительство сетевого издания **Страна.ru** (<http://northwest.strana.ru>), которое силами семи штатных журналистов и восьми собкоров в регионе создает до 20—25 сообщений в день о различных событиях. Помимо этого ежедневно представляется в среднем 7—8 аналитических статей, интервью и репортажей. По словам директора портала "Виртуальный Петербург" ([www.vpr.ru](http://www.vpr.ru)), его издание — это, прежде всего, PR-площадка для клиентов компании, т.е. средство сообщить о себе рекламную информацию для повышения престижа и известности каждой фирмы, представленной на этом портале.

Третья группа — это специализированные (тематические) интернет-издания. В первую очередь, это сетевые издания, освещающие высокотехнологический рынок. Петербургское издание Internet.ru ([www.internet.ru](http://www.internet.ru)) всегда оценивалось специалистами как одно из наиболее авторитетных. Ежемесячные затраты на издание журнала колебались в рамках \$5—7 тысяч, а сборы от рекламы достигали \$4 тысяч. Internet.ru, который входил в холдинг Port.ru, всегда существовал лишь за счет инвестиций. Современный проект ComNews.ru ([www.comnews.ru](http://www.comnews.ru)) также освещает события рынка высоких технологий. Более половины штата сотрудников составляет рекламный отдел, благодаря работе которого сайт приносит небольшой доход (несколько сот долларов в месяц). Владелец сайта ComNews "Северо-Западная Медиа Группа" издает также ежегодник «Энциклопедия связи "Петербург Телеком"» и журнал "Стандарт", выходящий один раз в два месяца. По мнению руководителей проекта ComNews ни технологически, ни коммерчески неотделим от прочих информационных средств компании, т.е. в этом случае конвергенция налицо. Именно она обеспечивает жизнеспособность сетевого издания.

Какие же факторы тормозят процесс конвергенции?

# Прежде всего, это отсутствие надежной модели получения доходов, связанных с организацией информационных сайтов.

# Недостаточность сетевого ресурса и низкая скорость передачи информации также служит тормозящим фактором, так как увеличивает время загрузки Web-страниц сайта на компьютер клиента и снижает скорость обновления информации.

# Консервативность большинства самих работников традиционных СМИ, которые не осознают то, что информация об одном и том же событии в различных СМИ должна оформляться по-разному.

Если первые два фактора носят объективный характер, то последний фактор достаточно субъективен. Его роль может быть снижена и преодолена с помощью правильной интеграционной политики в работе с творческим составом и техническими сотрудниками СМИ. Первые результаты в этой сфере обнадеживают.

В качестве примера можно рассмотреть отделение группы Independent. Это — одно из отделений газетной империи Тони О'Рейли — Independent News and Media. Отделение территориально находится в ЮАР и до 1999 г. насчитывало 14 изданий. В 1999 г. было дополнительно создано специализированное интернет-издание Independent Online ([www.iol.com](http://www.iol.com)). Этот сайт черпает сведения из газет группы, помещая, в частности, на своей заглавной странице последние новости. В результате появилось новое национальное интерактивное издание ЮАР, в котором объединяется в единое целое информация из многочисленных источников. Далее можно остановиться на СМИ в составе группы Media General в Майами, США. В нее входит газета Tampa Tribune, телестудия WFLA, отделение радиовещания и Web-сайт

www.TBO.com (Tampa Bay Online). Внешний вид лицевой страницы этого сайта показан на рис. 1.5. Все сотрудники группы Media General работают в одном здании (Центр новостей), используют единую интрасеть (корпоративную локальную сеть, одним из достоинств которой является возможность совместной работы нескольких пользователей над одним документом) и освещают события в тесном сотрудничестве. На достижение положительного результата было затрачено немало усилий и около 8 лет работы. Однако руководство Центра новостей считает, что до полной конвергенции СМИ в составе этой группы еще далеко. Реальной проверкой работы Центра новостей послужили трагические события 11 сентября 2001 г. в США. Журналисты телестудии WFLA первыми начали репортаж с места событий, затем подоспели на вертолете работники Webсайта. Они отсняли материал для размещения в сети и передали информацию журналистам, сотрудничающим в газете. Такое тесное и эффективное сотрудничество представителей различных СМИ в составе группы Media General оказалось неожиданным даже для ее руководства.

Интересно проследить процесс организации работ по интеграции газетной информации с данными, представленными на сайте www.TBO.com. Коллектив младших редакторов подбирает интересный фактический материал и иллюстрации к нему. Затем дежурный редактор просматривает все собранные материалы и определяет наиболее значимую его часть для использования в качестве передовицы на Web-сайте. На следующем этапе он распределяет наиболее интересные материалы по отдельным страницам сайта. В задачу дежурного редактора входит также просмотр телетайпных и интернетновостей, отбор и размещение их на странице новостей сайта (рис. 1.5).

Описанная организация интеграционного процесса стала возможной благодаря тому, что все газеты группы охвачены региональной сетью WAN (Wide Area Network), в которой используется эффективная система управления потоком данных. Интеграция газет с помощью этой сети положила начало объединению творческих работников, возможно в дальнейшем это значительно упростило конвергенцию СМИ в рамках группы Media General. Нетрадиционный подход к процессу обработки информации внутри этой группы проявляется даже в мелочах. Например, фотографы Independent News and Media не печатают на бумаге фотографии с фотопленок, а сканируют кадры пленки с помощью подключенных к сети сканеров (или используют для работы цифровые камеры). В результате графические кадры сохраняются в виде сетевых файлов и становятся доступными одновременно как для газет, так и для страниц сайта.





Рис. 1.5. Фрагмент лицевой страницы сайта Tampa Bay Online

Процесс конвергенции проявился не только в развитых странах, но даже в такой "классической" развивающейся стране, которой является Индия. В стране, где использование сети Интернет пока весьма ограничено фирме Mid-Day удалось создать англоязычную газету, выходящую в Бомбее (13 миллионов жителей) 3 раза в день, три радиовещательные станции FM-диапазона, интернет-портал Chalomumbay, который посещает в день свыше одного миллиона пользователей, и телефонную информационную службу.

Для реализации долговременного плана конвергенции все информационные сотрудники фирмы Mid-Day работают в составе единого "нюсрума" (newsroom). Они делятся на 2 большие группы: на тех, кто занимается сбором информации на различных языках, интересующей читателей (они здесь называются newser's – новообразование, полученное слиянием двух слов news user – потребитель новостей), и тех сотрудников, которые адаптируют contents (содержание) к особенностям специфического СМИ.

Технической основой успешного функционирования медиа-группы Mid-Day служит объединенная информационная база данных MIMID (Mid-Day Integrated Mumbai Information Database). Уникальная и единственная в Индии мультимедийная база данных рассчитана на газетные публикации, теле- и радиопрограммы, сеть Интернет и другие информационные средства. Наряду с текстовыми файлами – в том числе сверстанными в пакете QuarkXPress публикациями – в ней хранится графика в формате JPEG, аудиофайлы в формате MP3, цифровое видео в формате MPEG2 и MPEG4, а также HTML и XML-файлы. HTML и XML (Extensible Markup Language) форматы будут рассмотрены в главе 2, а гра-

фический формат JPEG и форматы цифрового аудио и видео (MP3, MPEG2 и MPTG4) – в главе 3. Пока в процессе конвергенции в медиа-группе Mid-Day доминируют печатные СМИ, которые охватывают около 70% деятельности группы и приблизительно такую же долю информационных ресурсов. Но постепенно и неуклонно происходит движение в направлении повышения роли сетевых и мультимедийных СМИ.

Процесс конвергенции постепенно распространяется и на мобильные средства сетевого доступа. Впрочем, это неудивительно, так как владельцы мобильных телефонов в крупных городах нашей страны составляют значительный процент общества. Например, в Петербурге их число существенно превысило 1 миллион. Имеются в виду абоненты операторов сотовой связи, предоставляющих услуги по цифровым протоколам, таким, как GSM 900/1800, так как протоколы WAP (Wireless Application Protocol), SMS (Simple Messaging Service – Служба передачи простых сообщений) и их аналоги доступны только для абонентов сетей GSM.

Основную часть пользователей сотовых телефонов составляют молодые (до 35 лет) люди, ведущие активный образ жизни и желающие постоянно быть в курсе событий, в том числе и в дороге, тогда, когда нет возможности воспользоваться традиционными источниками информации.

Какой же сетевой сервис может быть предоставлен для них сейчас? Пока это преимущественно возможность пользования электронной почтой в форме передачи SMS-сообщений объемом до 160 символов. В последнее время получил большое распространение протокол WAP (Wireless Application Protocol), позволяющий просматривать специально адаптированные Web-страницы прямо на экранах сотовых телефонов, но пока существует весьма немного действительно интересных и информативных WAP-страниц.

Лексику, которой пользуются для составления SMS-сообщений, внесли в Краткий оксфордский словарь английского языка. Как сообщает агентство Рейтер, решение об этом было принято 12 июля 2001 г. Словарь SMS-языка выделен в особый раздел. В нем будут содержаться аббревиатуры и сокращения, которые уже стали вторым языком для миллионов англоязычных пользователей. Он будет насчитывать несколько десятков слов и выражений. В него также включают так называемые "смайлики" (smile, т. е. улыбаться) – средства выражения эмоций в текстовом виде. Среди распространенных сокращений для SMS популярностью пользуются: BBLR ("be back later" – "вернусь позже") и HAND ("have a nice day" – "удачного дня"). А фразу "Great news – text messages are now in the Concise Oxford Dictionary" ("Классная новость – текстовые сообщения включены в Краткий оксфордский словарь") можно записать так: "GR8 news – txt MSGS r now in the COD". Редакторы словаря решили, что этот новый "подвид английского" уже настолько устоялся и получил такое широкое распространение, что пришла пора его узаконить.

Однако сами владельцы мобильных телефонов уже в настоящее время желают иметь возможность пересылки текста объемом в несколько тысяч символов, а также аудио- и видеосообщений. Практически все хотят, чтобы мобильный телефон стал многофункциональным, позволял прослушивать радио, просматривать фрагменты телевизионных передач, предоставлял возможность сетевых игр, прослушивания конкретных композиций и альбомов. Для деловых людей желательно иметь возможность осуществлять покупки и финансовые операции, заказывать билеты и просматривать новости. Большинство перечисленных услуг может быть оказано на основе удаленного доступа к сети Интернет. Процесс конвергенции должен предусматривать возможность адаптации Интернет – протоколов и средств визуализации Web-страниц для мобильных телефонов. Естественно, и мобильные телефоны должны стать многофункциональными, чтобы обеспечить возможность получения их владельцами соответствующих сервисов. Что же уже сделано в этом направлении и что еще предстоит сделать?

Японский телефонный гигант NTT DoCoMo представил в октябре 2001 г. высокоскоростную радиотелефонную сеть и три новых телефонных аппарата, при помощи которых можно проводить видеоконференции в режиме реального времени, посылать и принимать сообщения электронной почты объемом до 10 тыс. знаков и многое другое. Помимо прочего, абоненты сети получили возможность осуществлять покупки, резервировать концертные и авиабилеты и пополнять свои балансовые счета у оператора. В общем, они имеют возможность осуществлять практически любые финансовые операции (кроме, разве что, покупки акций) прямо с сотового телефона. Сеть начала функционировать с мая 2001 г. в Токио, где около пяти тысяч человек (включая как индивидуальных, так и корпоративных пользователей) тестировали ее, а в 2002 г. распространилась по остальной территории страны. Самый дешевый из "мобильников" для этой сети (модель FOMA P2401) стоит около \$250, а наиболее престижная модель (FOMA P2101V) – немного более \$500. Первые образцы мобильных телефонов третьего поколения (3G) с поддержкой видеорежима для сетей DoCoMo произведены фирмой Matsushita Communication Industrial.

Ведущий американский сотовый оператор Verizon Wireless также планировал запустить свою высокоскоростную сеть третьего поколения к концу 2001 года. Однако, события 11 сентября нарушили эти планы. Сейчас же в США практически нет свободных частот для введения в действие сетей 3G. По информации агентства Рейтер, Федеральная комиссия связи США и Министерство торговли занимаются поисками свободных диапазонов для 3G-связи, однако большинство частот оказалось занято Министерством обороны, Федеральным управлением гражданской авиации и другими правительственными структурами, причем вероятность передачи этих частот сотовым операторам весьма низка, так как проблемам общественной безопасности в США уделяется сейчас основное внимание.

В конце 2001 г. лаборатория Intel Nizhny Novgorod Lab (INNLL) в Нижнем Новгороде представила свою разработку – мультимедийный сотовый видеотелефон. Разработка нижегородской лаборатории основана на принципе создания изображения на основе одной из трех пространственных моделей, используемых в программном пакете 3ds max и рассмотренных в разд. 6.5. Те из читателей, кто знаком с этим программным пакетом, знает, что такие модели используются в нем для создания высококачественной цифровой видеоанимации для телевидения и кино. На экране видеотелефона изображение будет повторять движения реального абонента. При этом виртуальная модель собеседника может отличаться от оригинала. Например, виртуальный собеседник может быть представлен во фраке и при галстукке, в то время как реальный видеоабонент во время разговора принимает ванну. Создание анимационной модели вместо передачи реальной информации об объекте позволяет в 500 и даже 1000 раз уменьшить объем передаваемой по сети информации, что очень важно для мобильных сетей связи, одновременно обслуживающих большое количество абонентов.

По прогнозам экспертов к 2004 году число абонентов сотовой связи в мире превысит миллиард человек. Такие же прогнозы роста даются для сети Интернет. Более того, последние европейские исследования говорят о том, что более 50% абонентов сетей GSM являются пользователями Интернета. Уже сегодня многие абоненты во всем мире используют широкие возможности стандарта GSM для того, чтобы входить в Интернет, подключаться к корпоративным базам данных или к домашнему компьютеру, просматривать электронный почтовый ящик и отправлять факсимильные сообщения. Следующим шагом по усовершенствованию инфраструктуры "Мобильный телефон – Интернет" стало создание и внедрение протокола WAP. Многие интернет-сайты создали для своих клиентов специальные WAP-версии своих Web-страниц. На дисплеях WAP-телефонов оптимально размещаются текстовое содержание сайтов и небольшие графические вставки. Прогнозируется лавинообразный рост русскоязычных WAP-сайтов, представляющих самую разнообразную инфор-

мацию: новости политики и экономики, культуры и спорта, курсы валют, телепрограммы, прогноз погоды и многое другое (рис. 1.6 и 1.7).

Крупнейшие производители телекоммуникационного оборудования, компании Nokia, Motorola, Ericsson и Siemens, объявили о том, что они будут совместно разрабатывать универсальные решения для создания сетевых мобильных игр, основываясь на современных технологиях и платформах мобильной связи. С этой целью компании образовали форум Mobile Games Interoperability (MGI) Forum. Кстати, некоторые компании, такие как Jamdat Mobile и Unplugged Games, совместно с западными операторами уже предлагают сетевые WAP-игры: интерактивная (online) Fighting (бой, сражение – игра, связанная с уничтожением войск противника или противников), Gladiator (гладиатор, который борется с нападающими на него людьми или зверями) викторину Trivia (мелочи, в буквальном переводе с латыни), Soccer (футбол), Black Jack (азартная игра, аналогичная проводимым с игровыми автоматами) и многие другие. Демо-версии игр для мобильных телефонов можно посмотреть на сайтах Unplugged Games и JAMDAT ([www.ungames.com/tmpl.php/home/](http://www.ungames.com/tmpl.php/home/) и [www.jamdat.com/index.html](http://www.jamdat.com/index.html)).



Рис. 1.6. Пример получения на мобильном телефоне справочной информации о погоде в Санкт-Петербурге

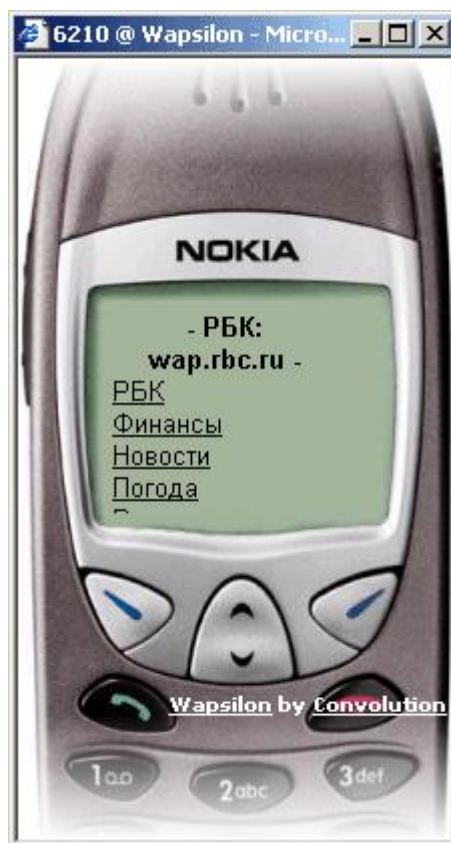


Рис. 1.7. Отображение специализированного Web-сайта wap.rbc.ru на экране мобильного телефона

Таким образом, владельцы "мобильников" имеют возможность не только пользоваться электронной почтой и просматривать новости в сети Интернет, но и пользоваться сетевыми развлекательными ресурсами, что особенно привлекательно для молодежи: старших школьников и студентов младших курсов. Однако сейчас интеграция мобильных средств связи в единое информационное пространство находится только в процессе становления, и, учитывая высокие скорости развития мобильных технологий, расширение сервисов для пользователей сотовой связи, а также рост разнообразных возможностей самих мобильных телефонов, рискнем предположить, что не так уж и далек тот день, когда наиболее "продвинутые" владельцы сотовых телефонов будут получать самые свежие новости на свой телефон в одном из мультимедиа-форматов.

## Контрольные вопросы

1. Попробуйте дать определение электронного издания. В чем основной недостаток такого издания?
2. Что такое Web-страница и Web-сайт? Чем Web-страница отличается от обычной страницы?
3. Работали ли вы в сети Интернет? Пользовались ли вы услугами некоммерческих библиотек? Какая из них понравилась вам больше других?
4. Отличаются ли электронные книги от печатных с точки зрения размещения иллюстраций. В чем состоит такое отличие? Постарайтесь вспомнить, в каких из прочитанных книг вы встречали качественные цветные иллюстрации?
5. Что может использоваться в электронном издании, но не может использоваться в печатном? Попробуйте перечислить все такие компоненты.
6. Какие основные форматы электронных изданий вы можете назвать? Чем они отличаются друг от друга?
7. Как могут распространяться электронные публикации?
8. С какими мультимедийными изданиями вы уже знакомились?
9. Видели ли вы электронные справочники и указатели и какие именно?
10. Что можно сказать о связи и интеграции печатной и электронной книги?
11. Читали ли вы электронные газеты? Какие именно? Чем они отличаются от печатных СМИ?
12. Как вы считаете, почему электронные СМИ пока, в основном, убыточны?
13. Что такое конвергенция СМИ? Зачем она нужна и кому?
14. Знаете ли вы примеры конвергенции СМИ в нашей стране? В чем они проявляются?
15. Знаете ли вы примеры порталов? А в Российской части сети Интернет? Каковы критерии того, что конкретный сайт может быть причислен к порталам? Что, по вашему мнению, следовало бы еще включить в функции и возможности портала?
16. У вас есть мобильный телефон? Если он есть (или когда он у вас будет) – хотели бы вы с его помощью получать информацию из сети Интернет? Играть в компьютерные сетевые игры со своими друзьями?
17. Как связаны друг с другом сеть Интернет и мобильная телефония? Укажите примеры. А вы сами пользовались информацией из сети Интернет, передаваемой на мобильный телефон? Знаете ли вы сайты, специально адаптированные под потребности мобильных телефонов?

## Глава 2

# Структура и элементы гипертекстовых документов



Эта глава целиком посвящена описанию методики создания HTML-документов с помощью средств языка HTML. В частности, рассматривается общая структура различных HTML-документов, включая издания с фреймовой структурой и HTML-формы. Рассмотрены основные разновидности тэгов для создания заголовочной части документа и его тела, а также форматирования текста и иллюстраций, организации списковых и табличных структур, включения в документ ссылок и комментариев. Описаны варианты электронного учебника на основе фреймовой структуры. Анализируется пример HTML-формы, используемой на сервере Rambler для предоставления клиентам бесплатного электронного почтового адреса. В заключение дается общая характеристика средств для создания HTML-страниц с динамически изменяющимся содержанием, а именно языка JavaScript и каскадных таблиц стилей.



## 2.1. Общая характеристика и структура HTML-документа

На заре компьютеризации пользователям приходилось работать с простыми текстовыми редакторами, такими как WordStar, для которых информация, отображаемая на экране, отличалась от выводимой на бумагу с помощью принтера. В текст документа вставлялись специальные управляющие символы, которые не отображались на экране, но обеспечивали вывод на печать в нужном формате отдельных фрагментов документа. При форматировании одна группа символов определяла начало фрагмента, затем следовал текст, к которому применялось это форматирование, а после него следовали символы конца фрагмента. Так задавалась разрядка, курсив, полужирный шрифт и пр. Аналогичные способы применялись и при работе с первыми версиями СУБД (системы управления базами данных) dBase (например, dBase II), о чем автор может судить по своему личному опыту.

Тот же принцип положен в основу структуры HTML-документа, однако управляющие символы, которые здесь называются тэгами, определяют особенности отображения информации, выводимой уже не на принтер, а на экран монитора. Значительная часть таких тэгов используется парами: вначале открывающий тэг, затем объект управления, а в конце – закрывающий тэг. Такая конструкция называется контейнером, так как объект форматирования размещается внутри нее. Некоторые тэги принципиально не нуждаются в паре. Примером может служить тэг принудительного перевода строки. Тэг может включать в себя некоторые параметры (атрибуты), которые размещаются непосредственно после имени тэга. Если параметров несколько, то в качестве разделителей используются пробелы.

Для просмотра HTML-документов используют специальные программы, которые называются браузерами (Browser, т. е. средство просмотра – в дословном переводе). Такие программы нуждаются в графической оболочке. В частности, значительное число браузеров (Microsoft Internet Explorer, Netscape Communicator, Opera и др.) работают в среде Windows 95, 98 и 2000.

Популярность языка HTML росла вместе с развитием сети Интернет. Отсутствие в течение некоторого времени стандартов на язык HTML привело к тому, что некоторые браузеры отображали должным образом не все тэги, а в версиях различных фирм одни и те же тэги могли трактоваться неодинаково или даже не поддерживались. Широкое использование HTML-документов в сети Интернет привело к появлению международных стандартов на этот язык, называемых спецификациями языка HTML. Этой работой занялась широко известная организация World Wide Web Consortium (W3C). Первая из спецификаций, признанная большинством компаний-разработчиков, явилась HTML 2.0, представленная в конце 1995 г. Вскоре была выпущена предварительная (рабочая) версия спецификации 3.0, а в мае 1996 г. – проект спецификации HTML 3.2. После длительного обсуждения и исправления в январе 1997 г. эта спецификация стала официальной рекомендацией для разработчиков HTML-документов и браузеров. В декабре 1997 года официальной рекомендацией стала спецификация HTML 4.0, которая остается в действии до настоящего времени.

Основой спецификации HTML 4.0 стало отделение параметров описания документов от параметров представления отдельных его фрагментов на экране монитора. Такое разделение облегчает адаптацию языка к различным платформам и средам (Windows, UNIX, DOS и пр.) и упрощает процесс внесения изменений в документы. В соответствии с такой концепцией для описания документа следует использовать таблицу стилей, причем это понятие сходно с используемым в современных текстовых редакторах, таких как Word 97 или 2000. Использование же данных о форме представления документа попеременно с содержанием самого документа не рекомендуется.



Спецификация HTML 4.0 отменяет ряд ранее использовавшихся тэгов. Отмена тэга означает, что этот тэг по-прежнему поддерживается браузерами, но его применение в современных документах не рекомендуется. В дальнейшем такие тэги могут быть переведены в разряд устаревших, которые уже могут не поддерживаться браузерами. Информация такого рода может быть получена на сайте <http://www.w3.org/TR/>.

Любой HTML-документ заключен в контейнер `<HTML>` и `</HTML>`.

#### Замечание

Здесь и далее по тексту запись "контейнер `<HTML>` и `</HTML>`" означает, что это контейнер с открывающим `<HTML>` и закрывающим `</HTML>` тэгами. Аналогичен смысл записи и для других возможных тэгов.

Сам документ, как уже говорилось выше, представляет собой обычный текстовый файл. Его можно просматривать в DOS с помощью стандартных редакторов и средств просмотра текстовых документов, при этом вместе с самим текстом мы будем видеть и тэги. Структурно документ распадается на 2 части: заголовочную и основную, или тело документа. Первая часть находится внутри контейнера `<HEAD>` и `</HEAD>`, а основная часть – внутри контейнера `<BODY>` и `</BODY>`.

Единственный обязательный тэг, который используется в заголовочной части, — `<TITLE>`, причем он образует контейнер, внутри которого размещается текст заголовка. При загрузке документа в браузер текст заголовка загружается в первую очередь в окно заголовка браузера. Если текст заголовка достаточно информативен, пользователь в процессе поиска релевантных документов уже на этапе загрузки документа может определить, действительно ли ему нужен этот документ, а если не нужен, то сразу отказаться от его загрузки и перейти к загрузке следующего. Текст названия используется и при создании закладки для данного документа, с этой точки зрения также очень важна его информативность.

#### Замечание

*Релевантность* (relevancy) – это мера, определяющая, насколько полно тот или иной документ отвечает критериям, указанным в запросе пользователя. Однако не все документы, признанные поисковой системой наиболее релевантными, будут таковыми по мнению пользователя.

Многие HTML-документы связаны друг с другом, и ссылки от одного документа на другие могут быть абсолютные и относительные, причем последние – формируются относительно того каталога, в котором расположен документ – источник ссылки. Абсолютные ссылки длинные, а относительные, хотя и короче, но перестают работать при перемещении основного документа. В HTML-документы можно включать сразу обе ссылки, чтобы связи между документами не нарушались в любых условиях.

Для формирования ссылок в заголовочной части используется тэг `<BASE>`, в котором используется один единственный параметр `HREF` (Hyper Reference – гиперссылка), в качестве значения которого указывается URL (Universal Resours Locator) – адрес файла, на который производится ссылка. Параметр `HREF` используется в том же качестве и в некоторых других тэгах, например, тэге-указателе ссылки (см. разд. 2.5). Адрес может быть локальным, т. е. задавать размещение файла в каталогах и подкаталогах на том же самом компьютере, или сетевым, например:

```
<BASE HREF="http://www.uprint.spb.ru/main.htm">
```

Для организации логической связи и соподчиненности электронных изданий может использоваться тэг <LINK> (Link – связь). В этом тэге могут использоваться 4 параметра: HREF, REL (от Relation – отношение), REV (от Reverse – противоположный, обратный) и TYPE (тип). С помощью первого из них задается URL связанного с данным документом. Параметр REL определяет вид отношения между текущим и связанным с ним документом, а REV – обратное отношение (между другим и текущим документами). Последний параметр TYPE задает тип и параметры присоединяемой к документу таблицы стилей. Примеры тэга <LINK>:

```
<LINK REL="contents" HREF="(адрес)">
<LINK HREF="mailto: (адрес автора)" REV="made">
```

Другими значениями параметра REL могут быть bookmark (закладка), copyright (авторское право), glossary (глоссарий, словарь специальных терминов в конце книги), help (помощь). Значение made параметра REV является признаком обратного отношения. Другие его значения: autor (автор), editor (редактор), publisher (издатель).

Позднее в состав заголовочной части был добавлен специальный тэг <META>, с помощью которого задается метаинформация, связанная с данным электронным изданием, или атрибуты для ускоренного поиска. Частные случаи метаинформации – это имя автора, издателя, редактора, название издательства, время публикации и другие характерные признаки издания. Тэг <META> имеет 2 параметра: NAME, которым задается имя атрибута и CONTENT (содержание), определяющий значение этого атрибута. Например:

```
<META NAME="author" CONTENT="В. А. Вуль">
<META NAME="description" (описание) CONTENT="Электронные издания">
```

Вероятно, приведенные примеры не нуждаются в каких-либо комментариях. Итак, внутри контейнера <HEAD> и </HEAD>, ограничивающего заголовочную часть HTML-документа, обычно используется один контейнер <TITLE> и

</TITLE>, внутри которого размещается текст заголовка, и могут быть включены тэги <BASE>, <LINK> и <META>. Отметим, что HTML-документ, содержащий только заголовочную часть, будет нормально отображаться браузерами как пустой документ. Пример такого документа представлен на рис. 2.1.



Рис. 2.1. Пример отображения в браузере MS Internet Explorer пустого HTML-документа

А ниже показан соответствующий этому документу HTML-код:

```
<HTML>
```

```
<HEAD> <TITLE>Пустой документ </TITLE> </HEAD>
```

```
<BODY> </BODY>
```

```
</HTML>
```

Замечание

Наличие пробелов между тэгами или внутри контейнеров никак не влияет на отображение HTML-документа.

## 2.2. Тело документа и оформление его основных фрагментов

### 2.2.1. Тело документа

Внутри контейнера <BODY> (BODY – тело) и </BODY> размещается содержательная часть электронного документа. Сам тэг <BODY> содержит ряд параметров, ни один из которых не является обязательным. Параметры данного тэга обеспечивают оформление всего тела документа. Перечислим их:

- # LINK – определяет исходный цвет ссылки;
- # BACKGROUND (фон) – задает URL изображения, определяющего фон тела документа;
- # BOTTOMMARGIN (нижнее поле) и TOPMARGIN (верхнее поле) – задает границу нижнего и верхнего полей документа в пикселах;
- # LEFTMARGIN (левое поле) и RIGHTMARGIN (правое поле) – границы левого и правого полей документа в пикселах;
- # BGCOLOR (Background Color – цвет фона) – задает цветовой оттенок фона документа (аналог тонирования бумаги, на которой печатается издание);
- # BGPROPERTIES (Background Properties – свойства фона) – определяет свойства фона, задаваемого предыдущим параметром;
- # LINK – задает цвет еще не просмотренной ссылки;
- # SCROLL (прокрутка) – определяет наличие полос прокрутки в документе, отображаемом в окне браузера;
- # TEXT – определяет цвет текста;
- # VLINK – цвет уже просмотренной ссылки.

Ряд параметров связаны с использованием цветовых оттенков в HTML-документах. Отметим, что по умолчанию в них используется RGB-цветовое пространство, причем значение каждого из основных цветов имеет 256 уровней и задается в виде 2 цифр в 16-ричной системе счисления от 00 до FF. Таким образом, цветовой тон задается последовательностью из шести 16-ричных цифр, которым предшествует символ #, например, последовательность #FFFFFF соответствует белому цвету максимальной интенсивности, а #800080 – фиолетовому тону, интенсивность которого равна половине от максимальной.

Приведем примеры использования параметров в тэге

```
<BODY>: <body link="#000080" background="image/background.gif">
<body bgcolor="#C0C0C0">
```

В первом примере задан синий цвет еще не просмотренной ссылки и в качестве фонового изображения используется файл background.gif, причем этот файл указан относительной ссылкой и находится в подкаталоге IMAGE. А во втором примере задан светло-серый фон для документа.

### 2.2.2. Тэги логического форматирования текста

Тэги логического форматирования текста позволяют выделить определенные типы текстовых фрагментов, т. е. осуществить структурную разметку текста. Их называют также тэгами уровня блока, так как они не применимы к отдельным словам, а воздействуют

лишь на фрагменты текста из нескольких слов. Так, два тэга <ABBR> (Abbreviation) и <ACRONYM> ("акроним", сокращение – заимствование из греческого языка) позволяют выделить аббревиатурные сокращения в пределах текста, причем второй используется для произносимой аббревиатуры. В тэге <ACRONYM> может использоваться параметр TITLE, значением которого является полный текст сокращения. Например, контейнер:

```
<ACRONYM TITLE="Московский Государственный Университет Печати"> МГУП
</ACRONYM>
```

позволяет при указании курсором манипулятора мышью на аббревиатуру выводить во всплывающем окне соответствующий ей полный текст названия этого Университета.

Тэг-контейнер <CITE> (цитата) используется для выделения цитат, названий книг, газет и журналов. Браузеры обычно выводят текст, находящийся внутри контейнера, курсивом. Тэг-контейнер <CODE> (код) выводит текст как фрагмент программного кода (строки листинга программы) моноширинным шрифтом, а <DFN> (Definition – определение) отмечает текстовый фрагмент как определяемый термин. Аналогичным образом тэг-контейнер <STRONG> (сильный, веский) выделяет содержимое как важный фрагмент текста.

Тэг <INS> (Insert – вставка) отмечает фрагмент текста, как вставку, т. е. с его помощью можно отслеживать изменения, вносимые в основной текст. Тэг может иметь 2 параметра: CITE (здесь в смысле – ссылаться, не путать с одноименным тэгом) и DATETIME (дата и время). С помощью первого параметра задается URL документа, поясняющего причины вставки, а второй параметр указывает дату и время вставки фрагмента с учетом часового пояса. Тэг-контейнер <DEL> (Delete – удалять, уничтожать) отмечает текст, как наметенный к удалению. У него имеются параметры CITE и DATETIME, назначение которых точно такое, как и в тэге <INS>. Ниже приводятся несколько примеров записи тэгов, которые не требуют каких-либо комментариев.

```
<CITE>"Известия"</CITE> – популярная Российская газета
```

```
<DFN>Netscape Communicator 4.5</DFN>– версия 4.5 популярного пакета фирмы Netscape, в который, в частности, входит браузер для просмотра HTML-файлов.
```

```
<INS CITE="www.uprint.spb.ru/mydoc.htm" DATETIME="2001-1-27 T16: 1:40+ 0.00">
(вставленный фрагмент текста) </INS>
```

#### Замечание

Кроме браузера Netscape Navigator в Netscape Communicator имеется HTML-редактор Netscape Composer и почтовый агент Netscape Messenger.

Ниже, на рис. 2.2, представлен пример отображения гипертекстового документа, в теле которого содержится ряд тэгов логического форматирования текста.

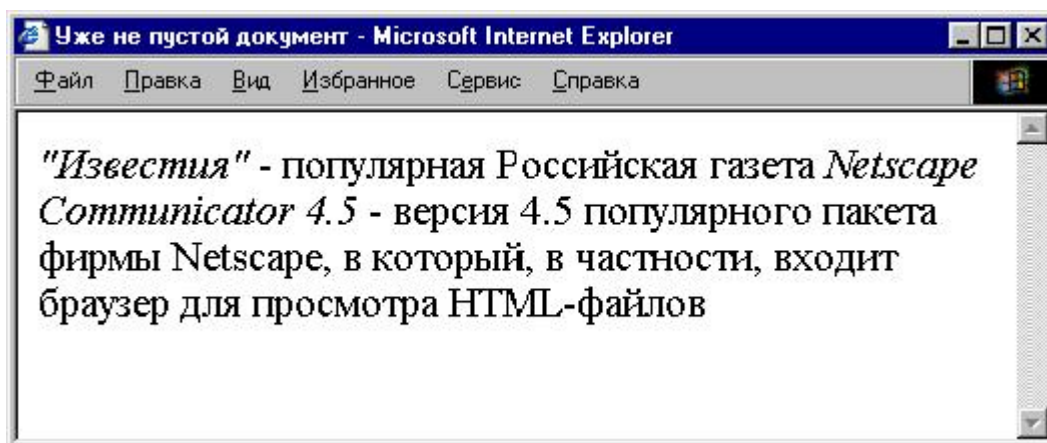


Рис. 2.2. Пример отображения в браузере MS Internet Explorer тэгов логического форматирования текста

### 2.2.3. Тэги физического форматирования текста

Тэги физического форматирования определяют формат заключенного внутри соответствующих контейнеров текста при отображении на экране монитора. Как уже говорилось в начале этой главы, вместо них желательно использовать тэги логического форматирования, если последние могут выполнить аналогичную функцию.

Так, тэг-контейнер **<B>** (Bold – жирный) выделяет заключенный внутри его текст полужирным начертанием. Аналогичным образом, тэги **<I>** (Italic – курсив) и **<U>** (Underline – подчеркивание) означают курсивное начертание и подчеркивание текста соответственно, а тэг **<TT>** – моноширинный шрифт. Тэги **<STRIKE>** (перечеркнуть) и **<S>** (аббревиатура от Strike) совпадают по назначению и обеспечивают перечеркивание фрагмента текста горизонтальной линией, как намеченного к удалению. Надо отметить, что все перечисленные тэги относятся к устаревшим и не рекомендуются к применению. Так, вместо тэга **<B>** предпочтительнее использовать рассмотренный выше тэг логического форматирования текста **<STRONG>**, вместо **<STRIKE>** или **<S>** – **<DEL>**, а вместо **<TT>** – **<CODE>**.

Из других тэгов "шрифтового оформления" отметим **<BIG>** (большой) и **<SMALL>** (малый), которые увеличивают или уменьшают размер шрифта на одну единицу, **<SUB>** (приставка, указывающая на положение ниже чеголибо) и **<SUP>** (от Super – приставки, переводимой как над– или сверх), размещающие символы выше или ниже уровня строки. Более универсальным в этой группе является тэг **<FONT>** (шрифт), который имеет 3 параметра: **FACE** (рисунок шрифта, гарнитура), **SIZE** (размер) и **COLOR**. Первый параметр позволяет указать используемую в текстовом фрагменте гарнитуру, причем, если такая гарнитура не установлена на компьютере, то параметр игнорируется. Можно задавать несколько наименований гарнитур, разделяя их запятыми, тогда будет использоваться первая по списку из имеющихся в наличии. Второй параметр задает размер (но не кегль) шрифта. Всего в языке HTML предусмотрено 7 размеров, из которых номер 3 используется по умолчанию. Последний параметр указывает цвет символов текста и может задаваться в виде группы из 7 символов, как указывалось выше, или в наименованиях цветовых тонов. Приведем несколько примеров:

```
<FONT FACE="Arial","Courier" SIZE=5 COLOR=red> текст </FONT>
<FONT FACE="Times" SIZE=2 COLOR=#ff0000 > текст </FONT>
<FONT FACE="Times"> 5 <SUP> <SMALL> 2 </SMALL> </SUP> </FONT>
```

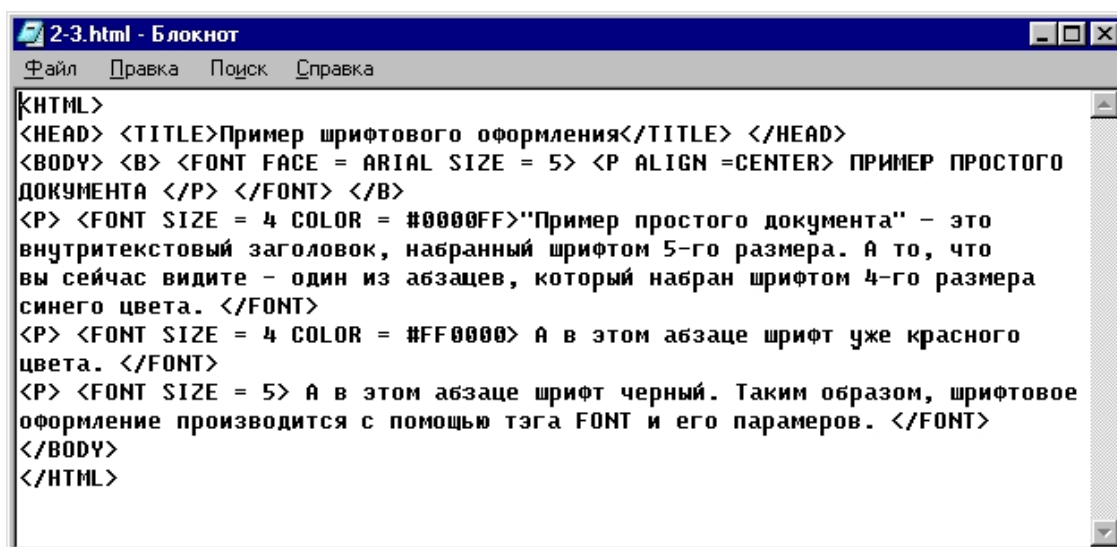
В первом примере символы текста, содержащиеся внутри контейнера будут отображаться гарнитурой Arial, если она установлена на компьютере, а в противном случае – гарнитурой Courier. Размер символов будет достаточно велик и они будут красного цвета. Во втором примере будет использована гарнитура Times, размер символов будет на 1 меньше нормального и символы будут того же самого красного цвета. В последнем случае в окне браузера будет отображаться  $5^2$ , причем размер показателя степени будет на единицу меньше, чем ее основания.

Третий пример иллюстрирует и другое важное свойство тэгов-контейнеров физического форматирования текстовых символов: контейнеры могут быть вложены один в другой, причем глубина вложений формально не ограничена. При написании HTML-документа вручную необходимо только тщательно следить, чтобы контейнеры строго размещались один внутри другого. Действительно, в примере контейнер `<SMALL>` и `</SMALL>` расположен внутри `<SUB>` и `</SUB>`, а последний вложен в контейнер `<FONT>` и `</FONT>`.

Все рассмотренные выше тэги физического форматирования, если их действие рассматривать подобно функциям текстовых редакторов, относятся к шрифтовому оформлению текстовых фрагментов. Поэтому далее следует остановиться на тэгах, выполняющих функции, подобные форматированию абзаца. К ним, в частности, относится тэг разделения документа на абзацы

`<P>` (Paragraph – абзац), который помещается перед началом очередного абзаца. Закрывающий тэг `</P>` не обязателен, так как абзацы следуют друг за другом и начало следующего абзаца указывает на конец предыдущего. Браузеры (см. главу 4) обычно отделяют абзацы друг от друга пустой строкой. Параметром этого тэга является ALIGN, т. е. выравнивание. Параметры выравнивания: LEFT (влево), CENTER (по центру), RIGHT (вправо) и JUSTIFY (выключка, выравнивание по ширине), причем последний только сравнительно недавно начал поддерживаться браузерами.

При отображении документов в окне браузера место перевода строки определяется автоматически, в соответствии с реальными размерами его окна. В тех случаях, когда требуется выполнить принудительный перевод строки в определенном месте, используется тэг `<BR>` (Break – разрыв), у которого нет закрывающего тэга. Тэг `<BR>`, в частности, используется при работе со стихотворными строками. В отличие от тэга `<P>`, при использовании тэга `<BR>` браузер не генерирует на экране пустую строку.



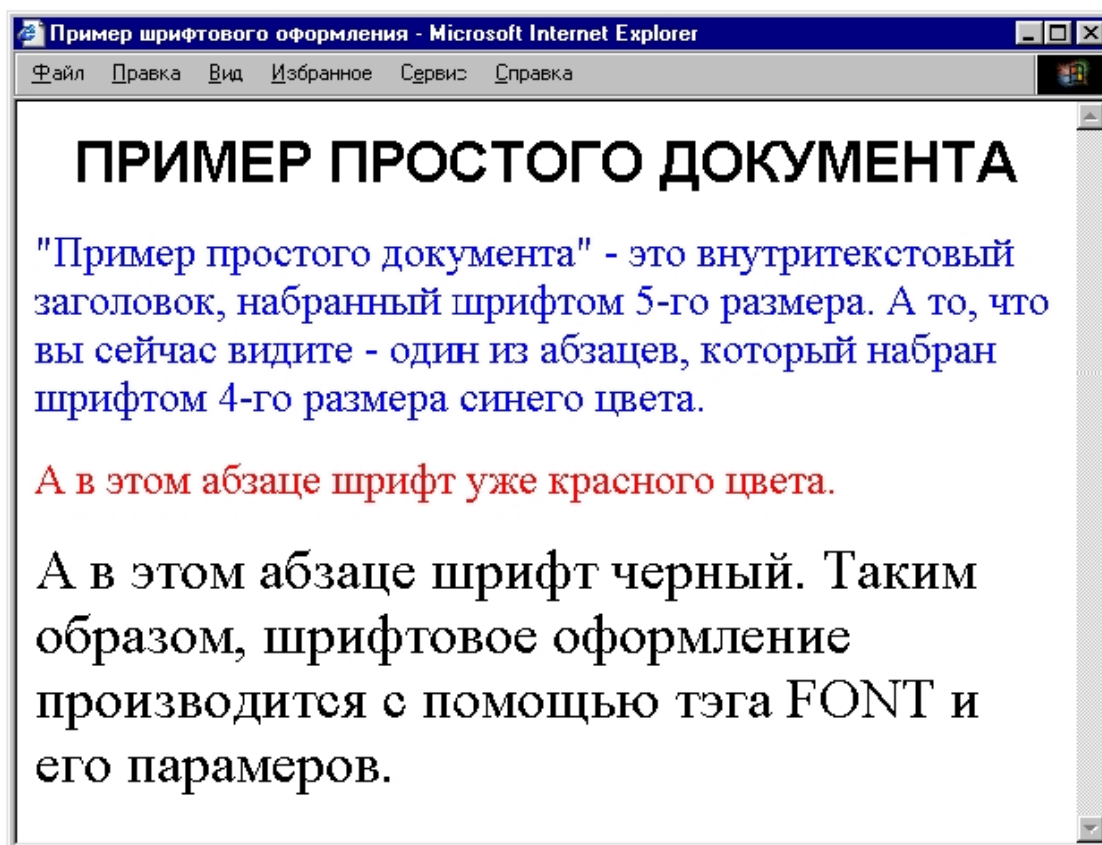


Рис. 2.3. Пример отображения шрифтового оформления документа

Наоборот, в тех случаях, когда перенос на другую строку внутри фрагмента текста недопустим, следует использовать тэг-контейнер `<NOBR>` (No Break – отсутствие разрыва) и `</NOBR>`. Находящийся внутри контейнера текст не переносится на другую строку, т. е. не разрывается. Если такая строка получается слишком длинной, то в окне просмотра браузера автоматически появляется полоса прокрутки. В качестве примера на рис. 2.3 показано отображение HTML-документа в окне браузера Internet Explorer, а выше – исходный HTML-код для этого документа.

Размер текста, помещенного внутри HTML-документа формально не ограничен, т. е. он может быть эквивалентен множеству печатных страниц. На практике это обстоятельство приводит к появлению длинных страниц, для просмотра которых требуется вертикальная полоса прокрутки. Поэтому предусмотрена возможность разбивки такого протяженного текста на отдельные главы и параграфы с помощью заголовков различного уровня. Для разметки заголовков служат 6 тэгов-контейнеров `<H1>...<H6>` (Head – заголовок), причем тэги с меньшими номерами определяют заголовки более высокого уровня. Все перечисленные тэги относятся к уровню блока, т. е. не могут использоваться для разметки отдельных слов. В качестве единственного параметра этих тэгов используется `ALIGN`, который задает способ выравнивания заголовка (точно так же, как в тэге `<P>`). Например,

```
<H1 ALIGN=CENTER> Заголовок 1-го уровня </H1>
<H3 ALIGN=LEFT> Заголовок 3-го уровня </H3>
```

В первом примере заголовок выравнивается по центру и в нем используется наибольший размер кегля шрифта, такой как в тэге `<FONT SIZE=7>`. Во втором случае текст заголовка выравнивается по левому краю, а размер шрифта соответствует значению параметра



SIZE=5. Во всех случаях заголовки выделяются осветленным пространством в виде пропущенной строки до и после строки заголовка и полужирным шрифтом самого заголовка. Наконец, тэг заголовка аккумулирует в себе функции тэга абзаца, т. е. нет необходимости в использовании одновременно с тэгом <H3> еще и тэга <P>.

Помимо заголовков, для разделения документа на отдельные разделы могут использоваться горизонтальные линии, играющие ту же роль, что и орнаментальные линейки в печатных изданиях, где такие линейки используются в конце определенных разделов, отделяя предыдущий фрагмент издания от последующего. Для генерации таких линий в электронном документе используется тэг <BP>, который не является контейнером. В тэге могут использоваться 5 различных параметров. Параметр выравнивания ALIGN может принимать значения LEFT, CENTER и RIGHT. Параметр WIDTH задает длину линии в пикселах или в процентах от ширины окна браузера, что обычно предпочтительнее, а SIZE определяет толщину линии в пикселах. Параметр COLOR определяет цветовые характеристики линии, т. е. реализует те же функции, что и аналогичный параметр в тэге <P>. Наконец, параметр NOSHADE отменяет рельефный характер линии, причем у него значение отсутствует, само наличие в тэге этого параметра характеризует отсутствие рельефа. Пример записи тэга:

```
<BP ALIGN=CENTER WIDTH=75% SIZE=8 COLOR=#800080>
```

На практике мы иногда сталкиваемся со случаями, когда в HTML-документ необходимо включить уже отформатированный текст, причем он должен отображаться браузером так, как он был подготовлен в текстовом редакторе. Для этой цели предусмотрен тэг-контейнер <PRE>. После этого тэга можно включить предварительно отформатированный текст, который должен в том же виде выводиться в окне браузера.

Для включения цитат большой протяженности (много сотен символов) предназначен тэг-контейнер <BLOCKQUOTE>. Это – тэг уровня блока. Браузеры выделяют текст, размещенный внутри такого контейнера, пустыми строками до и после и небольшим отступом слева по отношению к основной части документа. Поэтому тэг <BLOCKQUOTE> (цитата большого размера) можно использовать не только для длинных цитат, но и для выделения некоторых блоков текста.

Из других тэгов уровня блока, используемых для выделения фрагментов документа можно отметить <DIV> (Division – раздел). Используя контейнер с этим тэгом, можно управлять параметрами данного фрагмента путем задания параметра STYLE – назначения стилей, например:

```
<DIV STYLE="color: green"> (Фрагмент документа) </DIV>
```

выделит все текстовые элементы указанного фрагмента зеленым цветом. Тэг-контейнер уровня блока <CENTER> предназначен для горизонтального выравнивания всех заключенных в него элементов посередине окна просмотра браузера. В частности он используется для выравнивания по центру таблиц (см. разд. 2.6). Косвенное отношение ко всем рассмотренным в данном разделе тэгам имеет тэг-контейнер <ADDRESS> (адрес, обращение), который применяется для указания автора, владельца авторских прав, адреса автора, а иногда и даты создания и последнего обновления электронного издания. Текст внутри такого контейнера обычно отображается браузерами курсивом. Соответствующий пример приведен в конце главы, для сравнения исходного HTML-кода и отображения документа в окне браузера.

Специфическим видом текста, который иногда включают в электронное издание и ценность которого сомнительна – является бегущая строка. Зато на Web-страницах такие фраг-

менты нередко используются, преимущественно в рекламных целях. Для создания бегущей строки используется тэг-контейнер <MARQUEE> (буквальный перевод этого слова – большой навес, шатер). В тэге используется ряд параметров: DIRECTION (направление), BEHAVIOR (режим работы), WIDTH (ширина), HEIGHT (высота), LOOP (петля, повтор, цикл), SCROLLDELAY (задержка прокрутки), BGCOLOR. Первый параметр определяет направление движения строки и принимает значения LEFT или RIGHT. Параметры WIDTH и HEIGHT позволяют задать размеры окна, в котором движется строка текста. Размеры задаются в пикселах или в % от ширины и высоты страницы. Параметр LOOP задает число повторов цикла перемещения строки; если он вообще не упомянут, то число повторов бесконечно. Параметр SCROLLDELAY определяет период повторения изображения движущегося текста в миллисекундах. Наконец, значение BGCOLOR позволяет стандартным способом определить цвет фона окна, в котором движется строка текста. Если требуется изменить гарнитуру, кегль, цвет и другие параметры шрифта бегущей строки, то следует тэг-контейнер <MARQUEE> поместить внутрь тэга-контейнера <FONT>, в котором и задать все характеристики шрифта.

## 2.3. Оформление списков терминов и определений

Списки принадлежат к наиболее употребительным формам, как в электронных, так и в печатных изданиях. Возможность представления списковых структур предусмотрена практически во всех текстовых редакторах, в частности, развитая система для списковых структур имеется в MS Word, весьма популярном редакторе в нашей стране. В HTML-документах наряду с обычными маркированным и нумерованным списками предусмотрен отдельный список определений. Оформление этих списков посредством специальных тэгов составит предмет этого раздела.

### 2.3.1. Маркированные списки

Маркированные списки создаются с помощью тэга-контейнера `<UL>` (Unordered List – нумерованный список). В тэге могут использоваться два параметра: `COMPACT` (сжатый, компактный), который изначально предназначался для вывода элементов списка в компактной форме (уменьшенным кеглем и расстоянием между строками) и `TYPE` (тип), который используется для принудительного задания вида маркера. Параметр `COMPACT` не имеет значений. Современными браузерами он игнорируется. Параметр `TYPE` может принимать одно из трех значений: `disc` (круг), `circle` (окружность) и `square` (квадрат). В первом случае маркеры (bullets) имеют вид закрашенного круга, во втором – окружности малого диаметра, в третьем – закрашенного квадрата. По умолчанию значение `TYPE=disc`.

Каждому элементу списка предшествует тэг `<LI>` (List Item – элемент списка), который необязательно должен быть контейнером. В качестве параметра этого тэга также используется `TYPE`, который может принимать те же самые 3 значения. Таким образом, в списке принципиально могут использоваться и различные маркеры для представления отдельных элементов.

### 2.3.2. Создание нумерованных списков

Нумерованный список организуется с помощью тэга-контейнера `<OL>` (Ordered List – нумерованный список), внутри которого размещаются все элементы списка. Открывающий и закрывающий тэги обеспечивают пропуски строк до и после списка, выделяя его таким образом в документе. В составе тэга могут использоваться 3 параметра: `COMPACT`, `TYPE` и `START` (начало). Смысл первого параметра тот же самый, что и в тэге `<UL>`. Параметр `TYPE` по-прежнему определяет тип маркера, причем `TYPE=1` определяет маркеры в виде арабских цифр, `TYPE=A` задает маркеры в виде прописных букв латинского алфавита, `TYPE=a` – в виде строчных латинских букв, наконец, `TYPE=I` и `TYPE=i` определяют маркеры в виде римских цифр, больших и малых соответственно. Параметр `TYPE`, заданный в форме конкретного числа определяет начало отсчета для первого элемента списка. Начальный номер элемента может быть задан и путем задания значения параметра `START`.

Каждый элемент нумерованного списка предваряется тэгом `<LI>`, причем в этом случае в составе тэга может содержаться параметр `VALUE` (значение), который позволяет изменять номер очередного элемента списка, в результате чего соответственно изменяются и все последующие номера. В качестве примера можно привести фрагмент спортивных новостей, в которых содержатся результаты какого-либо соревнования:

```
<HTML> <HEAD> <TITLE> Пример нумерованного списка </TITLE> </HEAD>
<BODY> <OL> <B> Результаты чемпионата мира по гандболу 2001 </B>
```

```

<LI> Франция
<LI> Румыния
<LI> Германия <BR>
...
<LI VALUE=6> Россия
<LI> Украина
</OL> </BODY> </HTML>

```

При выводе в окне браузера документа, HTML-код которого представлен выше, мы увидим изображение, показанное на рис. 2.4. Название документа, размещенное в тэге-контейнере <TITLE> в заголовочной части документа, выведено в строке заголовка браузера. Название списка в теле документа показано полужирным шрифтом перед собственно списком. Нумерация списка по умолчанию производится арабскими цифрами, так как параметр TYPE в тэге <OL> отсутствует. Нумерация списка начинается по умолчанию с первого номера, так как и параметр START в тэге <OL> отсутствует. С помощью параметра VALUE=6 в одном из тэгов <LI> изменен номер очередного элемента списка, а дальнейшая нумерация продолжается с нового номера. Для создания строки, в которой содержится любое заданное количество точек использован тэг принудительного перевода строки <BR>.

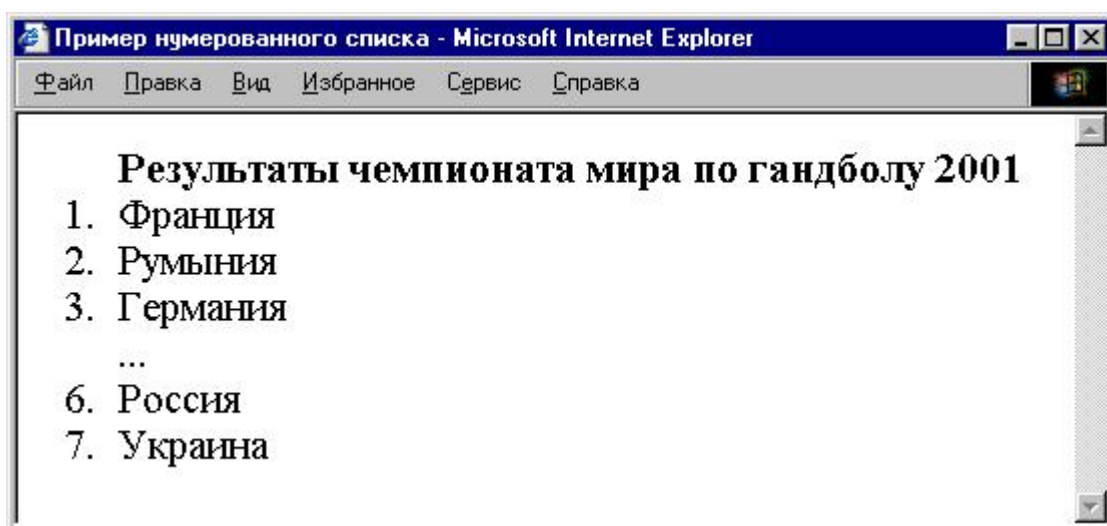


Рис. 2.4. Пример отображения в браузере MS Internet Explorer нумерованного списка

### 2.3.3. Списки определений

В научно-технических и учебных изданиях нередко используют списки или словари определений, которые в зарубежных книгах называют глоссариями. В качестве примеров можно сослаться на [32] и [24]. Каждый элемент такого списка начинается с определяемого термина, после которого следует его определение. Для создания списка определений служит тэг-контейнер <DL> (Definition List список определений), внутри которого тэг <DT> (Definition Term определяемое слово, термин) задает определяемый термин в форме единственной строки, а тэг <DD> (Definition Description описание определяемого термина) предшествует абзацу с определением этого термина. В качестве примера приведем фрагмент списка определений для компьютерной графики.

```
<html><head> <title>Список терминов и определений</title> </head>
```

```

<body bgcolor=#d5d5d5>
<DL> <DT>Пиксел
<DD>Наименьший элемент изображения, который характеризуется постоянством
параметров по всей своей поверхности
<DT>Разрешение монитора
<DD>Разрешение монитора определяется количеством пикселей на единицу длины и
составляет около 96 пикселей на дюйм
<DT>Линиатура растра
<DD> Частота размещения растровых точек в плоскости изображения, измеряется чис-
лом линий на дюйм
</DL> </body> </html>

```

Отображение этого файла в окне просмотра браузера показано на рис. 2.5. Как следует из записи исходного HTML-документа никаких попыток выравнивания текста и формирования отступов не производилось. Сами тэги <DT> и <DD> так форматируют текст, что обеспечивают оптимальное его восприятие пользователем. Фон документа задается определенным уровнем серого, в соответствии со значением параметра BGCOLOR в тэге <BODY>. Кстати, язык HTML, как следует из представленного выше фрагмента HTMLкода, нечувствителен к регистру.

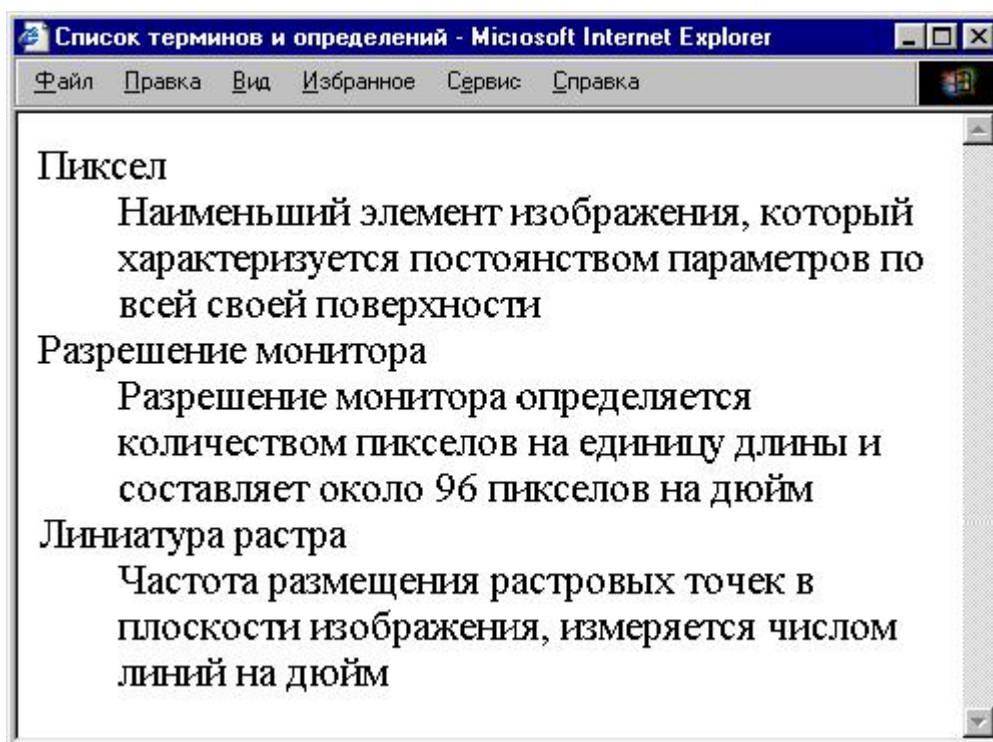


Рис. 2.5. Пример отображения в браузере документа – списка определений

### 2.3.4. Многоуровневые списковые структуры

Многоуровневые списки в HTML-документах могут создаваться путем вложения одного типа списка внутрь другого. Например, фрагменты нумерованного списка могут быть вложены внутрь маркированного или наоборот. Сказанное иллюстрируется простым примером, в котором перечисляются спутники некоторых планет нашей солнечной системы.

```
<HTML>
<HEAD>
<TITLE>Пример двухуровневого списка</TITLE> </HEAD>
<BODY> <UL> <B> Спутники планет </B>
<BR>....
<LI> ЗЕМЛЯ
<OL> <LI> Луна </OL>
<LI>МАРС
<OL> <LI> Фобос
<LI>Деймос </OL>
<BR>....
</UL> </BODY> </HTML>
```

Отображение двухуровневого списка в окне браузера показано на рис. 2.6. Внутри тэга-контейнера маркированного списка <UL> размещены 2 контейнера нумерованного списка <OL> и </OL> со своими тэгами <LI> каждый. Дополнительный отступ слева обеспечивает хорошую читаемость текста. По этой схеме можно создавать и многоуровневые списки.

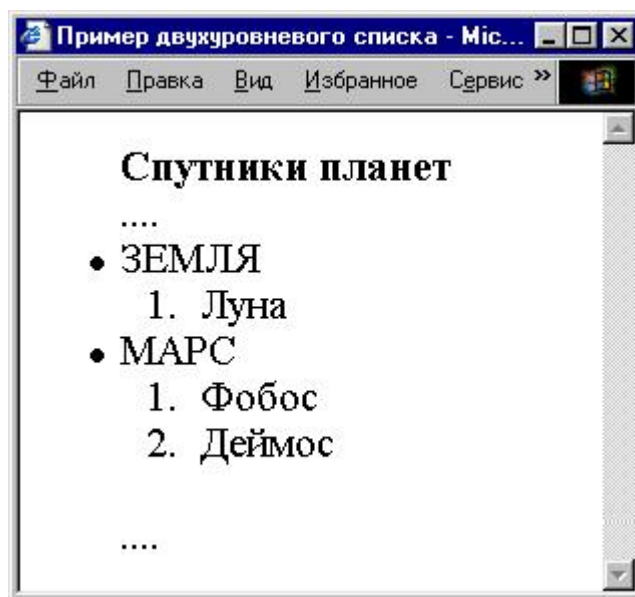


Рис. 2.6. Пример отображения в браузере двухуровневого списка

## 2.4. Использование графики при оформлении электронного издания

Известно, что с помощью графики можно точнее передать идеи и назначение любого издания, в том числе и электронного. Без иллюстрации полосы набора выглядят однообразными, а хорошо подобранная и умело размещенная графика делает издание гораздо более привлекательным. Ограничения в использовании графики с позиций эстетики связано с опасностью загромождения документа изображениями, в том числе и фоновыми, и навязчивой анимацией, зачастую не связанной непосредственно с тематикой издания.

С технической точки зрения изображения требуют значительно большего объема информации, чем простой текст, поэтому может возникнуть информационная перегрузка, особенно при использовании сетевых технологий с каналами недостаточной пропускной способности. Известно, что многие пользователи сети Интернет отключают загрузку графических файлов, чтобы быстрее загрузить документ и ознакомиться с ним. Несколько улучшило положение применение компактных форматов графических файлов, таких как JPEG (с частичной потерей информации), GIF и PNG. В качестве иллюстрации сказанного можно привести данные по размеру файлов различного формата, в каждом из которых хранятся одни и те же графические данные, представленные на рис.2.7, а именно TIFF – 46 Кбайт, JPEG – 31 Кбайт, PNG – 8 Кбайт и GIF – около 5 Кбайт.

Мы уже сталкивались с использованием изображения в качестве фонового при знакомстве с тэгом <BODY>. Для этого в параметре BACKGROUND указывался URL файла, хранящего изображения. Если в электронных изданиях, распространяемых на носителях, можно использовать практически любые графические файлы, то для сетевых изданий преимущественно используются GIF-файлы, реже файлы в формате JPEG.

Основным средством встраивания изображений в текст служит тэг <IMG> (Image изображение), обязательным параметром которого является SRC (Source источник), значение которого – адрес (URL) файла изображения. Параметр выравнивания изображения на полосе ALIGN может принимать 8 различных значений, определяющих различные способы его размещения на странице, два параметра позволяют установить размер изображения, еще 2 —расстояние от границы изображения до текста, отдельный параметр позволяет создать вокруг изображения рамку, а еще один – задать альтернативный текст, который будет выведен вместо изображения при отказе пользователя от его загрузки. Рассмотрим последовательно эти параметры.

Выравнивание изображения осуществляется двумя принципиально различными способами:

# по левому или правому краю наборной полосы (аналог оборочной иллюстрации в печатном тексте);

# изображение встраивается как элемент одной из строк текста (печатного аналога такого способа выравнивания нет, более того, в печатном издании такое расположение рисунка совершенно недопустимо).

Для выравнивания изображения по левому или правому краю окна браузера параметру ALIGN присваиваются значения LEFT или RIGHT соответственно, а текст обтекает рисунок с противоположной стороны. Для задания размеров изображения используются параметры WIDTH (ширина) и HEIGHT (высота), причем значения каждого параметра могут задаваться как в абсолютных единицах – пикселах, так и в относительных – процентах от соответствующего размера окна браузера. Реальные размеры изображения могут не совпадать со значениями параметров WIDTH и HEIGHT. В этом случае браузер автоматически масштабирует изображение по отдельным координатам.

Графические изображения могут задаваться в двух основных форматах: векторном и растровом. В соответствии с первым форматом создаются файлы сравнительно малого объема, легко масштабируемые – однако в них не может быть достигнуто высокое художественное качество. Кроме того, преобразование информации описания изображения, содержащейся в файле, в реальный рисунок требует значительных вычислительных ресурсов.

Растровый формат масштабируется с искажениями, файлы в нем имеют большой объем, но он обеспечивает достаточно высокое качество изображения. Кроме того, сама подготовка компьютерных графических файлов связана с оцифровкой (преобразованием в цифровой формат) реальных изображений, например, рисунков и фотографий. Для этой цели используются так называемые сканеры, которые преобразуют изображение в один из растровых форматов, чаще TIFF. С растровыми же форматами работают цифровые фото и видеокамеры, которые все шире используются в современных издательских технологиях для прямого преобразования изображения в его цифровой эквивалент.

Поэтому масштабирование оправдано только для векторной графики, для растровой же следует определить реальные размеры изображения по горизонтали и вертикали в пикселах и задать эти размеры в параметрах WIDTH и HEIGHT. При этом:

# изображение будет выведено в реальных размерах и без искажений;

# при сетевом распространении документа пользователь сразу будет знать размеры изображения, а в остальную часть документа будет без промедления загружаться текст.

Остановимся кратко на других способах выравнивания изображения в электронном издании. Соответствующие значения параметра ALIGN следующие: TOP (самое верхнее), TEXTTOP (самое верхнее для текста), BOTTOM (нижнее), BASELINE (базовая линия), ABSBOTTOM (Absolutely Bottom – самое нижнее), MIDDLE (среднее), ABSMIDDLE (Absolutely Middle – абсолютная средняя).

Значение TOP выравнивает верхнюю границу изображения по самому высокому элементу текущей строки, а TEXTTOP – по самому высокому текстовому элементу. Смысл различия в том, что элементом строки может быть и другое изображение. Значения BASELINE и BOTTOM действуют одинаково и выравнивают нижнюю границу изображения по базовой линии текущей строки, т. е. линии, на которой "стоят" шрифтовые элементы (нижние выносные элементы символов находятся под базовой линией), а ABSBOTTOM – выравнивает нижнюю границу изображения по нижней границе текущей строки, включая все элементы строки, в том числе и графические, если они имеются. Наконец, значения MIDDLE и ABSMIDDLE выравнивают середину изображения по базовой линии или посередине текущей строки соответственно. Еще раз отметим, что применение параметров выравнивания этой группы в большинстве случаев снижают качество оформления страницы и противоречат основным принципам дизайна полос.

Чтобы выделить изображение в текстовом документе, полезно предусмотреть незаполненное текстом пространство вокруг него. Это можно осуществить с помощью параметров HSPACE (Horizontal Space – расстояние по горизонтали) и VSPACE (Vertical Space – расстояние по вертикали), которые определяют отступы (осветленное пространство) по горизонтали и вертикали в пикселах. Для выделения изображения на странице можно заключить его в рамку, толщина которой в пикселах задается с помощью параметра BORDER (рамка). Наконец, для пользователей, работающих в режиме отключения загрузки графики, возможность вывода альтернативного текста дает параметр ALT (Alternate альтернативный). В случае же загрузки изображения при выводе на него курсора мыши появляется текст подсказки во всплывающем окне. Ниже приведен пример тэга <IMG> с различными параметрами:

```
<HTML><HEAD><TITLE>Пример вставки изображения в документ</TITLE> </HEAD>
```



```
<BODY><IMG SRC=globus.jpg ALIGN=LEFT WIDTH=40% HEIGHT=40%
HSPACE=8 VSPACE=10 BORDER=5
ALT="Рельеф земной поверхности">
```

<P ALIGN="Justify">В данном документе рисунок выравнивается влево, т. е. он должен размещаться в левом верхнем углу. Размеры рисунка 60 на 60 пикселей. Рисунок окружен рамкой шириной в 5 пикселей и отделен от текста полями по 10 пикселей с каждой стороны. Текст размещается справа от рисунка и, в дальнейшем, ниже его.</P>

```
</BODY> </HTML>
```

Приведенный выше HTML-документ отображался в браузере MS Internet Explorer и полученное в окне браузера изображение показано на рис. 2.7. В частности, на рисунке видна надпись "Рельеф земной поверхности", заданная в виде значения параметра ALT и возникающая на экране при подведении указателя мыши к изображению.

В заключение отметим, что фирма Netscape в своем браузере предусмотрела возможность использования еще одного параметра в тэге <IMG>, а именно LOWSRC. В качестве значения этого параметра указывается графический файл с альтернативным изображением того же самого объекта, но с гораздо более низким разрешением и глубиной цвета. В результате альтернативное изображение имеет на порядок меньший объем и быстро загружается на Web-страницу – а только позднее загружается основное изображение. Пример записи тэга с этим параметром представлен ниже:

```
<IMG SRC=globus.gif LOWSRC=low-globus.gif>
```



Рис. 2.7. Пример отображения в браузере MS Internet Explorer документа, содержащего изображение

Следует заметить, что размеры этих двух изображений должны совпадать. Если параметр LOWSRC задает изображение меньших размеров, то основное изображение будет

также трансформироваться к этим меньшим размерам. Чтобы избежать возникающих вследствие этого эффекта искажений, следует в тэге <IMG> указать действительные размеры основного изображения в пикселах с тем, чтобы при первом проходе и загрузке вспомогательного изображения размеры последнего приводились к размерам основного, а не наоборот.

## 2.5. Включение в издание ссылок и комментариев

Ссылки являются важнейшей составляющей HTML-документа, так как именно они и делают этот документ гипертекстовым, реализуя важнейшее преимущество электронного документа по сравнению с печатным, состоящее в широчайших возможностях перемещения по документу и вне его. Ссылка состоит из указателя и адресной части. Указатель представляет собой малый фрагмент документа (одно или два слова или изображение), который каким-либо образом выделен (например, цветом или подчеркиванием). Если подвести указатель мыши к ссылке и щелкнуть кнопкой мыши, то браузер загрузит новый документ или перейдет к той части данного документа, адрес которой указан во второй части ссылки.

Текстовые указатели чаще всего выделяются в окне браузера подчеркиванием. Если же в качестве ссылки использовано изображение, то такое изображение браузер самостоятельно заключает в рамку. Второй частью ссылки является адрес той страницы, которая должна быть загружена. Адрес, как уже указывалось ранее, может быть абсолютным или относительным. В последнем случае адрес формируется относительно того каталога, в котором расположен исходный документ. Пример предыдущего раздела содержит относительный адрес рисунка `globus.jpg`, который расположен в том же каталоге, что и сам HTML-документ.

Для организации ссылок используется тэг-контейнер `<A>` (Access Address указатель, ссылка), который имеет параметр `HREF`, причем значение этого параметра и является адресом загружаемой страницы, о чем говорилось выше. Конструкция данного контейнера выглядит следующим образом:

```
<A HREF=URL-адрес нужной страницы> Текст-указатель ссылки </A>
```

Когда курсор мыши подводится к текстовому указателю, во всплывающем окне появляется адрес ссылки, который записан в качестве значения параметра `HREF`. Щелкнув левой кнопкой мыши по этому указателю, мы инициируем процесс загрузки документа или страницы с этим адресом.

Запись ссылки, указателем в которой служит не текст, а изображение, выглядит следующим образом:

```
<A HREF=URL-адрес нужной страницы> <IMG SRC=(имя графического файла)>
</A>
```

Действительно, в представленном примере тэг `<IMG>` с параметром `SRC`, задающим имя графического файла, играет ту же самую роль, что и текст-указатель ссылки в предыдущем примере. Некоторым преимуществом последнего способа является то, что изображение занимает больше места в окне браузера, чем текст-указатель ссылки, и на него может быть быстрее выведен курсор мыши.

Ранее уже отмечалось, что относительная ссылка указывает не полный путь к файлу, а его местоположение относительно того документа, из которого производится ссылка. Для облегчения работы с относительными ссылками в заголовочной части документа может использоваться тэг `<BASE>`, в котором с помощью параметра `HREF` указывается абсолютный URL-адрес этого документа, относительно которого и строится вся адресация ссылок (см. также разд. 2.1).

Если ссылки на внешние документы позволяют свободно ориентироваться пользователю в безбрежном море сети Интернет, то для качественной навигации в каждом электронном издании необходимы ссылки внутренние (в пределах самого издания). Скажем, встре-

тив в издании незнакомый термин, вы захотите узнать его смысл и назначение, для чего необходимо обратиться к списку терминов и определений, т. е. глоссарию (см. разд. 2.3), а затем вернуться к исходному тексту. Любой большой документ требует также наличия ссылок из его оглавления к соответствующим разделам. Могут потребоваться и другие ссылки, например, в текущем и предыдущем абзацах, по существу, такие ссылки указаны в виде адресации пользователя к определенным разделам данной книги. Однако, в печатном издании пользователь должен заложить текущую страницу, найти раздел, на который указывает ссылка, прочесть его и вернуться к заложенному месту. В электронном же издании эти операции автоматизированы и их выполняет компьютерная программа браузер. И это их существенное преимущество.

Для организации переходов необходимо нечто, напоминающее закладки в печатном издании. Такие закладки, на которые может быть организована ссылка, реализуются с помощью параметра NAME (имя) уже упоминавшегося тэга <A>. Параметр NAME позволяет определить имя закладки с тем, чтобы в дальнейшем адресоваться именно к ней. Например, из данного раздела книги мы адресовались к разд. 2.1 и 2.3. Чтобы такие переходы можно было осуществить в электронном издании, к названиям таких разделов надо добавить закладки:

```
<A NAME=section_2-1>
```

```
<A NAME=section_2-3>
```

а в оглавлении записать названия этих параграфов внутри тэга-контейнера <A>:

```
<A HREF="#section_2-1"> 2.1. Общая структура гипертекстового документа, заголовок и тело документа, тэги оформления заголовка </A>
```

```
<A HREF="#section_2-3"> 2.3. Оформление списков терминов и определений </A>.
```

Когда такие закладки имеются, переходы к ним можно было бы оформить не только из оглавления, но и из любого другого места электронного издания. Вместо указания в скобках (см. разд. 2.1) можно было бы вставить конструкцию с тэгом-контейнером, а именно:

```
<A HREF="#section_2-1"> (см. разд. 2.1) </A>
```

Можно настроить браузер таким образом, что соответствующая любой ссылке страница будет демонстрироваться в отдельном окне. Тогда, просмотрев текст или рисунок на этой странице, мы можем ее закрыть или свернуть, возвращаясь, таким образом, к основному документу. Отрегулировав размер этого отдельного окна, мы можем одновременно рассматривать как ссылку, так и основной документ. Эти возможности будут подробнее рассмотрены в разд. 2.7 при изучении фреймовой структуры гипертекстового документа.

Ссылки могут указывать и на специальные виды документов, например, на аудио– или видеофайлы или анимационные файлы в GIF-формате. Браузер может уметь или же не уметь работать с указанным специальным типом документа. Скажем, все современные браузеры умеют работать с анимацией в формате GIF89a, часто встречающейся на страницах в сети Интернет. Ссылка на такой файл приводит к появлению на странице "движущегося" изображения. Если же браузер не умеет работать со специальным документом (не распознает тип принятого документа), он обратится к другим программам на компьютере пользователя. Если нужная программа будет обнаружена, браузер передаст полученный документ этой программе для дальнейшей его обработки. Скажем, ссылка на видео файл формата AVI (см.

разд. 3.5, 4.5, 5.5), может привести к запуску программы для демонстрации видеофайлов. В результате видеофайл будет показан в специальном дополнительном окне.

В заключение несколько слов о включении комментариев в HTML-документ. Комментарий может состоять из любого числа строк и слов, расположенных внутри контейнера из двух тэгов, а именно: открывающего текст комментария `<!--` и завершающего этот текст `-->`. Текст комментария не отображается в окне браузера, поэтому его можно рассматривать как личные заметки авторов или редакторов в процессе работы над изданием. В то же время, если не убрать комментарий из окончательного текста издания, то он принципиально становится доступным и для читателей. Другой тэг-контейнер для записи комментариев – `<COMMENT>` – поддерживается только браузером MS Internet Explorer и не нашел широкого применения.

## 2.6. Оформление таблиц

Таблицы достаточно широко применяются в электронной документации, причем для Web-страниц они используются не только в традиционном смысле, как метод упорядоченного представления данных, но и для форматирования самих этих страниц. Описание таблицы на языке HTML размещается внутри тела документа, т. е. в контейнере `<BODY>` и `</BODY>`. Внутри документа допускается любое число таблиц, причем некоторые из них могут быть вложенными. Каждая таблица создается в пределах контейнера `<TABLE>` (таблица) и `</TABLE>`, где размещается описание структуры самой таблицы и ее содержимое. Каждая строка таблицы размещается после тэга `<TR>` (Table Row – строка таблицы).

Каждая ячейка таблицы в пределах строки оформляется тэгом `<TH>` (Table Header – заголовок таблицы) – для заголовочной части таблицы или `<TD>` (Table Data – табличные данные) – для ячеек, в которых размещаются данные. В заголовочной части по умолчанию применяется полужирный шрифт и выравнивание по центру. Для отображения данных по умолчанию используется нормальное (светлое) начертание и выравнивание влево. Можно отметить, что для всех тэгов, перечисленных в этом абзаце, закрывающий тэг не обязателен, т. е. он может быть опущен.

Количество строк в таблице определяется количеством строчных тэгов `<TR>`, а число столбцов – максимальным количеством тэгов `<TH>` или `<TD>` в одной из строк. Строкой считается все то, что следует после очередного тэга `<TR>` и до следующего такого же тэга. Для ячейки таблицы, не содержащей данных, следует использовать пустой контейнер `<TD>` и `</TD>`. Если пустые ячейки расположены в конце строки, то их описание может быть опущено. В этом случае браузер самостоятельно оставит необходимое число ячеек пустыми.

Таблица может иметь название, т. е. то, что в редакционной практике называется тематическим заголовком (в отличие от нумерационного заголовка, в котором содержится слово "таблица" и порядковый номер этой таблицы в тексте издания), причем, если в печатном издании заголовок обязательно расположен над таблицей, то в электронном он может быть расположен как сверху, так и снизу. Заголовок расположен внутри контейнера `<CAPTION>` (заголовок) и `</CAPTION>`. Указанный контейнер должен быть помещен внутрь тэга-контейнера `<TABLE>`, но вне области описания тэгов `<TR>`, `<TH>` или `<TD>`. Последняя спецификация HTML рекомендует размещать тэгконтейнер `<CAPTION>` сразу после открытия таблицы, т. е. после тэга `<TABLE>` и до первого тэга `<TR>`.

### Замечание

Заголовок нумерационный – это заголовок, обозначаемый числом, определяющим порядковый номер рубрики, таблицы и пр.

В тэге `<CAPTION>` первоначально был предусмотрен один необязательный параметр `ALIGN`, который предназначался для вертикального выравнивания (размещения заголовка) и мог принимать одно из двух значений `TOP` (по умолчанию) или `BOTTOM`. Затем выяснилась необходимость и горизонтального выравнивания с тремя стандартными параметрами `LEFT`, `RIGHT` и `CENTER`. Однако нельзя в одном тэге дважды использовать один и тот же параметр. Поэтому в современных версиях языка HTML параметр `ALIGN` (по умолчанию `ALIGN=LEFT`) оставлен для выравнивания по горизонтали, а вертикальное выравнивание (точнее – размещение заголовка над или под таблицей) осуществляется с помощью параметра `VALIGN` (Vertical Allign – вертикальное выравнивание).

В тэге `<TABLE>` могут использоваться следующие параметры: `BORDER`, `CELLSPACING` (расстояние между ячейками таблицы), `CELLPADDING` (заполнение ячеек), `WIDTH`, `ALIGN`, `HEIGHT` и `BACKGROUND`. Параметр `BORDER` управляет отображением рамки вокруг каждой ячейки таблицы (т. е. задает вертикальные и горизонтальные линии сетки) и вокруг всей таблицы, причем его значение задает толщину рамки в пикселах вокруг всей таблицы, а само наличие этого параметра задает линии сетки. Значение параметра `BORDER` появилось лишь в версии 3.2 спецификации HTML, до этого толщина рамки вокруг таблицы не регулировалась.

Параметр `CELLSPACING` задает расстояние между смежными по горизонтали и вертикали ячейками, причем это расстояние задается в пикселах, т. е. внутри каждой ячейки создается нечто вроде рамки и лишь при `CELLSPACING=0` эти рамки отдельных ячеек сливаются в единую сетку. Параметр `CELLPADDING` определяет расстояние между рамкой вокруг ячейки и данными внутри нее, т. е. величину отступа символов от рамки. При значении `CELLPADDING=0` текст может касаться рамки, что в плане дизайна едва ли можно приветствовать. По умолчанию значение `CELLSPACING=2`, а `CELLPADDING=1` – в этом случае расстояние между данными в соседних ячейках одной строки будет равно 6 пикселям.

Параметры `WIDTH` и `HEIGHT` позволяют задать ширину и высоту таблицы как в абсолютных единицах – пикселах, так и в относительных – процентах относительно размера окна браузера. В большинстве случаев эти размеры не требуются браузеру, так как он автоматически вычисляет размеры таблицы, учитывая множество факторов, включая параметры документа в целом, количество ячеек в таблице и их заполнение. Браузер стремится установить ширину таблицы такой, чтобы она помещалась в окне просмотра браузера и не было необходимости в прокрутке таблицы по горизонтали, т. е. чтобы горизонтальный маркер прокрутки отсутствовал.

Надо отметить, что и при задании значений параметров `WIDTH` и `HEIGHT` нет гарантии в том, что они будут выдержаны браузером. Если ширина таблицы больше ширины окна просмотра, браузер сделает попытку уменьшить ширину до требуемой, пропорционально уменьшая размеры колонок, и только если это не удастся, установит заданную ширину таблицы, снабдив окно маркером прокрутки.

Параметр `ALIGN` задает горизонтальное выравнивание таблицы в окне просмотра браузера. Возможны 2 значения этого параметра: `LEFT` и `RIGHT`, каждый из которых обеспечивает обтекание таблицы текстом документа с противоположной стороны. Это соответствует клочковой таблице в печатном издании, т. е. таблице, заверстанной в оборку. По умолчанию параметр `ALIGN` принимает значение `LEFT`. Если параметр `ALIGN` опущен, то текста рядом с таблицей не будет вообще, т. е. таблица будет форматной или полосной (когда она занимает всю страницу по вертикали) – если использовать термины, принятые в печатных изданиях. Значение параметра `ALIGN=RIGHT` не предусмотрено. Однако, если мы хотим ориентировать таблицу по центру, можно пойти другим путем: заключить контейнер `<TABLE>` и `</TABLE>`, который описывает всю таблицу целиком, в контейнер `<CENTER>` и `</CENTER>`. Напомним, что последний имеет уровень блока, т. е. форматирует любое количество данных, размещенных внутри его (см. разд. 2.2).

Отметим, что параметр `ALIGN` может использоваться и для форматирования данных внутри каждой ячейки таблицы, например:

```
<TABLE ALIGN=LEFT WIDTH=40% BORDER=5 >
<TR> <TH ALIGN=RIGHT> Фамилия </TH> <TH ALIGN=RIGHT> Оценка </TH>
<TR> <TD ALIGN=RIGHT> Иванов А.Н. </TD> <TD ALIGN=RIGHT> 4 </TD>
<TR> <TD ALIGN=RIGHT> Сергеев И.Д. </TD> <TD ALIGN=RIGHT> 5 </TD>
```

```
<TR> <TD ALIGN=RIGHT > Лавров В.В. </TD> <TD ALIGN=RIGHT > 3 </TD> </TABLE>
```

Пример отображения браузером документа, включающего в себя приведенный выше фрагмент, показан на рис. 2.8. На рисунке хорошо видна рамка вокруг таблицы и то, что таблица выровнена влево, а справа обтекает текстом. В качестве текста использован фрагмент абзаца данной книги, расположенного чуть выше по тексту. Следует обратить внимание на то, что хотя вся таблица выровнена влево, данные в каждой строке выравниваются вправо. Формально можно было бы в каждой ячейке выравнивать текст посвоему. Однако, это визуально выглядело бы очень неприглядно.

Для качественного дизайна полосы данные во всех ячейках столбца таблицы должны выравниваться одинаково, т. е. в примере на рис. 2.8 возможно было бы оставить первый столбец неизменным, а данные во втором выравнивать влево (или даже по центру).

Параметр BACKGROUND, который уже был показан в примере, приведенном выше, также может использоваться как в тэге <TABLE>, так и в тэгах <TH> и <TD>, определяющих характеристики отдельных ячеек таблицы. Во всех случаях он определяет фоновый рисунок с помощью параметра HREF. В последней версии спецификации HTML в тэге <TABLE> появился параметр COLS (Column Splitting дробление колонок, разбиение столбцов), задающий общее число колонок в таблице. Эта дополнительная информация ускоряет процесс построения таблицы браузером.

При создании заголовочной части таблицы, реке боковика (левой графы таблицы, содержащей данные о строках таблицы) и прографки (составной части таблицы, содержащей сведения, относящиеся к заголовку и боковику), возникает потребность в объединении нескольких ячеек по горизонтали (в строке) или по вертикали в единое целое. Для этого используются параметры ROWSPAN (сцепление строк) и COLSPAN (сцепление столбцов), первый из которых объединяет несколько строк, а второй – столбцов, в одну ячейку. Пример описания такой таблицы приводится ниже.

```
<TABLE BORDER=10 SELLSPACING=3 CELLPADDING=4 WIDTH=80%>
<CAPTION> <H4> Объединение ячеек в заголовочной части</H4>
<TR> <TH ROWSPAN=2>Заголовок 1</TH><TH COLSPAN=2>Заголовок 2</TH>
<TR><TH>Заголовок 2-1</TH> <TH>Заголовок 2-2 </TH>
<TR ALIGN=CENTER> <TD> 1 <TD> 2 <TD> 3 </TABLE>
```



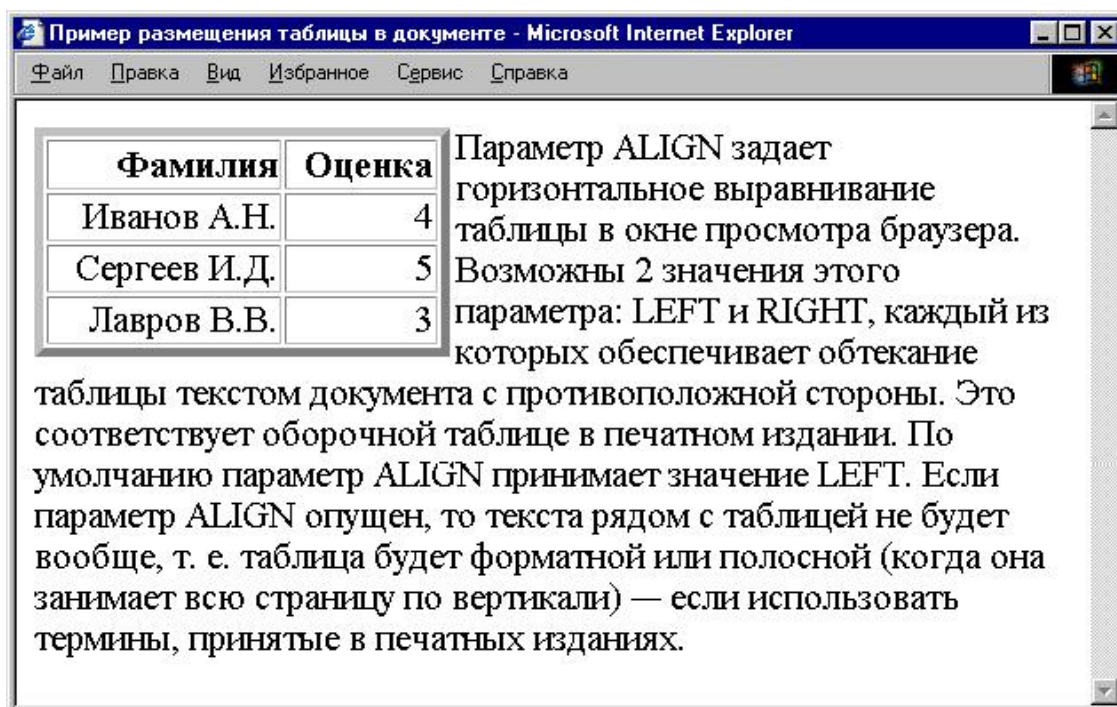


Рис. 2.8. Пример отображения таблицы в составе HTML-документа

Показанный ниже рис. 2.9 иллюстрирует отображение в браузере HTML-документа, в теле которого содержится приведенный выше контейнер `<TABLE>` и `</TABLE>` со всем его содержимым. По сравнению с таблицей, показанной на рис. 2.8, здесь с помощью тэга `<CAPTION>` добавлен тематический заголовок, выровненный по центру таблицы. В таблице хорошо видны рамки ячеек и рамка вокруг таблицы в целом. В ячейке "Заголовок 1" объединены 2 строки, а в ячейке "Заголовок 2" — два столбца, что соответствует исходному HTML-коду. В третьей строке, как это часто делают для длинных узких таблиц, указаны номера колонок, что позволяет при переносе на следующую полосу не повторять заголовка целиком. Отметим, что используя параметры `ROWSPAN` и `COLSPAN` можно составить многоуровневые заголовки таблиц любой степени сложности.



Рис. 2.9. Пример отображения сложной таблицы в составе HTML-документа



## 2.7. Фреймовая структура электронного издания

Для удобства работы с электронным изданием окно просмотра браузера может быть принудительно разбито на несколько прямоугольных областей, так называемых кадров (frames), непосредственно примыкающих друг к другу. В каждую такую область можно загружать отдельные HTML-документы, просматривая каждый из них независимо от других. Впервые возможность работы с фреймовыми структурами была реализована в браузере фирмы Netscape версии 2.0. Браузер MS Internet Explorer поддерживает фреймы начиная с версии 3.0.

Фреймовая структура целесообразна для применения в следующих случаях:

- # для представления информации в нескольких смежных областях окна просмотра браузера, каждая из которых просматривается независимо, чтобы иметь возможность сопоставить эту информацию;

- # для того, чтобы в определенной области экрана размещалась информация, которая должна постоянно находиться в поле зрения пользователя;

- # для того, чтобы иметь возможность управлять загрузкой документов в одну из областей, работая в другой области экрана.

В целом же фреймовая структура предоставляет пользователю удобные средства навигации в пределах электронного документа, практически нереализуемые при других вариантах. Видимо по этим причинам фреймовая структура достаточно часто используется для реализации Web-страниц, особенно входных или лицевых страниц сайтов, т. е. тех, на которые пользователь попадает при начальной загрузке этого сайта.

Принципиальное отличие HTML-кода, задающего фреймовую структуру, состоит в том, что вместо тэга <BODY> в нем используется тэг-контейнер <FRAMESET> (множество фреймов), параметры которого и определяют разбиение площади окна браузера на отдельные участки или области. Для этого у тэга <FRAMESET> используется два параметра: ROWS и COLS. Параметр ROWS делит окно браузера на горизонтальные полосы (строки), а COLS – на вертикальные (столбцы). Размеры этих полос могут устанавливаться в абсолютных единицах, т. е. пикселах и в относительных – процентах или пропорциональных единицах. Первые два способа задания уже рассматривались ранее и не требуют дополнительных комментариев, 3-й же способ использует некоторое количество звездочек (\*), каждая из которых представляет одну часть целого, величина этой части в данном случае определяется как высота окна браузера, деленная на суммарное количество звездочек под знаком параметра ROWS. Все 3 способа деления окна на полосы иллюстрируются приведенными ниже примерами:

```
<FRAMESET ROWS="10%,80%,10%">
```

```
<FRAMESET ROWS="60,480,60">
```

```
<FRAMESET ROWS="*,8*,*">
```

Все 3 варианта с тэгом <FRAMESET> описывают деление экрана на 3 горизонтальные полосы, из которых верхняя и нижняя – одинаковой ширины, а средняя – в 8 раз шире каждой из них. Аналогичным образом описывается и деление на вертикальные полосы с помощью параметра COLS. Для параметров ROWS и COLS возможен и смешанный вариант задания значений, когда используются два или даже все три способа одновременно. Например,

```
<FRAMESET COLS="70,4*,*,15%">
```

Тэг задает деление окна на 4 колонки, первая из которых имеет ширину в 70 пикселей, последняя – 15% от ширины окна, ширина 3-й колонки составляет 1/5, а 2-ой – 4/5 от оставшейся части ширины окна.

Внутри контейнера `<FRAMESET>` и `</FRAMESET>` могут использоваться лишь вложенные такие же контейнеры или тэги `<FRAME>` (фреймы, кадры), которые определяют каждый одиночный фрейм. Кстати в теле документа возможно использовать несколько расположенных последовательно тэгов контейнеров `<FRAMESET>`, в каждом из которых могут быть и вложенные аналогичные контейнеры. Практически же такие конструкции используются крайне редко. Тэг `<FRAME>` размещается внутри контейнера `<FRAMESET>` и `</FRAMESET>`, причем количество тэгов `<FRAME>` внутри этого контейнера должно в точности совпасть с числом отдельных фреймов, т. е. каждый фрейм должен быть описан своим тэгом `<FRAME>`. Тэг `<FRAME>` содержит 6 параметров, а именно: SRC, NAME, MARGINWIDTH (ширина поля, окаймляющего фрейм по ширине), MARGINHEIGHT (поле, обрамляющее фрейм по высоте), SCROLLING и NORESIZE (запрет изменения размера фреймов).

Наиболее важный из перечисленных параметров SRC определяет URL или адрес документа, который изначально загружается в данный фрейм. Обычно это HTML-документ, содержащий и заголовочную часть, и тело. Параметр NAME позволяет присвоить этому фрейму уникальное (не повторяющееся) имя, которое может использоваться для адресации к данному фрейму. В тэге `<A>`, который рассматривался в разд. 2.5, для адресации к фреймам используется специальный параметр TARGET (адресат, место назначения, приемник), значением которого является имя соответствующего фрейма. Проиллюстрируем сказанное примером:

```
<FRAME SRC="maintenance.htm" NAME="main_frame">
```

```
.....
```

```
<A HREF="main_doc.htm" TARGET="main_frame">Основной документ</A>
```

В представленном примере в основной фрейм (имя этого фрейма – main\_frame) исходно был загружен документ maintenance.htm, т. е. оглавление электронного издания. Если же подвести указатель мыши к словам "Основной документ", которые являются указателем ссылки, и щелкнуть левой кнопкой, то в этот фрейм будет загружен новый документ под названием main\_doc.htm, который находится в том же самом каталоге и представляет собой содержательную часть данного издания. Таким образом, присвоение фрейму определенного имени позволяет в дальнейшем загружать в него документы.

Параметры MARGINWIDTH и MARGINHEIGHT тэга `<FRAME>` позволяют установить ширину полей фрейма по ширине и высоте. Отметим, что левое и правое поле имеют одинаковую ширину. То же самое можно сказать и про верхнее и нижнее поля. Значение этих параметров задается в абсолютных единицах – пикселах. Под полями здесь подразумевается свободное пространство, где не могут присутствовать иллюстрации и текст (как и в печатном издании). Минимальное значение этих параметров, устанавливаемое по умолчанию, равно 1 пикселу.

Если содержимое фрейма не помещается в отведенной части окна браузера, для него будут автоматически создаваться полосы прокрутки по нужной координате. Для управления этим процессом служит параметр SCROLLING, который может принимать одно из трех допустимых значений: Yes (да), No (нет) и Auto (автоматически). Последнее значение устанавливается по умолчанию. Значение No запрещает создание полос прокрутки, а Yes приводит к обязательному их созданию, независимо от того, нужны они или нет.

Обычно пользователь может самостоятельно изменять размеры фреймов, так же точно, как и размеры окна браузера, где они размещаются. При таких операциях может нарушаться оптимальное соотношение между размерами фреймов и ухудшаться восприятие информации с экрана компьютера. Для того, чтобы предотвратить эти действия пользователя служит параметр NORESIZE. Для этого параметра не предусмотрено каких-либо значений, но само его использование в описании одного из фреймов запрещает не только изменение размера этого фрейма, но и любого смежного с ним. Приведем несколько примеров с использованием фреймовых структур. Первый пример задает структуру из 5-ти фреймов, которая образуется в результате деления окна браузера на 3 горизонтальные полосы, средняя из которых делится еще и на 2 колонки. Ширина нижней полосы, как следует из 3-й строки HTML-кода, составляет 50 пикселей, а верхней – составляет восьмую часть от высоты оставшейся области окна браузера. Средняя полоса разделена на узкую полосу слева и вчетверо более широкую оставшуюся часть (см. 5-я строка HTML-кода). Наконец, эта оставшаяся часть поделена пополам по горизонтали (см. 7-ая строка HTML-кода).

```
<HTML> <HEAD> <TITLE>Цифровая обработка изображений</TITLE>
<META NAME="Author" CONTENT="В. А. Вуль"> </HEAD>
<FRAMESET ROWS="*,7*,50">
<FRAME SRC="zag.html" scrolling="no">
<FRAMESET COLS="20%,80%">
<FRAME SRC="ogl.html">
<FRAMESET ROWS="*,*">
<FRAME SRC="1g.html" NAME="g1">
<FRAME SRC="2g.html" NAME="g2">
</FRAMESET></FRAMESET>
<FRAME SRC="inf.html" scrolling="no">
</FRAMESET>
</HTML>
```

Полученная в результате структура экрана показана на рис. 2.10. Оба фрейма, в которых выводится содержимое электронного издания имеют наибольшие размеры. В большинстве случаев содержание можно выводить в один большой фрейм, в этом же конкретном случае их два, что позволяет в одном фрейме изучать конкретный графический пакет PhotoShop, а в другом – общие методы обработки цифрового изображения, используемые в данном конкретном случае, например, сжатие изображения.

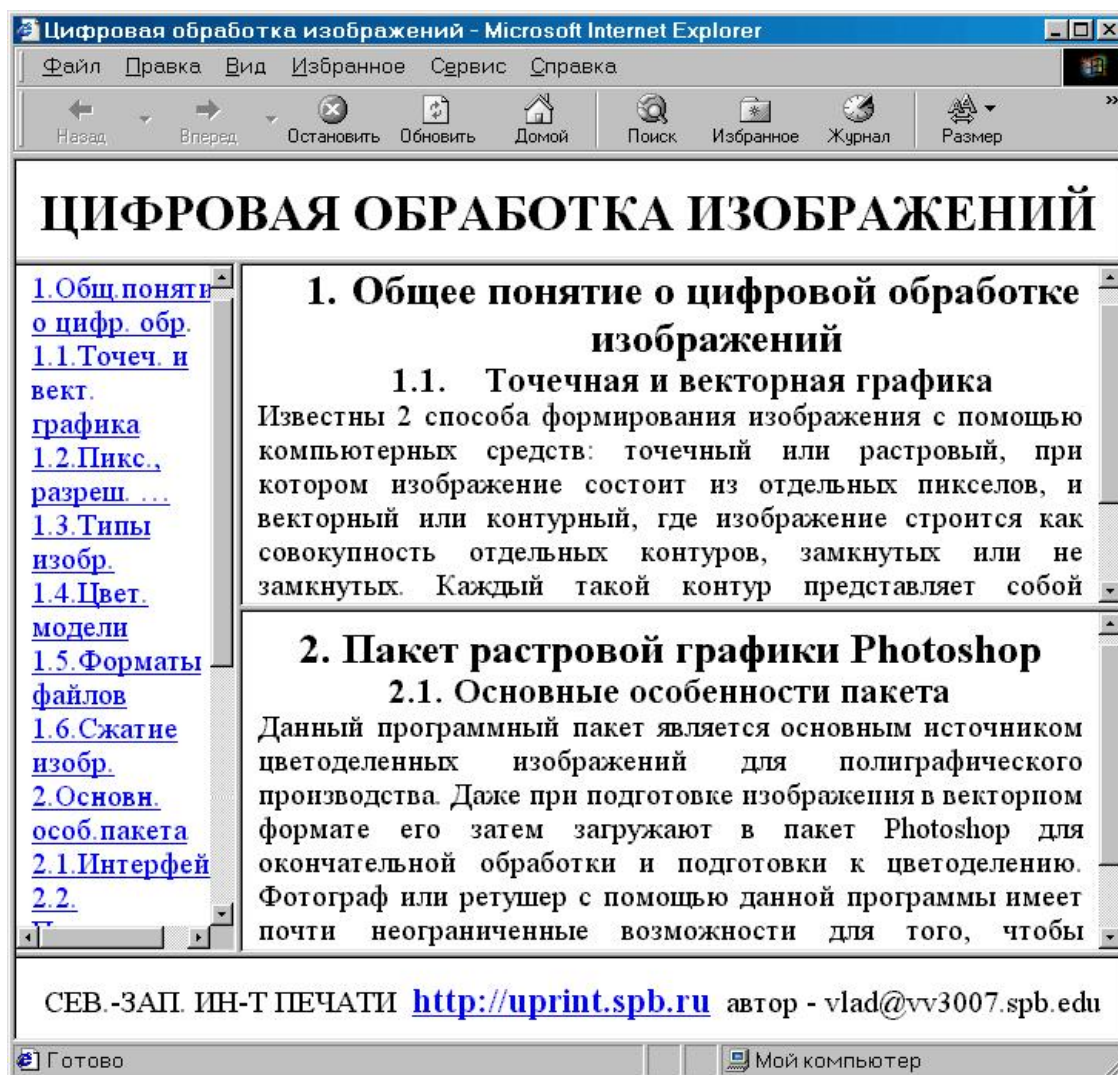


Рис. 2.10. Пример отображения документа с 5-фреймовой структурой

Если пронумеровать расположенные на экране фреймы, присвоив им номера с 1-го по 5-ый (верхний – 1, левый для 2-ой строки – 2, верхний правый для 2-ой строки – 3, нижний правый для 2-ой строки – 4, нижний – 5), то можно заметить, что содержание 1, 2 и 5 фреймов не изменяется: в 1-ый постоянно загружен HTML-файл, отображающий заголовок документа (см. строка 4 исходного кода), в 5-ый – информация об учебном заведении и авторе электронного пособия (чтобы можно было посмотреть другие учебники на сайте <http://uprint.spb.ru> или по электронной почте связаться с автором), а во второй – файл оглавления (см. строка 6 исходного кода). В 3-ем и 4-ом фреймах выводятся разделы 1-й и 2-й глав: в 3-ем фрейме выводится раздел 1.1. Точечная и векторная графика главы 1. Общее понятие о цифровой обработке изображений, а во фрейме 4 отображается раздел 2.1. Основные особенности пакета главы 2. Пакет растровой графики Photoshop. Таким образом, из одного фрейма (под номером 2, который можно назвать навигационным) с помощью указателей гиперссылок осуществляется управление информацией в двух других фреймах.

Отметим, что специалисты по педагогике и психологии образования считают, что не следует отображать на экране более 3–4 фреймов, в противном случае ухудшается процесс восприятия информации с экрана. Несколько подробнее на этом мы остановимся в главе 9, а здесь приведем следующий вариант того же самого электронного учебного пособия, но с лаконичной двухфреймовой структурой. Исходный HTML-код для него:

```
<html> <head> <title>Учебник по графике</title> </head>
<frameset cols="20%,*">
<frame name="Навигация" target="main" scrolling="auto"
src="navigagation.htm" noresize>
<frame name="main" src="content.htm" target="main">
<noframes> <body> <p>Ваш браузер не поддерживает фреймы</p>
</body> </noframes>
</frameset>
</html>
```

Соответствующая этому коду структура электронного издания представлена на рис. 2.11. На нем отсутствует фреймы заголовка и общей информации, а содержательная часть представлена в одном фрейме вместо двух. Каждый читатель может сам оценить, какая структура ему больше нравится и попытаться сформулировать причины этого. Здесь первый фрейм назван "Навигация" (см. строка 3 исходного кода), а второй – "main", т. е. "главный" (см. строку 5 исходного кода). При начальной загрузке в главном фрейме выводится оглавление, так как во фрейме "Навигация" названия разделов даны в сокращенной форме, чтобы уменьшить протяженность этого фрейма по вертикали.

Основная адресация производится из фрейма "Навигация", но ссылки работают и из оглавления в главном фрейме. Для возврата в оглавление во фрейме "Навигация" есть специальная кнопка "Назад в оглавление". Во фреймах сделан фоновый рисунок, улучшающий дизайн издания. Размеры фреймов нельзя изменить (см. строка 4 в исходном коде).

Из других отличий отметим вставку в исходный код строк 6 и 7 для того случая, когда браузер пользователя не поддерживает фреймовой структуры. В этом случае будет выдано сообщение "Ваш браузер не поддерживает фреймы", так как тэг-контейнер `<NOFRAMES>` (отсутствие поддержки фреймов) содержит внутри контейнер `<BODY>` и `</BODY>`, распознаваемый всеми без исключения браузерами, в результате чего они выводят в окне содержимое этого контейнера. При отсутствии этих двух контейнеров такой браузер выдал бы сообщение об ошибке или же, что еще хуже, не вывел бы ничего в своем окне. Тэг-контейнер `<NOFRAMES>` предусмотрен специально для такого случая.



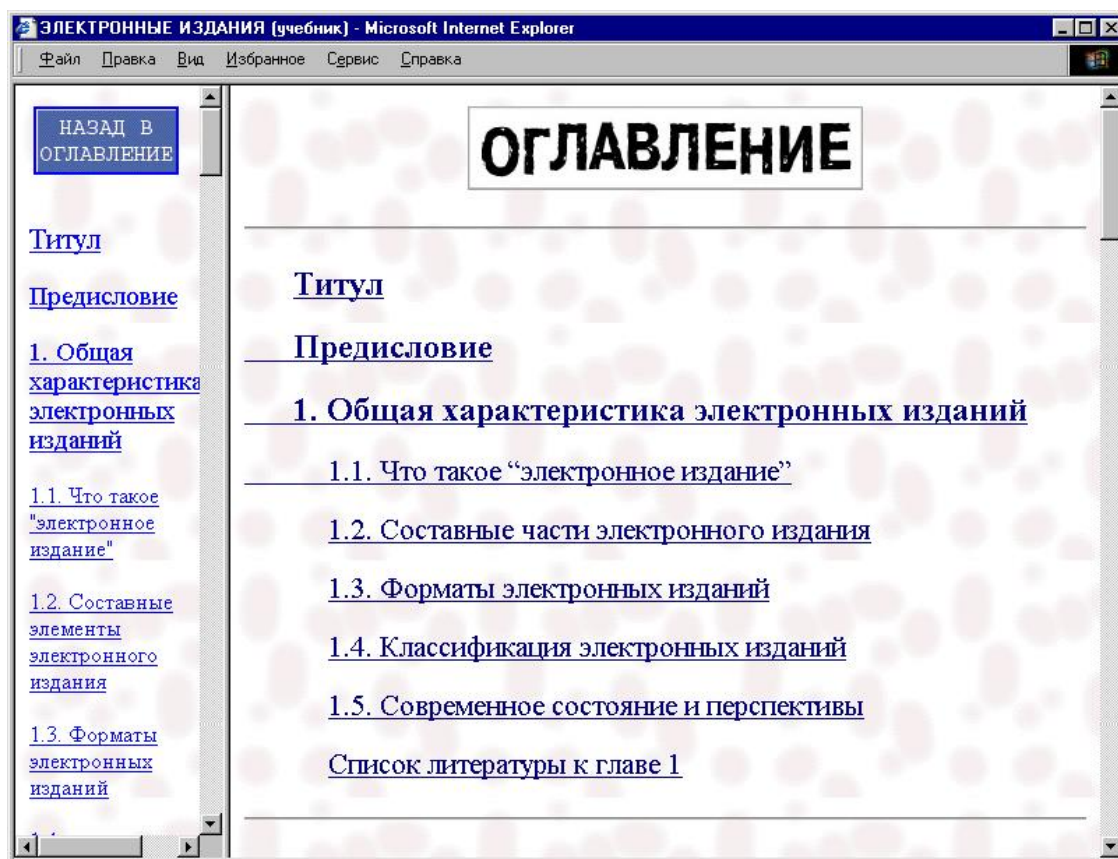


Рис. 2.11. Пример отображения документа с 2-фреймовой структурой

Из других замечаний можно указать, что звездочка (\*) во второй строке исходного кода интерпретируется как "оставшаяся часть", вместо нее можно было написать 80% и ничего бы не изменилось. Еще раз хочется обратить внимание на особенности навигации при использовании фреймов иначе – на взаимодействие фреймов. Так как во фрейме "Навигация" в качестве значения параметра target указано "main" (см. строка 3 исходного кода), то все ссылки из фрейма "Навигация" будут загружать документы во фрейм "main". Кнопка "Назад в оглавление" (изображение этой кнопки, используемой как указатель гиперссылки, содержится в документе navigation.htm) во фрейме "Навигация" играет ту же роль, что и изображение – указатель ссылки, использование которого было рассмотрено в разд. 2.5. Наконец, рисунки в этом учебнике выводятся в плавающих окнах, которые можно расположить так, как удобно пользователю. Пример вывода рисунка в плавающем окне в левой верхней части рабочего окна браузера показан на рис. 2.12.



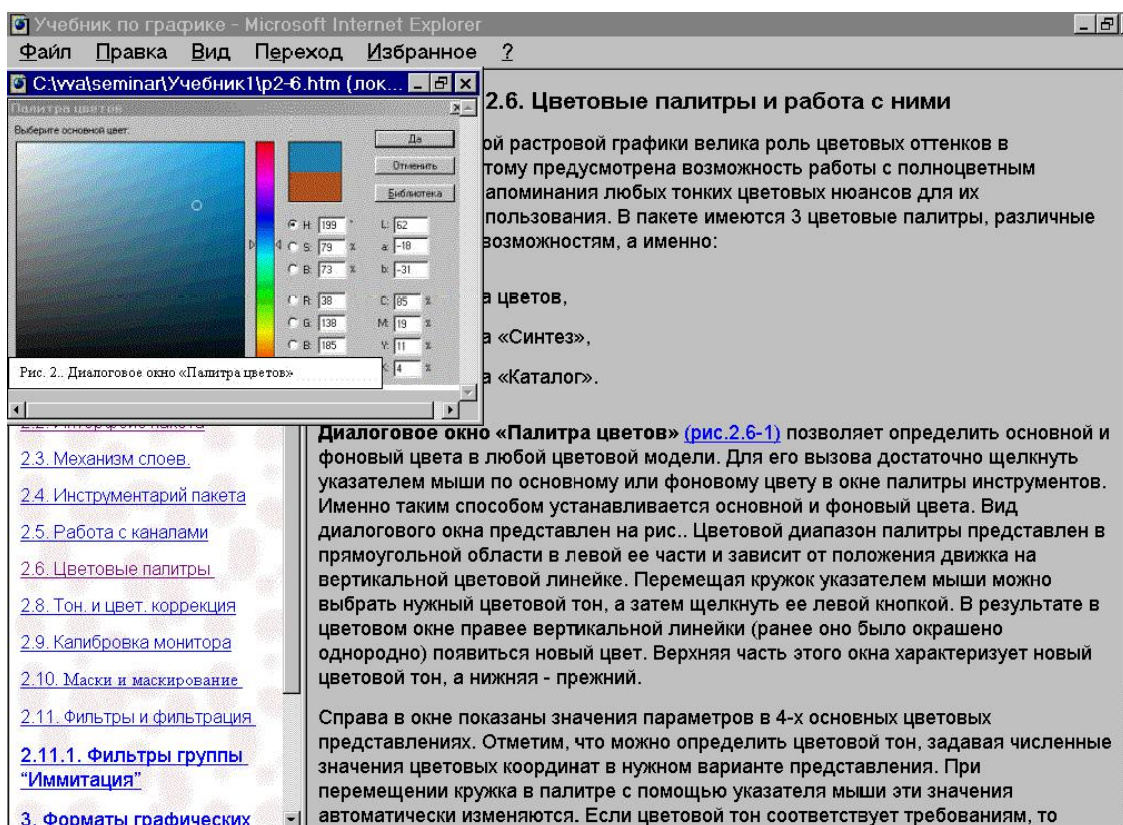


Рис. 2.12. Пример отображения документа с двухфреймовой структурой и изображением во всплывающем окне

## 2.8. HTML-формы

HTML-формы – это особый вид гипертекстового документа, предназначенный для организации интерактивного взаимодействия в электронных изданиях. Именно с помощью формы мы общаемся с поисковым сервером, начиная поиск нужных материалов (документов) для работы, учебы или развлечения. Эту форму мы получаем с сервера и в специальном ее окне задаем параметры для поиска, например, выражение в виде набора ключевых слов, т. е. тех слов, которые, по нашему мнению, наиболее важные и существенные в искомом документе. Таким образом, HTML-форма является средством пересылки данных от удаленного пользователя (клиента) к Web-серверу.

В языке HTML для задания форм используется тэги `<FORM>` (форма) и `</FORM>`. Этот контейнер размещается внутри тела HTML-документа, причем таких контейнеров может быть несколько, но вложений их один в другой не допускается. Тэг `<FORM>` может содержать несколько параметров, а именно: ACTION (действие, операция), METHOD (метод), ENCTYPE (медиа-тип, определяющий кодирование данных при передаче от браузера к серверу). Некоторые браузеры поддерживают и другие параметры, помимо перечисленных трех стандартных. Первый из параметров задает адрес (URL) CGI-программы, которая будет обрабатывать данные этой формы. Это единственный обязательный параметр тэга `<FORM>`.

Второй параметр METHOD определяет способ пересылки данных от браузера к Web-серверу, содержащихся в форме. Он принимает два возможных значения: GET (по умолчанию) и POST. Значение METHOD="POST" используется в случае необходимости пересылки на сервер определенного файла, который присоединяется к содержимому формы.

Параметр ENCTYPE определяет медиа-тип, используемый для кодирования и пересылки содержимого формы. Обычно его называют MIME-типом, где MIME – это аббревиатура, которая расшифровывается как Multipurpose Internet Mail Extension (многоцелевое расширение для интернет-почты). Именно кодирование в соответствии со стандартом MIME позволяет безошибочно пересылать файлы, присоединяемые к HTML-форме.

Для создания отдельных полей внутри контейнера `<FORM>` и `</FORM>` используются тэги `<INPUT>` (ввод данных), `<SELECT>` (выбор) и `<TEXTAREA>` (текстовая область). Первый из этих тэгов употребляется чаще других и обеспечивает многочисленные возможности ввода информации. Так, с его помощью можно осуществить ввод текста различного назначения, создать группу элементов-переключателей, одиночные кнопки для выполнения некоторых стандартных операций (например, сброс введенных данных), поля для установки флажков, поле для ввода имени файла, присоединяемого к форме.

Тэг `<SELECT>` используется для организации компактной структуры для выбора одного варианта из нескольких, задаваемых в виде списка прокрутки. А тэг `<TEXTAREA>` создает внутри формы поле для ввода многострочного текста в виде прямоугольной области, снабженной при необходимости полосами прокрутки. Кратко рассмотрим использование всех этих тэгов.

Тэг `<INPUT>` содержит 2 обязательных параметра: TYPE, который задает тип поля ввода и NAME, с помощью которого имя переменной, соответствующей этому полю ввода передается CGI-программе.

Параметр TYPE может принимать несколько значений. Одно из этих значений TYPE="TEXT". В этом случае в форме создается узкая прямоугольная область (фрагмент строки) для ввода текста. Причем при этом значении параметра TYPE в тэге INPUT могут использоваться 3 дополнительных параметра, а именно: MAXLENGTH (наибольшая длина) = <число>, который определяет максимальное число вводимых символов, SIZE=<число>,

задающий количество отображаемых в форме символов, и VALUE, который определяет значение текстового поля по умолчанию.

Другой вариант текстового поля – это TYPE="PASSWORD" (пароль). В этом случае вводимые символы отображаются в форме звездочками. Однако, реальная безопасность передаваемого пароля не обеспечивается, так как на сервер эта строка передается в незашифрованном виде, то есть открыто.

Еще одно значение TYPE="FILE", причем в этом случае одновременно с полем для ввода текста (имени локального, то есть размещенного на компьютере клиента, файла) в форме создается кнопка с подписью **BROWSE** (просмотр). Файл, выбираемый в процессе просмотра (локальный адрес его может быть сразу введен в поле), подсоединяется к содержимому формы при ее пересылке на сервер. Для корректного выполнения этой операции должны быть правильно заданы параметры формы: METHOD="POST" и ENCTYPE="MULTIPART/FORM-DATA". В противном случае вместо содержимого файла на сервер будет передано только его имя.

Значение TYPE="CHECKBOX" (переключатель) создает поле ввода одного символа–флажка или переключателя типа "включено" либо "выключено". Такие поля могут объединяться в группы путем задания для всех одного имени (NAME). Используется параметр VALUE, который передает на сервер выбранные переключатели. Может присутствовать также параметр CHECKED (установленное значение, обычно по умолчанию), указывающий, что определенное значение соответствующего элемента является выбранным по умолчанию.

Значение TYPE="RADIO" (селективная кнопка, зависимый переключатель) определяет структуру, подобную многопозиционному переключателю, то есть набор из нескольких выключателей (иначе радиокнопок), один (и только один) из которых включен в любой отрезок времени. Каждый выключатель оформлен в виде круга, внутри которого при его выборе появляется точка. Здесь строковая переменная VALUE передает на сервер значение выбранной кнопки, а параметр CHECKED устанавливает значение по умолчанию.

Значение параметра TYPE="RESET" (сброс, возврат в исходное состояние) или же "SUBMIT" (запустить, инициализировать выполнение) соответственно отменяет все введенные в форму значения или же пересылает содержимое формы на сервер. В обоих случаях в форме отображаются кнопки с соответствующими надписями. Надписи на кнопках можно русифицировать, задав значение параметра VALUE равным **Отмена** или же **ОК (Пересылка)** соответственно. Возможно использование значения параметра TYPE="IMAGE". В этом случае создается графическое изображение, аналогичное кнопке SUBMIT. Адрес изображения задается стандартным параметром SRC=<URL файла-изображения>, а способ выравнивания – значением параметра ALIGN.

Иногда вместо тэга INPUT используется контейнер <SELECT> и </SELECT>, который позволяет организовать выбор значений в пределах спискового окна. Это обеспечивает экономии пространства формы. Пример контейнера:

```
<SELECT NAME="имя поля" SIZE=[число видимых элементов списка] >
<OPTION VALUE="значение 1-го элемента"> значение элемента
<OPTION VALUE="значение 2-го элемента"> значение элемента
<OPTION VALUE="значение 3-го элемента"> значение элемента
</SELECT>
```

Если требуется возможность выбора нескольких значений элементов списка, то внутри тэга <SELECT> после параметра SIZE используется дополнительный параметр MULTIPLY

(множественный). Ему не требуется присваивать какого-либо значения, само его наличие в пределах тэга является достаточным.

Внутри формы может использоваться и еще один контейнер, а именно с тэгом <TEXTAREA>. Он обеспечивает ввод многострочного текста в прямоугольное окно. Параметры этого тэга: NAME – для задания имени соответствующей переменной, ROWS и COLS – для определения размера прямоугольного окна, внутри которого набирается нужный текст. Кроме того, может быть задан текст, выводимый внутри этого окна по умолчанию.

В качестве примера можно остановиться на форме сервера [www.rambler.ru](http://www.rambler.ru), которая выводится на браузере клиента, желающего получить бесплатный почтовый адрес. На рис. 2.13 представлена такая форма в несколько усеченном формате. Как видно на рисунке, большинство полей – текстовые однострочные, задаваемые тэгом INPUT и значениями параметра TYPE, равным TEXT (E-mail, "Имя", "Фамилия", "Возраст") и PASSWORD ("Введите пароль" и "Подтвердите пароль"). В состав формы включены 2 переключателя: для указания пола клиента и его юридического статуса (физическое лицо либо представитель юридического лица). Для сокращения размеров формы (она и так получилась довольно большой) в ней используются 4 списковых окна с одним единственным видимым элементом. Эти окна реализуются с помощью тэга-контейнера <SELECT> со значением параметра SIZE=1.

**Rambler-почта: Регистрация - Microsoft Internet Explorer**

Файл Правка Вид Избранное Сервис Справка

\* Ваш возраст:

\* Частота пользования Интернетом:

\* Образование:

\* Сфера деятельности:

\* Ваш социальный статус:

**Соглашение об оказании услуг:**

Соглашение "Об оказании услуг бесплатной службы электронной почты mail.Rambler.ru"

Публикуя настоящее Соглашение, Владелец бесплатной службы электронной почты mail.ОАО "Рамблер Интернет Холдинг" (далее - Владелец Службы) обязуется выполнять усл.указанные в нем и в Регламенте бесплатной службы электронной почты mail.Rambler.ru. Пользователь выражает согласие принять условия настоящего Соглашения регистрацией бесплатной службе электронной почты MAIL.RAMBLER.RU. При этом информация о данном волеизъявлении направляется Владелец, после чего настоящее соглашение вступает в.По настоящему Соглашению Пользователь получает, а Владелец Службы оказывает услуг.бесплатной службы электронной почты MAIL.RAMBLER.RU на условиях, указанных в наст.Соглашении и в Регламенте бесплатной службы электронной почты MAIL.RAMBLER.RU, опубликованном по адресу: <http://mail.rambler.ru/docs/Reglament.html> (далее - Рег.являющиеся неотъемлемой частью настоящего Соглашения. В случае недостижения разрешения споров по настоящему Соглашению путем претензион.

\* Вы регистрируетесь как: ☐ физическое лицо  
☐ представитель юридического л

**Зарегистрироваться**

Рис. 2.13. Пример отображения формы для регистрации клиента электронной почты

Наконец, в форме предусмотрена область для ввода текста, задаваемая с помощью тэга <TEXTAREA>. Размер этой области: 6 строк и около 100 столбцов. По умолчанию в этой области выводится текст юридического соглашения с клиентом бесплатной электронной почты, указывающий права и обязанности сторон. Клиент не может изменять этот текст, он должен его прочесть и принять условия соглашения, изложенные в этом тексте, если намерен получить бесплатный почтовый адрес на этом сервере. В конце формы предусмотрена кнопка типа SUBMIT, представленная в виде графического изображения с надписью "Зарегистрироваться". Если клиент согласен с текстом соглашения, он нажимает кнопку "Зарегистрироваться" и получает бесплатный почтовый адрес на сервере RAMBLER.

## 2.9. Понятие о динамическом HTML и его компонентах

Динамический HTML (DHTML) это не какой-то особый язык, а термин, который применяется для обозначения HTML-страниц с динамически изменяющимся содержанием. Такие страницы основаны на использовании помимо языка HTML каскадных таблиц стилей CSS (Cascade Style Sheets) и языка сценариев JavaScript или VBScript. Связь этих компонентов осуществляется с помощью так называемой объектной модели документа DOM (Document Object Model). При этом HTML-документ приобретает новое качество, а именно возможность динамического изменения без перезагрузки страницы.

С помощью каскадных таблиц стилей определяется внешний вид документа: шрифт, разбивка на абзацы, цвет фона и шрифта, причем свой вариант отображения в окно браузера с их помощью можно установить для содержимого каждого из контейнеров (тэгов).

Язык программирования JavaScript разработан фирмой Netscape для создания интерактивных HTML-документов. Язык позволяет разрабатывать как серверные приложения, так и клиентские приложения. Первые выполняются на Web-сервере, а вторые – браузером клиентского компьютера. Для обоих типов приложений используется общее ядро, включающее стандартные объекты и конструкции, и дополнения для конкретных приложений. Сказанное иллюстрируется рис. 2.14. Как видно на рисунке в состав ядра помимо переменных и функций LiveConnect входит средство взаимодействия с Java-апплетами. Клиентские приложения непосредственно встраиваются в HTML-страницы и интерпретируются браузером по мере отображения элементов страницы, серверные же приложения предварительно компилируются для увеличения производительности.

Язык JavaScript используется на клиентской стороне для создания HTML-документов с помощью сценариев, в том числе и совместно с CSS (каскадными таблицами стилей), для оперативной проверки HTML-форм до их передачи на сервер для последующей обработки и для взаимодействия с пользователем в процессе выполнения приложений этого языка. Трудно найти в сети Интернет хотя бы одну страницу, не содержащую операторов языка JavaScript. Чтобы браузер смог отобразить разработанную Web-страницу именно в том виде, каком она задумывалась, на странице обычно помещают вызов функции JavaScript для идентификации типа используемого браузера и его версии.

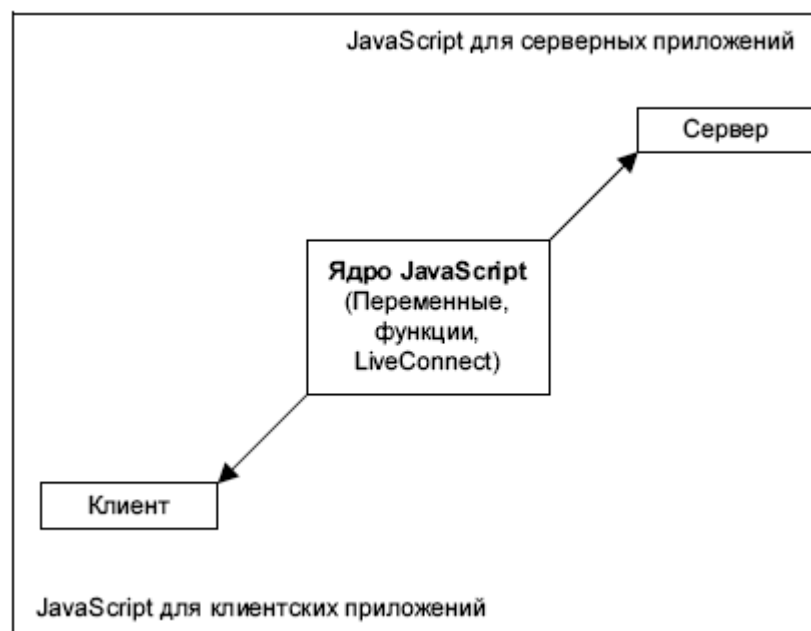


Рис. 2.14. Компоненты языка JavaScript

Встраивание сценариев JavaScript в HTML-страницу можно осуществить одним из 4 способов:

- # заданием операторов языка JavaScript внутри контейнера `<Script>` языка HTML;
- # заданием имени файла языка JavaScript в параметре SRC тэга `<Script>`;
- # использованием выражений языка JavaScript в качестве значений параметров HTML-тэгов;
- # путем определения обработчика событий в тэге HTML.

Первые два способа достаточно понятны и не нуждаются в дополнительных комментариях. Если переменные языка JavaScript желательно использовать в качестве параметров, то их следует заключать в фигурные скобки и размещать между символом амперсанда (&) и точкой с запятой. Например, в тэге

```
<TABLE ALIGN=LEFT WIDTH=40% BORDER=5>
```

значение параметра WIDTH можно сделать изменяемым, заменив константу 40% переменной языка JavaScript, как это сделано в выражении, представленном ниже:

```
<TABLE ALIGN=LEFT WIDTH="&{JWidth};" BORDER=5>
```

причем значение переменной Jwidth должно быть предварительно определено. Для совместимости с языками сценариев в элементы форм языка HTML (см. разд. 2.8) были введены специальные параметры обработки возникающих событий. Простейшим событием может быть нажатие левой кнопки мыши Click. Значениями этих параметров могут служить операторы языка JavaScript. Соответствующий любому событию параметр будет иметь дополнительную приставку on, например, событию Click будет соответствовать параметр onClick. Обработка событий задается в параметрах форм для того, чтобы проверить введенную в форму информацию до ее пересылки на сервер.

Обсуждение самого языка JavaScript потребовало бы отдельной главы. Поэтому мы ограничимся приведенными в данном коротком разделе сведениями, а всех желающих

глубже ознакомиться с ядром JavaScript отсылаем к многочисленным пособиям по этому языку, например, [27].

Отметим, что фирма Microsoft разработала язык создания сценариев VBScript (Visual Basic Script сценарии языка Visual Basic), функциональные возможности которого ничем не отличаются от JavaScript. Определенным достоинством этого языка является возможность использования с небольшими доделками ранее написанных процедур на языке Visual Basic.



## 2.10. Расширяемый язык разметки XML

### 2.10.1. Общая характеристика XML

Относительно недавно появился новый инструмент для форматирования Web-страниц – язык XML (eXtensible Markup Language расширяемый язык разметки). Этот компактный язык, разработанный Консорциумом W3C ([www.w3.org](http://www.w3.org)) специально для того, чтобы им можно было в дальнейшем заменить HTML. Язык XML не имеет никакого отношения к графическим языкам или к средствам визуальной разметки документов. Сам по себе XML не предоставляет даже тех средств форматирования, которые есть в HTML. Строго говоря, XML не язык, а метаязык, позволяющий создавать специализированные системы логической разметки для любых разновидностей документов.

Внешне XML-документ похож на HTML. XML отличается от HTML синтаксисом, используемым для представления пустых элементов. Это небольшое, но существенное отличие и является главной причиной того, что создать программу для обработки XML-документов намного проще, чем аналогичную программу для HTML-документов.

В идеале форматирующие, гипертекстовые и прочие функции тегов должны определяться самостоятельными, не зависящими от языка разметки формальными системами, а реализация этих функций в браузерах должна быть отделена от синтаксического выделения тегов. Шагом в этом направлении стала разработка каскадных таблиц стилей CSS (см. разд. 2.9) – независимой от HTML-системы и позволяющей в некоторых пределах изменять параметры форматирования, ассоциированные с тем или иным тегом. Однако, надо заметить, что это нововведение сильно запоздало и предоставляет довольно ограниченные возможности форматирования. Кроме того, данная система порождает вопросов больше, чем решает.

Главная и почти единственная задача языка логической разметки XML (eXtensible Markup Language) – разбить содержимое документа на элементы, причем теги для разграничения этих элементов пользователь может создавать сам. В языке XML нет ни одного заранее определенного тега с фиксированным значением.

Правильно структурированные документы годятся для большинства практических случаев. Это значит, что разработчик может сразу же начать описывать структуру документа на "почти человеческом" языке, заботясь лишь о правильном вложении тэгов и контейнеров.

Опыт языка HTML (Hypertext Markup Language) показал, сколь важна тщательная и своевременная интернационализация всех аспектов языка, претендующего на какую-то роль в сети Интернет. Поэтому создатели языка XML позаботились, в частности, о том, чтобы в именах тегов и атрибутов можно было пользоваться не только латинскими буквами, но и кириллицей, иероглифами и вообще всеми символами из набора Unicode (стандартная кодировка символов в операционной системе Windows).

Такая разметка позволит интерпретатору языка XML разделить документ на небольшие части в соответствии с его тэговой структурой. После этого в действие вступает другое приложение – его задачей может быть, например, автоматическое индексирование документа, занесение его в базу данных или (чаще всего) форматирование в соответствии с приложенной к документу стилевой спецификацией. (В нашем примере можно было бы, скажем, раскрасить разные части речи разными цветами.) Однако все эти задачи лежат уже за пределами собственно языка XML, который, таким образом, сфокусирован только на логической структуре документа.

Самое важное отличие языка XML состоит в использовании DTD (Document Type Definitions определение типа документа), или набора правил, определяющего имена

дескрипторов и их атрибутов в конкретном документе. Все HTML-документы отвечают единственному DTD – спецификации языка HTML. Однако язык XML позволяет каждому автору создать свое собственное DTD.

Возможность использования произвольных тэгов означает, в частности, что любой HTML-документ очень легко преобразовать в XML-документ. Изменения, требуемые для этого преобразования, немногочисленны и сугубо формальны – все тэги должны быть оформлены в виде контейнеров (содержать наряду с открывающим еще и закрывающий тэг), все значения атрибутов должны быть взяты в кавычки и т. п.

Возможности языка XML по манипулированию объектами мультимедиа, далеко превосходят подобные возможности HTML. У пользователя появляется возможность помещать в документ только часть графического изображения и начинать воспроизведение аудиоклипа с любого желаемого момента, а не с начала.

## 2.10.2. Надстройки для XML

Создатели языка XML понимали, что простота и изящество логического подхода к разметке имеют оборотную сторону – отсутствие мощных и притом стандартизированных средств определения семантики тэгов. Поэтому, с момента появления черновой спецификации языка XML в ноябре 1996 года, разработчики заняты в основном выбором и стандартизацией расширений языка – надстроек над XML, которые позволили бы формально описывать различные семантические аспекты тэгов.

В отличие от языка HTML, модульная структура XML является одним из важнейших преимуществ этого языка. Авторы XML прилагают все усилия к тому, чтобы логический "базис" и семантические "надстройки" удобно стыковались, не теряя при этом своей независимости друг от друга. Даже спецификация языка разбита на отдельные части, первая из которых – ([www.textuality.com/sgml-erb/WD-xml-lang.html](http://www.textuality.com/sgml-erb/WD-xml-lang.html)) – покрывает весь синтаксис логической разметки.

Вторая же (и пока последняя) часть стандарта ([www.textuality.com/sgmlerb/WD-xml-link.html](http://www.textuality.com/sgmlerb/WD-xml-link.html)) описывает механизм создания гипертекстовых ссылок в XML-документах. Этот аспект языка значительно усовершенствован в сравнении с HTML. Основные особенности гипертекстовой модели языка XML представлены ниже:

1. XML-ссылки реализованы не на уровне тэгов (тег <A> языка HTML), а с помощью зарезервированных имен атрибутов. Это позволяет с легкостью превратить в гипертекстовую ссылку любой элемент документа, просто расширив его список атрибутов.

2. Для XML-ссылки можно указать, будет ли она обычной ссылкой, активизируемой пользователем, или же браузер, встретив в документе эту ссылку, должен активизировать ее сам.

3. Для ссылки можно указывать результат ее активации, а именно: выводить ли документ, на который она ссылается, вместо текущего (например, в том же окне браузера), создать ли для него новое окно или же содержимое нового документа нужно вставить прямо внутрь текущего документа.

4. Важные усовершенствования внесены в синтаксис URL-адресов, использующихся в ссылках. Язык XML расширяет синтаксис этих конструкций, благодаря чему, не теряя обратной совместимости с существующими адресами, они позволяют адресовать практически любой фрагмент любого

5. XML– или HTML-файла. При этом не требуется, чтобы автор файла, на который ссылаются, как-то по-особому разметил этот фрагмент (оформил для него "закладку").

6. Что же касается визуального форматирования, то здесь наиболее вероятным кандидатом в партнеры XML является язык DSSSL (Document Style Semantics and Specification

Language, язык стилистических и семантических спецификаций документов), описание которого приведено на сайте [www.jclark.com/dsssl/](http://www.jclark.com/dsssl/). По сравнению с CSS язык DSSSL является гораздо более мощным и разносторонним инструментом; с его помощью можно описывать не только визуальное форматирование XML-документов, но и преобразование документов от одного варианта DTD к другому. Для сетевого представления документов разрабатывается упрощенная версия DSSSL, получившая название DSSSL Online или DSSSL-O. Вполне возможно, что спецификация DSSSL-O станет третьей частью стандарта языка XML.

### 2.10.3. Перспективы развития XML

Прежде чем приступать к выработке стандарта, разработчики языка XML составили список требований, которым должен удовлетворять новый язык. Видимо, этот список можно рассматривать как долгосрочную перспективу развития этого языка. Содержимое списка требований представлено ниже.

1. Язык XML должен наилучшим образом подходить для документов, распространяемых через Интернет. Современные Web-серверы требуют лишь минимальных изменений конфигурации для обработки XML-документов. Стандартный способ связи XML-документов и DTD-файлов посредством URL также уже использовался ранее в HTML-документах.

2. Программы для работы с языком XML должны быть просты в написании и отладке. Некоторые из экспериментальных интерпретаторов языка написаны в виде Java-классов и занимают объем буквально в несколько килобайт.

3. Число необязательных функций должно быть минимально. В идеале они должны отсутствовать.

4. XML-документы должны быть достаточно просты для чтения и понимания человеком. Благодаря встроенной возможности создания новых тэгов и их записи на любом языке мира, XML-документы должны в будущем читаться как обычный текст.

5. Стандарт языка XML должен быть четким и сжатым.

Технология языка XML находится пока еще в зачаточном состоянии, однако растущий интерес к XML определяет возможность того, что в будущем язык XML заменит HTML [16].

## Контрольные вопросы

1. Что такое тэг и для чего он используется? Чем контейнер отличается от тэга?
2. Какие 3 контейнера из HTML-тэгов должны содержаться в некотором документе для того, чтобы этот документ мог считаться гипертекстовым или HTML-документом?
3. Как определяется заголовочная часть HTML-документа?
4. Какие тэги и контейнеры могут содержаться в заголовочной части? В какой контейнер следует поместить заголовок? Для чего следует выделить заголовок в HTML-документе?
5. Зачем нужен тэг <META>? Какие параметры могут быть в нем?
6. Каким образом можно оформить тон фона всего HTML-документа? Каким образом задаются его цветовые координаты? Можно ли оформить фоновый рисунок? Если можно – то как это выполнить?
7. Чем принципиально отличаются друг от друга тэги физического и логического форматирования текста? Какие примеры тэгов физического форматирования вы можете привести? Для чего служит каждый из них?
8. Какой тэг обеспечивает "шрифтовое оформление" произвольного текстового блока? Сколько в нем различных параметров и для чего служит каждый из них?
9. Как обеспечить выбор одной из шрифтовых гарнитур, установленных на вашем компьютере?
10. Каким образом можно изменять размер шрифта в HTML-документе? Сколько таких возможных размеров существует?
11. Какими способами можно изменить цвет фрагмента текста?
12. Как оформить в HTML-документе математическое выражение с показателями степени и индексами? Попробуйте привести примеры.
13. Какой тэг служит для оформления абзаца? Сколько у него параметров и какие? Как при отображении HTML-документа в браузере один абзац отделяется от другого?
14. Как оформить рубрики в HTML-документе? Как определить уровень заголовка, размещенного внутри текста HTML-документа?
15. Какие виды списков, используемых в HTML-документах, вы можете назвать? Как такие списки оформляются (общие принципы оформления)?
16. Попробуйте записать пример оформления маркированного списка. Как задается конкретный тип маркера?
17. Прodelайте то, что сказано в п.16, но для нумерованного списка.
18. Чем отличается список определений от двух предыдущих видов списков? Как оформляется определяемый термин и как сам текст определения?
19. С помощью какого тэга можно вставить рисунок в текст документа? Какие обязательные и необязательные параметры у этого тэга?
20. Какими способами можно выравнивать рисунок в тексте документа? Какие способы для вас предпочтительнее?
21. Рисунки в каких форматах могут быть помещены в текст HTML-документа?
22. Что такое "Альтернативный текст" для изображения? В каком случае он требуется и для каких целей служит?
23. Какой тэг используется для задания ссылки? Чем отличаются ссылки внутренние и внешние? Что такое указатель гиперссылки?
24. Может ли изображение служить указателем гиперссылки?
25. Какой обязательный параметр у тэга, задающего ссылку? Какие еще могут быть у него параметры?
26. Как определяется закладка внутри документа? Для чего она используется?

27. Как записывается тэг, задающий внутреннюю гиперссылку? Что является признаком внутренней ссылки?

28. На какие типы файлов могут указывать гиперссылки?

29. Какие тэги используются для оформления таблиц? Как задается заголовочная и информационная части таблицы?

30. Попробуйте записать HTML-код для задания простой таблицы, содержащей одну строку заголовка, 2 строки для занесения данных и 2 столбца.

31. Какие параметры используются для оформления (улучшения внешнего вида) таблиц и что они определяют?

32. Как задается обтекание таблицы текстом документа?

33. Какие параметры для объединения смежных ячеек в строке таблицы?

Смежных ячеек в ее столбце?

34. Что такое "фрейм" и для чего он используется? В каких случаях целесообразно использовать фреймы в гипертекстовых документах?

35. Какие тэги используются для определения фреймов? Попробуйте определить (задать) фреймовую структуру из 4 фреймов: 3 горизонтальные строки 10%, 80% и 10%, средняя из которых разделена по вертикали в отношении 20% и 80%.

36. Для чего служит параметр "TARGET" при работе с фреймами?

37. Что такое HTML-формы и для чего они служат? Встречали ли вы такие конструкции и сети и где именно?

38. Какие разновидности форм вы знаете? Для чего служит каждая из них?

39. Какой тэг используется для ввода многострочного текста? Какие параметры в нем используются?

40. Что такое "Динамический HTML"? Попробуйте описать, для чего служат каскадные таблицы стилей, сценарии различного типа.

41. Для чего служит язык XML? Чем он сходен и в чем отличается от HTML? Каковы его перспективы?

## Самостоятельные (лабораторные) работы

### Работа №2.1. Преобразование текстового документа txt-формата в HTML-документ

#### Порядок выполнения работы

1. Скопируйте в одну из папок (например, c:\el-books\lab-21) на жестком магнитном диске вашего ПК папки Исходные тексты и HTML-документы, размещенные в папке К главам 2 и 4 – тексты на гибком магнитном диске, прилагаемом к пособию. В первом папке находятся 10 текстовых файлов с именами от 1.txt до 10.txt, представляющие собой фрагменты 1 и 3 глав книги. Во второй папке размещаются вложенные папки с рисунками к этим главам, а именно Ris-g1 и Ris-g3.

2. Загрузите один из этих текстовых файлов в любой текстовый редактор, предпочтительно Блокнот, предусмотренный в оболочке Windows (в примере на рис. 2.15 показан загруженный в редактор Блокнот документ). Заключите нужные части документа в контейнеры <HTML> и </HTML>, <HEAD> и </HEAD>, <BODY> и </BODY>. Контейнер <HEAD> и </HEAD> пока можно оставить пустым. Затем сохраните документ в новом формате в папке HTML-документы. Диалоговое окно **Сохранение** также показано на рис. 2.15. В строку, отведенную для имени файла, обязательно впишите расширение html или htm. Откройте папку, в которую помещен полученный HTML-документ, и обратите внимание на изменение внешнего вида его пиктограммы (рис. 2.16) по сравнению с пиктограммой текстового документа. Щелкните указателем мыши по пиктограмме HTML-документа и проследите за его загрузкой в браузере, установленном на вашем компьютере. Обратите внимание на содержимое строки заголовка браузера.

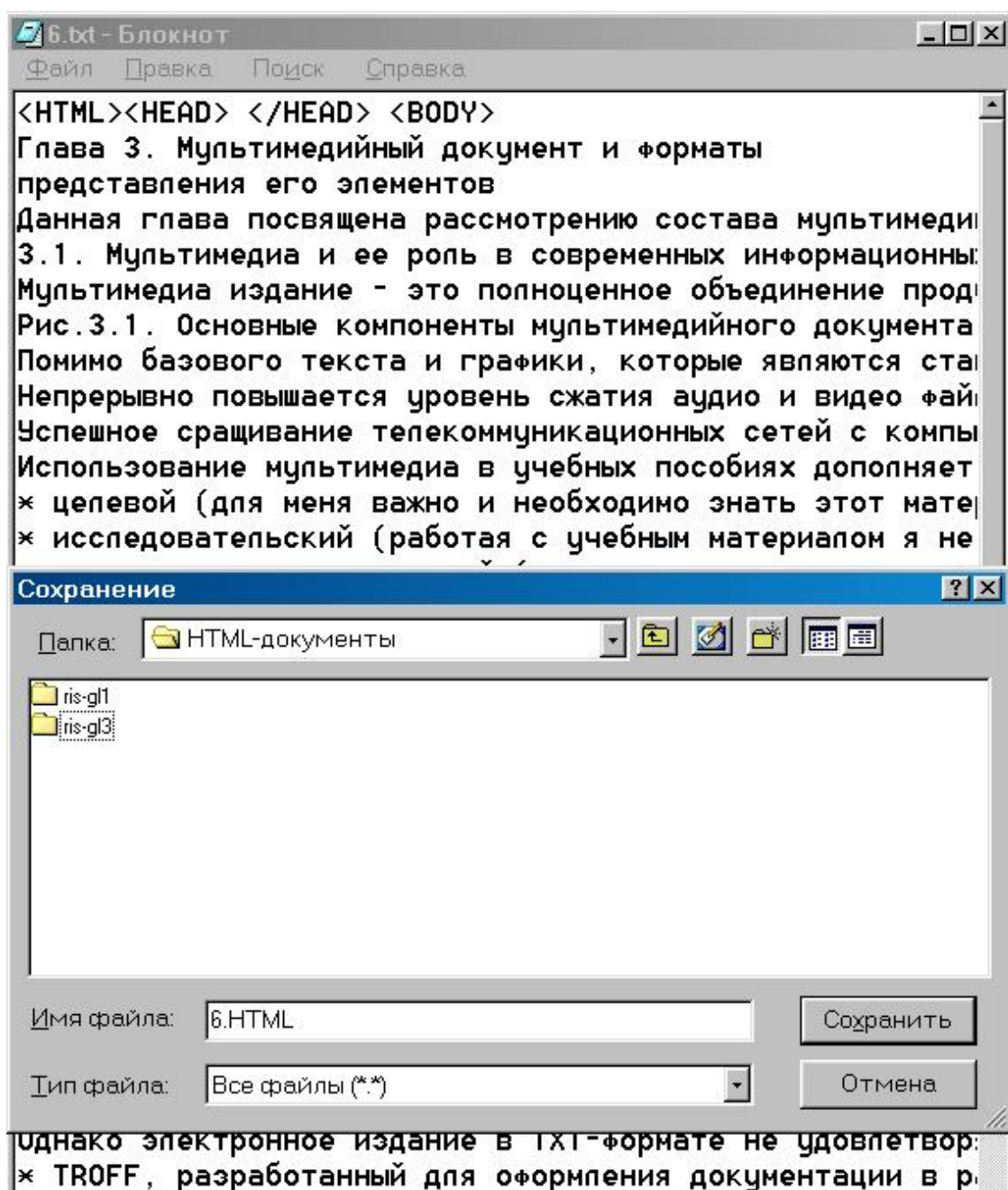


Рис. 2.15. Пример сохранения текстового документа в формате HTML

3. Теперь следует занести содержимое в контейнер `<HEAD>` и `</HEAD>`. Для этого разместите внутри него контейнер `<TITLE>` и `</TITLE>`, в который поместите название документа "Электронные издания". Попробуйте также вставить в заголовочную часть документа тэг `<META>`, указав в нем автора и название документа. Если встретились трудности, прочтите вновь разд. 2.1, обратив особое внимание на приведенные там примеры.

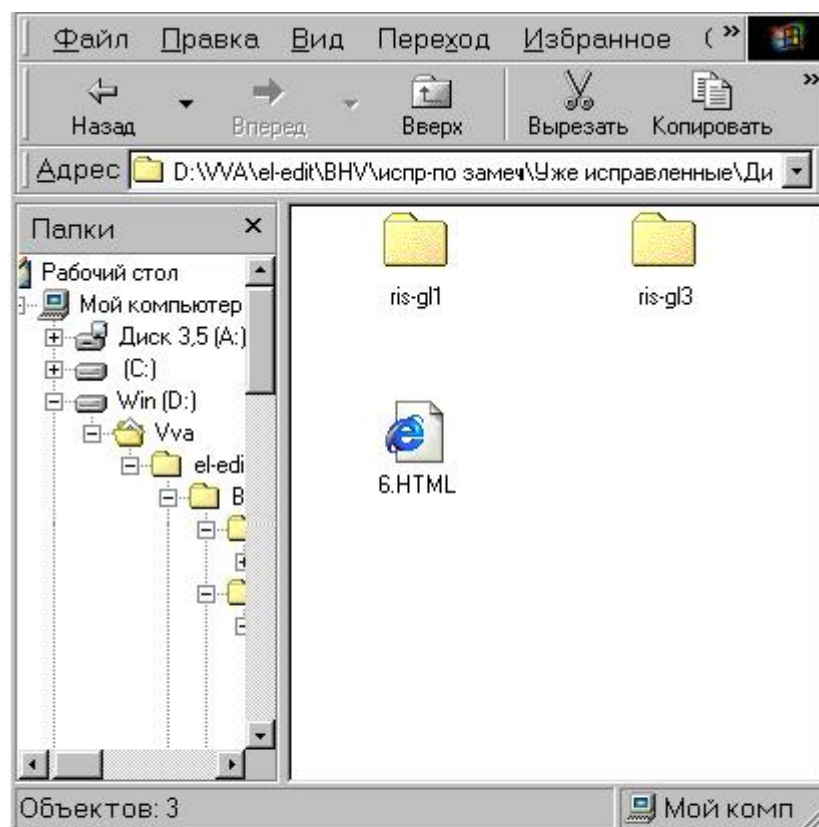


Рис. 2.16. Внешний вид пиктограммы HTML-документа в "Проводнике"

4. Сохраните внесенные изменения в редакторе Блокнот (команда **Сохранить** в меню **Файл**), откройте вновь папку, в которой храниться HTML-документ и, щелкнув по его пиктограмме, загрузите его в браузере, установленном на вашем компьютере. Отметьте, что изменилось в интерфейсе по сравнению с предыдущим случае в п. 2. Постарайтесь объяснить причины этих изменений.

5. Теперь попробуйте оформить весь документ в целом, задав с помощью параметра BACKGROUND фоновый рисунок или же тонирование с помощью параметра BGCOLOR. Можно попробовать несколько вариантов, сохранив тот, который покажется вам самым удачным. Тэги следует вносить, работая в редакторе Блокнот. Каждый раз сохраняйте измененный документ в HTML-формате, а затем – загружайте его в браузер, чтобы проверить полученный результат. Для перехода к работе в редакторе Блокнот можно использовать команду **В виде HTML** меню **Вид** браузера.

6. Далее рекомендуем оформить все внутритекстовые заголовки. Уровень заголовка определяется количеством групп чисел, предшествующих тексту заголовка и разделенных точками. В наших примерах преимущественно используются заголовки 1-го и 2-го, реже – 3-го уровней.

7. Следующий этап заключается в оформлении отдельных абзацев текста. Симпатии автора принадлежат способу выравнивания по ширине, но, если вы сможете обосновать целесообразность других вариантов – пользуйтесь также и ими.

8. Теперь внимательно прочтите текст и выделите с помощью соответствующих тегов шрифтового оформления отдельные его фрагменты. Цель этого пункта – сделать текст более



легко читаемым, облегчить читателю восприятие наиболее существенных его фрагментов. Мы рекомендуем пользоваться для этого тегом <FONT> (см. разд. 2.2) с различными значениями параметров. Допустимо оформление текста различными гарнитурами, но делать это следует очень аккуратно, используя близкие по рисунку гарнитуры. Для выделения важных фрагментов текста мы рекомендуем пользоваться увеличением размера шрифта или полужирным его начертанием.

9. Если в документе имеются какие-либо списки – оформите их должным образом с помощью контейнеров <UL> и </UL> или <OL> и </OL> и тэгов <LI>.

10. Вставьте в текст рисунки, ссылки на которые имеются в тексте. Обязательно добейтесь, чтобы рисунки были видны в тексте и правильно отформатированы. При необходимости еще раз прочтите раздел 2.4, обратив особое внимание на приведенные в нем примеры. Учитывайте, что рисунки находятся в отдельной папке, которая размещается рядом с HTML-документом, из которого производятся ссылки. Таким образом соответствующие контейнеры должны выглядеть так: <IMG SRC=risG1\r1-1.gif>

11. Далее следует включить в текст ссылки на 2—3 рисунка. В качестве указателя гиперссылки используйте упоминание в тексте о рисунках (например, Рис. 1.2. Классификация электронных изданий). От этого указателя организуйте переход к самому рисунку, расположенному в папке. Проверьте, работают ли эти ссылки и выводятся ли рисунки. Учтите, что на вашем компьютере должен быть установлен viewer (программа просмотра рисунков) или один из графических пакетов, например, Photoshop, позволяющий просматривать и редактировать рисунки. Рисунки в форматах GIF или JPEG можно просматривать с помощью браузера. Если у вас что-то не получается, вновь прочтите внимательно разд. 2.5, где рассматривается тэг <A>, с помощью которого создаются гиперссылки.



Рис. 2.17. Пример отображения HTML-документа в браузере

12.Теперь постарайтесь организовать качественную навигацию в пределах вашего документа. Для этого, прежде всего, вынесите все внутритекстовые заголовки в самое начало документа, организовав что-то вроде оглавления. Если заголовки различных уровней, постарайтесь, чтобы это было видно из оформления этих заголовков. Затем установите скрытые закладки перед каждым внутритекстовым заголовком в виде: [[A NAME=p12](#) </A> 1.2. ...]. В качестве примера здесь выбран разд. 1.2. Перед заголовком этого параграфа помещен пустой контейнер, в котором определено имя закладки (p12), на которую в дальнейшем можно будет сослаться. Для того, чтобы создать гиперссылку, ведущую из заголо-

вочной части документа к разд. 1.2 в тексте, следует поместить этот параграф в оглавлении, которое вы недавно создали в начале документа, в контейнер вида: `<A HREF=#p1-2>1.2....</A>`. Теперь строка оглавления будет играть роль указателя гиперссылки, которая приведет к нужному месту в тексте документа.

13.Теперь попробуйте создать переходы из начала раздела в оглавление. Сам автор книги в таких случаях в качестве указателя перехода использует цифровую часть заголовка, а в начале документа делает закладку.

14.Попробуйте создать дополнительные навигационные элементы – переходы из начала документа в его конец и, наоборот, из конца – в начало. Для этого вначале поместите закладки в начале и конце документа, затем в самой верхней его строке (до оглавления) поместите фразу "В конец документа", в самом конце – "В начало документа". Эти фразы используйте как указатели соответствующих гиперссылок.

15.Просмотрите созданный вами гипертекстовый документ в браузере. Проверьте работу всех указателей гиперссылок, в том числе – переходов к рисункам. Пример отображения в браузере MS Internet Explorer HTMLфайла показан на рис. 2.17.

## Работа №2.2. Преобразование набора тематически связанных текстовых документов в сетевой HTML-документ

Работу целесообразно выполнять на нескольких компьютерах, соединенных в локальную сеть, работающую под управлением файлового сервера. В отличие от предыдущей работы здесь из отдельных фрагментов составляется единый документ в HTML-формате, локализованный на сетевом диске. Связь отдельных фрагментов осуществляется одним из 2-х способов (в обоих случаях все отдельные документы или фрагменты сохраняются на сетевом диске):

- # путем организации переходов от конкретного документа к предыдущему и последующему;

- # с помощью отдельного дополнительного документа, содержащего оглавление, в котором упомянут каждый документ, и, кроме того, дополнительных указаний по навигации (помощи).

В первом случае все документы нумеруются по порядку, причем в начале каждого из них предусматривается гиперссылка **Переход к предыдущему**, а в конце – **Переход к следующему**. Все они сохраняются в одной папке на сетевом диске. В результате мы получим единый документ, фрагменты которого могут просматриваться только последовательно, друг за другом.

Второй способ предоставляет пользователю возможность произвольной навигации, т. е. он может переμεжаться между фрагментами единого документа в произвольном порядке. Следует лишь предусмотреть в каждом фрагменте гиперссылку **Возврат в оглавление**.

Рекомендуется выполнить оба варианта, чтобы представить себе и полностью разобраться в отличиях между ними. Еще раз обратим внимание на то, что данную работу целесообразно выполнять группой. Помимо всего прочего, это позволяет оценить результаты и преимущества коллективной работы, возможности которой в наибольшей степени реализуются в интрасетях, краткая характеристика которых дается в главе 7.

Если работа выполняется одним человеком, мы рекомендуем ограничиться первым вариантом, преобразовав в гипертекст 2—3 исходных текстовых документа с последовательными номерами.

## Порядок выполнения работы

1. Скопируйте в одну из папок (например, c:\el-books\lab-22) на жестком магнитном диске вашего ПК папки Исходные тексты и HTML-документы, размещенные в папке К главам 2 и 4 – тексты на гибком магнитном диске, прилагаемом к книге. В первой папке находятся 10 текстовых файлов с именами от 1.txt до 10.txt, представляющие собой фрагменты 1 и 3 глав книги. Во второй папке размещаются вложенные папки с рисунками к этим главам, а именно Ris-g1 и Ris-g3. Выберите для работы любой текстовый документ, например, 2.txt.

2. Выполните все те действия, которые прописаны в пп. 2—15 самостоятельной работы №1.

3. В начало документа 2.HTML вставьте текстовый фрагмент **Переход к предыдущему** и оформите его как указатель гиперссылки на документ 1.HTML.

4. В конец документа 2.HTML вставьте текстовый фрагмент **Переход к следующему** и оформите его как указатель гиперссылки на документ 3.HTML.

5. Повторите все действия п.2 для документов 1.HTML и 3.HTML.

6. Загрузите документ 2.HTML в браузер и проверьте, как работают переходы к предыдущему и последующему документам.

7. Если самостоятельную работу выполняют 10 человек, каждый из которых работает за своим компьютером в сети, то можно проверить на любом рабочем месте доступность всех частей документа. Для этого следует последовательно задействовать переходы к предыдущему документу в каждом из них (или к последующему).

## **Глава 3**

### **Мультимедийный документ и форматы представления его элементов**



Данная глава посвящена рассмотрению состава мультимедийного документа и представления отдельных его компонентов. В частности, описываются различные варианты представления текстовой и графической информации. Указаны наиболее популярные графические форматы, используемые в электронных изданиях. Описаны возможные форматы представления аудиофайлов, особенности кодирования и сжатия соответствующих данных. В заключение анализируются форматы анимационных файлов и цифрового видео и варианты компрессии соответствующих данных.

### 3.1. Мультимедиа и ее роль в современных информационных технологиях

Мультимедиа-издание – это полноценное объединение продуктов современных информационных технологий: текста, графики, видео, аудио, фото, кино и телекоммуникаций (телефон, телевидение, радиосвязь). С помощью мультимедиа-приложений текст, графика, аудио– и видеоинформация объединяются в единое информационное поле, подобно тому, как в кинофильме объединяются звук и "движущееся" изображение. Однако, в отличие от кинофильма мультимедиа представляет собой интерактивную среду, т. е. пользователь может управлять процессом представления мультимедиа с помощью различных средств ввода – таких как клавиатура и манипулятор. Выполняемые сценарии и встраиваемые компоненты "оживляют" документы и заставляют их "реагировать" на действия пользователя. Состав мультимедиадокумента представлен на рис. 3.1.

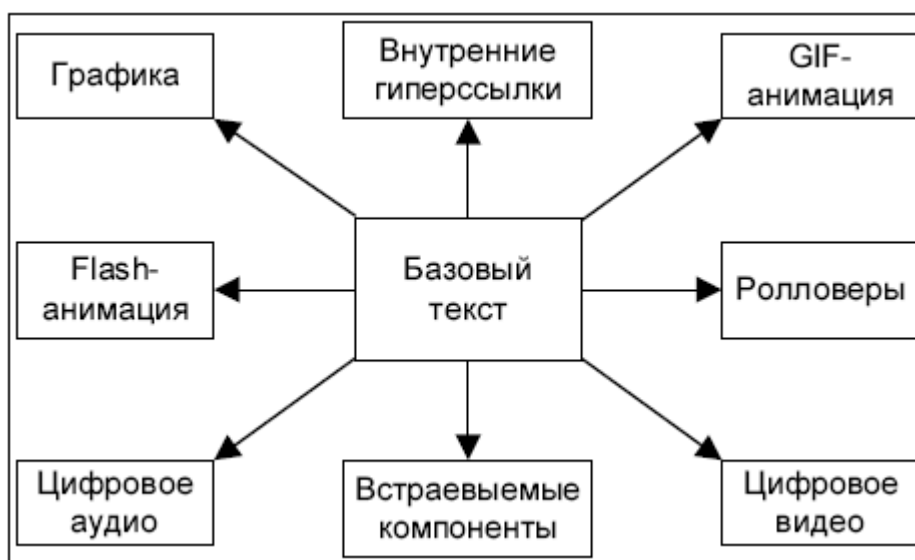


Рис. 3.1. Основные компоненты мультимедийного документа

Помимо базового текста и графики, которые являются стандартными компонентами печатных изданий, в состав публикации мультимедиа входит ряд дополнительных элементов. Прежде всего это анимация. Наряду со стандартной анимацией в формате GIF в последнее время как в сетевых документах, так и в документах, локализованных на носителях, все чаще используется анимация в формате Flash (или Flash-фильмы). Flash-анимация, которая будет подробнее рассмотрена в 6-ой главе, весьма компактна за счет использования векторного формата. В то же время она позволяет создавать множество новых интересных эффектов как с точки зрения графики, так и повышения уровня интерактивности. Появились также интерактивные элементы, основанные на применении Java-скриптов – ролловеры, которые также позволяют резко повысить уровень интерактивности графической среды гипертекстового документа. Средства для создания ролловеров будут рассмотрены в главе. 4.

Замечание

*Ролловер* представляет собой динамический элемент, изменяющий внешний вид, когда на нем оказывается курсор мыши или делается щелчок ею. Ролловеры предназначены для отображения различных состояний графических элементов Web-страницы (как правило, кнопок и ссылок).

Непрерывно повышается уровень сжатия аудио– и видеофайлов при неизменном их качестве. Как уже упоминалось, практически к любому гипертекстовому документу могут быть присоединены такие файлы. Создаются новые программные и аппаратные средства для производства, воспроизведения и передачи таких файлов, в том числе – в сетевой среде с низкой пропускной способностью – что позволяет надеяться на постепенное расширение их практического использования.

Успешное сращивание телекоммуникационных сетей с компьютерами, стремительный рост их качества и количества преобразует вещательные сети в интерактивные, создает единое мировое информационное мультимедиапространство. Важнейшей частью этого пространства является сеть Интернет и, особенно, ее гипермедиа-система World Wide Web. Распространение мультимедиа-технологий (в сочетании с развитием электронной коммерции) в дальнейшем наложит жесткие ограничения на конкурентоспособность издательско-полиграфических фирм, ориентированных на широкий спрос. Преимущества в продаже даже самой высококачественной продукции получают те, кто быстрее и эффективней освоил электронные способы коммерции и обслуживания.

Использование мультимедиа в учебных пособиях дополняет аналитические (вычислительные и логические) и навигационные возможности компьютеров способностью к образному, синтетическому описанию изучаемого предмета или объекта. Многочисленные исследования показали, что обучаемый с первого раза запоминает лишь четверть услышанного и треть увиденного, при комбинированном воздействии на слух и зрение запоминается приблизительно половина информации, а при вовлечении обучаемого еще и в активные действия (например, при использовании интерактивных мультимедиа-технологий доля усвоенного достигает 75%). Мультимедиа, особенно интерактивное, активизирует индивидуальные, личностные мотивы обучаемого (в частности, студента) при усвоении материала, в том числе:

# *целевой* (для меня важно и необходимо знать этот материал и уметь выполнять такую работу);

# *исследовательский* (работая с учебным материалом, я не только узнаю что-то новое, но и чувствую себя активным участником процесса познания, сам участвую в творческом процессе);

# *эмоционально-эстетический* (в процессе изучения материала я испытываю удовольствие, как от получаемых результатов, так и от самого процесса изучения этого материала);

# *игровой* (эта форма обучения интересна, начав изучать материал, я не могу остановиться, мне интересно и хочется довести до конца изучение материала);

# *инициационный* (предполагает органичное сочетание в мультимедийном учебнике информационной и эстетически-эмоциональной глубины).

По уровню творческих мотивов и степени воздействия на человека мультимедиа следует отнести к новому виду синтетического искусства, отличительной особенностью которого является высокая информативность и интерактивность. Поэтому, в будущем следует ожидать создания теории педагогики мультимедиа, учитывающей психофизиологические и эстетические законы восприятия и усвоения большого объема информации. Не исключая традиционной формы обучения, предполагающей творческое и воспитательное общение с преподавателем, мультимедиа создает новые позитивные факторы, в частности, значитель-

ный рост эффективности обучения за счет повышения качества самостоятельной работы обучаемого с электронными учебными материалами.

Специалисты считают, что самую сложную систему автоматизированного управления было бы гораздо легче освоить в том случае, если она реализована на основе стандартного мультимедиа-интерфейса. В будущем, видимо, будут созданы эвристические алгоритмы мультимедиа, которые позволят не только человеку адаптироваться в компьютерной системе, но и компьютеру адаптироваться к уровню восприятия человека, т. е. сделать процесс адаптации двусторонним.

В последующих разделах излагается характеристика и роль отдельных мультимедиа-компонент.



### 3.2. Форматы представления текстовых блоков электронного издания

Еще несколько лет тому назад ответ на поставленный в заголовке данного раздела вопрос был предельно прост: текстовые блоки должны быть в гипертекстовом (HTML) формате или же в формате PDF, так как только эти форматы поддерживали возможность включения в электронное издание мультимедийных компонентов. В настоящее время практически все верстальные пакеты поддерживают не только преобразование издания в формат PDF (или HTML), но и подключение к изданию мультимедиа-компонентов. В частности, в программном пакете PageMaker фирмы Adobe предусмотрена команда **Дополнения | QuickTime Media** в меню **Сервис** (рис. 3.2), которая обеспечивает подключение к электронному документу объекта в универсальном формате QuickTime, позволяющим работать с любой времязависимой информацией, начиная от аудиоданных и кончая фильмами с несколькими видео- и аудиодорожками.

В широко распространенном в нашей стране текстовом редакторе Microsoft Word, начиная с версии Word 97, предусмотрена возможность включения в состав документа не только анимации в формате GIF, но также и видеофильма в формате QuickTime, видеоклипа в формате AVI, клипа мультимедиа. На рис. 3.3 (слева) показано диалоговое окно **Вставка объекта** этого редактора, которое иллюстрирует сделанное утверждение. Кроме того, принятый в этой версии редактора формат DOC стал в полной мере гипертекстовым, так как в нем появилась возможность включать в документ гипертекстовые ссылки как внутреннее, для чего в документе делаются специальные закладки, так и внешние – по URL-адресу любого другого документа. В редакторе добавлена также возможность преобразования исходного документа в формат HTML, а также создания специальных HTML-форм. Часть этих возможностей представлена и усовершенствована в новых версиях редактора – MS Word 2000 и Word 2002. В частности, на рис. 3.3 (справа) показано аналогичное диалоговое окно **Вставка объекта** для этих последних версий редактора Word. Там также имеется возможность вставки в документ видео в формате QuickTime (**QuickTime Movie**) и в формате AVI (**Видеозапись**).

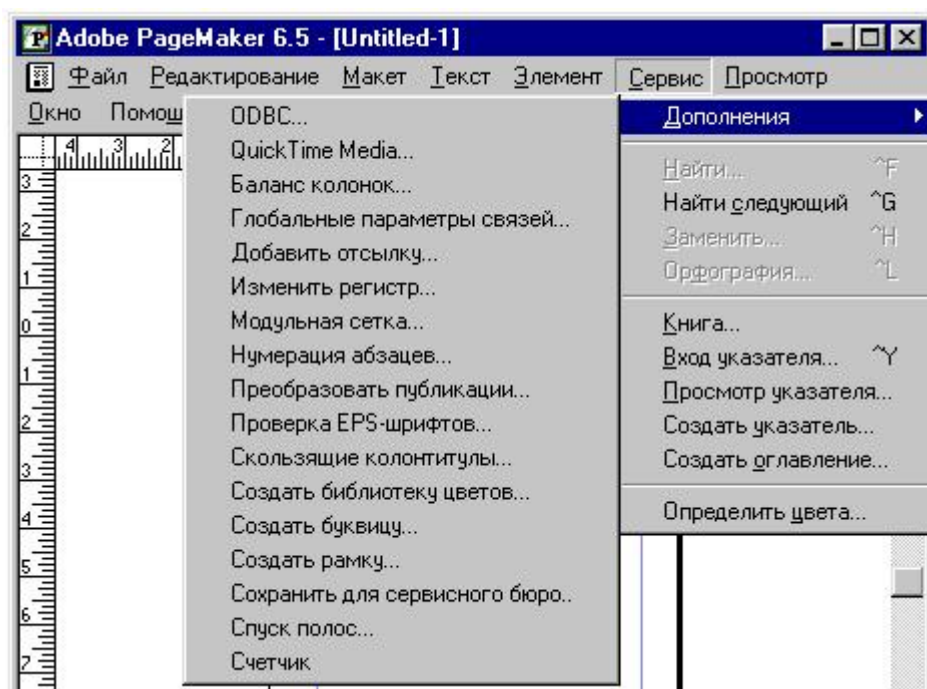


Рис. 3.2. Интерфейс программного пакета PageMaker с раскрытой командой **Дополнения** меню **Сервис**

Таким образом, наряду с форматом HTML и PDF (последний, строго говоря, хранит текст в графическом формате), текстовые блоки электронных изданий могут быть представлены в форматах DOC (MS Word), P65 (для Adobe PageMaker 6.5) и многих других.

В том случае, когда электронное издание не содержит мультимедиакомпонентов, то оно может храниться в формате любого текстового редактора или верстального пакета, единственное дополнительное требование к текстовому редактору состоит в том, что он должен поддерживать графические форматы рисунков, если они включены в текст издания.

Для чисто текстовых изданий ограничений еще меньше. Их можно хранить и распространять в любом текстовом формате, используемом в современных персональных компьютерах. В частности, может использоваться формат TXT (в том числе "простой текст" или plain text), гораздо более экономичный, чем формат DOC. Для кодирования любого символа такого текста используется всего один байт. Пример такой кодировки представляет код ASCII (American Standard Code for Information Interchange, Американский стандартный код для обмена информацией). Для языков на основе латиницы и кириллицы такое кодирование вполне удовлетворительно.

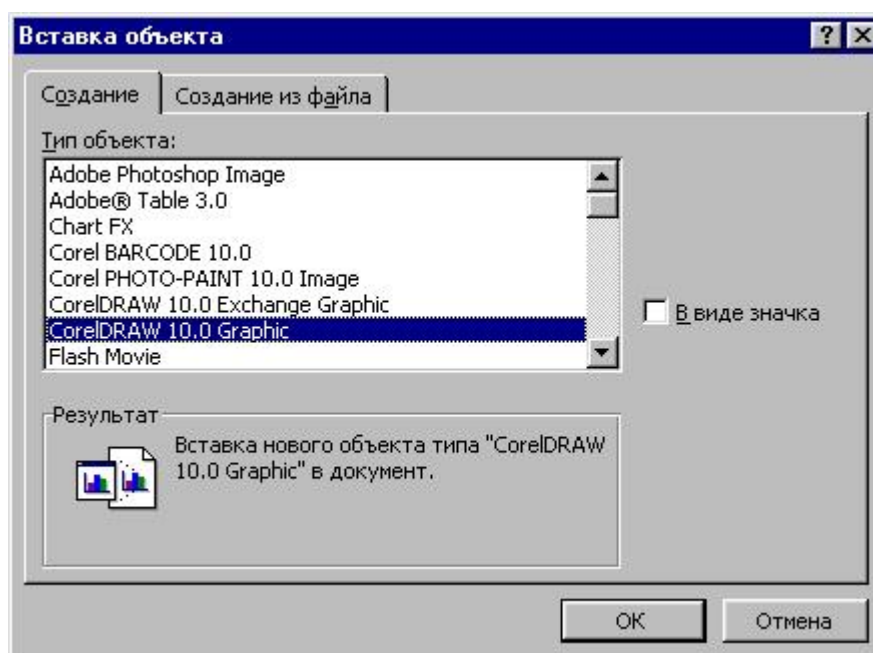
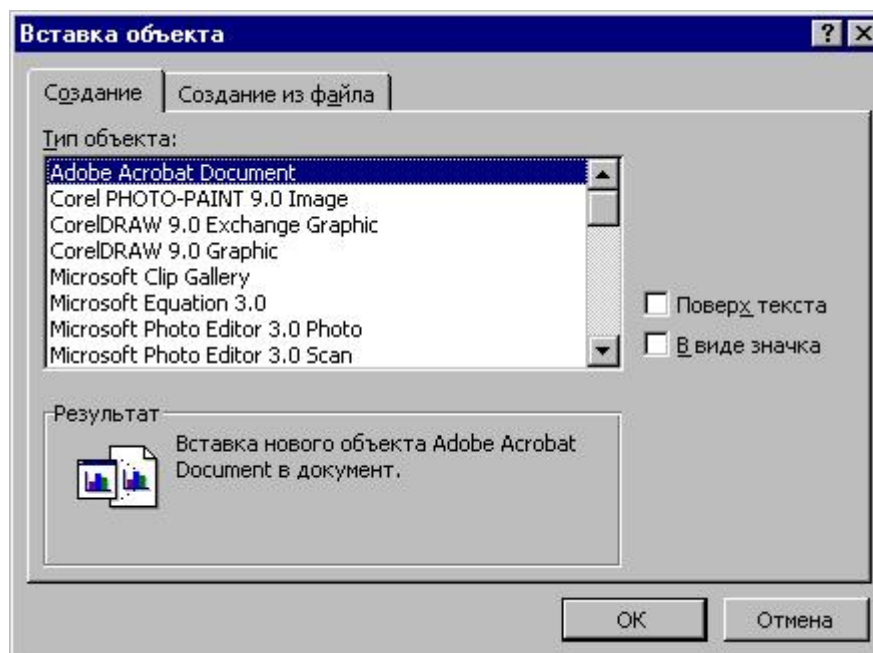


Рис. 3.3. Интерфейс и диалоговое окно Вставка объекта редактора MS Word 2002

Однако, для некоторых восточных языков, например, китайского или японского, такой подход неприменим, так как разнообразие символов в этих языках многократно превышает 256 – предельного значения этого параметра в ASCII-стандарте. В последние годы все более прочные позиции приобретает стандарт Unicode или ISO 10646, т. е. стандарт Международной организации по стандартизации (International Organization for Standardization) под номером 10646. В этом стандарте каждый символ кодируется уже 2 байтами, т. е. предельное разнообразие символов достигает значения 65536. Этот стандарт часто называют стандартом многоязыковой поддержки, так как он позволяет кодировать символы государственных языков всех стран нашей планеты.

Однако электронное издание в TXT-формате не удовлетворяет даже самым скромным эстетическим запросам, так как в нем нет возможности использования не только графики, но даже шрифтов различного начертания, заголовков и подзаголовков, примечаний и других элементов, которые в совокупности называют "разметкой текста" (markup). Из языков разметки текста помимо HTML, рассмотренного в предыдущей главе, наибольшее распространение получили:

- # TROFF, разработанный для оформления документации в рамках операционной системы UNIX и различных ее версий, включая LINUX;

- # TEX, который широко используется для подготовки изданий с большим количеством математических формул;

- # SGML (Standart Generic Markup Language – стандартный обобщенный язык описания документов), разработанный для длительного хранения документов большого объема.

Исходная программа форматирования электронных документов в системе UNIX называлась ROFF (от Run OFF – тиражирование). TROFF означает Typesetting ROFF, т. е. форматирование текста для принтеров с высоким разрешением и фотонаборных устройств. Хотя область распространения этого языка и поддерживающих его программ постепенно сужается, в США он продолжает использоваться для создания электронных отчетов, которые могут одинаково успешно выводиться на терминалах с низким разрешением и распечатываться с полиграфическим качеством.

Система верстки и язык TEX был разработан хорошо известным среди программистов и математиков профессором Дональдом Кнудом для подготовки книг и пособий по математическим дисциплинам. TEX работает на различных аппаратных и программных платформах. Его можно отыскать в сети Интернет и бесплатно перенести на свой компьютер. Имеются и коммерческие версии этого продукта, в частности для платформы Macintosh фирмы Apple. Отметим, что помимо математических книг на этом языке издаются различные академические журналы, в том числе и в нашей стране.

Язык SGML реализует принцип логической разметки текста, который позволяет разграничить содержимое издания и его электронное представление. Именно этим принципом руководствовались специалисты фирмы IBM, создавшие этот язык, который с 1986 года получил статус международного стандарта. Кстати, HTML был создан именно на основе SGML. Основное достоинство языка SGML состоит в его универсальности, независимости от программных средств для его интерпретации. Этот формат может быть конвертирован в форматы TROFF или TEX. Язык изначально создан для производственных нужд, связанных с длительным хранением электронных документов большого объема, таких как описания крупных проектов или их документация.

Большинство из перечисленных текстовых форматов можно встретить в многочисленных "Электронных библиотеках", представленных в сети Интернет, о чем подробнее будет сказано в главе 9. Для ускорения загрузки таких изданий с сайта на компьютер пользователя они нередко представлены в архивированном виде, для чего чаще всего используются программы-архиваторы ARJ, ZIP и RAR, работающие под управлением DOS, а так же WINZIP и WINRAR, предназначенные для работы в оболочке Windows.

Специальные языки разметки страниц в будущем будут активно развиваться. Одна из причин этого связана с автоматизацией извлечения информации из подготовленных электронных изданий. Такая операция обязательно производится для облегчения поиска информации, в том числе – в сети Интернет. Для публикаций, в производстве которых использованы принципы логической разметки и языки разметки страниц, многократно повышается эффективность поиска ключевых слов и выражений, адекватно отражающих содержание этих изданий.

### 3.3. Форматы представления графической информации

Без иллюстраций любое издание выглядит однообразно. Хорошо подобранная и рационально размещенная в издании графика не только улучшает дизайн издания, но и делает его значительно более информативным для читателя, помогают лучше передать его содержание и даже суть. Однако графика требует значительного информационного пространства для своего размещения, с чем связаны основные ограничения ее использования.

#### 3.3.1. Общая характеристика графических форматов

Известны два способа описания компьютерного изображения: точечный (растровый) и векторный (контурный). В первом случае изображение формируется из отдельных точек или пикселей. Векторное изображение состоит из отдельных объектов, ограниченных замкнутыми или незамкнутыми контурами, каждый из которых представляет собой сочетание отдельных отрезков прямых линий (векторов) и кривых линий (дуг окружностей, фрагментов параболических кривых и кубических сплайнов). Каждый такой графический объект можно перемещать, масштабировать, вращать без потери качества изображения и независимо от любых других объектов.

##### Замечание

Сплайн – это гладкая линия, проходящая через заданные точки.

Векторные файлы содержат математическое описание всех элементов изображения, которое используется программой визуализации для их отображения на экране монитора. Таким образом, сам процесс отображения информации требует определенных вычислительных мощностей для преобразования математического описания объектов в растровый формат монитора. Векторная графика характеризуется рядом положительных черт, к числу которых можно отнести:

1. Экономичность хранения изображений, т. е. сравнительно небольшие размеры графических файлов, хранящих изображение в векторном формате.
2. Легкость трансформации и манипулирования отдельными графическими объектами (и всем изображением в целом).
3. Максимальное использование разрешающей способности выводного устройства, с помощью которого осуществляется визуализация цифрового изображения, так как величина разрешения обычно в графическом файле непосредственно не задана.
4. Простота интеграции с текстом, который состоит из отдельных символов, формируемых преимущественно контурным методом (например, элементы TrueType-шрифтов и шрифтовые объекты в формате PostScript).

Простейшие форматы векторного типа реализованы в электронных таблицах, используемых в пакетах Lotus и Excel. Большинство же векторных форматов разработано для хранения чертежей, созданных программами САПР (Систем автоматизированного проектирования). В издательском деле и полиграфии к наиболее распространенным можно отнести формат и язык PostScript, относящийся к группе языков описания страниц PDL (Page Description Language). Язык широко используется для описания сверстанных страниц и их последующего вывода на печать в фотонаборных автоматах и лазерных принтерах. Ряд векторных форматов используются преимущественно как внутренние в графических программ-

ных пакетах, таких как Corel Draw, Adobe Illustrator и пр. Векторный формат используется в так называемых Flash-фильмах, которые все чаще используются вместо анимационных GIF-файлов.

С точки зрения живописности и реалистичности изображения векторная графика имеет весьма ограниченные возможности, поэтому в издательском деле шире используется растровое представление. В случае применения векторной графики определенные трудности возникают и с автоматизацией ввода графического изображения в компьютер или оцифровкой изображения. Сканеры, цифровые фото- и видеокамеры хранят оцифрованное изображение в растровых форматах.

Точечная или растровая графика исторически стала применяться гораздо раньше векторной. К ней можно отнести художественные изображения мозаичного типа: смальта, мозаика и даже вышивка. Таким образом, к растровой графике относят изображения, полученные из мельчайших отдельных элементов, каждый из которых неделим и характеризуется постоянством тона на всем своем протяжении. Такие элементы принято называть пикселями (это понятие мы уже упоминали во 2-ой главе). Каждый такой пиксел формально независим от соседних пикселей, т. е. может иметь различные характеристики: яркость, цветовой тон, насыщенность цвета и прочее.

К **достоинствам** точечной графики можно отнести следующие факторы:

1. Простота и легкость ввода (оцифровки) изображений.
2. Удобство технической реализации вывода информации (на монитор, лазерный или струйный принтер и так далее).
3. Реалистичность изображения.
4. Возможность получения тонких живописных эффектов, таких как туман, тонкие цветовые переходы и нюансы цвета, перспектива изображения, размытость и нерезкость и пр.

Однако и **недостатки** точечной графики существенны. К основным из них относятся:

1. Необходимость точных установок параметров до начала создания графического изображения. Следует задать количество точек на единицу длины изображения, размер изображения по каждой координате, а также глубину цвета – количество бит для представления каждого отдельного пикселя.

1. Большой информационный объем получаемого графического файла, который определяется произведением трех величин: площади изображения, разрешающей способности и глубины цвета в согласованных единицах измерения. Например, максимальное разрешение в пакете PhotoShop составляет 10000 пиксел на дюйм при максимальном значении 30000 пикселей по каждой координате, чему соответствуют размеры файлов до нескольких сотен Мбайт.

3. Неизбежное появление искажений при трансформациях изображения, т. е. когда при повороте и других трансформациях изображения, входящие в его состав горизонтальные и вертикальные линии превращаются в ступенчатые.

Важной характеристикой любого изображения, в частности растрового, является глубина цвета. Самое простое изображение использует 2 уровня серого, т. е. черный и белый цвета. На цветное описание элемента такого изображения (пикселя) требуется лишь 1 бит. Следующий вариант использует множество уровней серого, обычно 256, в результате чего каждый элемент изображения кодируется 1 байтом ( $2^8 = 256$ ). Цветные изображения также могут быть различных типов.

В некоторых графических файлах используют так называемые индексированные цвета. В этом случае количество цветовых оттенков обычно не превышает 256, причем все они хранятся в самом графическом файле в виде палитры цветовых оттенков и каждый возможный цветовой тон в изображении соответствует одному из элементов этой палитры. Общее разнообразие или глубина цвета равна, как и в предыдущем случае, 8 битам или 1

байту. Кстати, индексированные цвета используются и в оболочке Windows в виде встроенной палитры цветов, с которой мы еще не раз встретимся в рамках данной книги.

Наконец, так называемое полноцветное изображение (True color) чаще всего работает в RGB-цветовом пространстве и использует 1 байт на каждый из 3-х основных цветовых компонент (красная, зеленая и синяя), т. е. общая глубина цвета равна 24 битам или 3 байтам. При таком представлении количество различных цветовых оттенков превышает 16 миллионов. В некоторых графических файлах используется даже 48-битная глубина цвета. В этом случае каждый основной цвет представляется 16 битами или 65536 различными уровнями, а общее число различных цветовых оттенков выражается фантастической величиной, превышающей  $2,6 \times 10^{14}$ .

#### Замечание

RGB – это аббревиатура от Red, Green и Blue (красный, зеленый и синий). Цветовая модель RGB чаще всего используется для показа изображений на экране монитора или с помощью широкоформатных проекторов.

### 3.3.2. Сжатие графических данных

Для электронных изданий, зачастую распространяемых по сетям, объем занимаемый файлом представляет собой очень важную характеристику. Понятно поэтому стремление использовать векторные форматы или же специальные растровые форматы с внутренним сжатием информации для представления графических файлов. Кратко остановимся на методах и средствах сжатия изображений.

Существует 2 группы методов сжатия изображений: без потерь и с потерями. В первом случае при распаковке сжатого графического файла полностью восстанавливается вся исходная информация, в том числе, цветовой оттенок каждого отдельного пиксела. Во втором же – часть информации теряется, т. е. изображение становится несколько менее качественным, некоторые мелкие его детали утрачиваются. Во многих случаях это вполне допустимо, так как человеческий глаз различает, в лучшем случае, лишь несколько тысяч оттенков цвета. Кроме того, он не реагирует на мелкие детали изображения (разрешение глаза близко к одной угловой минуте, откуда при нормальном расстоянии до изображения в 25—30 см можно подсчитать величину линейного разрешения глаза, которая близка к 90—100 мкм).

Большинство методов сжатия без потерь основано на варианте группового кодирования RLE (Run-Length Encoding). Идея такого метода заключается в том, что последовательности повторяющихся значений заменяются на пару чисел, первое из которых дает количество повторяющихся значений, а второе – само это значение. В описаниях многоцветных изображений очень часто соседние пикселы характеризуются одними и теми же тоновыми и цветовыми характеристиками, что и обеспечивает эффективность такого сжатия [29].

Схема сжатия без потерь Лемпела-Зива-Велча (LZW) в последние годы используется все шире и шире. Она позволяет работать с данными любого типа, обеспечивая достаточно быстрое сжатие и распаковку данных. Этот алгоритм называют алгоритмом подстановок или алгоритмом сжатия словарного типа. На основе входного потока данных алгоритм формирует словарь данных (его также называют переводной таблицей или таблицей строк). Образцы новых данных сравниваются с записями словаря. Если они там не представлены, то создается новая кодовая фраза. Если строка повторно встречается во входном потоке, то

в выходной поток записывается ссылка на соответствующую строку словаря, которая имеет меньшую величину, чем исходный фрагмент данных. Так реализуется сжатие информации.

Декодирование LZW-данных производится в обратном порядке. Декомпрессор читает код из потока данных и, если этого кода еще нет в словаре, добавляет его туда. Затем этот код переводится в строку, которую он представляет, и заносится в выходной поток несжатых данных. Ряд графических форматов, в том числе и один из базовых – TIFF – используют в современных своих версиях встроенное LZW-сжатие. В частности, этот формат использован для представления рисунков в данной книге. Достоинством этого метода для графических файлов является хорошее сжатие данных для любой глубины представления цвета, начиная со штриховых и кончая полноцветными изображениями. В частности, такое сжатие успешно используется в формате GIF с индексированными цветами (глубина цвета 8 бит). Оба эти формата будут описаны в данном разделе.

В других случаях используется сжатие с регулируемой величиной потерь и переменным коэффициентом сжатия. Чем больше величина потерь, тем больше и коэффициент сжатия. Программа сжатия обычно делит все изображение на блоки размером  $8 \times 8$  пикселей каждый. Уменьшение сжимаемого фрагмента позволяет уменьшить пропорционально квадрату его линейных размеров время обработки, т. е. деление на фрагменты эффективно увеличивает скорость преобразования. Далее к значениям пикселей применяется формула, называемая дискретным косинусным преобразованием. Оно преобразует матрицу пикселей в матрицу значений амплитуд пространственного спектра изображения.

Значения элементов полученной матрицы характеризуют различные составляющие спектра: левый верхний угол результирующей матрицы соответствует самым низким частотам пространственного спектра, а правый нижний – самым высоким. Коэффициент качества преобразования, введенный предварительно пользователем, используется при получении значений элементов матрицы квантования. Чем ниже коэффициент качества, тем большие будут значения у элементов последней матрицы. Далее каждый элемент матрицы амплитуд делится на соответствующий элемент матрицы квантования. Полученные в итоговой матрице значения округляются до ближайшего целого числа. В результате таких операций в правой нижней части итоговой матрицы будет тем больше нулевых элементов, чем ниже заданный пользователем коэффициент качества. Затем программа сжатия кодирует элементы последней матрицы, начиная от левого верхнего по строке до правого нижнего одним из методов кодирования без потерь, причем чем больше нулей в последней матрице, тем меньше окажется информационный объем сжатого файла. Величина коэффициента сжатия для этого метода изменяется в пределах от 10 до 100 в зависимости от заданного значения коэффициента качества.

Декодирование сжатого описанным методом файла изображения начинается с шага обратного преобразования без потерь в результате чего восстанавливается заключительная матрица, в которой содержится ряд нулевых элементов в правой нижней части. Затем значения элементов этой матрицы домножаются на элементы матрицы квантования, хранимой в самом сжатом файле. В результате получим восстановленную матрицу амплитуд пространственного спектра, значения которой отличаются от элементов исходной (они округлялись до целого значения в процессе сжатия с потерями), что и определяет отличия восстановленного изображения от исходного. Далее применяется обратное косинусное преобразование, в результате чего получим восстановленную матрицу значений пикселей, размер которой по-прежнему  $8 \times 8$ . В результате потери высокочастотных составляющих восстановленное изображение будет выглядеть более блеклым и размытым по сравнению с исходным.



### 3.3.3. Описание наиболее распространенных графических форматов

Далее остановимся на основных форматах графических файлов. Самым универсальным из них является TIFF (Tag Image File Format). Спецификация этого формата была представлена фирмой Aldus Corporation в 1986 г. А в 1996 г. была выпущена современная его версия 6.0. Формат предназначался для хранения оцифрованных изображений больших размеров и высокого разрешения. В 1999 г. фирма Aldus вошла в состав фирмы Adobe, известной среди дизайнеров, издателей и полиграфистов всего мира своими программными пакетами для профессиональной работы с растровой и векторной графикой. Формат TIFF подходит как для профессиональной работы художников с графикой, так и для факсимильной связи и передачи изображений большого размера. Формат обладает универсальностью и высокой гибкостью и хранит графические данные в структурированном виде, что позволяет графическим приложениям осуществлять быстрый поиск и загрузку нужных фрагментов изображения.

В нем может использоваться как 24– и 32-битное цветовое представление (True Color), так и представление с индексированными цветами. Он часто применяется как промежуточный при преобразовании из одного графического формата в другой. Однако размер графических файлов, представленных в этом формате, велик, что долгое время препятствовало его использованию в электронных публикациях. Этот недостаток компенсируется, начиная с 5-ой версии формата, применением эффективного встроенного LZWсжатия, о котором было сказано выше. В 6-ой версии формата TIFF стандартное представление данных в RGB-цветовом пространстве дополнено возможностью представления в пространстве CMYK, которое широко используется при выводе графических данных на бумагу или иной носитель для тиражирования оттисков. Кроме того, в этой же версии предусмотрена возможность хранения описания фрагментов изображения в формате JPEG, о котором мы поговорим позднее.

#### Замечание

CMYK – это аббревиатура от Cyan, Magenta, Yellow и Black (голубой, пурпурный, желтый и черный). Именно эти цвета составляют основу модели CMYK, которая предназначается для печати цветоделенных материалов и чаще всего используется для отображения на бумаге цветных изображений в полиграфии.

Общая структура графического файла в формате TIFF показана на рис. 3.4. Как следует из представленных на рисунке данных, в TIFF-файле принципиально может храниться несколько изображений, для каждого из которых предусмотрен свой собственный директорий (или каталог) и своя структура для хранения данных изображения. Следует отметить гибкость этого формата, выражающуюся в том, что директории и данные могут располагаться друг относительно друга 3-мя различными способами:

1. Как показано на рис. 3.4, т. е. вначале – все директории, а затем все данные изображений.
2. Парно: директорий и данные 1-го изображения, затем – то же самое для 2-го и так далее.
3. Вначале все данные изображений, а уже затем – директории этих изображений.

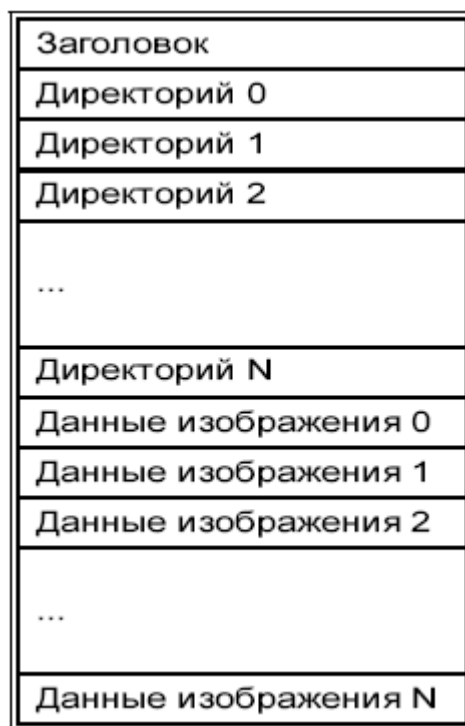


Рис. 3.4. Общая структура TIFF-файлов

Замечание

Используя устаревший термин "директорий" автор следует стандартному описанию формата TIFF, предложенному фирмой Aldus в середине 90-х годов.

Гибкость формата проявляется и в структуре директория для каждого изображения, а также в размещении графических данных, служащих для описания изображений. Эта структура подробнее представлена на рис. 3.5. Заголовок файла содержит информацию, необходимую для идентификации типа файла. Это собственно идентификатор, номер версии и смещение первого директория (IFD0) относительно заголовка файла. В директории изображения (Image File Directory, IFD) содержится несколько структур данных, изначально названных тэгами. В первом из них хранится информации об общем количестве элементов описания изображения – тэгов, а затем размещается каждый такой тэг. В самом конце каждого IFD помещена ссылка на следующий IFD в виде величины его смещения от начала TIFF-файла. Структуры данных, называемые здесь тэгами уже совсем не те тэги, с которыми мы сталкивались в предыдущей главе. Здесь тэг представляет собой элемент данных, служащий для описания изображения. В современных версиях формата TIFF их чаще называют полями. В каждом таком тэге или поле могут содержаться данные изображения или же ссылка на то место в памяти, где они помещены. Все изложенное иллюстрируется рис. 3.5.

Каждый тэг или поле представляет собой стандартную 12-байтовую структуру, где в первых 4 байтах хранится идентификатор тэга (поля) и тип элементов данных. Следующие 4 байта содержат количество элементов данных в этом поле и лишь последние 4 байта отведены для хранения собственно данных этого поля или же смещения того места, где хранятся эти графические данные.

Универсальность формата TIFF проявляется также в том, что данные изображения могут быть организованы как в виде полос, так и в виде фрагментов. Понятие полоса для

растрового изображения соответствует одной строке или последовательно расположенным нескольким строкам раstra. Протяженность полосы измеряется количеством строк раstra, которые в нее входят. Таким образом, полоса представляет собой одномерную структуру данных. Каждая полоса задается в некотором тэге (или поле) с помощью смещения относительно начала файла. Это позволяет программам отображения легко находить в файле нужные данные и отображать их независимо от других данных изображения.

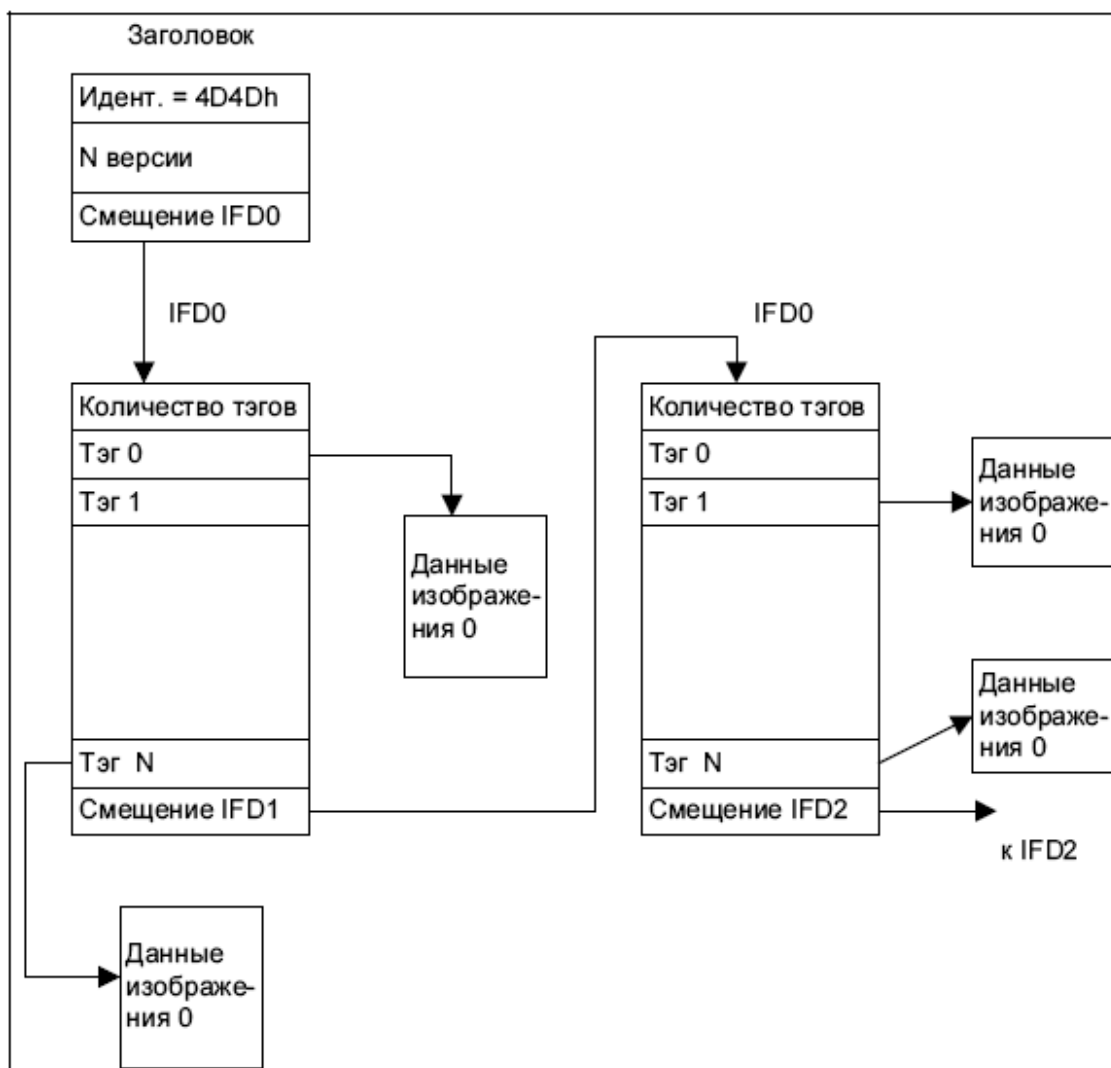


Рис. 3.5. Структура размещения данных в TIFF-формате

Фрагмент представляет собой двумерную структуру данных, в нем задается как горизонтальная протяженность (количество пикселей вдоль строки развертки), так и вертикальный размер, т. е. число строк развертки, которые присутствуют (хотя бы частично) внутри фрагмента. Фрагмент – структура прямоугольная. Это позволяет выводить на экран части изображений в том случае, если разрешение компьютера не позволяет вывести изображение целиком. Индивидуальная адресация фрагментов (в виде величины смещения) позволяет программам легко и быстро находить нужный фрагмент изображения и отображать его на экране.

Нелишним будет также упомянуть, что формат TIFF обладает широкими возможностями развития. Действительно, для того чтобы использовать новый тип данных достаточно определить его в соответствующих тэгах или полях и предусмотреть возможность

правильного распознавания этого типа данных в программах просмотра файлов формата TIFF. Именно так были определены сжатые с помощью алгоритма Лемпела-Зива-Велча графические данные в версии 5.0 и с помощью алгоритма JPEG-сжатия – в формате 6.0.

Формат GIF (Graphic Interchange Format) создан компьютерной информационной службой CompuServe. Он является одним из наиболее употребительных растровых форматов в электронных, в особенности, в сетевых изданиях. Формат изначально был создан для упрощения обмена данными в локальных компьютерных сетях, при возможности отображения этих данных. Основных достоинств у формата 3:

- # пригодность для различных платформ, т. е. формат GIF является платформно-независимым;

- # возможность хранения нескольких изображений в одном файле и их представление в форме анимационного GIF-файла;

- # малый размер файлов благодаря использованию мощного алгоритма сжатия без потерь.

Изображение записывается в этом формате с использованием RGB-цветовой модели и данных встроенной в файл палитры индексированных цветов. К сожалению, серьезным ограничением для этого формата является малая глубина цвета, не превышающая 8 бит на пиксел. Важное достоинство этого формата состоит в том, что он позволяет хранить в одном файле несколько изображений. Современная версия GIF89a решила проблему обработки таких изображений, размещенных в одном файле, с помощью дополнительно включенного в файл блока управления графикой. Этот блок позволяет программе просмотра организовать взаимодействие каждого последующего изображения с текущим, что и обеспечило создание широко распространенных анимационных GIF-файлов. Структура анимационного GIF-файла и взаимодействие хранимых в нем различных изображений подробнее рассмотрено в разд. 3.5.1. Там же мы остановимся на структурной организации файлов этого типа. Графические данные в формате GIF могут храниться как последовательно, строка за строкой, так и с различными вариантами чередования строк. В последнем случае рисунок первоначально может создаваться, например, на основе 1/8 части от общего описания, а в дальнейшем к нему добавляются все новые строки, и изображение постепенно "прорисовывается" все четче и четче.

Растровый формат с глубиной представления цвета до 48 бит и с использованием той же самой RGB-модели создан специальной группой разработчиков и получил название PNG (Portable Network Graphic – переносимый сетевой формат), что произносят как "пинг". Формат PNG изначально планировался как замена формату GIF, но с улучшенными возможностями представления цвета. Он, как и GIF, поддерживает чередование строк и ускоренную начальную загрузку файла. В нем используется еще более эффективный алгоритм сжатия информации. Кроме того, в формате поддерживается режим полупрозрачных корректирующих слоев, аналогичный используемому в векторном AI (Illustrator) и растровом PSD (Photoshop) форматах графических пакетов фирмы Adobe. Единственное ограничение формата PNG по сравнению с GIF состоит в его непригодности для хранения в одном файле нескольких изображений и, вследствие этого, отсутствии анимационных возможностей.

Рассмотренные выше графические форматы содержали внутреннюю компрессию без потерь информации. Еще один формат, который относят к числу наиболее употребительных, характеризуется регулируемой величиной сжатия в зависимости от допустимой потери качества изображения. Этот формат разработан объединенной группой экспертов в области фотографии JPEG (Joint Photographic Experts Group) и назван аббревиатурой JPEG (расширение файлов, созданных в этом формате – JPG). Этот формат также растровый с глубиной цвета, равной 24 битам. Преимущественно используется цветовая модель HSL (Hue-Saturation-Lightness или ОттенокНасыщенность-Яркость). Алгоритм сжатия, используемый

в таких файлах, носит названия "алгоритм сжатия JPEG". Он был описан ранее в этом же разделе в качестве примера технологии сжатия с потерями в графических файлах. Различные его варианты использованы также при организации сжатия видеоданных (см. разд. 3.5).

Графические программы, которые позволяют хранить данные в этом формате, обычно выводят специальную линейку, на которой устанавливаются значения параметра качества, изменяющегося в пределах от 0 до 10 (см. рис. 3.6). Одновременно с непрерывным изменением коэффициента качества на линейке появляется дискретный параметр качества в форме целого числа в поле ввода **Качество**, а рядом в раскрывающемся списке расположена соответствующая характеристика этого параметра. При значении от 0 до 4 качество "Низкое", от 5 до 7 – "Среднее", 8 и 9 – "Высокое" и от 10 до 12 – "Максимально". На рисунке значение коэффициента качества равно 6 и качество "Среднее". При сохранении изображения можно установить переключатель **Разновидность формата** в положение **Progressive**, при котором величина чередования строк устанавливается в пределах от 3 до 5, что обеспечивает быструю начальную загрузку изображения низкого качества в сетевых структурах.

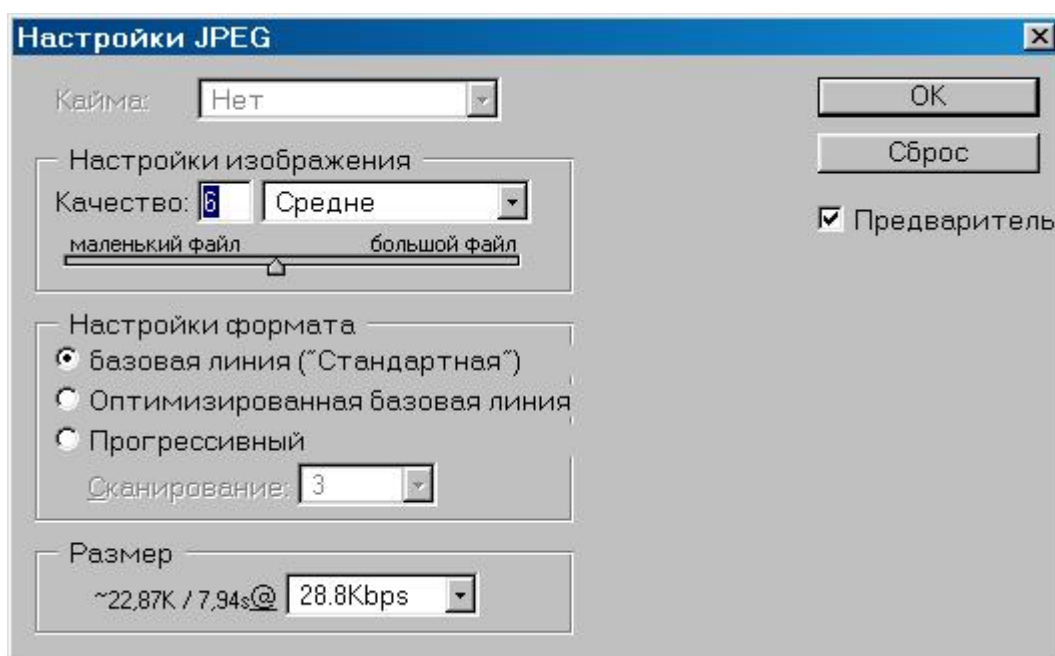


Рис. 3.6. Диалоговое окно для установки параметров файла в формате JPEG

Используемый в формате JPEG подход "сжатие с потерями" частично идентифицирует и удаляет ту информацию, которая несущественна для восприятия изображения. Лишь при сжатии изображения с резко выраженными контурами линии начинают "дрожать". При высоких значениях коэффициента качества изображения этот эффект не проявляется. Возможно, в дальнейшем появится вариант формата с избирательной установкой коэффициента качества для различных фрагментов изображения, что позволит достигнуть высокого качества при очень высоких коэффициентах сжатия.

В заключение этого раздела следует отметить последовательность графических форматов в порядке убывания их популярности (или частоты применения) для электронных изданий и документов: GIF, JPEG, PNG, TIFF.

### 3.4. Форматы представления аудиофайлов

В этом разделе речь пойдет о цифровых форматах файлов, предназначенных для хранения звука. Иначе говоря, – о хранении оцифрованного звука. Напряжение, передаваемое по телефонным каналам и несущее звук, представляет собой аналоговый сигнал достаточно сложной формы. Чтобы преобразовать такой сигнал в цифровую форму необходимо выполнить последовательно две операции: дискретизацию и квантование. Дискретизация состоит в периодическом измерении значений напряжения (на рис. 3.7 дискретизация сигнала производится в моменты времени  $t_1, \dots, t_4, \dots$ ), а квантование – в преобразовании аналоговых значений в дискретный цифровой код. На рисунке показан пример 3-разрядного квантования аналогового сигнала, при котором имеется  $2^3 = 8$  уровней квантования. Соответственно на качество оцифрованного звукового сигнала оказывают влияние 2 фактора: частота дискретизации и разрядность цифрового кода, получаемого при квантовании. При увеличении частоты дискретизации и разрядности кода качество оцифрованного звука улучшается, но пропорционально возрастает объем информации, которая должна храниться в файле. Таким образом необходимо найти компромиссное решение между качеством и размерами файла.

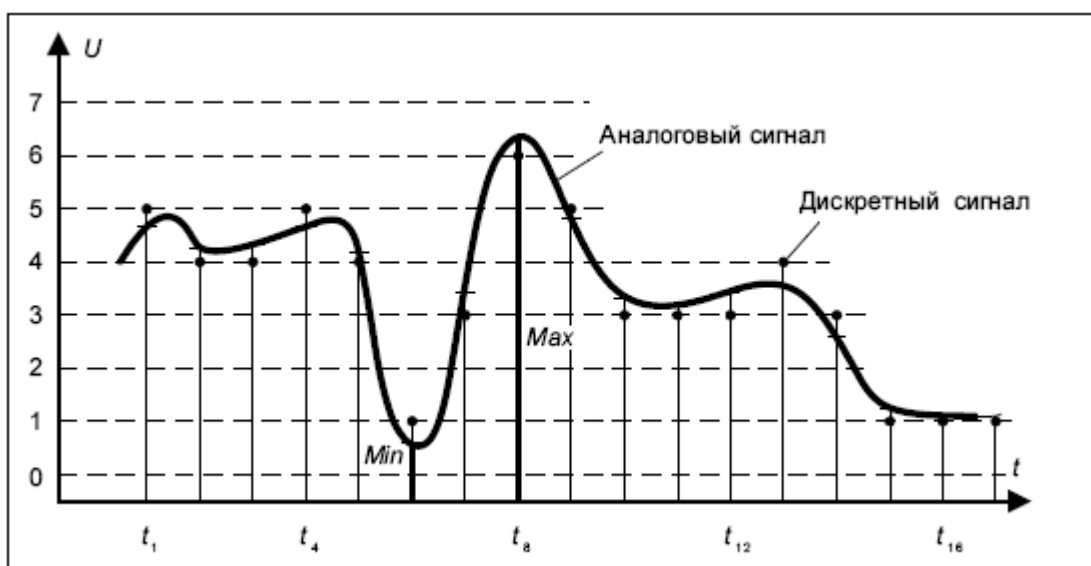


Рис. 3.7. Дискретизация и квантование аналогового звукового сигнала

Частоту дискретизации определить достаточно просто. В соответствии с теоремой Котельникова частота дискретизации должна быть вдвое больше максимальной частоты спектра звукового сигнала. Принятая в настоящее время частота дискретизации для аудио-CD составляет 44100 Гц, т. е. максимальная воспроизводимая частота звукового спектра составляет 20050 Гц, что превышает диапазон звуков, воспринимаемых ухом человека. Это обеспечивает идеальное звучание таких устройств. Частота дискретизации в телефонных сетях составляет 8000 Гц, что более чем в 2 раза превышает полосу пропускания телефонного канала, равную 3000 Гц.

Количество уровней или разрядность квантования характеризует точность передачи уровня звукового сигнала. Действительно, при 256 уровнях квантования или представлении уровня звука с помощью 8 бит информации, величина погрешности квантования равна половине расстояния между соседними уровнями, так как именно с этой точностью значение электрического напряжения может быть преобразовано в цифровой код (при том условии,

что наименьшему уровню сигнала, обозначенному min на рис. 3.7, соответствует цифровой код 000000002 или 00h, а наибольшему, обозначенному max, – 111111112 или 0ffh в 16-ричной форме представления), т. е. отнесено к одному из двух соседних уровней, между которыми находится реальное значение сигнала.

Нетрудно сообразить, что при частоте дискретизации в 44100 Гц и квантовании каждого такого уровня 16 двоичными разрядами (65536 уровней квантования) хранение 1 минуты цифрового аудио потребует около 5 Мбайт информационного пространства, а 30 минут стереозвучания – около 300 Мбайт.

Закодированные описанным способом цифровые аудио-данные характеризуются значительной избыточностью, т. е. они могут быть упакованы, а затем восстановлены без всякой потери качества. Однако применение для сжатия цифрового аудио архиваторов обычного типа, таких как ARJ или ZIP позволяет сжать исходный файл лишь приблизительно на 20%, т. е. такое сжатие является неэффективным.

Основная идея сжатия аудиосигнала с потерями – пренебрежение теми фрагментами звука, которые лежат вне пределов восприятия человеческого уха. Первая такая возможность определяется маскирующим эффектом, в соответствии с которым сильные звуки приводят к невосприимчивости уха к слабым звукам в том же самом частотном диапазоне. Поэтому слабые звуки можно кодировать с малым количеством уровней, в результате чего сокращается количество информации, используемое при кодировании звука.

Далее, весь частотный диапазон делится на поддиапазоны, каждый из которых обрабатывается отдельно, причем маскирующий эффект используется как внутри каждого поддиапазона, так и между ними, т. е. очень мощный звук в одном из поддиапазонов приводит к маскированию во всех остальных. Затем используются особенности психоакустической модели человеческого слуха, в соответствии с которой тщательно сохраняются звуки хорошо воспринимаемых частот и удаляются звуки тех частот, которые не воспринимаются.

Для стереозвучания используется дополнительный прием, связанный с тем, что стереоэффект воспринимается человеком только в области средних звуковых частот. Поэтому сигнал низких и высоких частот передается в монофоническом звучании.

Наконец, используются специальные алгоритмы сжатия, основанные на высокой предсказуемости звукового сигнала, т. е. большом значении его коэффициента автокорреляции. Все перечисленные выше методы и алгоритмы позволяют получить десятикратный и более высокий коэффициент сжатия практически без потери качества звучания, что реализуется в формате MP3, разработанном комитетом MPEG (Motion Picture Expert Group – группа экспертов в области движущихся изображений).

Для воспроизведения звуковых файлов формата MP3 существует целая группа программ-плееров. Список большинства из них можно найти на сайте <http://www.dailymp3.com>. Самые популярные из них это Winamp, которая включена в дистрибутив браузера Netscape Communicator, начиная с версии 4.7, K-Jofol, которая является самым быстрым декодером звука, и NAD (или NADDY), которая является лидером по качеству воспроизведения звука. Другие плееры используются гораздо реже.

Из форматов звуковых файлов следует упомянуть AU (от слова Audio, предложен фирмой Sun Microsystems) для UNIX-подобных систем и платформ, WAV (Microsoft Waveform – стандарт звуковых файлов для операционной системы Windows), AIFF (Audio Interchange File Format – интерактивный формат аудио файлов) – стандарт звуковых файлов для платформы Apple Macintosh и MIDI (Musical Instrument Digital Interface – цифровой интерфейс для музыкальных инструментов) – межплатформный формат электронных музыкальных инструментов. Кратко остановимся на каждом из них.

Формат AU – один из наиболее распространенных в сети Интернет. В заголовочной части файла определяются параметры звуковых данных: частота дискретизации и разряд-

ность квантования, число звуковых каналов и метод кодирования. Используется разновидность этого формата с 16-разрядным квантованием стереозвука и частотами дискретизации 20050 и 44100 Гц с одним или несколькими звуковыми каналами. Наиболее распространенные файлы этого формата носят подзаголовок -Law. Они рассчитаны на один звуковой канал с полосой 8000 Гц. Подзаголовок -Law означает преобразование выбранных в процессе дискретизации значений в двоичный код по логарифмическому закону. Преобразование производится в соответствии с выражением:

$$Y_{\mu} = \text{sign}(m) \ln(1 + \mu|m/m_p|) / \ln(1 + \mu), \quad (3.1)$$

где  $Y_{\mu}$  – значение в логарифмической шкале,  $m$  – исходное квантованное значение,  $m_p$  – максимальная величина последнего значения,  $\mu$  – постоянное значение, величина которого определяет область, в которой обеспечивается наиболее высокое качество звучания.

В формате AU наряду с 8-разрядным логарифмическим кодированием и 16-разрядным линейным кодированием стереозвука, предусмотрена возможность представления 24- и 32-разрядного линейного стереозвука, имеющего частоту дискретизации 22050 и 44100 Гц.

Формат WAV является основным на платформе Windows. Фактически это специальная разновидность файла формата RIFF (Resource Interchange File Format – файловый формат взаимодействия ресурсов), который предназначен для хранения произвольных структурированных данных. Полное название такого формата – WAVE RIFF Microsoft Windows. Звуковые данные в таком файле обычно хранятся в PCM-форме (PCM – Pulse Code Modulation или импульсно-кодовая модуляция). Это означает запись в файле значений квантованного кода в последовательных точках дискретизации. В заголовочной части файла содержится основная информация об оцифрованном звуке, например, число каналов и частота дискретизации, а также среднее число передаваемых в секунду байтов. Последняя характеристика позволяет программе воспроизведения звука выбрать требуемые размеры буфера для хранения звуковых данных. Большинство программ воспроизведения звука буферизуют количество данных, соответствующее одной секунде непрерывного звучания.

Формат WAV поддерживает также ряд дополнительных блоков данных. К ним относят дополнительную информацию о сжатых звуковых данных. В частности, фирма IBM зарегистрировала специальные коды форматирования для сжатия в формате -Law. Специальный блок предназначен для того чтобы пометить определенные позиции в потоке звуковых данных, что позволяет синхронизировать звуковой ряд с видеорядом. Предусмотрены также блоки для размещения дополнительной текстовой информации.

Формат AIFF (Audio Interchange File Format) преимущественно предназначен для работы на платформе Macintosh. Он во многом напоминает WAV, но позволяет, в отличие от последнего, хранить еще и шаблоны, т. е. образцы оцифрованного звука, которые можно использовать как шаблоны для отдельных нот. Специальная версия формата AIFF-C поддерживает сжатие.

MIDI (Musical Instrument Digital Interface) – старейший звуковой формат, который позволил стандартизировать работу с различными электронными музыкальными инструментами. Стандарт базируется на использовании пакетов данных, каждый из которых соответствует определенному MIDI-событию. Эти события можно разделить по каналам. Сложная среда такого файла может включать различную аппаратуру на каждом канале, причем отдельная ее часть будет отвечать за события на определенном канале. Такие файлы хранят не запись оцифрованного звука, а только ноты. В результате они гораздо компактнее других типов звуковых файлов. Недостатком такого формата является то, что он не определяет в явном виде всех тонкостей воспроизведения звука.



## **3.5. Форматы представления анимации и цифрового видео**

### **3.5.1. Анимационные GIF-файлы**

Известно, что анимационные файлы в формате GIF занимают почетное место на Web-страницах. Надо сказать, что в электронных изданиях любого типа анимационные файлы также используются достаточно широко. Это объясняется тем обстоятельством, что GIF-файлы непосредственно воспроизводятся большинством браузеров, причем объем памяти, занимаемый этими файлами, сравнительно невелик. Для учебных электронных изданий анимационные файлы обеспечивают уникальную возможность демонстрации последовательности действий, протекания физических, электрических, технологических и других процессов в простой и наглядной форме. Следует особо отметить, что затрата вычислительных ресурсов на такие демонстрации – минимальна.

Искусственный мир компьютерной анимации лежит где-то посередине между миром неподвижных изображений, форматы которых были рассмотрены в разд. 3.3, и реальным миром видеоизображений. Обычные мультипликационные фильмы состоят из множества рисованных изображений-кадров, в которых последовательно изменяются позиции объекта анимации. В результате, при отображении с достаточной скоростью такой последовательности изображений у зрителя возникает впечатление движения объектов.

Возможности GIF-анимации связаны с тем, что этот формат позволяет хранить в одном файле несколько различных изображений. Единственный существенный недостаток GIF-файлов связан с применением индексированных цветов, для чего в файле используется глобальная и локальные цветовые палитры. Глобальная цветовая палитра хранит до 256 различных цветовых оттенков, каждый из которых может быть использован в любом из изображений, которое хранится в данном файле. Локальные палитры относятся к каждому отдельному изображению, т. е. хранимые в них цветовые оттенки не могут использоваться в других (не своих) изображениях. Каждое такое изображение формирует отдельный кадр, причем задержка следующего кадра и его линейное смещение относительно предыдущего по каждой координате может регулироваться. Разрешение для всех изображений, входящих в данный файл, или количество пикселей по каждой координате должно в таком файле поддерживаться постоянным.

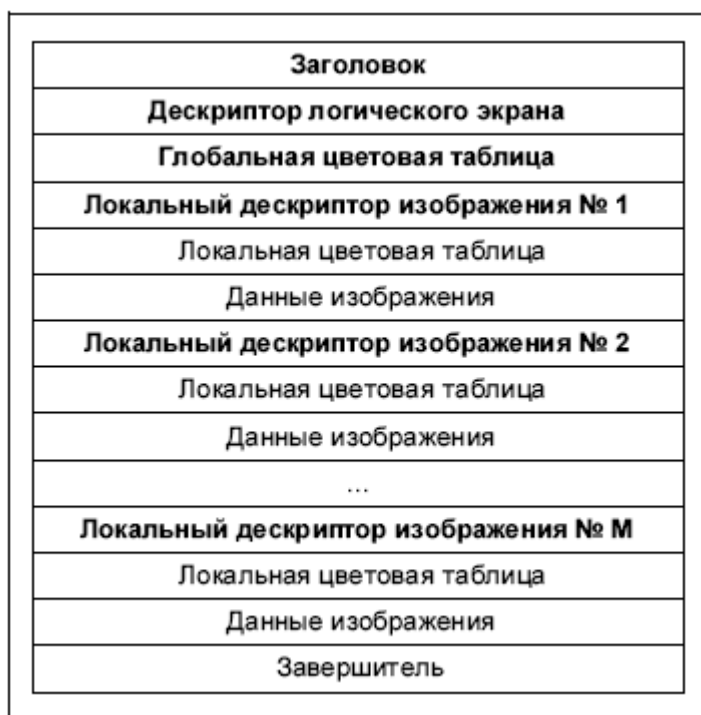


Рис. 3.8. Структура анимационного файла формата GIF 87a

Структура файлового формата GIF 87a представлена на рис. 3.8. Файл начинается с общего заголовка и дескриптора логического экрана, причем в последнем хранится ширина и высота каждого изображения в пикселах, индекс цвета фона и значение коэффициента сжатия. Там же задается размер глобальной цветовой таблицы, которая может и отсутствовать. В этом случае обязательно используется для каждого отдельного изображения локальная палитра. В большинстве случаев рекомендуется пользоваться именно глобальной палитрой, что экономит общее информационное пространство, занимаемое файлом.

После указанных трех элементов следуют наборы данных, характеризующие каждое из входящих в файл изображений. Каждое изображение в отдельности описывается локальным дескриптором и локальной цветовой палитрой, после которых следуют данные изображения. Данные обычно состоят из последовательностей пакетов данных, называемых блоками, причем в состав отдельных блоков могут входить и подблоки.

## **Конец ознакомительного фрагмента.**

Текст предоставлен ООО «ЛитРес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на ЛитРес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.