

情報基礎実習 第7回 画像ファイルと Web の仕組み

2011年6月9日, 6月10日

担当教員: 逸村裕

TF: 佐藤翔

1. 本日のメニュー

画像ファイルの扱い方に慣れてもらい、そのあとで Web の仕組みを理解してもらう。

- ・ 画像ファイルを加工して、見栄えやファイルの大きさを比べる (ペイント)
- ・ ファイルとフォルダの操作を体得する
- ・ 簡単な HTML 文書を作って Web で公開する (メモ帳)

課題 3種類の画像 (PICT0003.bmp, scanin.bmp, 画面のコピー) と 4種類の画像形式 (BMP, JPG, PNG, GIF) について、画像の品質 (見栄え) とデータサイズを比較し、画像データについての**実験レポート** (試したこと・調べたこと・考察・感想やコメントなど) を Form.doc を利用して画像入りの Word 文書にまとめなさい。片面印刷とし、ステープラで左上部をとめて、次回の実習開始時に提出しなさい。以下の点に注意すること:

- (1) 結果の画像があるか (モノクロでじゅうぶん)
- (2) 具体的な数値が示してあるか
- (3) 考察はあるか (たとえば、どのような場合にどの画像形式を使えばよいか?)
- (4) 単なる感想文になっていないか

Word 文書に画像を入れるには、「挿入」タブの「図」をクリック、挿入したい画像ファイルを選択する。画像データ形式によって、品質が劣化することがある。画像の大きさは、挿入した画像の上で右クリック、「図の書式設定」で変更する。画像の左右位置は、画像を選択してから、「ホーム」タブの「段落」で、文字列の左・右・中央揃えするときと同様に操作する。

(出席)単純な HTML ファイルを <http://www.u.tsukuba.ac.jp/~s1111987/first.html> から公開し、そのページのプリントアウトを TA に提出しなさい (手書きでよいので、学籍番号と氏名を記入すること。「1111987」の部分は自分の学籍番号の下7桁とする)。うまくできていた場合にのみ TA が受け取ってくれる。

2. 画像ファイルの扱い

画像ファイルの大きさは

たてのドット数×横のドット数×色数×p

で決まる。色数は Red, Green, Blue それぞれが 8 bit の場合を TrueColor といい、

$2^{\{24\}}$, すなわち 1600 万色が表せる。200 万画素のデジタルカメラだと、画像の大きさが 1600×1200 、1つの画素 (pixel) に 3 Byte (24 bit) を使っているのでデータサイズは約 5.7 MB になる。 $0 \leq p \leq 1$ は圧縮率で、規則的な繰り返しなどを符号化して記録することでデータサイズを小さくできる。圧縮・展開には手間 (演算時間) がかかるが、一般にはデータを格納するスペースや通信する時間を節約できるメリットのほうが大きい。

画像 (静止画、動画) には色々なデータ形式があり、用途に合わせて使い分けられている。代表的な画像ファイル形式には、24bit ビットマップ圧縮なし ($p=1$) の BMP 形式 (Microsoft Windows Bitmap)、圧縮あり ($p<1$) の JPEG 形式 (Joint Photographic Experts Group format) や PNG 形式 (Portable Network Graphics) がある。いっぽう、1画素を 256 色 (1 Byte) で表現するのが GIF 形式 (Graphics Interchange Format) であり、データ量が TrueColor の 3分の1になる。 p の値は、データの形式、データ内容、色数によって異なるので予測は難しい。

データの圧縮には可逆圧縮と非可逆圧縮があり、JPEG のような非可逆圧縮の場合は、完全な圧縮前の状態に復元できない。非可逆圧縮の場合、画像の修正・保存を繰り返すと画質は劣化する (アナログ方式の VHS ビデオを何度もコピーするような感覚)。いっぽう、PNG や GIF は可逆圧縮である。これらの画像データ形式は Internet Explorer などの標準的なインターネットブラウザで表示できる (特殊なデータ形式には Plug in と呼ばれるソフトウェアが必要)。

たとえば、<http://klis.tsukuba.ac.jp/jk11/> の Pictures にある PICT0003.bmp は幅 1600、高さ 1200 の BMP 形式の画像であり、ファイルサイズは約 5.7 MB である。これを JPEG 形式にすると約 250KB ($p=0.04$)、PNG 形式にすると約 4.2 MB ($p=0.73$) になる。教育用計算機システムのパソコンの画面は幅 1280、高さ 1024 なので、画面のハードコピーは BMP 形式なら約 3.7 MB になる (計算してみよ)。実際にハードコピーを作り、JPEG 形式にすると 48KB、PNG 形式にすると 35KB になった (画面に何が表示されているかで異なる)。このように、画像の内容によって圧縮率 p は変化する。

3. ペイントを用いた画像の加工とデータ形式の変換

(1) <http://klis.tsukuba.ac.jp/jk11/> の Pictures から、ファイル PICT0003.bmp をダウンロードしてデスクトップに保存する。ふつうにクリックすれば画像が表示される (試してみよ!) ので、右クリックして、「名前をつけて画像を保存」を選択する。ファイルの名前は変えてもよい。

(2) 画像左上に写っている指を取り去って、建物の看板と花水木 (Dogwood) の部分を取りだしたい。以下の操作を組み合わせると目的の画像を作りなさい。

- ・スタート、「すべてのプログラム」⇒「Accessories」フォルダの中から「ペイント」を選択、起動する
- ・ファイル(F)、開く(O) とし、デスクトップから PICT0003.bmp を選択し、開く(O)
- ・変形(I) にあるサイズ変更/傾斜(S) を用いてサイズを約半分にする
- ・変形(I) にある反転と回転(F) を用いて取り去る部分を右側に移動する
- ・変形(I) にあるキャンパスの色とサイズ(A) を用いて不要な部分を取り除く (マウスを動

- かすと画像の上をペンのアイコンが移動する。このときの座標が下端に表示される)
- (3) できあがった画像を BMP 形式のまま保存する。ファイル(F), 上書き保存(S)
 - (4) 画像を JPG 形式で保存する。ファイル(F), 名前を付けて保存(A) でファイルの種類(T) を JPEG とする。同じファイル名でよい。
 - (5) いったん保存した BMP 形式の画像を再び読み込む。
 - (6) この画像を PNG 形式で保存する。
 - (7) 保存してある BMP 形式の画像を読み込む。
 - (8) この画像を GIF 形式で保存する。
 - (9) ペイントを終了する。ファイル(F), ペイントの終了(X) または右上の×印

これで、PICT0003.bmp, PICT0003.JPG, PICT0003.PNG, PICT0003.GIF という4つの画像ファイルができた。これらの画像ファイル (のアイコン) をクリックすると、それぞれに対応したアプリケーションが起動し、それぞれの画像とファイル名 (左上を見よ) が表示される。拡大などを使って、4つのファイルの見栄え (画像の品質) とプロパティ (サイズ、幅、高さ、深さなど) の関係を表にまとめてみよう。スキャナで読み込んだ画像ファイル scanin.bmp やあなたの Windows 画面のハードコピーではどうなるか?

Windows 画面全体のハードコピーをとるには、PrtScreen キーを押し (画面がクリップボードという場所にコピーされる)、ペイントを起動して編集(E)、貼り付け(P) とする。特定のウインドウの画像をとるには Alt キーを押しながら PrtScreen キーを押して同様にすればよい。

4. ファイルとフォルダ

コンピュータでは、情報を“ファイル”に格納して扱う。Word で作った文章 (文字情報)、ペイントで描いた絵 (画像情報) もそれぞれファイルとして保存される。また、Word やペイントというアプリケーションのプログラムもファイルに保存されている。ファイルは名前によって識別され、この名前を“ファイル名”という。

複数のファイルをまとめて保存するために、コンピュータ上には“フォルダ”(ディレクトリということもある) が用意されている。ファイルを紙や写真とすれば、フォルダはそれらを整理してしまっておく箱や引き出しのようなものである (詳しいことは授業で!)。フォルダの中にフォルダを入れてもよい (箱の中に箱!)。フォルダの名前を“フォルダ名”という。

コンピュータでは、名前 (表札、ラベル) と内容 (住人、箱の中身) を区別して考えることが重要である。また、ファイル名やフォルダ名は、コンピュータや人間(ユーザ)がデータを識別するのに用いる。中にどのような情報が入っているのかがわかるよう、名前の付け方には十分注意しよう。

ファイルの中身の種類(属性、型)をプロパティ(性質)という。ファイル名、ファイルサイズ(大きさ)などもプロパティのひとつである。プロパティを調べるには、アイコン(後述)の上で右クリックする。

4.1 ファイルの拡張子とアイコン

Word で作成した文書ファイルのファイル名の末尾は “.docx” または “.doc” という文字列になっているし、画像ファイルのファイル名の末尾は “.bmp” とか “.JPG” になっている。このようなファイル名の末尾の文字列 doc や bmp をファイルの拡張子という。ファイルの中味は 0-1 の列 (bit 列) であり、それがどのようなフォーマット (文字なのか、画像なのか、あるいは音声なのかなど) で記号化されているかは拡張子に記述される。コンピュータは、拡張子に記述されたルールでファイルの内容を識別し、そのファイルの内容にふさわしいアプリケーションを起動する。コンピュータの知らない拡張子や拡張子のないファイルに対しては起動すべきアプリケーションが判断できないため、ユーザは起動すべきアプリケーションを指示する必要がある。認知されていない拡張子のデータ形式とか、文書だとはわかっているけど日本語もあれば中国語もあるなど、拡張子だけでは決定できない問題もある。

Windows の画面上 (正確には Windows が動作しているコンピュータの画面上) には、ファイル名を表す文字列と小さな絵が表示されている。この絵をアイコン (icon) という (知識情報・図書館学類の学生なら、icon の原語についても知っていてほしい!)。拡張子によってどのようなアイコンになるかは決まっている。アイコンを見ればおおよそ拡張子がわかることから、画面上に拡張子が表示されないこともある。

拡張子を変更しても名前が変わるだけで、ファイルの中身は変わらない (ラベルを貼り替えただけ!)。文書ファイル memo.doc を memo.bmp にしても、名前が変わっただけで、中身が画像になるわけではない。多くの場合、名前を変えるだけでなく、アプリケーションを用いて中身を変換する必要がある (なんか暗的?)。

Windows 画面上にファイルの拡張子を表示させるには、どこかのフォルダで

1. フォルダの整理タブからフォルダと検索のオプションを選択する
2. [表示]タブを選択する
3. 詳細設定の一覧中の[登録されているファイルの拡張子は表示しない]という項目のチェックを外す。
4. [適用(A)]、[OK]をクリックする

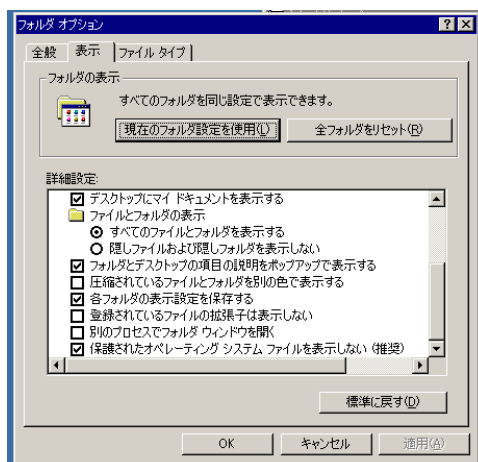


図 1 : 拡張子の表示方法

4.2 ファイルとフォルダの操作

大事なデータはこまめに保存することが重要であり、ある時点でのファイルの内容を別のファイルにしまっておく（スナップショットとかバックアップという）ことがよく行われる。これらを整然と保つには、ファイルをフォルダにまとめて置いたり、後で内容がわかりやすいように名前を付け替えたりといった作業が必要になる。フォルダとファイルを合わせてオブジェクトということもある。

ファイルやフォルダは、アイコンとしてディスプレイに表示されている。Windows にログインしたとき画面に表示されているのは「デスクトップ」という特別なフォルダである。フォルダの中身を表示するには、そのフォルダをダブルクリックすればよい。どのファイルがどのフォルダにあるとかといった、ファイルやフォルダの構成を表示するにはエクスプローラを用いる。エクスプローラを表示するには、フォルダのアイコンで右クリックして、エクスプローラ(X) を選択すればよい。ファイルやフォルダについての情報の表示形式を変えるには、表示タブで、特大～小アイコン、並べて表示、一覧、詳細などをクリックすればよい。

以下の操作をマスターしなさい（実習用手引きを参照）：

- ① ファイルのコピーを作成する
- ② ファイルの名称を変更する（ファイル名は半角英数字を使うこと）
- ③ フォルダを作成する（フォルダ名は半角英数字を使うこと）
- ④ ファイルをフォルダ内に移動する
- ⑤ フォルダのコピーを作成する
- ⑥ ファイルを削除し、ゴミ箱の内容を確認する
- ⑦ 複数のオブジェクト、あるいはすべてのオブジェクトを選択する
- ⑧ 選択したオブジェクトまとめて削除（または移動）する
- ⑨ 画像ファイルを Word で開く
- ⑩ Word で作成した文書ファイルをペイントで開く

5. Web ページ公開のための手続き

Web ページを公開するためにはひとつ事務的な手続きをしなければならない。まずはブラウザで <http://www.u.tsukuba.ac.jp/user/user.html> にアクセスし（パスワード必要）、「Web サーバ利用申請」をクリックする。出てきた内容をよく読んで「同意」にチェックし次へ進む（「同意」しないと演習が進まないの、細かいことはおいて、「同意」することにしよう）。次のような画面で、「Web サーバの利用を申請します」にチェックを入れ、次の画面で「上記の通り申請します」にチェックを入れる。

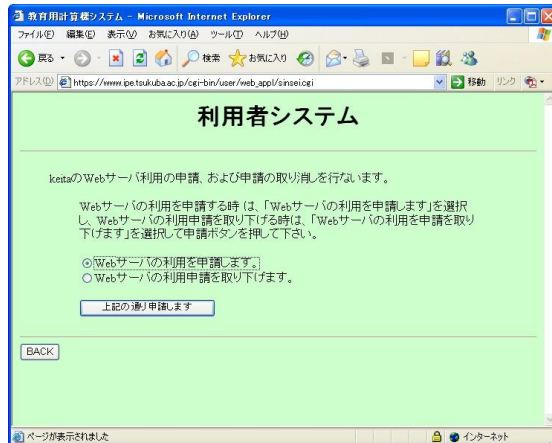


図 2 : Web サーバの利用申請画面

Web ページとして公開するには、ファイルを Web ページ用のスペース `www` に置けばよい。教育用計算機システムの場合は、コンピュータを開いて「ネットワークの場所」の自分の学籍番号がついたドライブ（例えば `s1111987(Z:)` など）に `www` という名前のフォルダを作成し、ここにファイルを保存するだけでよい。`www` フォルダは目の前のパソコンと Web サーバとが共有している特別なフォルダになる。

実習で作成した画像ファイル `abcd.bmp` をペイントで開いて、「名前をつけて保存」で `www` フォルダに保存しよう。たとえば、`xyz.bmp` として保存したのなら、ブラウザ Internet Explorer の URL に：

<http://www.u.tsukuba.ac.jp/~s1111987/xyz.bmp>

と指定する（半角文字で入力すること。“~” はキーボード右上にある）。大文字・小文字の違いに注意すること。正確に指定できれば、Web ブラウザの画面に画像が表示されるはずである。念のため、友人の PC の画面でも見えることを確認してもらおう。友人に見えるなら、その Web ページは世界中から見えるようになっている。同じ結果でも、自分以外からも見えるというのが大違いである。

これで Web サーバの仕組みがわかった。とても単純だが、基本はこれだけである。

注意しなければならないのは、Web ページは世界へ公開するため、URL ^{*用語説明} を構成する文字は半角英数字（`a, b, c, 1, 2, 3` など）を用いなければならないことである。全角文字（2 バイトコード。`a, b, c, 1, 2, 3, あ, い, う`）を使ってはいけない。したがって、「山田太郎作.bmp」のような指定をしてはいけない。

6. 簡単な HTML 文書を作る

HTML（Hyper Text Markup Language）を使えば、文書、画像、他のファイルへのリンクなどを組み込んだ Web ページが作成できる。ここでは、単純な HTML ファイルを作成しよう（単純でないものは次週）。

メモ帳を使って次のような内容の文書ファイルを作成しなさい。メモ帳は Note Pad とも言う。<html> などはいずれも 1 バイトコード（半角）で入力すること。

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">
<html lang="ja">
<head>
<title>はじめての HTML</title>
</head>
<body>
情報基礎実習は簡単だ。
<hr>
もっと難しいことがやりたい。
</body>
</html>
```

ファイル名を半角英数字 (1 バイトコード) で first.html、ファイルの種類(T) は「すべてのファイル」、文字コード(E)は「ANSI」として www フォルダに保存しなさい。Windows ではどのようなアイコンが表示されるか? Internet Explorer の URL を入力する欄に <http://www.u.tsukuba.ac.jp/~s1111987/first.html> と入力すれば、このファイルの内容がブラウザをとおして表示される。

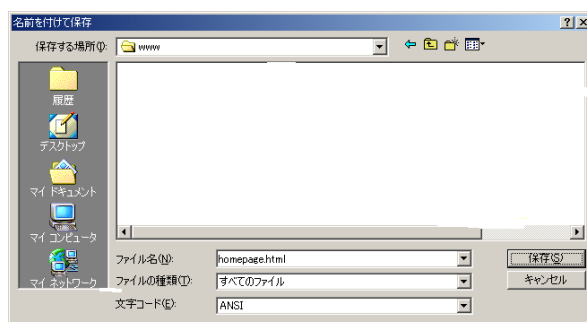


図 3 : www フォルダへのファイルの保存

自分の PC 上なら、ブラウザの URL 欄に Z:\www と入力し、「参照」でアクセスもできる。これで見えてもインターネット上で見るとは限らないので注意がいる。ただ、HTML のデバッグにはなる。時間に余裕があるなら、次ページの HTML マークアップ要素を試してみよう！

(重要な注意): これだけは守ること。あなたの「常識」期待されています！

- ・ 氏名、住所、電話番号、プロフィールなどの個人情報を不用意に公開しないこと
- ・ よそから無断で拝借してきた画像や文書を使わないこと
- ・ 中傷、外部に対する無責任な発言を慎むこと
- ・ iPhone 等、位置情報を記録する機器で撮影した画像の扱いには注意すること
(撮影場所の位置情報がプロパティに記録されている場合がある)

* 様々な HTML マークアップ要素の例（詳細は次回）

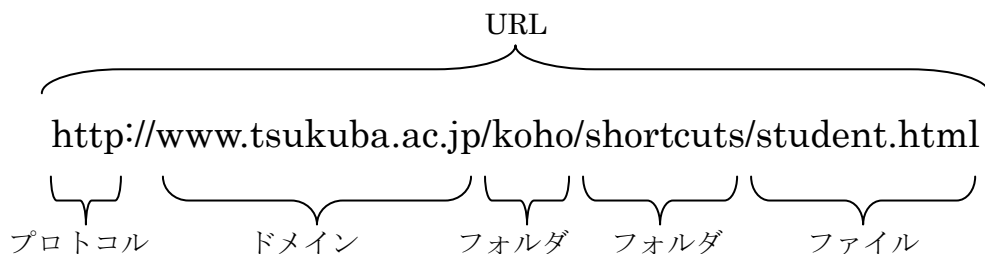
Internet Explorer のようなブラウザは制御語（HTML マークアップ要素）が入ったファイルを読み込むと、それを HTML (Hyper Text Markup Language) とみなして整形してくれる。HTML マークアップ要素は大文字・小文字を問わないが、半角英数字で入力すること（決して全角文字で入力してはならない）。メモ帳を用いて、以下のマークアップ要素をいくつか追加してはファイルに保存し、Internet Explorer でファイルを表示するという操作を繰り返し、マークアップ要素の働きを確かめなさい。

<code><html> </html></code>	HTML 文書であることを宣言
<code><title> </title></code>	文書タイトル（ブラウザのウインドウに表示される）
<code><head> </head></code>	文書の前文
<code><body> </body></code>	文書の本体
<code>
</code>	改行 ¹
<code><p></code>	段落分け。次の段落との間に空行が入る
<code><hr></code>	罫線
<code> </code>	ボールド（太字体）
<code><i> </i></code>	イタリック（斜体）
<code><tt> </tt></code>	タイプライタスタイル
<code> </code>	強調スタイル
<code><h1> </h1></code>	レベル 1 ヘッディング（文字のサイズに注意）
<code><h2> </h2></code>	レベル 2 ヘッディング
<code> </code>	番号付きリスト（箇条書き）
<code> </code>	番号なしリスト
<code><dl> </dl></code>	用語解説用の定義リスト
<code><dt> </dt></code>	定義項目
<code><dd> </dd></code>	定義記述
<code></code>	<code> </code> の項目
<code><pre> </pre></code>	プリフォーマットテキスト（そのままの表示）

¹ `
` で改行できるが、そのようにして「1行」を決めたとしても、他の人の画面でちゃんと「1行」になる保証はないことに注意。画面が小さい人、ブラウザの画面サイズを小さくしている人のところでは複数行になってしまう。

* 用語説明

- URL (Uniform Resource Locator) : インターネット上に存在する情報資源 (文書や画像など) につける識別子。インターネットにおける情報の住所のようなもの。



- プロトコル : ネットワークを介してコンピュータ同士が通信を行う上で、相互に決められた約束事。通信手順、通信規約などと呼ばれることもある。例えば人間同士が意思疎通を行なう場合には、どの言語を使うか (日本語か英語か)、どんな媒体を使って伝達するか (電話か手紙か)、といった取り決めが暗黙のうちに成されているが、コンピュータ通信においてそれを決めたものがプロトコルである。プロトコルとしては `http`、`ftp`、`smtp`、`pop` などがある。
- ドメイン : インターネット上に存在するコンピュータやネットワークにつけられる識別子。右端の `jp` や `com` をトップレベルドメイン、右から2番目の `ac` や `co` をセカンドレベルドメインという。
- Web ページ : WWW システムを使ってインターネット上で公開されている文書。Web ブラウザに一度に表示されるデータのまとまりで、テキストデータや HTML によるレイアウト情報、文書中に埋め込まれた画像や音声、動画などから構成される。たとえば、知識情報・図書館学類の概要が記された `http://klis.tsukuba.ac.jp/index.html` などは Web ページである。
- Web サイト : 1 冊の本のように、ひとまとまりに公開されている Web ページ群。例えば筑波大学のページ群 `http://www.tsukuba.ac.jp/....` などである。
- トップページ : Web サイトの入り口の Web ページ。筑波大学で言うと `http://www.tsukuba.ac.jp/index.html` がトップページに当たる。
- スタートページ : ブラウザを起動した時最初に表示されるページ。ただし Internet Explorer では昔の名残か「ホームページ」という名称で設定するようになっている。
- ホームページ : 元々はブラウザ起動時に最初に表示されるページの意味だったが、現在では様々な意味で用いられている。トップページや、Web サイト全体、1 枚の Web ページなどの意味で用いられることもある。
- Plug in : 特別な形式のデータを処理するための補足的なプログラム