
基礎演習テキスト

椋山女学園大学現代マネジメント学部 三木 邦弘
令和4年9月27日版

1	はじめに	1	4.6	プレートの作成	25
2	Linux (コンピュータ管理ソフト)	3	5	3D プリンター	26
2.1	ユーザー登録	3	5.1	3D プリンターの使い方	26
2.2	リモートデスクトップによる接続	3	5.2	3D プリンター使用上のヒント	29
2.3	パスワードの変更	6	6	拡張現実	30
2.4	インストール済みのソフト	6	6.1	拡張現実 (Argumented Reality) とは何か	30
3	文書清書システム	7	6.2	AR テストシステムの使い方	30
3.1	Texmaker の起動	7	6.3	ブラウザによる AR システム	32
3.2	簡単な例	7	6.4	テキストチャ付きモデルの作成方法	34
3.3	印刷する方法	8	6.5	マーカー例	35
3.4	LaTeX のコマンド	8	7	深層学習	38
3.5	使用できる文字	9	7.1	コンピュータによる問題解決	38
3.6	文章のスタイル	9	7.2	機械学習	38
3.7	タイトルページ	9	7.3	ニューラルネットワーク	39
3.8	文章構成	10	7.4	様々なフレームワーク	41
3.9	空白	10	7.5	画像認識プログラムの例	41
3.10	文字のスタイル	10	7.6	画像認識プログラムの実行	44
3.11	行、段落	11	7.7	学習プログラムの例	45
3.12	行の空け方	12	7.8	学習プログラムの実行	48
3.13	中揃え・右揃え	12	7.9	学習用データを増やす	50
3.14	箇条書	12	7.10	水増し学習プログラムの実行	51
3.15	そのままの形	13	7.11	YOLO による物体検出	51
3.16	表	13	7.12	AI に対抗するには	57
3.17	脚注の入れ方	14	8	動画プレゼン	59
3.18	数式	14	8.1	動画作成の手順	60
3.19	画像の入れ方	15	8.2	動画作成の例	61
3.20	図表の参照	15	8.3	タイトルの挿入	63
3.21	参考文献リスト	16	8.4	動画や画像の挿入	64
4	三次元モデルの作成	18	8.5	動画や画像の切り替えの設定	66
4.1	Blender の起動方法	18	8.6	説明文章の挿入	67
4.2	シーンに対する操作	19	8.7	音声の設定	67
4.3	オブジェクトに対する操作	20	8.8	動画の保存	68
4.4	オブジェクトの変形などの操作	23	8.9	動画編集のまとめ	68
4.5	オブジェクトなどの保存	24			

1. はじめに

私の基礎演習では、様々な IT (情報技術) を学ぶことを目標にしています。コロナの感染拡大のために、我々は様々な面で行動の制約を受けることになりました。それを補うために様々な IT が活用されることが期待されたのですが、日本はどうやらだめな国だったらしい、ということが明らかになりました。仕事がなくなり収入がなくなった方にお金を支給したいと思ったが、結局そういう方々を特定することができず、国民全員にお金を支給することになりました。その作業にもかなり時間がかかりました。他の国では申請してから一週間もかからずに支給された、と言う話もありました。コロナの感染拡大の状況も、保健所から Fax で送られてきた情報を手入力していると言う有様でした。その後オンラインで感染者情報を現場で入力することができるようになりましたが、入力すべき情報があまりにも多く、感染者の対応に忙しい現場ではとても入力する余裕がない、と言う状況も招きました。

IT を全て学ぶのはとても無理です。細かいところや難しいところは、理系の専門の方々にお任せするしかないでしょう。ただ「入力すべき情報が多すぎて使えないシステム」を作らせたのは理系の専門の方々ではないでしょう。現場を知らない、IT を知らない、そう言う人たちが大金を投じて役に立たないシステムを作らせる、と言うのは実は日本のあちこちで見られる風景です。

IT を利用して色々簡単にできる、ということを理解し、さらに自分で IT システムを作る人を育てたいと考えています。ささやかなシステムでも自分で作ることができれば、役に立つものか、使いやすいものかなどの検証ができます。そういうことができる人が増えれば、この国ももっとましになるのではないかと思います。基礎演習では IT で色々できることを学びつつ、「プログラミング基礎」などで IT システム作りの基礎を学んでもらいます。そして来年度の「展開演習」では実際に IT システムを作ります。案外簡単にできるものだなあ、と感じてもらえれば良いのですが。

技術の細かいところや難しいところは、理系の専門の方々にお任せするとして、代わりに磨いてほしいのは「利益」に対する感覚です。前世紀の末の頃の日本の大企業は世界的にも大企業でしたが、現在では見る影もありません。日本の大企業は儲けられなくて成長しない会社ばかりになってしまいました。利益は「売上」から「変動費」と「固定費」を引いたものです。まず利益にプラスとして関係する売上は、お客の数に比例するという考え方はだめです。つまりお客の数が 2 倍になれば、売上も 2 倍になるだろうというビジネスモデルでは足りません。Amazon や楽天などのオンラインモールでは、お客が増えると、売れるので出店も増えます。店が増えれば一人のお客が買うものも増えます。よってお客の数が 2 倍になれば、売上は 2 倍以上になります。この時注意しなければいけないのは、出店したお店は売上が 2 倍にしかならないところです。むしろ店が増えて同業者が増えると、売上の伸びが悪くなるでしょう。よって Amazon や楽天に出店して喜んではいけません。

売るのが情報だと変動費がほぼゼロになります。SNS の類はお客が情報を持ち込んでくれるので調達費もかかりません。一方パクられやすいのも情報です。音楽や動画配信ではコピー防止のために色々苦労しています。ただデジタル情報は劣化しませんが、新鮮さは次第に失われます。今後はコピー防止より、新しい情報の提供を重視する方が良いのではないかと思います。ただ調達費が十分用意できる規模がないと無理ですが。

売るのが物であれば、お客によるコピーの心配はなくなりますが、大量生産可能なものだと海外からコピー商品がどっとやってくるかもしれません。コピーされない高品質の実現がまず無理ならば、大量生産は諦めましょう。お客の個別の要求に合ったものを提供するという形態を実現するには、3D プリンターのような色々なものが作れる機械や、オンライン受注のシステムなどが必要になってきます。大量生産品に比べると変動費が高くなりますが、それ以上に販売価格を上げたり、直接お客に提供して中間マージンを削減できればなんとかなるでしょう。

固定費の削減は昔からの大きなテーマでした。IT による自動化による人件費の削減は、工場などでは大きな成果がありましたが、事務的な面ではまだまだです。せっかく入った会社からリストラされるのは誰でも嫌なことです。自動化がリストラにつながるとなると、当然協力したくないでしょう。仕事が楽になり、早く帰れ

るようになり、給料が上がるような自動化でなければ成功しないと思います。

IT系の固定費としてはWebサーバーやデータベースサーバーなどの各種サーバーの維持費があります。大きな仕事には大きなサーバーが必要になり、費用もかかります。最近ではAmazon、Googleやマイクロソフトのクラウドサービスを利用することにより、仕事量に合わせた規模のサーバーを使えるようになりました。クラウドサービスは様々な内容のものがあり、料金も仕事量にほぼ比例するため変動費として扱えます。またシステム管理の部分も任せられるのも大きな利点です。

ITを技術の面から捉えるだけでは、理系の学生に対して現代マネジメント学部の学生に勝ち目はありません。他の授業で習った経営・会計・経済・法律・政治などの知識もうまく組み合わせて行ける人を育てたいと思っています。

2. Linux (コンピュータ管理ソフト)

単一機能の電子機器に組み込まれる超小型のコンピュータを除いたほとんどのコンピュータで、OS (Operating System) と呼ばれるシステム管理ソフトが動いています。OS はコンピュータの持つ資源 (resource: 例えば CPU、メモリ、ファイル、入出力機器) を管理し、そのコンピュータで動くアプリケーションソフトの要求に従い、管理している資源を提供します。パソコンでは Windows が最もよく使われ、それ以外では Unix 系の OS がよく使われています。Unix 系の中では Linux が一番よく使われており、スマホの Android も Linux の一種です。

Windows については大学に入る前から使う機会があったと思いますが、Linux に関しては Android のスマホくらいしか縁がないと思います。そして実際 Android のスマホをさわっても、iPhone とあまり変わらないなあ、という印象を受けるだけでしょう。見た目で言えば Linux は色々です。Windows、iOS (iPhone の OS)、Android などはバージョンが決まればほぼ同じ見た目や操作方法になりますが、Linux は組み合わせ方によって Windows や iOS とよく似た形にもできるし、全くグラフィカルな機能なしでキー入力だけでコマンドを入れて使うような形にもできます。そのような自由度があるため、スマホでも使われるし、Web サーバーのような、ひたすら送られてきた URL に対してデータを送り返すサービスをするコンピュータでも使われるのです。

2.1 ユーザー登録

複数の人が利用する OS では通常利用する人を登録するユーザー登録が必要です。そして利用者ごとに様々な設定が行われます。システムの管理者であれば、何でもできる権限が与えられ、一般の利用者はシステム関係の設定変更や他の利用者のファイルに触れることができないように設定されます。この演習では mars^{*1} という名前のサーバーをみんなで共有して利用します。よって演習を受講する学生を登録するための Web ページを作成しましたので、それを利用してユーザー登録を行います。ブラウザで「<https://mars.mgt.sugiyama-u.ac.jp/AC/>」へアクセスすると図 2.1 のようなページが表示されます。

ここで自分の希望する登録名、自分の学籍番号と氏名を入力します。登録名の衝突はあまり考えられませんので、無理に長い複雑なものにする必要はありません。以下は昔の卒論生が使用した登録名です。

```
aki ako aoki asaka asako bluecat chizuko fujipon fusae happy hiromi ishii itomaki kaori kikumi kitty18
kozue kyoko lovin machi maki mame masami mayo miho mika misa miwako miyabi miyuki momoko
naochan natu oida okada rena rie risa saori sasai sasara shiori taki tomo tomoko toshiko toya yuki yukko
yuko yunke
```

2.2 リモートデスクトップによる接続

実は mars 本体には通常マウスやディスプレイが接続されていません。もし接続されていても、複数の学生が一つのディスプレイを覗き込んで作業をするのは大変です。Unix 系の OS は基本的に複数利用者の同時利用が可能になっていて、通常利用者はネットワーク経由で利用します。

昔はネットワーク経由でサーバーに接続し、キーボードでコマンドを入力し、その結果をまた文字出力で得るという形 (CUI: Character User Interface) のみでした。実は現在でもこのような形で管理されているサーバーは少なくありません。一方パソコンなどでは、アイコンやメニューなどを利用した形 (GUI: Graphical User Interface) でのコンピュータ利用が一般的で、同様な形でサーバーの操作ができれば便利です。その代わり CUI であればネットワークを経由して文字のやり取りで済みましたが、GUI では画面の状態なども常時送らないとマウスポインターも動きません。

^{*1} 正式な名称は mars.mgt.sugiyama-u.ac.jp です。なぜ mars かと言えば既に venus があったからです。

利用者登録受付

これはmarsの利用者登録を受け付けるものです。

希望する登録名などを入力して、申請ボタンをクリックしてください。登録名は最大8文字で以下のような文字から構成されるものです。

- 数字、アルファベット、アンダーバー(_)が使えます
- 漢字や大文字は使えません
- 先頭に数字は使えません

なお、既に登録されたものと同じ名前は使用することはできません。

登録名	<input type="text" value="maki"/>	登録済みの名前： _apt, avahi, backup, bin, boinc, colord, cups-pk-helper, daemon, dhcpd, dnsmasq, games, geoclue, gnats, irc, landscape, lightdm, list, lp, lxd, mail, man, messagebus, miki, mosquito, mysql, news, nobody, nvidia-persistenced, pollinate, postfix, proxy, pulse, root, rtkit, saned, sshd, sync, sys, syslog, systemd-network, systemd-resolve, usbmux, uucp, uuid, www-data, xrdp
学籍番号	<input type="text" value="A18EA999"/>	ex. A18EA999
氏名(漢字)	<input type="text" value="梶山 まき"/>	ex. 梶山 花子
氏名(ローマ字)	<input type="text" value="Maki SUGIYAMA"/>	ex. Hanako SUGIYAMA

図 2.1 ユーザー登録の画面

Windows では、WindowsXP 以降^{*2}に「リモートデスクトップ接続」というソフトが標準でインストールされています。本来は Windows によるサーバーへの接続のためのものですが、サーバーとのやり取りの部分を真似て、Linux であれば xrdp というソフトをインストールすれば、Windows のパソコンから Linux サーバーに接続して使用することができます。さらにスマホ用の「リモートデスクトップ」ソフトもマイクロソフト社から無料で提供されています。よって使いやすいとは言えませんが、スマホから Linux や Windows サーバーを使用することも可能です。

Windows10 では、「スタート」「Windows アクセサリ」「リモートデスクトップ接続」で起動することができます。起動すると図 2.2 のようなダイアログが表示されるので、コンピュータのところに「mars.mgt.sugiyama-u.ac.jp」を入力し接続ボタンをクリックします。



図 2.2 リモートデスクトップ接続の画面

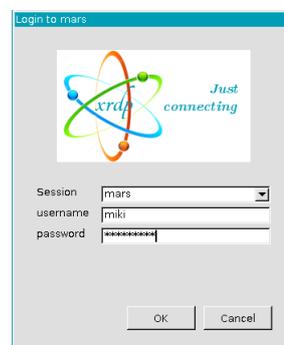


図 2.3 login の画面

^{*2} WindowsXP、Windows vista、Windows7、Windows10 の他、Windows NT 4.0 以降の Windows サーバー

無事に mars に接続できると、図 2.3 のような mars の認証のダイアログが表示されますので、mars での自分の登録名とパスワードを入力し **OK** ボタンをクリックします。

すると図 2.4 のような Windows7 以前の Windows に似たデスクトップが表示されます。mars では Linux の中でも ubuntu*³ と呼ばれるディストリビューションを使っていますが、GUI としては標準の GNOME*⁴ ではなく、より軽量で Windows に似ていると言われる LXDE*⁵ を利用しています。ただ KDE*⁶ や GNOME のソフトも一部取り込んでいるので、本来の LXDE とは少し見た目が異なります。



図 2.4 LXDE の画面

LXDE の基本的な操作は、

- 画面の左下角にあるスタートボタン? () をクリックするとメニューが出てくるので、そこで起動したいソフトを選択する。
- メニューの一番下の「ログアウト」を選択するか、画面右下角にある  をクリックすると終了することができる。ログアウトせずにリモートデスクトップ接続を終わると、次回接続した際にすぐに続きをすることができます。
- 漢字を入力する際には、キーボードの **半角/全角** キーを押すとかな漢字変換が起動される。再度同じキーを押せば元に戻る。

となっています。デスクトップにある「教材フォルダ」と「共有フォルダ」の内容は現マネ棟のパソコン演習

*³ <https://ja.wikipedia.org/wiki/Ubuntu> 「誰にでも使いやすい最新かつ安定した OS」を開発目標としている。最初から画像編集ソフトやワープロソフト、ゲームなどが入っている。よってデスクトップ OS としてよく使われているが、サーバー用としても使われている。

*⁴ <https://ja.wikipedia.org/wiki/GNOME> Unix 系では X-Window という画面表示の基本システムがあり、その上に GNOME、KDE、LXDE などの様々なデスクトップ環境があります。

*⁵ <https://ja.wikipedia.org/wiki/LXDE>

*⁶ <https://ja.wikipedia.org/wiki/KDE>

室の同名のフォルダと同じ内容で、同じように使用することができます。

2.3 パスワードの変更

システムが適当に設定したパスワードでは、入力しにくい、覚えにくいとか、友人に知られてしまったような場合はパスワードを変更しましょう。ウインドウ左下角の「スタート」「設定」「パスワード変更」で図 2.5 のようウインドウが出ますので、現在のパスワード、新しいパスワードを 2 回入力して「パスワード変更」をクリックします。入力に間違いがなければ図 2.6 のようになりますので、「OK」をクリックします。



図 2.5 パスワード変更のウインドウ

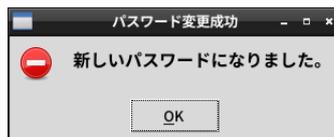


図 2.6 パスワード変更成功のウインドウ

2.4 インストール済みのソフト

このテキストでは特に取り上げないソフトを簡単に紹介します。「アクセサリ：Calculator」は「アクセサリ」サブメニューにある「Calculator」という意味です。

- アクセサリ：Calculator 電卓ソフト
- アクセサリ：Kate 文書ファイル編集ソフト Windows の「メモ帳」を少し高級にしたもの。
- アクセサリ：nitrogen 壁紙選択ソフト
- アクセサリ：スクリーンショット 画面を画像ファイルに変換する。
- インターネット：firefox ウェブブラウザ Web ページ閲覧ソフト
- インターネット：Sylpheed メールを読み書きするソフト
- オフィス：LibreOffice
 - Base データベースソフト。Microsoft Access のようなもの。
 - Calc 表計算ソフト。Microsoft Excel のようなもの。
 - Draw 作図ソフト
 - Impress 発表用資料作成ソフト。Microsoft PowerPoint のようなもの。
 - Math 数式記述ソフト
 - Writer ワードプロソフト。Microsoft Word のようなもの。
- グラフィックス：inkscape お絵かきソフト (SVG 形式対応)
- グラフィックス：Pinta お絵かきソフト Windows の「ペイント」のようなもの。
- システムツール：LXTerminal Linux のコマンドを入力するためのソフト

3. 文書清書システム

ここでは文書清書システムである \LaTeX の使い方を説明します。 \LaTeX は同じく文書清書システムである \TeX をより使いやすくしたのですが、ワープロソフトの生まれる前に誕生したもので、それでも大変使いにくいものです。イメージとしては、HTML のタグを文章に挿入して Web ページを作成するのに似ています*7。HTML のタグは何かして欲しいところに挿入します。 \LaTeX のコマンドも同様です。よって多少のコマンドは覚えなければなりません、レポートや論文ではそれほど多くのコマンドは必要ないので、あまり問題にならないでしょう。

同じコマンドは同じ結果をもたらすために、文書の見た目を揃えることができます。Word でその日の気分で見出しのフォントの大きさを変えたりしていると、簡単に統一の取れていない文書ができてしまいますが、 \LaTeX ではそのような問題を避けることができます。また HTML のタグと違い、 \LaTeX のコマンドはカスタマイズが可能です。標準の大きさや形が気に入らない場合は変更することができます。

システムの処理結果を示す場合、ブラウザに表示するのであれば HTML を使用し、体裁を整えるならさらに Style Sheet を使用するのが普通でしょう。紙に出す場合、プリンターごとに指定した文字を指定した場所に出す方法が異なるので、様々なプリンターに対応するのは困難です。そこでよく使われているのが処理結果を PDF の形にして出すと言う方法です。どのようなソフトで作成した文書も、PDF であればパソコンのプリンターで印刷できます。あとは処理結果をきれいな PDF の形にするところですが、処理結果に \LaTeX のコマンドを追加して清書し、その結果を PDF にすることができます。これならばブラウザに表示するのと余り変わらない手順で、紙に印刷ができます。

3.1 Texmaker の起動

\LaTeX の原稿となるファイルは、Windows 付属の「メモ帖」ソフトでも入力が可能です。HTML のタグと同様に \LaTeX のコマンドも普通のキーボードで入力できる文字で構成されているからです。しかし、入力が終わればコンパイルという処理やその結果の表示などの操作が必要なので、これらを簡単に実行できるソフトが色々作られています。mars では Texmaker というソフトをインストールしていますのでこれを利用します。ウインドウの左下角の「スタート」「オフィス」「Texmaker」で起動できます。また拡張子が tex のファイルをダブルクリックしても、Texmaker の起動・読み込みができます。

\LaTeX はコンパイル処理の際に複数のファイルを生成します。文書を変換したファイルの他、目次や参照などのデータが別のファイルになります。ファイルがばらばらにならないように、文書ごとにディレクトリ(フォルダー)を作成することをお勧めします。

3.2 簡単な例

Texmaker を起動しただけでは文章の入力できません。「ファイル」「新規作成」をすると入力できるようになります。簡単な例として以下の文章を入力してください。入力する際に \backslash は \yen (半角) のキーを押します。また \backslash で始まる横文字の部分 (\LaTeX のコマンド) は「LaTeX」メニューに入れることができます。また \LaTeX のコマンドの最初の数文字を入れると、コマンドの候補が出てきます。複数の候補が出てきた場合は更に入力するか、矢印キーで選択してから \Tab を押すと確定されます。

*7 HTML が Tex の真似をしたという方が正しいが。

```
\documentclass[10pt]{jsarticle}
\title{簡単な例}
\author{A20EA999 梶山花子}
```

```
\begin{document}
\maketitle
```

```
\section{はじめに}
```

これは大変簡単な例です。文章中の改行は無視されるので、入力の際には適当に改行してください。1 行空けると段落の終わりともみなされます。

これは 2 つ目の段落です。先頭は勝手に字下げされるので空ける必要はありません。

```
\end{document}
```

日本語の部分はこの通り入れなくても大丈夫です。無事入力ができたら「ファイル」メニューの「名前をつけて保存」で保存します*⁸。そして[F1]のキーを押すとコンパイルの処理が行われて、問題がなければ PDF 形式の結果ファイルが生成されて、表示されます。

誤りがあった場合には、画面の下に残念ながら英語でメッセージが表示されます。誤りのあった行番号が示されるのでそれを参考に修正をしてください。なお、修正のあとは上書き保存をしなくても、[F1]のキーを押すとコンパイル処理の前に上書き保存をしてくれます。

3.3 印刷する方法

通常インターネットのどこからでも mars にアクセスすることができます。つまりどこからでも文書作成が可能です。ところが作成した文書を印刷するとすると、mars の側にあるプリンターから出てくるのでは大抵の場合困ります。現在自分のいる場所に近いプリンターから出てくるのが望ましいのですが、それがどれなのかは mars には分かりません。コンパイル処理された文書は PDF 形式のファイルになっているので、これを自分の使っているパソコンにコピーして、パソコンから印刷することによって、身近なプリンターから出すことができます。

1. PDF 形式の文書ファイル (~.pdf) のアイコンを右クリックする。
2. 「コピー」を選択する。
3. パソコンの適当な場所 (例えばデスクトップの空いたところ) で右クリックする。
4. 「貼り付け」を選択すると、文書ファイルがパソコンに送られる。卒論のような大きなファイルや画像を多数含むファイルは多少時間がかかるかもしれません。
5. パソコン上の文書ファイルをダブルクリックして開き、そこで印刷の指示をする。

3.4 LaTeX のコマンド

基本的なコマンドの一部を紹介します。なおこの部分は鈴木裕信氏 (hironobu@sra.co.jp) が書かれた「ひろのぶの LaTeX 入門 Version 0.23」を元はかなり省略したものになっています。

LaTeX には局所的なコマンド (command) と広域的な環境 (environment) があります。コマンドは \ で始まります。そして {} や [] で囲まれた引数を取ることができます。たとえば、節の見出しのコマンドは \section であり、この章の原稿には、\section{文書清書システム} と書いてあります。

一方環境は \begin{name} と \end{name} で囲まれた範囲が、name という規則にすべて従うという感じになります。

*⁸ 拡張子として .tex が必要ですが自動的に補われます。

3.5 使用できる文字

文中で使用できる文字は全角文字、半角の英数字や記号などです。ただし、次の 10 個の特殊文字は、 \LaTeX の命令に使用されます。ですからそのまま使用することはできません。

\$ % & _ { } ~ ^ \

\ 以外は \#, \\$ とすることによって文字として表現できます。 \ は次のような方法で表現します。

\backslash ⇒ \

\LaTeX はテキスト中に空白があっても自動的につめてしまいます。意図的に空白を挿入する時は、 \ する 必要があります。改行の場合は \\ を使います。マイナス記号 (-) は特別な働きをします。連続して書くこ とによってハイフンなどの線 (- - -) を引くことができます。

3.6 文章のスタイル

一番最初は必ず \documentclass[option]{style} で始めなければなりません。そして文章の本体は \begin{document} と、 \end{document} で囲まれた範囲内に書きます。

文書のスタイル (style) の標準的なものには次のようなものがあります。

	欧文	和文 (旧)	和文	和文 (縦書き)
論文	article	jarticle	jsarticle	tarticle
本	book	jbook	jsbook	tbook
報告書	report	jreport	-	treport

「論文」スタイルは最もよく使われます。標準では用紙の下部の中央にページ番号が付きま す。このテキス トもこのスタイルに手を入れて使っています。 \section のレベルから書くことができます。

「本」のスタイルでは、奇数ページの上部 (ヘッダ) では節 (セクション) の番号とセクション名が左側、 ページ番号が右側に出ます。偶数ページヘッダでは右側に章 (チャプタ) 名、ページ番号が左側に出ます。章の最初 のページが奇数になるようにページが送られます。章立ては必ず \chapter から始めなければなりません。

「報告書」はレポートの形式です。ページ番号は用紙の下部の中心にくるような形になります。章立ては必 ず \chapter から始めなければなりません。もし、 \section レベルから始めた場合、エラーとなります。

オプション (option) はスタイルの補助的な役割をします。基本となるフォントサイズの変更や見開きに変 更する事ができます。

	指定なしの場合	指定できるもの
文字の大きさ	10pt	9pt、10pt、11pt、12pt、14pt、17pt、 21pt、25pt、30pt、36pt、43pt
用紙サイズ	A4 サイズ	a4paper、a5paper、b4paper、b5paper
本文の段組	onecolumn (一段組)	onecolumn、twocolumn
見開きの設定	本のスタイルでは twoside、 それ以外では oneseide	oneside、twoside
表題を別ページにする	論文スタイルでは本文と同じページに	titlepage

3.7 タイトルページ

タイトルページは \maketitle があらわれた時点で作成されます。よってその前に \title や \author を 使って表題や著者を指定する必要があります。 \date は指定しなければ、コンパイルを行った日時が使われ

ます。

```
\title{HOW TO USE \LaTeX}
\author{Hironobu Suzuki}
\date{4 Oct 1988}
.....
\maketitle
```



HOW TO USE L^AT_EX
Hironobu Suzuki
4 Oct 1988

タイトルページに要約に付ける時は次のようにします。もちろん `\maketitle` コマンドより前にこれを使う必要があります。

```
\begin{abstract}
This article is for the beginner.
This abstract is printed onto top page.
\end{abstract}
```



Abstract
This article is for the beginner. This
abstract is printed onto top page.

3.8 文章構成

文章の構成で、章や節を表すものとして、次のようなものが用意されています。

```
\part          \section          \paragraph
\chapter       \subsection       \subparagraph
              \subsubsection
```

各々 `\subsection{文章構成}` のように見出しの文字を指定します。通常の `jsarticle` スタイルでは、節 (section) の中が項 (subsection) に分かれ、さらにその中が部分項 (subsubsection) に分かれます。 `jbook` や `jreport` は章 (chapter) の中が節に分かれ、以下 `jarticle` と同様に細かく分かれて行きます。

見出しの番号は `LaTeX` が自動的に付けてくれます。見出しの文字の大きさなども自動的に設定されます。`\section` を飛ばして `\subsection` を使用すると番号が変になるのでご注意ください。

3.9 空白

半角の空白は複数あっても、1つの空白と見なされます。逆に全角の空白は目に見えませんが、行の前後に残っていると歯抜けの段落ができたりするので注意が必要です。

意図的に空白を挿入したい場合は `\□` とします。これは `LaTeX on UNIX` のような書き方をしたとき `LATEX on UNIX` というような結果になってしまいます。このような結果を避けたい時は `LaTeX\□on□UNIX` とすれば `LATEX on UNIX` となります。

`\hspace` を使用すると、きっちり指定した幅だけ空けることができます。ただしその空白がちょうど行の先頭や末尾に来たときは無視されますのでご注意ください。

```
1cm \hspace{1cm} space.
1cm \hspace{1cm} space.
1cm \hspace{1cm} space.

This command can accept minus
paramater like this \hspace{-1cm}////
```



```
1cm      space.  1cm      space.  1cm
space.
This command can accept minus paramater
like\this
```

3.10 文字のスタイル

文字の大きさは直接ポイント数を与えて変えるような野蛮なことはできません。 `large` や `small` などのコマンドによって文字のポイント数が変わりますが、原稿の基準フォントのポイントサイズに合わせて、相対的に

large であり、small であります。このように相対的に指定する事により、後になって基本となるサイズを変更した場合にも、本文の指定を変更する必要がなくなります。

文字の大きさの有効範囲は同一ブロック内です。ブロックとは、環境の中や { と } で囲んだ範囲の事です。例えば、{\Large abc}d と行なうと abc が大きくなり d は元のサイズのままです。実際に行なってみると abcd となります。

通常の文書で頻繁に文字の大きさを変えると見やすさよりも見にくさを助長するだけなので、次の指定はあまり使う機会はないでしょう。

文字 \Huge	文字 \huge	文字 \LARGE
文字 \Large	文字 \large	文字 \normalsize
文字 \small	文字 \footnotesize	文字 \scriptsize
文字 \tiny		

英数字の文字のスタイルは色々ありますが、全角に関しては太字と下線のみが有効です。影響を与える範囲などは文字の大きさと同じです。

```
{\bf Bold type style. 全角は太字です。}
{\sf Sans serif type style. 全角は同じです。}
{\sl Slanted type style. 全角は同じです。}
{\tt Typewriter type style. 全角は同じです。}
{\sc Small caps type style. 全角は同じです。}
{\em A emphasized line. 全角は太字です。}
\underline{\underline{underlined text}} \underline{下線付き文字です。}
```

⇒

```
Bold type style. 全角は太字です。 Sans serif type style. 全角は同じです。 Slanted type style. 全角は同じです。 Typewriter type style. 全角は同じです。 SMALL CAPS TYPE STYLE. 全角は同じです。 A emphasized line. 全角は太字です。 underlined text 下線付き文字です。
```

下線はコマンドなので指定の仕方が異なりますのでご注意ください。

3.11 行、段落

原稿の文字列は自動的に詰められ改行されたりします。空行は段落の切れ目を意味しますが、複数の空行は 1 行分の空行と解釈されます。段落の先頭は 1 字下げがなされるので、原稿では字下げは不要です。

文字列や単語の途中で改行が入ってしまった場合は \mbox を使います。例えば、*Segmentation fault* の *Segmentation* と *fault* を離れたくない場合は \mbox{Segmentation fault} とします。

左につめて書きます。ただし、前の行は空白行です。

連続した行です。
連続した行です。
連続した行です。
連続した行です。
連続した行です。

⇒

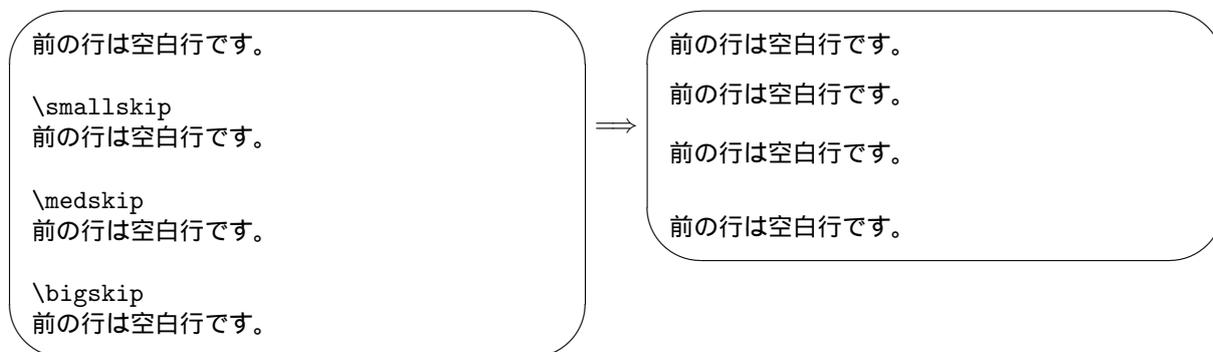
左につめて書きます。ただし、前の行は空白行です。

連続した行です。連続した行です。連続した行です。連続した行です。連続した行です。

3.12 行の空き方

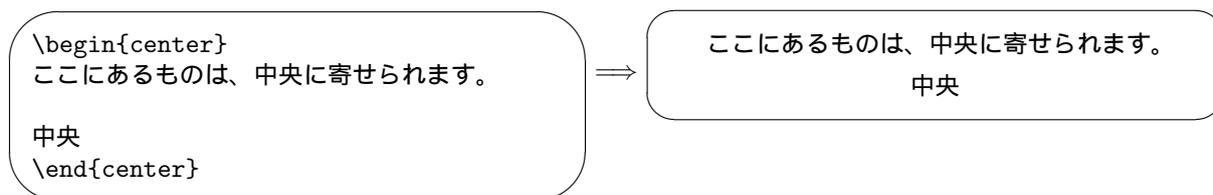
LaTeX では段落、表や画像などがくっつくことがよくあります。これらの間を空けたいときは、`\smallskip`、`\medskip`、`\bigskip` を使用します。実際の間隔はページのレイアウトに左右されますので、同じコマンドでもいつも同じ間隔が空くとは限りません。

行間や段落間の間隔を調整する直接指定する方法として `\vspace` などのコマンドがありますが、既に書式が決まっているものに無理矢理合わせるような場合以外は使わない方が無難ですので例は省略します。

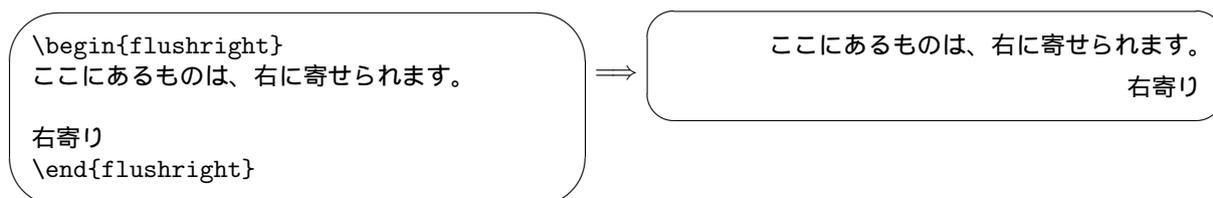


3.13 中揃え・右揃え

通常のワープロソフトを使用する際でも、中揃えや右揃えは空白を使って位置を調節するのではなく、そういう指示で行うのが常識です。LaTeX では次のようにして中揃えを実現します。



一方 `flushright` を使用すると右揃えになります。



3.14 箇条書

箇条書の書式として LaTeX では 3 種類の方法があります。先頭に `•` をつけるもの、先頭に番号を振るものと、先頭にラベルを指定できるものです。それぞれ、`itemize`、`enumerate`、`description` です。

```
\begin{itemize}
\item 各々の行はあたみにポチを付けられてな
らんでいるでしょ。
\item
書き出しは次の行でも大丈夫です。
\end{itemize}
```

```
\begin{enumerate}
\item この行の頭には数字がついているはず
です。
\item ほらね、2行目でしょ。
\end{enumerate}
```

```
\begin{description}
\item [ラベル] この行の頭にはラベルが
ついているはず。
\item [Label] ラベルは bold 体になってい
ます。
\end{description}
```

⇒

- 各々の行はあたみにポチを付けられてな
らんでいるでしょ。
- 書き出しは次の行でも大丈夫です。

1. この行の頭には数字がついているはず
です。
2. ほらね、2行目でしょ。

ラベル この行の頭にはラベルがついている
はずです。

Label ラベルは bold 体になっています。

3.15 そのままの形

`verbatim` はテキストを入力したそのままの形で表すための環境です。例えば、`%\}` はそのまま打っても $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ では正しく表示してくれません。プログラムのリストをそのまま引用するような場合、この環境を使用します。

```
\begin{verbatim}
次のキャラクターは通常、特殊文字として扱わ
れるので表示できない。
バックスラッシュを使うか verbatim を使うか
どちらかである。
\end{verbatim}
\verb+\ \ \ \ ??? ~~~ $$$$ ###+
```

⇒

次のキャラクターは通常、特殊文字として扱わ
れるので表示できない。
バックスラッシュを使うか `verbatim` を使うか
どちらかである。

```
\ \ \ \ ??? ~~~ $$$$ ###
```

行中にこれらの特殊文字を組み込みたいときは、`\verb` を使用します。`\verb+??###$+` とすると、`??###$` と出力されます。

3.16 表

$\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ では、表のようなものを作るのに、`tabbing`、`tabular`、`array` という環境が利用できます。それぞれ想定された用途があり、特徴がありますがここでは `tabular` のみ説明をします。HTML の表も様々なタグが多数必要で面倒でしたが $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ はそれ以上に面倒です。ただ `Texmaker` の「作成支援」メニューの中に「表の速成」というものがあります。とりあえずこれで表を作成し、後で微調整すると言うやり方も可能です。

`tabular` の引数の中での命令の意味は、`l` は左詰め、`r` は右詰め、`c` は中央に各項目をそろえます。`p{wd}` は列の幅を `wd` に指定 (たとえば `p{1cm}`) することにより、固定することができます。

| をつけると垂直方向の線が書けます。`\`の後の `\hline` は水平方向の線を引きます。`\cline{i-j}` は `i` 列から `j` 列までの水平方向の線を引きます。

```

\begin{tabular}{|l|l|c|r|}
\hline
sra & sra-b1-net & 12098 packets & \\
\cline{2-3}
& sra-2-net & 290873 packets & \\
\hline
ntt & sun-loop-back & 2839 packets & \\
\cline{1-1} \cline{3-3}
etl & & 287 packets & \\
\hline
\end{tabular}

```

⇒

sra	sra-b1-net	12098 packets
	sra-2-net	290873 packets
ntt	sun-loop-back	2839 packets
etl		287 packets

一つ一つの項目欄は`\multicolumn{n}{pos}{item}`によって調整することができます。n は使用する列数、pos は項目の水平方向の位置、item は項目の文字列です。引数である pos は現在すでに設定されているものを無効とし新たに設定します。この時有効となる引数 pos は l r c | p からなるものでなくてはなりません。

```

\begin{tabular}{|l|l|l|r|}
\hline \hline
{\em team } & & & \\
\multicolumn{2}{c|}{\em driver} & & \\
\hline
マクラーレン & セナ & プロスト & \\
ロータス & ピケ & 中島 & \\
\hline \hline
\end{tabular}

```

⇒

<i>team</i>	<i>driver</i>	
マクラーレン	セナ	プロスト
ロータス	ピケ	中島

3.17 脚注の入れ方

ページの下の部分に脚注を入れることができます。本文の補足的な情報をここに入れます。例えば URL^{*9}などは参考文献として挙げても良いですが、それほど重要でなければ脚注に入れた方が本文が読みやすくなります。脚注の参照番号が入るところに `\footnote{...}` を入れます。「...」の部分に脚注の文章が入ります。

3.18 数式

数学モード (math mode) は数式を表現するために使用するモードです。TeX は元もと数学の本を美しく作るために生まれたために数式の表現は大変得意です。しかし、とりあえずは x^2 、 $\sqrt{2}$ と $\frac{2}{3}$ ぐらいが書ければ十分でしょう。

文中とは独立して数式を示したいときには、`\begin{math} ... \end{math}` を使います。文中に数式を入れる場合には `$...$` を使います。例えば、`$ \int \sin(2x+3)\cos(x)^2 dx $` と記述すると、 $\int \sin(2x+3)\cos(x)^2 dx$ のように出力されます。

数式中での英字は xy というように斜体になります。一方関数名などは例えば、`log` は通常の字体でなければなりません。そこで `\log` を使用します。例えば、`$\log x^y = y \log x$` とすれば $\log x^y = y \log x$ となります。

\pm	<code>\pm</code>	\times	<code>\times</code>
A_b^c	<code>A_{b}^c</code>	$\frac{x}{y}$	<code>\frac{x}{y}</code>
\sqrt{A}	<code>\sqrt{A}</code>	$\sqrt[b]{A}$	<code>\sqrt[b]{A}</code>
$gcd(x,y)$	<code>\gcd(x,y)</code>	$\log(x)$	<code>\log(x)</code>
$\max(x,y)$	<code>\max(x,y)</code>	$\min(x,y)$	<code>\min(x,y)</code>
A^b	<code>A^b</code>	A_b	<code>A_b</code>

*9 Uniform Resource Locator: 例えば https://ja.wikipedia.org/wiki/Uniform_Resource_Locator

3.19 画像の入れ方

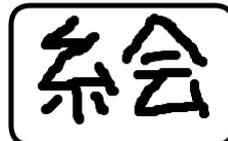
まずいちばん最初の`\documentclass`と`\begin{document}`の間に次のパッケージの設定を入れます。パッケージは LaTeX の機能を拡張するためのもので、これ以外にも様々なもの^{*10}があります。

```
\usepackage[dvipdfmx]{graphicx,color}
\usepackage{bmpsizer}
```

そして画像を入れたいところに次のようなコマンドを入れます。

```
\includegraphics[width=3cm]{sample.png}
```

⇒



この例の場合、「sample.png」という画像ファイルの内容が、「width=3cm」によって幅 3cm に整形されて取り込まれます。また「scale=0.5」のように相対的な大きさ(倍率)で指定する事もできます。例えば画面から同じように取った複数の画像は、同じ倍率で指定した方が良いでしょう。画像ファイルの形式は、JPEG と PNG のものが可能です。また PDF 形式の文書もこれで取り込むことができます。

画像ファイルは様々な方法で作成することが可能ですが、最終的に原稿のファイルがあるところへ持ってくる必要があります。パソコンまで持ってこれれば「印刷する方法」で PDF ファイルをパソコンに送った方法の逆をすれば、mars に画像ファイルをコピーすることができます。また mars にも画像を編集するソフト「Pinta」^{*11}があるので、Windows の「ペイント」のようにお絵かきも可能です。

スマホで撮影した画像を mars に持ってくる場合は、メールに添付して mars へ送る方法があります。またパソコンの画面を画像として取り込む事も可能です。Windows の場合(Print Screen)をクリックすると全画面が取り込まれますので、「Pinta」を起動して貼り付けるか、Windows の「ペイント」を起動して貼り付けます。アクティブなウィンドウのみを取り込む場合は(Alt)を押しながら(Print Screen)をクリックします。mars の中のものを画像にする場合はこれらのボタンを利用しても良いですが、「スクリーンショット」^{*12}と言うソフトが入っていますので、こちらを利用すると直接画像ファイルになります。

3.20 図表の参照

文書の中に図や表が一つしかない場合はどのような書き方をしても間違いがありませんが、複数の場合はどの図表を指しているのかわかりにくくなります。また図や表のような比較的大きなものの場所を厳密に指定してしまうと、図表がページの残りに入り切らない時に文中に大きな空気ができてしまうこともあります。このような問題を生じないように図や表は次のようにしましょう。

図の場合

画像の場合は次のようにします。

^{*10} 例えば <https://www.biwako.shiga-u.ac.jp/sensei/kumazawa/texindex.html> には 300 種類以上のパッケージが紹介されています。

^{*11} 画面の左下の「スタート」「グラフィクス」「Pinta」と選択して起動できる。

^{*12} 画面の左下の「スタート」「アクセサリ」「スクリーンショット」と選択して起動できる。

```

\begin{figure}[htbp]
\centering
\includegraphics[width=3cm]{sample.png}
\caption{図の例}
\label{zu-rei}
\end{figure}

```



図 3.1 図の例

最初の行の「htbp」は図をどこに置くかの指定です。「原稿の位置のとおりに表示 (here)」、「ページの上の方に表示 (top)」、「ページの下の方に表示 (bottom)」、「別のページに表示 (page)」の順番に検討するという意味です。`\centering` は図の位置を中央揃えにする指示です。`\caption` で図の見出しを指定し、`\label` でこの図を参照する際の名前を指定します。文中でこの図について触れる箇所に「`\ref{zu-rei}`」を入れると、「図 3.1」のように表示されます。また「`\pageref{zu-rei}` ページ」を入れると「16 ページ」と言うように図のあるページも表示できます。

図を同じ場所に複数並べる際には `minipage` 環境を使用する必要があり、結構複雑になります。

表の場合

表の場合も `tabular` 環境を `table` 環境の中に入れます。`figure` の部分が `table` になるだけで、後は図の場合と変わりません。

```

\begin{table}[htbp]
\centering
\caption{表の例}
\label{hyou-rei}
\begin{tabular}{|l|l|l|}
\hline
aaa & bbb & ccc \\
\hline
xxx & yyy & zzz \\
\hline
\end{tabular}
\end{table}

```



表 1 表の例

aaa	bbb	ccc
xxx	yyy	zzz

表 1 のように `caption` コマンドを先に使用すると、表の上に見出しを付けることができます。図と表の番号はそれぞれ独立してカウントされます。

3.21 参考文献リスト

文書内で、他の本や Web ページなどに掲載された事実や意見などを引用した場合、どこに掲載されていたものか示さないと問題になることがあります。また自分のオリジナルでない意見などは、誰によるものかを明らかにした方が信じてもらえます。

簡易的に出典を示すには既に出てきた脚注を使用しても良いでしょう。もう少しちゃんとしたやり方としては、次のように `thebibliography` 環境の中で `bibitem` コマンドで文献を示し、`cite` コマンドでそれを参照します。

参考文献 `\cite{latex}` を参照のこと。

```
\begin{thebibliography}{99}
  \bibitem{latex} Leslie Lamport 著、Edgar Cooke・倉沢良一監訳、
  「文書処理システム  $\LaTeX$ 」 p.191-192、株式会社アスキー、1990 年
\end{thebibliography}
```

`thebibliography` 環境で「99」を指定しているのは、参考文献の数が 2 桁という意味です。大抵の場合 99 で十分でしょう。`bibitem` コマンドで書く内容は、図書の場合、著者名、図書名、出版社、出版年が最低限必要です。これによって次のように示されます。ただし `cite` コマンドで表示される番号は 2 回コンパイルしないと出てこないので注意が必要です。参考文献 [1] を参照のこと。

参考文献

- [1] Leslie Lamport 著、Edgar Cooke・倉沢良一監訳、「文書処理システム \LaTeX 」 p.191-192、株式会社アスキー、1990 年

4. 三次元モデルの作成

ここでは、三次元モデル作成ソフトである Blender^{*13} ver.2.82a の使い方^{*14}を説明します。これで作成した三次元モデルは、3D プリンターで実物を作成したり、AR(拡張現実)を利用して現実世界へ投影することができます。三次元モデルを作成するためのソフトは様々なものがあります。Blender は Windows だけでなく Linux などでも使用可能な高機能なフリーソフトです。

Blender は高機能のため様々なことができますが、逆に機能が豊富すぎて困る場合も少なくありません。例えばメニューも多数あり、かつ多段階になっているのでなかなか望みの機能にたどり着けません。また二次元のお絵描きと比べると三次元モデルの作成はかなり複雑と言う点があります。お絵描きソフトで絵を描く場合、二次元のディスプレイの上に見えたものがそのままお絵描きの結果となります。三次元の場合、二次元のディスプレイに見えるものはある方向から見た場合の姿ですので、別の方から見ると全然予定と異なる形になっていることも少なくありません。また色についてもお絵描きでは、色を選択して塗りつぶすという感じですが、三次元では材質(マテリアル)の色(ディフューズ)を設定し、さらに鏡面反射(スペキュラー)する色、反射する割合なども設定します。光線の位置や強さも別に設定することにより、他の物体の映り込みなどは自動的に計算されます。逆にお絵描きでは映り込みなどは自分で描き込まなければいけません。

さらに三次元モデルでは動かす事も可能になっています。アニメの登場人物を三次元モデルで作成する際には、モデルの内部に骨や関節を設定し、体の各部分をシーンに合わせて動かせるようにします。お絵描きでアニメを作成する場合は、少しずつ動いたシーンを描くことになりませんが、三次元の場合は一つのモデルの変形で対応できます。さらに MikuMikuDance^{*15}のような動画作成ソフトでは、最初と最後の姿勢を指定すると、その間の動きを自動的に補間してくれますが、二次元の絵で同様のものはありません。

4.1 Blender の起動方法

パソコンで Blender を使用する場合は、そのパソコンに Blender を自分でインストールする必要があります。mars には既にインストールしてありますので、すぐに使うことができます。ただし複数の人が同時に使うと処理が追いつかなくなって遅くなる可能性があります。mars ではウインドウの左下角の「スタート」「グラフィックス」「Blender」で起動できます。すると図 4.1 のような Blender の画面が出てきますが、かなり複雑です。中央のオブジェクトの表示画面(3D ビューウィンドウ)の周りを様々なメニューやボタンが取り囲んでいます。(起動したときには中央にバージョンなどを示したパネルが表示されますが、パネルの外をクリックする消えます。)

Blender で作成する三次元モデルを構成するものをオブジェクトと呼びます。各オブジェクトは平面で構成されています。曲面ではありません。球のような丸いものは、実用上問題のないレベルまで細かく分割した平面にします。

Blender では様々な操作をメニューで選択して行うことができますが、キー入力でもかなりの操作を指示することができます。メニューで一々選択するよりもその方が早いのでキー入力で行う方を勧めますが、マウスポインターが適切な場所がないとうまく行きませんのでご注意ください。また操作を間違えたらすぐ **[Ctrl]+[Z]** で戻しましょう。

^{*13} <https://www.blender.org/>

^{*14} 他のソフトでもよくあることですが、バージョンが異なると操作法が変わってしまうので、検索して使い方を探す場合はどのバージョンの説明なのかに注意しましょう。Blender は ver.2.80 からかなり変わりました。

^{*15} <https://sites.google.com/view/vpvp/>

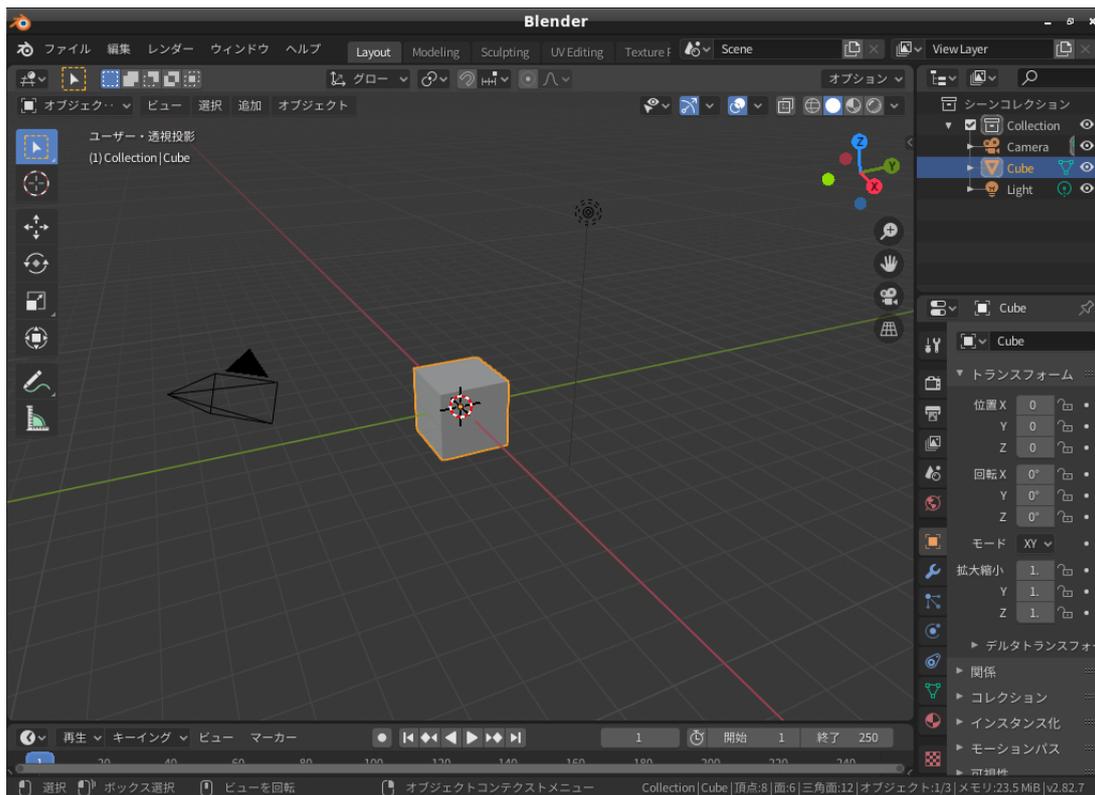


図 4.1 Blender の画面

4.2 シーンに対する操作

全体をシーンと呼びます。シーンには、複数のオブジェクトだけでなく、照明やカメラが含まれます。マテリアルにいくら色を設定しても、発光する設定がなく、照明がないと真っ黒です。同じ色でも照明の当て方によって明るさが変わりますし、影も生じます。よって Blender で綺麗なオブジェクトの画像を作成したい場合は、照明をどのように当てるかは大きな問題となります。一方オブジェクトを取り出して 3D プリンターで造形する場合は、Blender の外で様々な光に当たる訳ですから、照明の設定は不要です。カメラの位置や向きなどの設定もオブジェクトの画像を作成したい場合の際は重要ですが、オブジェクトを外へ取り出す場合は関係ありません。

4.2.1 寸法の単位の設定

オブジェクトを 3D プリンターで実物にする場合や AR(拡張現実) で設定した大きさで表示したい場合、まず単位を設定する必要があります。右側のパネルのところで図 4.2 のように設定をします。まず①の部分が変わったものになったらクリックして「プロパティ」を選択します。次に②をクリックし、③の「単位系」が「メートル」になっているのを確認します。そしてその下の「単位の倍率」を「0.001」に変更してから④の「長さ」を「ミリメートル」にします。大きめのオブジェクトを作成するならば「単位の倍率」を「0.01」に変更してから④は「Centimeters」(cm) でも良いでしょう。

4.2.2 シーンを見る向きの変更

マウスポインターを中央のオブジェクトの表示の部分へ持って行って、テンキーの数字を押すと表 2 のようにオブジェクトの見える向き(視点)が変わります。キーボードのアルファベットの上に横一列に並んでいる数字のキーではだめなのでご注意ください。テンキーを押す際に(Ctrl)キーを押しながらだと回転などの向きが逆になります。(Shift)キーを押しながらだと選択対象基準で回転などをするようになります。視点を色々変更

していると、時々オブジェクトが何も見えなくなります。そういう場合はテンキーの「.」を押すと見えるようになります。

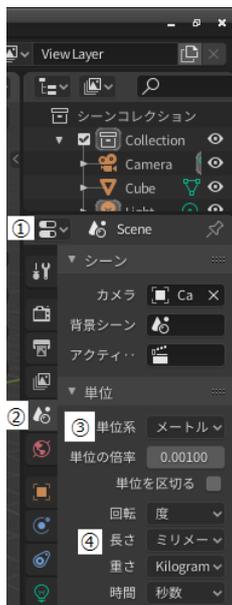


図 4.2 寸法の単位の設定

表 2 テンキーによる視点の変更

キー	視点
2	下方方向に 15° 回転
8	上方方向に 15° 回転
4	左方向に 15° 回転
6	右方向に 15° 回転
9	180° 回転
1	正面からにする
3	右正面からにする
7	真上からにする
0	カメラからの視点にする 元にもどす
5	投影方法の切り替え： 透視投影 平行投影
.	オブジェクトの全てが見えるようにする
/	選択しているオブジェクトが見えるようにする

マウスのホイールで視点からの距離を変えることができます。近くにすればオブジェクトが大きく見え、遠くにすれば小さく見えます。また`(Shift)`を押し、マウスのホイールを押しながらマウスを動かすと視点を上下左右に動かすことができます。

4.3 オブジェクトに対する操作

クリックしてオブジェクトを選択した上で、`(Tab)`を押すと「オブジェクトモード」と「編集モード」の切り替えができます。「オブジェクトモード」では、次に述べるようなオブジェクトの位置などオブジェクト全体の操作ができます。またオブジェクトの形を変えるような場合は「編集モード」にまず切り替えます。

4.3.1 オブジェクトの追加

図 4.3 のように①の「追加」のメニューをクリックし、②の「メッシュ」のサブメニューをクリックすると追加するオブジェクトの一覧が出てくるので、「平面」～「モンキー」の中の追加したいものをクリックします。すると  (3D カーソル) の位置に新しいオブジェクトが追加されます。場合によっては大きすぎたり、小さすぎるオブジェクトが出てきます。追加後すぐであれば、画面の左下に「▶ 立方体の追加」というがあるので、それをクリックすると大きさなどの設定のパネルが出てくるので数字を変更して直します。

3D カーソルの位置は、画面の左端にある図 4.4 のツールバー^{*16}の②をクリックすると、マウスでドラッグしたり、画面をクリックすることにより変更することができます。

4.3.2 文字オブジェクトの追加

立体的な文字を作成することができます。なお、説明文のような文章は後述のテキストチャを利用して、文章の画像を貼り付けた方が良いでしょう。図 4.3 の②の下の方にある「テキスト」をクリックすると、3D カーソルのあった位置に「Text」という厚みのない文字オブジェクトが追加されます。`(Tab)`で「編集モード」に変更すると`(Back Space)`で文字を消したり、キー入力で文字を追加することができます。

*16 `(T)`を押すとツールバーの表示、非表示を切り替えることができます。

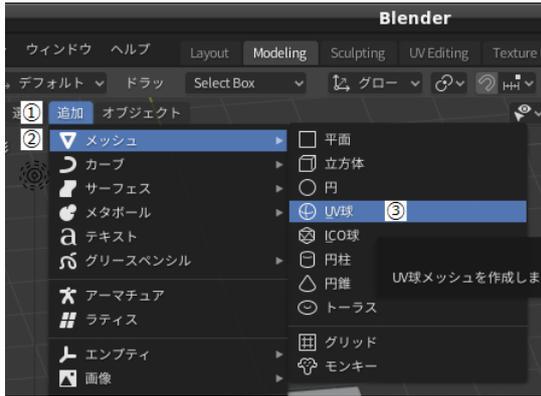


図 4.3 オブジェクトの追加



図 4.4 ツールバー

文字に漢字などが含まれる場合は、フォントを変更しないと表示されません。「オブジェクトモード」で文字オブジェクトを選択して、右のパネルのところで、図 4.5 の①の下の「a」をクリックし、②の「フォント」のところを開き、③の下のフォルダーボタンをクリックするとフォントファイルの指定画面になります。ここで、ディレクトリーを順番に usr、share、fonts、truetype とダブルクリックで開いていくと fonts-japanaese-mincho.ttf や gonts-japanese-gothic.ttf が出てくるので、このどちらかを選択すると日本語が表示されるようになります。



図 4.5 文字オブジェクトの設定

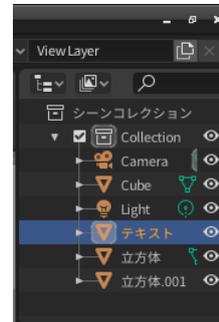


図 4.6 オブジェクトの名前の変更

図 4.5 の②の「フォント」の上の「ジオメトリ」を開いて、その中の「押出」に 0 より大きな値を設定すると厚みのある立体的な文字になります。文字間隔を変更する際には、図 4.5 の②の「フォント」の下の「段落」の中の「間隔」の中の「文字の間」で設定できます。

形状などの設定が終わったら、文字オブジェクトの形式をメッシュに変更します。これをしないと立方体などの他のオブジェクトとの結合などができませんし、3D プリンターで造形もできません。メッシュに直すには、文字オブジェクトを右クリックして出てきたメニューで、「メッシュ」を選択します。するとこれまで固まりだった文字オブジェクトが平面で囲われた形になります。一度メッシュにすると元に戻すことはできません。フォントや文字の変更が必要となったら、もう一度文字オブジェクトの追加からします。

4.3.3 オブジェクトの名前の設定

新しく追加されたオブジェクトには自動的に名前が付いています。機械的に立方体ならば「Cube001」のような名前が付くので、どんどん変形をして全く違う形になった場合識別する役に立ちません。オブジェクトをクリックして選択すると、画面の右上の図 4.6 のような「シーンコレクション」のところにある、対応するオブジェクトのところが白くなるので、名前のところをダブルクリックして変更します。

4.3.4 オブジェクトの大きさや位置などの変更

移動：図 4.4 のツールバーの③をクリックすると、選択したオブジェクトに赤、緑、青の矢印が出てきます。これを使ってそのオブジェクトの原点を x 軸、y 軸、z 軸の方向に動かすことができます。青の矢印の向いている方向が z 軸なので上方向になります。この矢印をドラッグするとオブジェクトが動きます。矢印の根本にある黄色い点をドラッグすると自由な移動は可能ですが、立体的に正しい位置へ持っていくのはなかなか難しいと思います。ツールバーをクリックしなくても、**G**を押すと移動モードになります。マウスを動かすとそれに従ってオブジェクトが動きます。最後にクリックすると移動モードが終わります。

回転：図 4.4 のツールバーの④をクリックすると、選択したオブジェクトに赤、緑、青の円弧が表示されます。この円弧をドラッグするとオブジェクトを回転させることができます。ツールバーをクリックしなくても、**R**を押すと回転モードになります。マウスを動かすとそれに従ってオブジェクトが回転します。最後にクリックすると回転モードが終わります。

拡大・縮小：図 4.4 のツールバーの⑤をクリックすると、選択したオブジェクトに赤、緑、青の先が丸い矢印が出てきます。この矢印をドラッグするとオブジェクトの大きさが変わります。矢印をドラッグした場合、その方向だけオブジェクトの大きさが変わります。そのためにオブジェクトの形状は歪みます。形状を変えずに大きさを変更したい場合は、矢印の外側にある白い円をドラッグします。ツールバーをクリックしなくても、**S**を押すと拡大・縮小モードになります。マウスを動かすとそれに従ってオブジェクトの大きさが変わります。最後にクリックすると拡大・縮小モードが終わります。

4.3.5 数字でオブジェクトの大きさや位置などを設定

オブジェクトを 3D プリンターで実物にする場合、その大きさの設定をマウスで適当に済ますわけには行きません。数字できっちり設定するためにはまず単位を設定する必要があります。そしてオブジェクトを選択して**N**を押すと右側に設定用のパネルが出てきます。そこで位置、回転、寸法を数字で指定できます。「拡大・縮小」のところは寸法を変更すると自動的に元の大きさに対する倍率が設定されます。逆にここで倍率を入力すると「寸法」のところの数字が自動的に設定されます。

なおこのパネルが不要になった場合は、もう一度**N**を押すと消すことができます。

4.3.6 オブジェクトの削除

オブジェクトを選択して、**Delete**を押すと、オブジェクトが削除されます。うっかり違うオブジェクトを削除してしまった場合は、**Ctrl**+**Z**で元に戻します。

4.3.7 オブジェクトの複製

オブジェクトを選択して、**Shift**+**D**でオブジェクトの複製ができます。できたオブジェクトは元のオブジェクトとぴったり重なっていますが、マウスをそのまま動かすと出てきます。適当な位置まで動かしてクリックすると位置が確定します。

4.3.8 オブジェクトの結合

Shiftキーを押しながらクリックすると、複数のオブジェクトを選択することができます。そして**Ctrl**+**J**で選択した複数のオブジェクトを一つのオブジェクトにすることができます。なお、この方法で結合したオブ

ジェクトは、あとで元のオブジェクトに分けることができます。

4.3.9 オブジェクトの計算

オブジェクトに丸い穴を開けたい場合、オブジェクトを後述の変形操作で丸く削るのは困難です。一方丸い穴の形状は通常円筒で、この形のオブジェクトは簡単に作ることができます。よって元のオブジェクトから円筒オブジェクトの部分を引き算すれば、丸い穴の空いたオブジェクトを簡単に作ることができます。このようにあるオブジェクトに、別のオブジェクトを加える(結合)、切り取る(差分)、重なり部分だけにする(交差)ことができます。

まず元のオブジェクトを選択し、右のパネルのところで、図 4.5 の①の 2 つ上の  (モディファイア) をクリックすると、**モディファイアを...**が出てくるのでクリックすると様々な修正の一覧が表示されます。その中から「ブーリアン」を選択すると図 4.7 のようになります。「演算:」の下で、「差分」、「結合」、「交差」を選択し、「オブジェ」の下にある「」をクリックするともう一つのオブジェクトが選択できます。最後に**適用**をクリックすると実行されます。結果は元のオブジェクトと重なっているなので、移動させないと結果が見えないと思います。

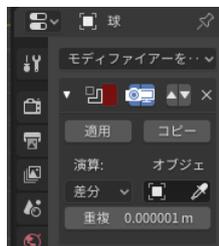


図 4.7 ブーリアンの設定



図 4.8 編集対象の選択

4.4 オブジェクトの変形などの操作

Tabで「オブジェクトモード」と「編集モード」の切り替えができます。「オブジェクトモード」ではオブジェクトの位置などのオブジェクト全体の操作ができましたが、「編集モード」ではオブジェクトの部分に対する操作ができます。オブジェクトを編集モードにすると、画面左上が図 4.8 のようになり、「編集モード」の横の 3 つのボタンで編集対象を点、線、面に切り替えることができます。図 4.8 は一番左が青くなっているので、点が編集対象になっています。ここでオブジェクトの適当な点をクリックして選択し、その点に対してオブジェクトの移動と同じ操作を行うと、選択した点のみが動くので、オブジェクトの形が変わる事になります。

4.4.1 オブジェクトの原点を底面に移動

3D プリンターで出力する 3 次元モデルは宙に浮いては困ります。またオブジェクト同士の位置をきちり数字で設定するような場合、オブジェクトの原点がわかりやすい位置にないと面倒です。オブジェクトの底面となる面を選択し、**Shift+S**で円環状のメニューが出るので「カーソル 選択物」を選択して 3D カーソルを面の中央に移動します。そして、**Tab**で「オブジェクトモード」に変更して、図 4.9 のように画面の左上の「オブジェクト」の横の方にある「オブジェクト」メニューの中の「原点を設定」をクリックし、「原点を 3D カーソルへ移動」を選択します。この場合原点が 3D カーソルの方にやってくるので、見た目は変わりません。**N**を押して位置の数字を全てゼロにすると、底面が床? に着いて正しく設定できたかどうかわかります。

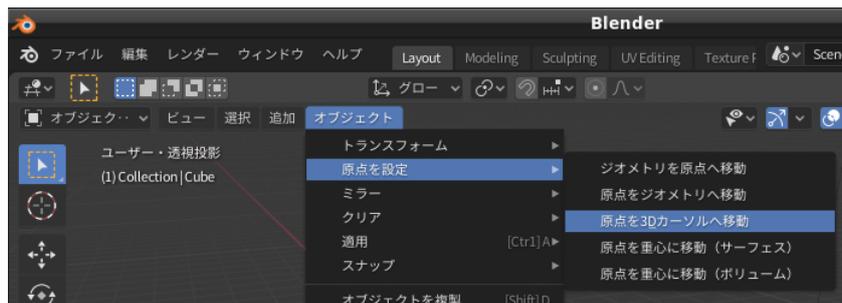


図 4.9 オブジェクトの原点の設定

4.4.2 面の細分化

元となるオブジェクトが少ない数の面で構成されていると、形を細かく設定することができません。そのような場合は面を細分化するとやりやすくなります。オブジェクトの細分化したい面をクリックして選択し、右クリックでメニューを出して「細分化」を選択します。この場合四角形に分割されますが、メニューで「扇状に分割」にすれば扇形に分割することもできます。

4.4.3 角を丸める

我々が手にするものは大抵角が丸く作られています。そうしないと手などに当たった時に怪我をする恐れがあるからです。オブジェクトの角を丸くするには、面や辺を選択した上で、**[Ctrl]+[B]**で「ベベル」という面取りの状態になります。マウスを動かすとどのくらいの幅を丸くするか(量)が変わり、ホイールを動かすと分割数が変わります。適当なところでクリックすると確定します。すると画面の左下に「ベベル」というものが現れるのでクリックするとパネルが開きます。パネルの一番上の「オフセット」の数字で元の角を押し込む幅、「セグメント数」の数字で分割数を設定することができます。

4.4.4 テクスチャの貼り付け

オブジェクトの表面に画像を貼り付けることによりオブジェクトの質感(ザラザラしているとか)や模様を付けることができます。この貼り付けられた画像のことをテクスチャと呼びます。ただ画像は平面ですが、オブジェクトの表面は多角形です。よって通常のオブジェクトにテクスチャを設定するには、そのオブジェクトの展開図を作成し、それに合った画像を用意するような手順となります。画像の作成は Blender で行うこともできますし、一旦別のソフトで編集したものを取り込むことも可能です。なおテクスチャはオブジェクトに直接結びつくのではなく、マテリアルに付く形になるので、もしオブジェクトにマテリアルが設定されていない場合は、まずマテリアルを設定する必要があります。

4.5 オブジェクトなどの保存

三次元モデルが完成したら、画面最上部にある「ファイル」メニューの中の「名前を付けて保存」を選択します。そして保存場所やファイル名を設定して保存します。拡張子は「.blend」が付きます。これはシーン全体の保存となりオブジェクトだけでなく照明やカメラなども含まれます。

3D プリンターで使用する STL 形式で保存する際は、クリックして保存するオブジェクトを選択してから、「ファイル」メニューの中の「エクスポート」の中の「Stl (.stl)」を選択します。拡張子は「.stl」が付きます。AR で使用する OBJ 形式も同様にオブジェクトを選択してから、「エクスポート」の中の「Wavefront (.obj)」を選択して保存します。拡張子は「.obj」が付きます。

4.6 プレートの作成

個々の操作の説明だけでは、なかなか実際のオブジェクトの作成方法は分かりません。ここでは図 4.10 のような名前プレートの作成を例にして、Blender でいかに三次元モデルを作成するかを説明します。なお () の間の数字はやり方を説明している章の番号です。

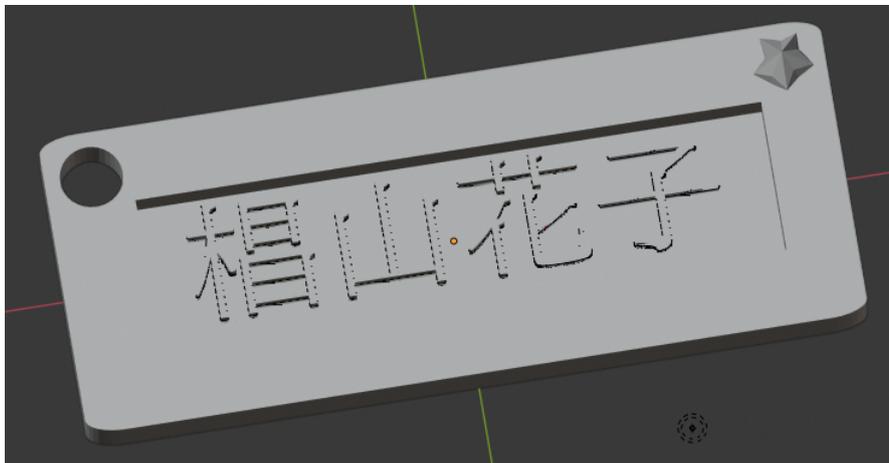


図 4.10 名前プレート

1. Blender を起動して、単位を mm に設定する。(4.2.1)
2. 最初からある立方体 (Cube) の大きさを 50mm×20mm×3mm にする。(各々 xyz の値)(4.3.5)
3. Cube の下の面の中央を原点にする。(4.4.1)
4. Cube の 4 つの角を丸くする。オフセットは 0.1mm、セグメントは 8 にする。(多少異なる値でも構わないが 4 つの角が全て同じ丸みになるようにする)(4.4.3)
5. 立方体を追加しその大きさを 40mm×10mm×4mm にする。(4.3.1)
6. 立方体を動かして Cube に重ねる。立方体の位置を $x=0.5\text{mm}$ 、 $y=0$ 、 $z=3\text{mm}$ にする。(4.3.4)
7. Cube から立方体と重なった部分を切り取る。(差分) そして立方体をずらしてちゃんと凹みがついていれば、立方体を削除する。(4.3.9)
8. 半径 2mm、長さ (深度)10mm 程度の円柱を追加する。これを Cube の左上の部分に重ねた ($x=-22\text{mm}$ 、 $y=7\text{mm}$) 上で、重なった部分を切り取る。うまく Cube に丸穴が空いたら円柱を削除する。(4.3.9)
9. 星形を付ける。半径 2mm、長さ (深度)1mm、頂点数 10 の円柱を追加する。円柱の上面の頂点を一つおきに内側に動かし星形にする。さらに上面を扇形に分割し、中心点を少し持ち上げて立体的な星形にする。これを Cube に埋め込んで結合させる。(4.4.2)
10. 文字オブジェクト (相山花子) を追加する。押し出しを 2mm にする。また文字の大きさは Cube の凹みに入り、丸穴や星形と重ならない程度にする。メッシュ形式に変換した上で $z=1\text{mm}$ にして Cube に埋め込み、結合し、文字オブジェクトを削除する。(4.3.2)
11. 名前を付けて保存する。(4.5)
12. stl 形式で保存して 3D プリンターで造形する。(宙に浮いた部分はないがサポートは追加する)(4.5)
13. obj 形式で保存して AR で表示する。(4.5)

5. 3D プリンター

通常のプリンターは平面の紙の上に字や絵を書きます。3D プリンターは立体を作ることができます。様々な方式のプリンターがあり、立体を作成するために使用する物質によってもできるものが変わります。この授業で使用する 3D プリンターは、FFF 方式^{*17}と呼ばれる比較的低い温度で柔らかくなるプラスチックを利用して、紐状の粘土で土器を作るように立体を造形します。色はプラスチックの色になりますので、通常は単色の立体になります。高級な 3D プリンターでは途中で別の色のプラスチックに切り替えることによって別の色の部分を作ることができますが、何種類もの色というわけには行きません。しかし、比較的構造が簡単で、プラスチックもそれほど高価ではないため、個人的な利用や立体の試作などに使われています。

用途によってはフィギュアのように色がないと色気がない場合もあります。そのような場合は石膏を利用したものが用いられます。石膏の粉に色付きののりを吹き付けて造形します。インクジェットのプリンターのインクにのりが混ざっている感じで、平面が一つ分描いたら、その上に石膏の粉を薄くまいてまた描くという繰り返しで造形します。最終的には石膏の中からのりで固まった立体を掘り出し、さらに石膏は柔らかいので表面を薄くコーティングします。これも石膏が比較的安いので安価ですが、プリンターの大きさはかなり大きくなります。

工業的に利用するものは、プラスチックや石膏では強度が足りない場合も多いので、金属による造形が可能な 3D プリンターが用いられます。チタンによるものは、石膏と同様な方式で造形が可能なのでチタンそのものは安くありませんが、一番安価に手に入るようです。ただしチタンの粉を固めるためにはのりでなく、レーザー光線で熱するため色は付けられません。それ以外の金属として、鉄、アルミニウム、銀、白金、金なども可能ですが、これらは一旦石膏などで型を作成し、そこへ溶けた金属を流し込んで作成するという手間のかかる方法を使用するため、時間や費用がかかります。

5.1 3D プリンターの使い方

mars には「Flash Forge Adventure III」という比較的低温で柔らかくなるプラスチックを利用した 3D プリンターが接続されており、三次元モデルさえあれば、簡単に 15cm×15cm×15cm より小さい立体を造形することができます。材質は ABS(アクリロニトリル・ブタジエン・スチレン) と PLA(ポリ乳酸) が使用可能で、現在は白色の PLA がセットされています。PLA は ABS よりも低温で柔らかくなり、固くて脆いので加工しにくい、塗料が付きにくいというような特徴があります。逆に言えば ABS の方が高温に耐えて、柔軟性があるので落としても割れにくいものことができます。以下に造形の手順を示します。

1. Blender で三次元モデルを作成し、上部の「ファイル」メニュー 「エクスポート」 「Stl (.stl)」と選択し、STL と呼ばれる形式で保存します。
2. 画面左下の「スタート」 「アクセサリ」 「FlashForge 3D Printer Control Software v5」(以下プリンター制御ソフトと呼ぶ) と選択します。
3. 「ファイル」メニューの「ロードファイル...」で STL 形式で保存したファイルを読み込みます。すると図 5.1 のようになります。
4. 「編集」メニューの「サポート材」で空中に浮く部分の足場を生成します。図 5.2 のような設定画面で **自動サポート** をクリックし、サポート材が生成されたら **もどる** をクリックします。
5. 画面の上部に出てくる **スライス実行** ボタンで立体モデルを輪切りにします。図 5.3 のような設定画面が出ます。「標準」、「細かい」、「速い」のどれかをクリックします。「細かく」にするとより綺麗な造形が可能ですが時間がかかります。「速い」にすると短い時間でできますが、荒いものになります。**スライス** ボタンをクリックすると輪切りを実行します。上部に **スライスのプレビュー** が表示されるのでこれをク

^{*17} Fused Deposition Fabrication、FDM 方式とも呼ばれる。

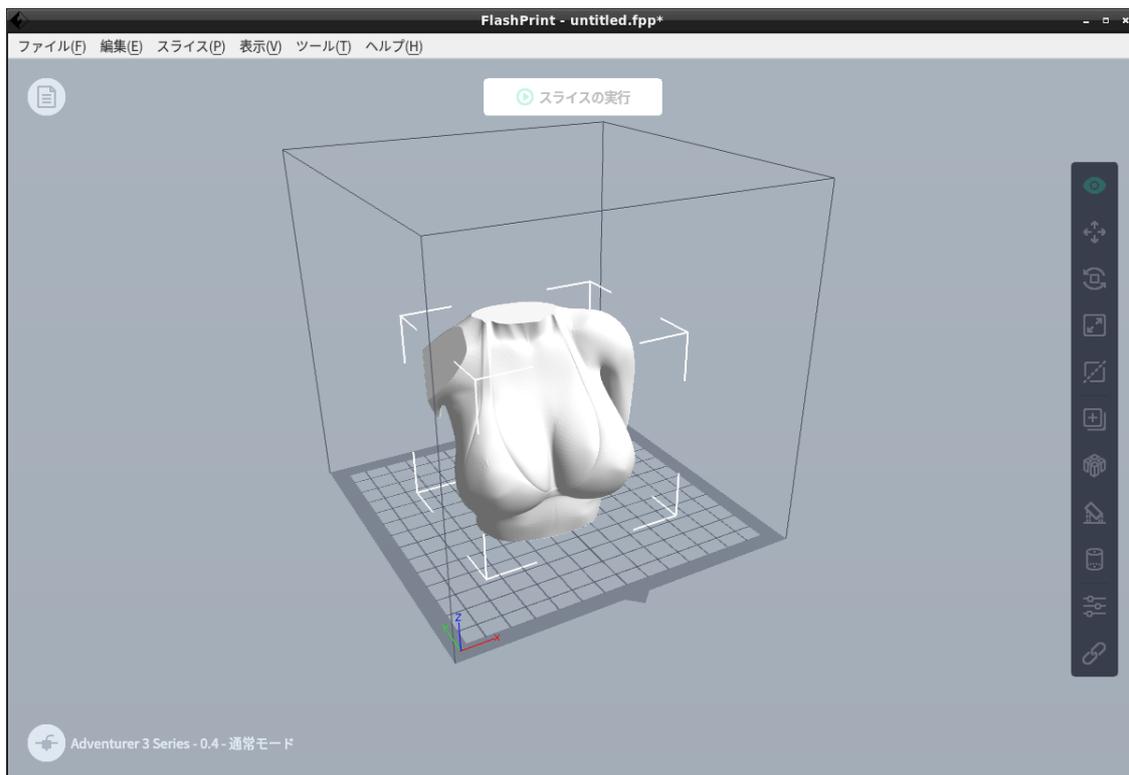


図 5.1 三次元モデルを読み込んだ状態



図 5.2 サポート材設定画面



図 5.3 スライスの設定画面

リックするとどのくらいの樹脂を使用するかや出力にかかる時間を見ることができます。

- 「スライス」メニューの「プリンターに接続する」を選択すると、接続先を尋ねてくるので、「自動スキャン」でプリンターを表示させて「全て接続」をクリックします。すると「マルチ制御機器」の画面に代わるので、プリンターのアイコンの前のチェックをクリックして、「G コードを送信」ボタン^{*18}をクリックします。すると「G コードファイル名」を尋ねてくるので通常そのまま「OK」をクリックします。
- 印刷が開始されると、ヘッドなどの温度や進み具合が表示されます。プリンターの動いている様子を、図 5.4 のように <http://mars.mgt.sugiyama-u.ac.jp/3dp.htm> で見ることができます。このページはパソコンやスマホからでも見えます。



図 5.4 3D プリンターのモニタ画面

- 印刷が終了したら、4 階の 414 室に行って造形された立体モデルを取り外します。取り外さないと次の人が出力できないので、終了しただけに取り外してください。立体モデルの乗っている黒いプレートの手前の部分を少し下に押しと手前にプレートを引き出すことができます。プレートを少しひねるとパリと言う感じで立体モデルが剥がれます。

複数の人が同時に 3D プリンターを使用した場合、まぜこぜの立体モデルができる事はありません。最初の人在使用中は、次の人は 3D プリンターに接続できません。造形が終わると次の人が 3D プリンターに接続できるようになります。そこで最初の人造形された立体モデルを取り外していない場合、3D プリンターヘッドが衝突してモデルか 3D プリンターが壊れることとなります。よって印刷開始の前には、必ず 3D プリンターの内部を確認してください。

^{*18} G コードは元々工場で行われる工作機械の制御に使われるものです。工作機械では金属の塊から作りたいものを削り出して作りますが、3D プリンターでは逆に積み重ねていく感じで作るため、ヘッドの動きなどが外側からでなく内側からと言うように逆になります。

5.2 3D プリンター使用上のヒント

インターネットで検索すればたくさんの STL 形式のファイルを見つけることが可能です。これらを頂いてきて、3D プリンターにかけることも可能ですが、大きさがプリンターの造形可能なサイズを超えている事もあります。そういう場合は、ファイルを読み込んだあとで、画面の右側にある縦に並んだボタンの上から 4 番目の「スケール」ボタンで大きさを変更することができます。

この授業で使用する 3D プリンターは、立体モデルの下の方から造形します。途中では宙に浮いてしまう部分があるとサポート材による支えが必要になります。全て造形が済めば宙に浮いた部分もつながるので、サポート材を物理的に切り離して完成となります。複雑な立体モデルでは内部にこのサポート材が必要となり、後でそれを除去できなくなるという事もよくあります。支柱の部分に使われたプラスチックは無駄になりますし、除去の手間を考えるとできるだけサポート材は少ない方が良いでしょう。例えば蓋のない箱の場合、蓋のない面を上に向ければサポート材は不要ですが、蓋の無い面を下にすると箱の中はサポート材だらけになります。この辺りを考慮して立体モデルの作成の際に向きを決めましょう。

画面の右側にある縦に並んだボタンの上から 5 番目の「カット」を利用して立体モデルを分割することができます。例えば球はどのような向きにしても、床面と点で接触するのでサポート材が必要になります。しかし半分に切り分けて断面を下にすれば、単なる丸い山形ですからサポート材は不要です。造形した後で貼り合わせる必要が生じますが、サポート材を外すのよりは楽でしょう。

6. 拡張現実

6.1 拡張現実 (Argumented Reality) とは何か

仮想現実 (VR: Virtual Reality) というものがあります。コンピュータが作り出した仮想世界をゴーグルの形をしたディスプレイ (HMD: Head Mounted Display) で見るもので、頭や体の動きに合わせて見せる仮想世界を動かすので、本当にその世界の中に居るような感じになります。コンピュータの性能向上のおかげで本当に綺麗な仮想世界を見せることができるようになりました。一方問題点は次のとおりです。

- 綺麗な仮想世界を作るためには膨大な三次元データが必要となる。見渡す限りに存在する全ての物を用意する必要があります。全て空想の世界であれば、単純な立体で構成することも可能でしょうが、現実に近い世界を作るとなると細かいところまで忠実な三次元データを作成しなければなりません。
- 頭や体の動きと仮想世界の動きにはどうしても遅れやズレがあり、その結果長時間使用すると酔いのような症状をもたらすことがあります。

これに対して拡張現実 (AR: Argumented Reality) は現実世界の画像の中に、コンピュータが作り出した仮想物体を表示するものです。例えば「ポケモン GO」では、スマホのカメラで写し出した現実の風景の中に、コンピュータの作り出したポケモンが登場します。周りのものや背景は全て現実のものを借りるので、仮想物体だけ用意すれば済みますので VR よりも楽です。また VR のように頭や体の動きの検出の必要もありません。もちろん全く現実と異なる世界を提供することはできません。現実世界のどこに仮想物体を出現させるかによって AR は大きく 2 つの方式に分かれます。

- マーカーと呼ばれる図のあるところに出す方式：マーカーのあるところを基準にして、大抵その上に仮想物体が居るように表示されます。マーカーを動かしたり、マーカーに近寄るとその動きに合わせて仮想物体が動いたり大きさが変わったりします。マーカーが必要ですが、マーカーさえあればどこにでも仮想物体を出すことができます。現在はマーカーとして普通の絵も利用可能です。例えば「日経 AR」は日経新聞記事中のグラフや写真をマーカーとしています。
- GPS など求めた位置を基準に仮想物体を出す方式：「ポケモン GO」がこの方式で、おかげでポケモンを探してあちこち歩く必要があります。現在の GPS では位置の精度には数 m の誤差がありますので、出現する位置もその影響を受けます。また屋内などの GPS の電波が十分届かない場所では使えません。GPS に頼らず、スマホの向きとカメラに写った内容から地面を検出して位置を特定して仮想物体を出すことも可能になりました。こちらはスマホからの相対的な位置にマーカー無しで仮想物体を出すことができます。

両者を活かしたものもあります。2020 年 10 月に発売された「マリオカートライブホームサーキット」では無線操縦で模型のマリオカートを室内で走らせます。マリオカートにはカメラが付いていて、ディスプレイでマリオカートの前方の風景を見ることができます。ディスプレイには現実の室内にある家具などに加えて、競争相手や障害物、床の上にはコースのラインなどが表示されます。さらに組み立て式のゴールなどがマーカーになっています。

この演習では大量の三次元データや HMD を必要としない AR を取り上げます。またマーカーさえあればどこでも使える方式を採用します。作成した三次元モデルを 3D プリンターを使えば現実の物にすることができましたが、AR を使えば三次元モデルが現実にあるように見せることができるようになります。

6.2 AR テストシステムの使い方

スマホのアプリで AR を検索すると多数のアプリが出てきます。それぞれ出てくる仮想物体が違います。要するに表示される物の種類だけアプリがあります。Blender で自分が作成した三次元モデルを出せるようなス

マホアプリを作りかけたのですが、間に合わなかった(できなかった?)のでブラウザで見るシステムを借りました。以下のような手順で使います。なお、単純な図形の表示で良ければ obj 形式の三次元モデルなしでも使えます。

1. Blender で三次元モデルを作成する。

マーカーとの位置関係を設定する際に、ややこしくならないように原点の位置をわかりやすい外部の点に設定します。実は obj 形式のファイルには単位情報が含まれません。よって Blender で単位をメートルにして作成することを仮定しています。

一方、3D プリンターでの出力を想定して単位を mm にして作成したモデルは、後から単位をメートルに変更しても、内部での扱いが変わらないので、次の obj 形式で保存を行うと、大きさが千倍のものとして保存されてしまいます。これを避けるには、一旦「ファイル」メニューの中の「エクスポート」にある「Collada (デフォルト) (.dae)」形式で保存します。そして Blender を起動し直して最初状態で、立方体を削除し、単位をメートルにしてから Collada 形式で保存したものを「ファイル」メニューの中の「インポート」の中の「Collada (デフォルト) (.dae)」で取り込みます。

2. 三次元モデルを obj 形式で保存する。

- a. 対象をクリックして選択する。
- b. 「ファイル」メニューの中の「エクスポート」の中の「Wave front (.obj)」を選択する。
- c. 保存場所とファイル名を指定してから「OBJ をエクスポート」をクリックする。これで指定した「~.obj」ファイルと同じところに「~.mtl」ファイルが保存される。

3. AR テストシステムへブラウザからアクセスする

「https://mars.mgt.sugiyama-u.ac.jp/AR/」へアクセスすると図 6.1 のように表示されます。



図 6.1 AR テストシステムの最初の画面

4. 「新規登録」をクリックすると図 6.2 のように表示されますので、「タイトル」、「パスワード」を入力し、「マーカーの種類」は適当に選択し、「マーカーの大きさ」は作成したモデルと近い大きさを入力します。「ターゲットの種類」は必ず「Blender などで作成した obj ファイル (要アップロード)」を選択します。さらにその下で、「ターゲットの 3D モデルの OBJ ファイル」と「ターゲットの 3D モデルの MTL ファイル」を指定します。全て入力し、選択し、指定できたら「新規登録」をクリックします。

ARデータの登録

AR用マーカー

https://mars.mgt.sugiyama-u.ac.jp/AR/new.htm

ARデータの登録

マーカーやそれに対して表示するものを指定してください。パスワードは登録内容の修正や削除の際に必要になりますので、何を入力したか忘れないでください。入力が終わったら「新規登録」をクリックしてください。 [新規登録せずに戻る](#)

ページタイトル	プレートテスト	(ARのWebページのタイトル)
パスワード	9999	(修正や削除の際のパスワード)
マーカー	マーカーの種類: hiro	マーカーの大きさ: 4 cm
	マーカーの画像ファイル: 参照...	ファイルが選択されていません。
ターゲット	ターゲットの種類: Blenderなどで作成したobjファイル(要アップロード)	
	A-Frameのコマンド: (ターゲットがobjファイル指定の場合は追加設定など)	A-Frameコマンド例
	OBJファイル: 参照... untitled.obj	MTLファイル: 参照... untitled.mtl
	IMGファイル: 参照...	ファイルが選択されていません。(必要があれば)

[新規登録](#)

by K.Miki 2019/09/03, 09/13, 09/17, 2020/11/15

図 6.2 新規登録の画面

5. 「新規登録」をクリックすると図 6.3 のように URL や QR コードが表示されるので、これらをもとにスマホやタブレットでアクセスします。
6. タブレットやスマホでアクセスすると、背面カメラによる映像が表示されるので、カメラをマーカーの方に向けると図 6.4 のように三次元モデルが出現します。映像が表示される前にカメラの使用の許可を求めてきた場合は、毎回許可しないと表示されません。

6.3 ブラウザによる AR システム

予め AR アプリをインストールする必要のないブラウザによる AR システムは、手軽な AR システムとしてもっと普及しても良いのではないかと思います。ただ全ての処理を JavaScript で行っており、マーカーの検出などの処理に不安を感じる他、毎回 Web サーバーから全てを受け取らなければならないため、時間がかかったり、ネットワークのないところでは使えなかったりします。

前節で使い方を説明した AR テストシステムは、Jerome Etienne 氏の 2017 年 7 月 11 日のブログの記事「Creating Augmented Reality with AR.js and A-Frame」^{*19}をもとにしたものです。「A-Frame」は仮想現実の世界を HTML のようなタグで記述し、ブラウザでそれを見ることができるようになろうとする試みです。

*19 <https://aframe.io/blog/arjs/>



図 6.3 登録結果の画面

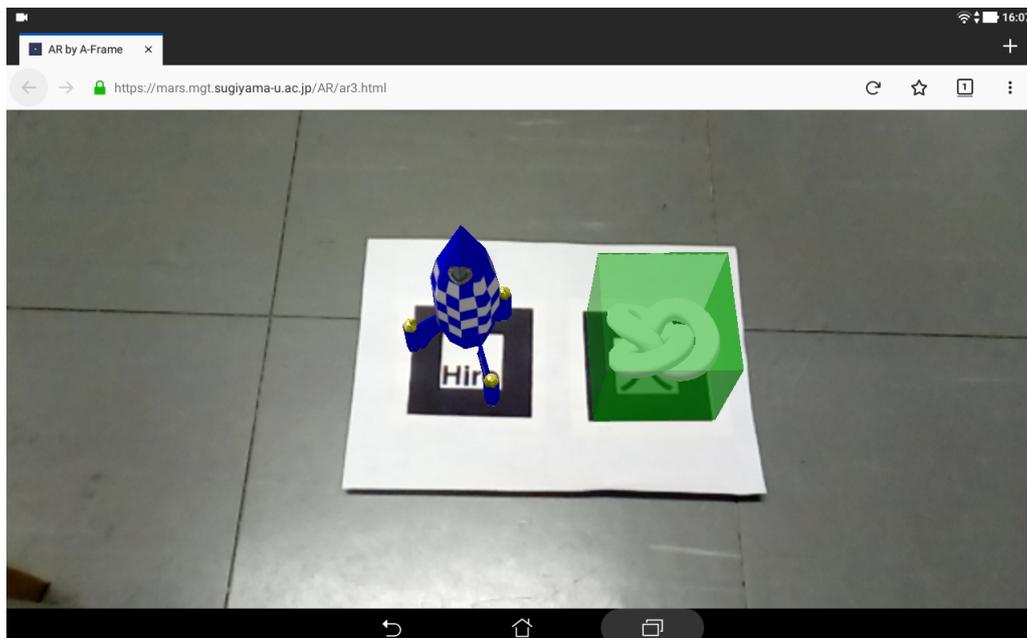


図 6.4 タブレットでの表示

例えば次のような記述のみで、水色の立法体、赤い球、黄色の円筒が薄い水色の面の上にあるように表示されるだけでなく、マウスでドラッグすると見る方向を変えることもできます*20。

```
<html>
  <head>
    <script src="https://aframe.io/releases/1.0.4/aframe.min.js"></script>
  </head>
  <body>
    <a-scene>
      <a-box position="-1 0.5 -3" rotation="0 45 0" color="#4CC3D9"></a-box>
      <a-sphere position="0 1.25 -5" radius="1.25" color="#EF2D5E"></a-sphere>
      <a-cylinder position="1 0.75 -3" radius="0.5" height="1.5" color="#FFC65D">
      </a-cylinder>
      <a-plane position="0 0 -4" rotation="-90 0 0" width="4" height="4"
        color="#7BC8A4"></a-plane>
      <a-sky color="#ECECEC"></a-sky>
    </a-scene>
  </body>
</html>
```

JavaScript を使用して独自のタグが使えるようにしています。そしてこれで記述した三次元モデルが、カメラで撮影したマーカーの位置に表示されるようにするのが AR.js です。こちらも独自のタグでマーカーの種類などを指定します。以下はこれらの独自のタグの使い方の一部を紹介します。

まず単独の赤い箱です。

```
<a-box position='0 0.5 0' material='opacity: 0.5; side: double; color: red;'>
</a-box>
```

position で箱の位置を指定し、material ではこの材質を指定しています。opacity は不透明度で数字が大きくなるほど中は見えなくなります。

次は青い箱の中で紐がぐるぐるするものです。

```
<a-box position='0 0.5 0' material='opacity: 0.5; side: double;color: blue;'>
  <a-torus-knot radius='0.26' radius-tubular='0.05'
    animation='property: rotation; to: 360 0 0; dur: 5000;
    easing: linear; loop: true'>
  </a-torus-knot>
</a-box>
```

a-box のタグの間に紐のタグである a-torus-knot が入っています。radius は半径のことでこれで紐の大きさを設定します。animation で動きの設定をします。property で動き方を回転 (rotation) としています。to でどこまで回転するかを指定しています。三次元なので回転軸も3つあるので数字が3つ並んでいます。dur は回転周期ですので、小さな値にすると早く回転します。loop が true になっているのでずっと回転します。

OBJ 形式の三次元のモデルを動かす場合は、a-obj-model のタグに animation の指定を追加します。

6.4 テクスチャ付きモデルの作成方法

テクスチャとはモデルの表面の模様のことです。本来は3章で説明すべきことですが、Blender でテクスチャ付きモデルの作成方法をここで説明します。現在本学部にある3Dプリンターでは色付きのモデルを造形

*20 <https://mars.mgt.sugiyama-u.ac.jp/AR/aframe.htm> へアクセスすると見ることができます。

することができませんが、AR では簡単に表示することができます。むしろ色ぐらい付いていないと寂しいのと、立体モデルとしては簡単な形状だがテクスチャによっては AR で使えるものも考えられるからです。

モデルの材質を示す材料には色の設定があるため、これを利用すれば色付きモデルを作成することができます。ところが同じ材料の部分と同じ色になるので細かい模様のようなものは表現できません。そこで色付きのモデルでは大抵材料の上に模様であるテクスチャを貼り付けています。大抵のフィギュアの顔の部分も目玉などはテクスチャで描いています。テクスチャのデータとなる画像は普通の平面画像です。これを立体に貼り付けるために結構面倒な手順が必要になります。特に顔のような凹凸が色々ある場合、展開図を作成してそこに色を付けるような形になります。ここでは一番簡単な平面にテクスチャを貼り付ける方法を説明します。

1. 貼り付ける画像を用意する。画像は PNG 形式のものとし、後で OBJ 形式のファイルを保存するところ、例えばデスクトップに置く。
2. Blender を起動し、初期設定の立方体を削除する。単位はメートルでなければメートルに変更する。
3. 21 ページの図 4.3 のようにメニューを開き、最後に「UV 球」でなく「平面」をクリックして平面を追加する。
4. N を押して寸法設定画面を呼び出して大きさを設定する。画像の元となったものと同じ大きさにする。それが無理な場合は、せめて縦横比は画像と同じにしないとテクスチャが歪む。
5. 右側のところで図 6.5 の①をクリックして材料の設定を出して、②の「新規」をクリックして材料を追加する。

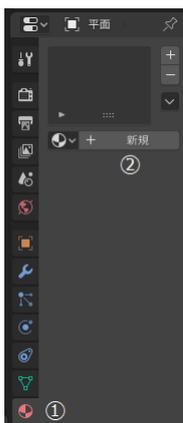


図 6.5 材料の追加



図 6.6 材料に画像を設定



図 6.7 材料プレビューモード

6. 材料を追加するとさらに様々な設定ができるようになる。図 6.6 の赤の矢印の先にある「ベースカラー」の をクリックするとメニューが表示されるのでその中の「画像テクスチャ」を選択する。するとフォルダの形をしたボタンが現れるので、それをクリックして画像ファイルを指定する。
7. 画面右上の図 6.7 のような表示のメニューを、赤の矢印の先にある「材料プレビューモード」をクリックして変更すると、貼り付けられた画像の様子が見えるようになる。
8. 「ファイル」メニューの中の「エクスポート」で OBJ 形式で保存する。保存先は貼り付けた画像があるディレクトリ (フォルダ) にする。
9. AR システムに登録する際には、OBJ ファイル、MTL ファイル、画像ファイルの 3 つを必ず指定する。

6.5 マーカー例

以下はマーカーの例です。



図 6.8 4cm 角 (内側は 2cm 角)



図 6.9 12cm 角 (内側は 6cm 角)

演習問題

以下のようなものを AR で実現せよ。ただし誰が登録したものかわかるように、タイトルに自分の名前か id を追加すること。

1. 黄色の箱の中で青い箱がぐるぐる回るもの
2. 前章で作成した名前プレートが、ゆっくり回転して裏も見えるようなもの
3. 有名絵画を実物大で見ることができるもの



図 6.10 19cm 角 (内側は 9.5cm 角)

7. 深層学習

この章では近年流行の AI (Artificial Intelligence: 人工知能) でよく用いられている深層学習 (Deep Learning) について学びます。

7.1 コンピュータによる問題解決

コンピュータは、何かを入れると何かを返してくれる機械、と言う捉え方をすることができます。パソコンならばキーボードやマウスで指示を入れます。スマホならばタッチして指示します。お返しは画面への表示であったり、音声だったりします。入れたものに対して、出てきたもの間には何らかの関連があります。この関連を表すものとして代表的なものがアルゴリズムです。最大公約数を求めるアルゴリズムである「ユークリッドの互除法」が有名ですが、このアルゴリズムに従えば必ず最大公約数を求めることができます。コンピュータの動き方を決めるプログラミングは、このアルゴリズムをコンピュータで実行できる形に書き直す作業です。

どのような問題に対しても解決するアルゴリズムが存在するのだろうか？という疑問はかなり昔に理論的に否定されています。つまり理論的にアルゴリズムが存在しない問題が世の中にはあります。また良いアルゴリズムが見つからない問題、一応アルゴリズムはあるのだが、ものすごく時間がかかる問題などが多数あります。とは言うものの、コンピュータに人の代わりに働いてもらおうとすると何らかのアルゴリズムが必要です。

AI はその全てがそうではありませんが、アルゴリズム無しで問題の解決をします。アルゴリズムの代わりにデータで学習して、コンピュータが自分で問題を解く方法を発見します。人の神経組織をまねた構造を作り、人の学習のまねをするという発想はかなり昔からありました。ただこの学習がなかなかうまく行かない、学習のために膨大なデータと膨大な計算量が必要ということで、長い間お蔵入りしてました。最近になってコンピュータの計算力が超強力になり、効率よく学習させる方法が見つかってきたのでこの技術が復活して、様々な分野で使われようとしています。

AI による解決の良い点は、アルゴリズムを考えなくて良いところです。問題点は、

- 適切な学習用データが大量に必要
- 結果がきっちり出るわけではない
- なぜその結果になるのかわからない

と言うような点です。逆にアルゴリズムの良い点は、結果がきっちり出ると、なぜその結果になるのかが明確なところです。AI を利用したシステムは極端な言い方をすれば、独断と偏見で結果を出します。よって従来のアルゴリズムによって動いているシステムと同じように考えると危険です。いくら正しい結論が出ても、なぜそのような結論になるのかわからない場合、結論を出したのが人であれば疑いを持つ人が出ます。一方コンピュータが出したとか AI が出したとなると、信じる人がたくさん出てきます。よく SF に出てくるコンピュータに支配された未来世界は、もうすぐ実現するかもしれません。

7.2 機械学習

AI が学習するためのデータは、「教師あり」のもの、「教師なし」のものに分けられます。大抵は「教師あり」です。これは正解付きデータと言っても良いでしょう。教師ありデータを学習して、学習していないデータに対して正解が答えられるようになることを目指します。人の勉強とは違って、AI は学習の際に使用したデータに対しては必ず正解を答えます。学習していないデータに対してどのくらい正しい答えを出せるかが評価のポイントになります。

「教師なし」の例としては、2012 年に Google がインターネット上の様々な写真画像を処理した結果、猫を認識できるようになった、と言うものがあります。これは個々の画像に対して「猫である」とか「猫でない」と

いう情報を与えることなく、猫を識別できるようになりました。もちろん AI が行ったのは単なる画像の分類です。その分類された画像グループの一つを人が見たら猫ばかりだった、ということです。個々のデータに対して正解を付ける作業は大変な手間がかかることが多いので、教師なしの学習は困難ですが様々な研究が行われています。Google の作った Alpha GO^{*21} というシステムが囲碁の世界チャンピオンに勝ったと言う話は有名なのでご存知だと思います。この世界チャンピオンに勝ったシステムは、過去の対戦記録を 3,000 万局分も学習していたそうです。ところが現在このシステムにポロ勝ちをするシステムがあります。囲碁などのゲームは勝ち負けが明確です。その性質を利用して、AI システム同士で対戦させて勝ち負けを自分らで判定して学習するようにして、過去の対戦記録なしに大量の学習を行った結果です。このようにシステムが正解を判定できるようにして実現した教師なし学習も有望です。

7.3 ニューラルネットワーク

ニューラルネットワークは、図 7.1 のような人などの生物の神経組織をモデル化したものです。神経細胞（ニューロン）は、樹状突起と呼ばれる部分で他の神経細胞などからの信号を受け取り、数 mm ~ 数十 cm の細長い軸索で他の細胞に信号を送り出します。神経細胞では化学物質を利用して信号伝達を行いますが、これをモデルにしたニューラルネットワークでは数値を使います。神経細胞は通常複数の細胞から信号を受け取りますが、どの細胞からの信号も同じ扱いにしている訳ではありません。同様にニューラルネットワークでも、受け取った数値に適当な数を掛け算してから合計します。その結果を活性化関数というもので計算した結果を次へ送ります。ニューラルネットワークでの学習の実体は、この掛け算する数（係数）を求めることです。

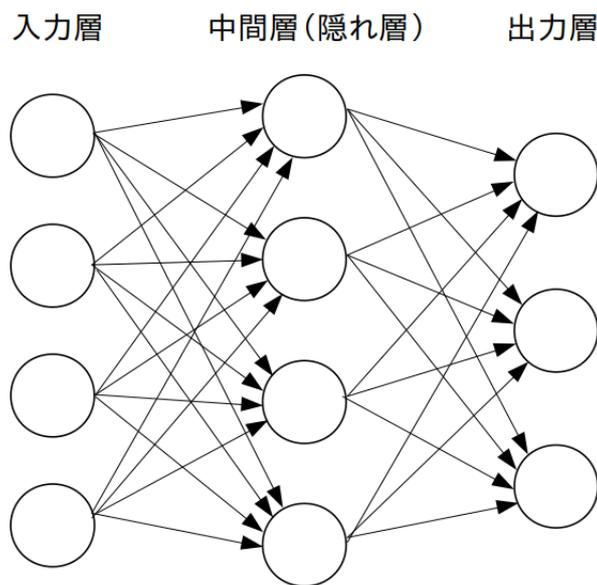


図 7.1 ニューラルネットワーク

図 7.1 では、入力層に 4 つのニューロンがありますが、この数は入力の数に対応しています。数が多い例としては、画像の入力があります。我々が目で物を見るとき、目の中の視神経が見えているものを細かく分割して脳に送っています。同様にニューラルネットワークで物体の認識をさせる場合は、画素ごとにニューロンを割り当てます。800 万画素のカメラにそのまま対応させようとする入力層に 800 万個のニューロンが必要となりますが、さすがに多すぎるので 100×100 程度に間引いて扱うのが普通です。処理の都合から考えると少

^{*21} 私が恐ろしいと思ったのは、Google はこの囲碁プログラムの開発に 900 億円かけたという話を聞いたときです。それまでに囲碁対戦プログラムの研究や開発にかけられた費用を遥かに超える額を、世界チャンピオンに勝てるめどもないに出すと言う Google に、勝てる日本の会社はあるのだろうか。

ないほどよいのですが、間引きしすぎると何の画像か人が見てもわからなくなります。人が見てもわからないものは、ニューラルネットワークで処理してもたいてい識別できません。

図 7.1 では、中間層は 1 段で 4 つのニューロンから構成されていますが、ここを何段にするか、ニューロンを何個にするかは決まっています。段数や個数を増やすと処理に時間がかかり、学習がうまく行かない可能性が高くなりますが、段数や個数が足りないといくら学習させても正答率が上がりません。似たような事例を参考に段数と個数を決めてやってみて、だめならば少し変更してやってみるを繰り返すしかありません。

図 7.1 では、出力層に 3 個のニューロンがありますが、これは結果の種類の数で決まります。例えば学生さんの顔写真を入れて、可愛さのみを求めるならば、出力層はニューロンが 1 個になります。同時に美しさ、賢そうか、優しそうかも求めるならば 4 個になります。

図 7.1 のニューラルネットワークには矢印が 28 本 ($= 4 \times 4 + 4 \times 3$) あります。この矢印の数だけ係数があり、学習によってこれを求めなければなりません。残念ながら二次方程式のような解法はありません。とりあえず係数に適当な数値を設定してやってみて、少しずつ正しい答えに近づくように修正していくと言うような解法になります。山登りに例えると、足元の地面を見てより高い方へ向かって進んで行けば、そのうち山頂にたどり着けるだろうと言う感じです。繰り返してもあまり改善が見られなくなったところで学習終了とします。残念な結果になったら、データを増やすとか中間層の構成を変更することになります。データや中間層の構成に問題がなくても失敗することがあります。山で遭難しそうになったら、麓 (ふもと) を目指して低い方へと進むのですが、盆地にたどり着くとどちらを向いても上り坂になってしまい、進む方向がわからなくなります。同様なことがニューラルネットワークの学習にも生じることがあります。現実の山と違って地図もないので盆地にハマっているかどうかすら通常はわかりません。

この問題に対して例えば確率勾配降下法と言う手法を用いて盆地にハマらないようにしています。学習用データをランダムに小分けして学習させる方法です。全員で行き先を決めると、行き先が平均化されて盆地にハマるかもしれないが、多数の小グループであればその中にはとんでもない方向へ行くグループもあるだろうし、もしかすると麓への道を見つけ出すグループが出てくるかもしれないと言う感じです。

画像認識などでより良い認識率を実現するために畳み込みニューラルネットワーク (CNN) というものも考えられました。確か猫で行われた実験だったと思いますが、生まれたときから縦縞ばかりの環境で育てたところ縦縞にしか反応しなくなり、頭の中の神経を調べたところ縦縞のみに反応する部分が見つかったが、横縞に反応する部分はなかったというものです。我々は様々な形のものを見て育ちますから、頭の中には縦縞に反応する神経だけでなく、横縞や角などの各種部分に対応する神経があるでしょう。そういうものの組み合わせとして物を認識している可能性があります。畳み込みニューラルネットワークでは画像から複数のフィルターを使って複数の特徴的な部分を探しその結果を利用して学習を行います。またどのような特徴が有効なのかは分かりませんので、並行してフィルターの内容も学習します。

英語を日本語に翻訳させる、株価の予想をさせたい、と言うような場合も図 7.1 のようなニューラルネットワークでは対応が難しくなります。明日の株価を予想させるために、前日の株価だけではなかなか当たりません。ある程度の幅 (例えば一週間) を持った過去の株価を元に予想した方が良いでしょう。そして、単純に過去 7 日分の株価データではだめで、株価の推移も考慮してもらわないといけません。英語や日本語もどの単語があったでは意味が取れません。同じ単語でも先頭に出てくれば大抵主語ですので語順は重要です。再帰的ニューラルネットワーク (RNN) では、7 日分の株価から明日の株価を予想する場合、図 7.1 のようなニューラルネットワークを 7 つ用意する代わりに、一つのニューラルネットワークに日にちと株価のセットのデータを与えて学習させます。このやり方ならば 100 日前の株価からだって入れられますが、あまり欲張ると学習できなくなります。

Deep Q Network (DQN) は強化学習用のニューラルネットワークです。強化学習は教師なし学習の一種と言われていますが、例えば迷路から抜け出す道筋を学習するような場合に使われます。現在位置の様子を見て、前進・後退・右へ・左へと選択をします。そのうち出口にたどり着いたらこれが正解として学習します。一番遠回りの道を選ぶようになっては困りますから、より短いコースを選択するように学習させます。囲碁の

学習も同様に強化学習で行いました。現在これを車の自動運転に利用することが検討されています。車が自分の周囲を見て、ちゃんと道路の上を走るようにするためです。

7.4 様々なフレームワーク

深層学習のために様々なフレームワーク (枠組み) が作られています。フレームワークを利用することにより容易に深層学習を行うシステムを作ることができます。代表的なものとして、Google が作った Tensor Flow、Amazon が作った MxNet、Microsoft が作った CNTK、Preferred Networks が作った Chainer などがあります。どのフレームワークを使うにしろ Python というプログラミング言語に関する知識が必要になります。Python 自体は深層学習のためのプログラミング言語と言うものではなく、普通のプログラムが書けるものです。将来 AI システムを作るお仕事をするならば Python を勉強しておくべきですが、他のプログラミング言語とは少し書き方が違うので、将来未定の人に勧めるのはどうかなあ、と現在は思っています。

有名な会社がそれぞれ別のフレームワークを打ち出しているのは、AI の中核とも言える深層学習での覇権を狙っているからです。多くの利用があれば、多くの使用例が生まれ、大学などの授業で使用例が紹介され、そういう授業を受けた学生が会社で同じようにやってみようとする訳です。このテキストでは Chainer を使用することにしました。他のフレームワークが超有名会社が作っているのに対して、Preferred Networks なんて聞いたことがない会社でしょうが、実はこの会社は日本の会社です。よってトヨタ自動車など日本のメーカーや大学などとの共同開発の例などがあるようです。こうして令和元年度は国産のフレームワークに賭けてみたのですが、その後 Chainer の開発が終了しました。来年度は別のフレームワークにしようと思います。

7.5 画像認識プログラムの例

株式会社明治が 40 年以上前から販売している「きのこの山」というお菓子があります。姉妹品として「たけのこの里」というお菓子もあります。この両者を見分けるのは幼児でも簡単な問題ですが、コンピュータではなかなか面倒なプログラムを書かなければなりません。この 2 種類のお菓子を深層学習で学習させたモデルを利用して見分けるプログラム^{*22}の例を元に、小物の文房具 4 種類を見分けるものに直したものがリスト 1 です。

リスト 1 見分けるプログラムの例

```
1 # -*- coding: utf-8 -*-
2 # CQ 出版「インターフェース」2017 年 8 月号 p.41-43 のものを一部改変
3 import argparse
4 import os
5 import numpy as np
6 from PIL import Image
7 from PIL import ImageEnhance
8
9 import chainer
10 import chainer.functions as F
11 import chainer.links as L
12 import chainer.initializers as I
13 from chainer import training
14 from chainer.training import extensions
15
16 # 畳み込みニューラルネットワークの定義
```

^{*22} オリジナルは、CQ 出版社「インターフェース」2018 年 8 月号 p.41-43 の牧野浩二、西崎博光による「ラズパイでディープ・ラーニング初体験 (2)・・・画像認識」による。入力部分など一部改変している。

```
17 class CNN(chainer.Chain):
18     def __init__(self, n_units, n_out):
19         w = I.Normal(scale=0.05) # モデルパラメータの初期化
20         super(CNN, self).__init__(
21             conv1=L.Convolution2D( 3, 16, 5, 1), # 1層目の畳み込み層
22             conv2=L.Convolution2D(16, 32, 5, 1), # 2層目の畳み込み層
23             conv3=L.Convolution2D(32, 64, 5, 1), # 3層目の畳み込み層
24             l4=L.Linear(None, n_out, initialW=w), # クラス分類用
25         )
26
27     def __call__(self, x):
28         # 最大値プーリングは 2x2, 活性化関数はReLU
29         h1 = F.max_pooling_2d(F.relu(self.conv1(x)), ksize=2, stride=2)
30         h2 = F.max_pooling_2d(F.relu(self.conv2(h1)), ksize=2, stride=2)
31         h3 = F.max_pooling_2d(F.relu(self.conv3(h2)), ksize=2, stride=2)
32         # 9x9, 64ch
33         return self.l4(h3)
34
35 # 画像の大きさを指定サイズに切り詰める
36 def crop(img, size):
37     w, h = img.size
38     assert w >= size or h >= size, "画像サイズが小さすぎます"
39     if w > h :
40         p = (w - h) / 2
41         box = (p, 0, p+h, h)
42     else :
43         p = (h - w) / 2
44         box = (0, p, w, p+w)
45     return img.crop(box).resize((size, size))
46
47 # 画像の4つの角の明るさを元に、全体の明るさを調整する
48 def doBright(img):
49     imgrgb = img.convert('RGB')
50     r,g,b = imgrgb.getpixel((0,0))
51     s=r+g+b
52     r,g,b = imgrgb.getpixel((99,99))
53     s+=r+g+b
54     r,g,b = imgrgb.getpixel((0,99))
55     s+=r+g+b
56     r,g,b = imgrgb.getpixel((99,0))
57     s+=r+g+b
58     er=1+(2400-s)/1600
59     return ImageEnhance.Brightness(img).enhance(er)
60
61 def main():
62     # オプション処理
63     parser = argparse.ArgumentParser(description='Chainer example: MNIST')
64     parser.add_argument('--model', '-m', default='model',
65                         help='Resume the training from snapshot')
66     parser.add_argument('--unit', '-u', type=int, default=1000,
```

```

67             help='Number of units')
68     args = parser.parse_args()
69
70     model = L.Classifier(CNN(args.unit, 4))
71     chainer.serializers.load_npz(args.model, model)
72     hantei = ['だるま型', '画鋏', 'クリップ', '割りピン']
73     np.set_printoptions(precision=6, floatmode='fixed', suppress=True)
74     for imgname in [f for f in sorted(os.listdir('testdata')) if ('jpg' in f)]:
75         img = Image.open(os.path.join('testdata', imgname))
76         img = np.asarray(doBright(crop(img, 100)))
77         img = img.transpose(2, 0, 1)
78         x = chainer.Variable(np.array([img]).astype(np.float32))
79         y = model.predictor(x)
80         c = F.softmax(y).data.argmax()
81         s = F.softmax(y)
82         print('{:<15} は{}です。{}'.format(imgname, hantei[c], s.data))
83
84 # main()関数を実行する
85 main()

```

このプログラムは Python というプログラミング言語で記述されています。さらに細かいことを言えば ver.3 の Python です。ver.2 の Python ので書かれたプログラムもよく使われているのですが、微妙に書き方が異なります。

3-14 行 `import` という語が使われていますが、これは既にあるライブラリの読み込みを指示しています。ライブラリは Excel の関数のようなものの集まりで、Python には様々な用途のライブラリがあります。ここでは数値処理、画像処理、そして `chainer` などのライブラリを読み込んでいます。

17-33 行 ここで `chainer` で使用するニューラルネットの定義をしています。17 行目から `CNN` というクラス の指定が始まり、その中身は 18-25 行目と 27-33 行目の部分に分かれています。意外に短いですが この書き方を理解できれば、自分の好きなニューラルネットを作ることができるようになります。Python ってすごいなあと思われるかもしれませんが、この程度の行数で書けるのは、最初の部分で `chainer` のライブラリを読み込んでいるためで、Python が偉いわけではありません。プログラムの行の先頭が段々と右へ後退し、また左へもどる様子が見られると思いますが、これは Python の特徴的なところ です。通常のプログラミング言語では `()` や `{}` で囲うところを、Python ではその部分を右へ後退させて示します。

36-45 行 ここで `crop` という関数を定義しています。画像処理ライブラリの助けを借りて、この関数は一つの画像データを $size \times size$ に修正します。 $size$ の具体的な値は 76 行目で `crop` を呼び出しているところで 100 だとわかります。

48-59 行 またここで `doBright` という関数を定義しています。画像処理ライブラリの助けを借りて、縮小した画像の四隅の明るさを調べて、その平均をもとに全体の明るさの修正を行っています。

61-82 行 ここでプログラムの本体となる `main` 関数を定義しています。関数は定義しただけでは動かないので、85 行目で呼び出しています。

63-68 行 これはプログラム起動時に指定した設定の取り込みを行っています。学習したモデルのファイル名と、`unit` の数を取り込んでいます。

70-71 行 ここで既に定義した `CNN` クラスを実体化し、指定したモデルをそこに読み込んでいます。

72 行 モデルの出力するカテゴリー (分類) の数が 4 つなので、それに対応するメッセージを 4 つここで用意しています。

- 74-78 行 ここで「testdata」ディレクトリ内の jpg のファイルを順番に取り出し、大きさを揃え、ニューラルネットの入力用の形式に変換しています。取り出された画像が最終的に変数 x に入ります。
- 79 行 ここで変数 x の内容についてニューラルネットの処理を行い結果を変数 y に入れています。
- 80-81 行 得られた結果から、一番大きな値となったカテゴリーの番号を変数 c に、全てのカテゴリーの値を変数 s に入れています。
- 82 行 変数 c の値を元にメッセージを表示しています。

結局文房具の画像を識別するプログラムとして特有の部分は、72 行目しかありません。ここさえ直して、「かわいい女子大生」、「きれいな女子大生」、「賢そうな女子大生」、「優しそうな女子大生」の画像を大量に用意して学習したモデルを使えば、女子大生識別プログラムの出来上がりとなります。

7.6 画像認識プログラムの実行

ここでは前節で説明したプログラムを実際に動かしてみます。コマンドをキーボードから入力して実行するという一番原始的な方法を説明します。

1. 「教材フォルダ」から「deep1」のフォルダを丸ごとデスクトップにコピーします。単にドラッグすればコピーされます。
2. コピーした「deep1」フォルダをダブルクリックして開きます。開いたウィンドウの「ツール」メニューの中の「現在のフォルダで端末を開く」を選択すると「LXTerminal」が起動します。(メニューを使う代わりに `F4` キーを押しても同様になります。)「miki@mars:~/Desktop/deep1\$」のようなプロンプトが表示されますので、以下のようなコマンドを入力します。コマンドの最後には `Enter` を押します。
3. 「python3 test.py -m dgkw.model」と入力します。python3 が ver.3 の python を起動することを意味しています。test.py は前節で説明したプログラムが入っているファイルの名前です。-m dgkw.model で学習済みモデルのファイル名 (dgkw.model) を指定しています。すると、リスト 2 のようになります。

リスト 2 画像認識の結果

```
miki@mars:~/Desktop/deep1$ python3 test.py -m dgkw.model
1 d001.jpg は「だるま型」です。 [[0.999946 0.000000 0.000054 0.000000]]
2 d002.jpg は「だるま型」です。 [[0.999999 0.000000 0.000001 0.000000]]
3 d003.jpg は「だるま型」です。 [[0.998352 0.000000 0.001648 0.000000]]
4 g001.jpg は「画鋏」です。 [[0.000000 1.000000 0.000000 0.000000]]
5 g002.jpg は「画鋏」です。 [[0.001159 0.989272 0.009569 0.000000]]
6 g003.jpg は「画鋏」です。 [[0.000000 1.000000 0.000000 0.000000]]
7 k001.jpg は「クリップ」です。 [[0.000000 0.000000 1.000000 0.000000]]
8 k002.jpg は「クリップ」です。 [[0.000000 0.000000 1.000000 0.000000]]
9 k003.jpg は「クリップ」です。 [[0.212270 0.115372 0.672358 0.000000]]
10 w001.jpg は「クリップ」です。 [[0.000000 0.000000 1.000000 0.000000]]
11 w002.jpg は「画鋏」です。 [[0.000000 0.999996 0.000004 0.000000]]
12 w003.jpg は「画鋏」です。 [[0.000000 0.999986 0.000014 0.000000]]
```

2-13 行がプログラムの出力で、「testdata」ディレクトリに入っていたテスト画像を認識した結果が並んでいます。行の先頭が画像ファイル名で、表 3 の通りに付けています。ファイル名の後に認識結果があり、その後にある `[]` に囲まれた 4 つの数字は、与えられた画像をどのくらいの精度で認識したかを示しています。残念ながら 11-13 行目は誤認識していますが、数字を見ても思いっきり間違えていることがわかります。

このようにして、事前に用意した学習済みモデルで事前に用意した画像の認識はできましたが、自分で用意した画像を認識させる際は、次のような手順となります。

表 3 画像認識の認識対象とファイル名

名称	だるま型画鋏	画鋏	クリップ	割りピン
画像例				
ファイル名	d~.jpg	g~.jpg	k~.jpg	w~.jpg

1. スマホやタブレットで認識させたい文房具を撮影します。一つ一つ撮影するのも手間ですので、混ぜて複数個まとめて撮影しても構いません。ただ後で画像を切り分ける必要があるので重ならないように撮影します。また余分なものが入らないように、白い紙の上で、影があまり映らないようにします。この辺りの条件は学習させた画像と同じ条件にすれば良いので、白っぽいものを学習させるために背景は全て黒色にした場合は、背景は黒色にしないと認識率が下がります。
2. 撮影した写真画像をパソコンに送ります。USB ケーブルで接続して送る、メールに添付して送るなどの方法があります。
3. 写真画像を Windows に標準装備されている「ペイント」で開きます。
4. 文房具を一つだけ  (四角形選択) で囲みます。できるだけ正方形になるように「ペイント」のウインドウの一番下のところに囲いの大きさがほぼ同じ (10px ぐらいは違ってても良い) になるようにします。縦と横の大きさは 100px 以上にしないと認識プログラムにはねられますのでご注意ください。そして **[Ctrl]+X** で切り取ります。(コピーでは後でどれを取り出したのかわからなくなる。) もう一つ「ペイント」を起動し、そちらを選択してから **[Ctrl]+V** で貼り付けます。**[トリミング]** をクリックすると貼り付けた画像以外の部分は削除されます。そしてこれを「ファイル」メニューの「名前を付けて保存」で保存します。形式は「JPEG 画像」にします。名前はどの文房具なのか簡単に見分けられるようなものにします。
5. mars のデスクトップにある「deep1」を開き、その中の「testdata」を開いたところへ、作成した画像ファイルをコピーします。
6. LXTerminal が先程の実行した後のままであれば、 で pytho3 で始まる行を呼び出し **[Enter]** を押せば再度認識プログラムを動かすことができます。

7.7 学習プログラムの例

ここでは 4 種類の文房具を学習させる際に使用したプログラム (リスト 3) の説明をします。

リスト 3 画像学習プログラムの例

```

1 # -*- coding: utf-8 -*-
2 import argparse
3 import os
4 import sys
5
6 import chainer
7 import chainer.functions as F
8 import chainer.links as L
9 import chainer.initializers as I
10 from chainer import training

```

```
11 from chainer.training import extensions
12
13 # 畳み込みニューラルネットワークの定義
14 class CNN(chainer.Chain):
15     def __init__(self, n_units, n_out):
16         w = I.Normal(scale=0.05) # モデルパラメータの初期化
17         super(CNN, self).__init__(
18             conv1=L.Convolution2D(3, 16, 5, 1, 0), # 1層目の畳み込み層(フィルタ数は16)
19             conv2=L.Convolution2D(16, 32, 5, 1, 0), # 2層目の畳み込み層(フィルタ数は32)
20             conv3=L.Convolution2D(32, 64, 5, 1, 0), # 3層目の畳み込み層(フィルタ数は64)
21             l4=L.Linear(None, n_out, initialW=w), # クラス分類用
22         )
23
24     def __call__(self, x):
25         # 最大値プーリングは2x2 活性化関数はReLU
26         h1 = F.max_pooling_2d(F.relu(self.conv1(x)), ksize=2, stride=2)
27         h2 = F.max_pooling_2d(F.relu(self.conv2(h1)), ksize=2, stride=2)
28         h3 = F.max_pooling_2d(F.relu(self.conv3(h2)), ksize=2, stride=2)
29         # 9x9,64ch
30         return self.l4(h3)
31
32 def main():
33     # オプション処理
34     parser = argparse.ArgumentParser(description='Chainer example: MNIST')
35     parser.add_argument('--batchsize', '-b', type=int, default=100,
36                         help='Number of images in each mini-batch')
37     parser.add_argument('--epoch', '-e', type=int, default=20,
38                         help='Number of sweeps over the dataset to train')
39     parser.add_argument('--gpu', '-g', action='store_true',
40                         help='GPU ID (negative value indicates CPU)')
41     parser.add_argument('--model', '-m', default='test.model',
42                         help='Name of model file')
43     parser.add_argument('--unit', '-u', type=int, default=1000,
44                         help='Number of units')
45     args = parser.parse_args()
46
47     print('# unit: {}'.format(args.unit))
48     print('# Minibatch-size: {}'.format(args.batchsize))
49     print('# epoch: {}'.format(args.epoch))
50     print('# GPU: {}'.format(args.gpu))
51     print('')
52
53     train = []
54     label = 0
55     print('loading dataset')
56     for c in sorted(os.listdir('train')):
57         print('class: {}, class id: {}'.format(c, label))
58         d = os.path.join('train', c)
59         imgs = os.listdir(d)
60         for i in [f for f in imgs if ('jpg' in f)]:
```

```
61         train.append([os.path.join(d, i), label])
62         label += 1
63     print('')
64
65     train = chainer.datasets.LabeledImageDataset(train, '.')
66     train, test = chainer.datasets.split_dataset_random(train, args.unit)
67
68     model = L.Classifier(CNN(args.unit, 4))
69     if args.gpu:
70         gpu = 0
71         chainer.cuda.get_device(gpu).use()
72         model.to_gpu()
73     else :
74         gpu=-1
75
76     # optimizer の設定
77     optimizer = chainer.optimizers.Adam()
78     optimizer.setup(model)
79
80     # 学習データ用のイタレータ
81     train_iter = chainer.iterators.SerialIterator(train, args.batchsize)
82     # 評価データ用のイタレータ
83     test_iter = chainer.iterators.SerialIterator(test, args.batchsize,
84                                                  repeat=False, shuffle=False)
85
86     updater = training.StandardUpdater(train_iter, optimizer, device=gpu)
87     trainer = training.Trainer(updater, (args.epoch, 'epoch'), out='logs')
88     trainer.extend(extensions.Evaluator(test_iter, model, device=gpu))
89
90     # 計算グラフ
91     trainer.extend(extensions.dump_graph('main/loss'))
92     trainer.extend(extensions.LogReport())
93     # エポック経過ごとの損失の値を表示したりグラフ化する
94     trainer.extend(
95         extensions.PlotReport(['main/loss', 'validation/main/loss'], 'epoch',
96                               file_name='loss.png'))
97     trainer.extend(
98         extensions.PlotReport(['main/accuracy', 'validation/main/accuracy'],
99                               'epoch', file_name='accuracy.png'))
100    trainer.extend(extensions.PrintReport(
101        ['epoch', 'main/loss', 'validation/main/loss',
102         'main/accuracy', 'validation/main/accuracy', 'elapsed_time']))
103
104    # 学習開始
105    trainer.run()
106
107    # モデルをCPU 対応へ
108    if args.gpu:
109        model.to_cpu()
110    # 保存
```

```
111     print('save the trained model: {}'.format(args.model))
112     chainer.serializers.save_npz(args.model, model)
113
114 main()
```

画像認識プログラムに比べるとこちらの方が 1.5 倍ぐらい長くなります。学習の過程をグラフ化する部分の他、画像認識プログラムと共通する部分もかなりあります。

2-30 行 この部分は同じです。つまり同じライブラリーを取り込んで、同じモデルを定義しています。

32-113 行 プログラムの本体となる main を定義しています。画像は適切な大きさに変換済みとしているので、画像認識プログラムにあった crop 関数の定義部分がありません。

34-51 行 ここでプログラム起動時に指定した設定を取り込んで表示しています。学習した結果を保存するファイル名その他、学習のやり方にかかわる設定を取り込みます。

53-63 行 「train」ディレクトリーにある学習用データを変数 train に取り込みます。学習する画像は同じカテゴリのものは同じディレクトリーに入っていると想定しています。

68-74 行 ここで既に定義した CNN クラスを実体化し、GPU (Graphical Processing Unit) が指定されている場合は、さらに GPU 用の形式に変換しています。

77-88 行 ここでさらに学習用の設定を行っています。

91-102 行 学習がどのくらい行われたかをグラフ化したり、学習中に途中経過を表示するための設定をしています。

105 行 設定が全て終わったのでここで学習をします。

108-109 行 GPU を利用して学習した場合は、ここで GPU を使用しない形式のモデルに戻します。

111-112 行 学習済みのモデルを保存します。

こちらのプログラムも学習する画像によって変わるところはありません。この例では画像を 4 つに分類していますが、より多くの分類を行う場合も変更箇所はわずかです。

7.8 学習プログラムの実行

ここでは前節のプログラムを実際に動かしてみます。やり方は画像認識プログラムの実行とほぼ同じです。

1. 「教材フォルダ」から「deep2」のフォルダを丸ごと mars のデスクトップにコピーします。
2. コピーした「deep2」フォルダをダブルクリックして開き、「LXTerminal」を起動します。
3. 「python3 learn.py -u 12 -e 200 -m dgkw.model」と入力して学習プログラムを実行します。するとリスト 4 のようになります。

リスト 4 画像の学習

```
1 miki@mars:~/Desktop/deep2$ python3 learn.py -u 12 -e 200 -m dgkw.model
2 # unit: 12
3 # Minibatch-size: 100
4 # epoch: 200
5 # GPU: False
6
7 loading dataset
8 class: d, class id: 0
9 class: g, class id: 1
10 class: k, class id: 2
11 class: w, class id: 3
12
```

```

13 epoch      main/loss  validation/main/loss  main/accuracy  validation/main/
    accuracy  elapsed_time
14 8          408.788    592.26              0.24          0.166667
    3.27823
15 16         494.481    445.394             0.4           0.291667
    6.97513
16 25         515.182    464.031             0.08          0.291667
    10.4623
17 33         487.03     286.122             0.18          0.166667
    13.9718
18 41         234.264    212.101             0.44          0.291667
    17.4836
19
20      中略
21
22 158        0.0200752  11.429              1             0.5
    69.1469
23 166        0.927973   12.6139             0.77          0.458333
    72.6224
24 175        0.444112   13.1676             0.82          0.333333
    76.1169
25 183        0.0705336  13.542              0.92          0.25
    79.6127
26 191        0.122254   10.8558             0.92          0.25
    83.1162
27 200        0.0330382  9.01501             1             0.375
    86.6048
28 save the trained model: dgkw.model

```

約 87 秒でこれらの結果が得られます。なお LXTerminal の画面の幅を広げないとこのように途中で改行が入りますので、ちょっと見にくい表示になります。「loss」(損失、エラーの量)の値が小さく、「accuracy」(正確さ)の値が 1 に近いほど良いのですが、この例のように途中でまた悪い方へ変わることもあります。名称に validation がついた方は、学習用データを別の形で取り出した評価用データで測定した結果です。この例の場合、最終的には学習データに対しては 100% の正確さになっていますが、評価用データに対しては 37.5% の正確さですから外している方が多くなっています。この学習の様子をグラフ化したものが「logs」ディレクトリーの中にあります。例えば「accuracy.png」をダブルクリックするとグラフが表示され、学習用データでの正確さは上がっていますが、評価用の方はあまり上がっていない様子が見られると思います。

4. 「python3 test.py -m dgkw.model」で学習したモデルで画像の認識を行うことができます。結果はリスト 5 のようになります。

リスト 5 画像学習の結果の確認

```

1 miki@mars:~/Desktop/deep2$ python3 test.py -m dgkw.model
2 d001.jpg      は「だるま型」です。[[0.949683 0.000001 0.037190 0.013126]]
3 d002.jpg      は「クリップ」です。[[0.044127 0.000000 0.946383 0.009490]]
4 d003.jpg      は「だるま型」です。[[0.996112 0.000000 0.003855 0.000033]]
5 g001.jpg      は「画鋸」です。[[0.000000 1.000000 0.000000 0.000000]]
6 g002.jpg      は「割りピン」です。[[0.000000 0.000000 0.000000 1.000000]]
7 g003.jpg      は「だるま型」です。[[0.999649 0.000001 0.000000 0.000350]]
8 k001.jpg      は「クリップ」です。[[0.000000 0.000000 1.000000 0.000000]]

```

9 k002.jpg	は「クリップ」です。[[0.000000 0.000000 1.000000 0.000000]]
10 k003.jpg	は「だるま型」です。[[1.000000 0.000000 0.000000 0.000000]]
11 w001.jpg	は「割りピン」です。[[0.000000 0.002091 0.003251 0.994658]]
12 w002.jpg	は「クリップ」です。[[0.000704 0.048079 0.942931 0.008286]]
13 w003.jpg	は「割りピン」です。[[0.000000 0.000000 0.000000 1.000000]]

残念な結果になっていますが、学習データが少ないのか学習が足りないのでしょうか。ランダムで与えた初期値の影響も大きく、同じデータかつ同じパラメタで学習しても、認識結果がかなり違うこともあります。

矢印キーで前に入力したコマンドを呼び出して一部修正することもできます。学習の際の「-u 12 -e 200」の数字を変更して、再度学習をさせることができます。u の値はある程度大きい方が良いでしょうが、学習データの数より大きな値にすることができません。e の値を大きくすると学習の回数を増やすことができます。回数が足りないとだめですが、増やせば良くなるとは限りません。過学習と言う状態になると、学習用データに適応しすぎて、それ以外のデータに対しての応答がだめになってしまいます。また回数を増やせばそれに比例して時間もかかるようになります。

学習にかかる時間を短縮するために GPU などが用いられます。GPU は本来動画の処理 (特に動きの激しいリアルなゲームなど) のために作られたものですが、比較的安価に入手できるために最近のスーパーコンピュータでも使われています。さらに学習用であれば、それほど高精度の計算は必要ないことも判ってきたので、専用の処理装置も使われるようになってきました。mars には Nvidia 社製の GeForce GTX 1080 Ti という GPU が付いています。その内部には 3,584 個の演算器があり、全体で 11.3TFLOPS^{*23} の計算ができます。これを学習の際に使う場合は、「-g」を追加します。先程は約 87 秒かかりましたが、8 秒程度で終わるようになります。

自分で用意した学習データを用いて学習させる場合には、「moto」ディレクトリの中にある「d」ディレクトリにだるま型の画像、「g」ディレクトリに画紙の画像、「k」ディレクトリにクリップの画像、「w」ディレクトリに割りピンの画像を入れます。既に入っている画像はそのままにしても良いし、置き換えても良いです。そして「python3 crop.py -i moto -o train」を実行すると、「moto」ディレクトリに入っていた画像ファイルが、100×100 の大きさに整形され、明るさの調整もされて「train」ディレクトリに入ります。既に「train」ディレクトリに入っていたファイルは削除されますのでご注意ください。そして learn.py を実行すれば学習が行われます。

7.9 学習用データを増やす

とにかく学習用データを増やさないと精度良く認識することができません。一方学習用データを大量に用意するのは費用や手間がかかることは確かで、いかに効率よく用意するかが重要になります。効率の良い収集方法はデータによって異なります。消費者の行動データのようなものは、スーパーやコンビニのレジの情報から得られます。これらは既に商品発注や在庫管理システムで利用されていますので、関係者であれば容易に入手することができるでしょう。一部のコンビニで行われていた、レジで入力したお客様の年齢のデータもあれば、より興味深い結果が得られるかもしれません。ただ個人情報保護には注意しなければなりません。かわいい女子大生の顔写真を集めるために、現マネ棟の入り口に密かにカメラを設置して撮影などすると問題になるでしょう。ただ密かに撮影して、分類して、学習させて、学習結果だけ残した場合、どのような画像を用いたかは分かりません。つまり、いつの間にか自分の個人情報を学習データとして使われても、本人は気が付けないので注意が必要です。と言われても、対策は AI の得意な企業にはできるだけ近寄らない、でしょうか？

^{*23} FLOPS は 1 秒間に何回小数点付きの数の計算ができるかというもの。11.3TFLOPS は 1 秒間に 11.3 兆回計算できることになります。

画像認識の場合、認識する対象は様々な向きや大きさであることが普通です。深層学習が使われる以前の画像認識では、認識処理に入る前の前処理として、撮影した画像に映るものの向きや大きさなどを揃えていました。深層学習を使う際にも前処理をした方が効果的ですが、学習データに逆の処理を行うことがよくあります。つまり撮影した学習用データを回転させたり、縮小したのも学習させることにより、回転や縮小に強くなりますし、このようなデータの水増しは機械的にできるので手間もかかりません。学習用データの機械的な収集と効率的な学習データの水増しは、深層学習の精度を上げるための大事なポイントです。

7.10 水増し学習プログラムの実行

ここでは実際に水増ししたデータをもとに学習プログラムを実行し、その効果を確認します。

1. 「教材フォルダ」から「deep3」のフォルダを丸ごと mars のデスクトップにコピーします。
2. コピーした「deep3」フォルダをダブルクリックして開きます。そして「LXTerminal」を起動します。
3. 「python3 cropmod.py -i moto -o train」と入力すると moto ディレクトリーの中の画像ファイルを、100×100 の大きさに整形し、明るさを調整した後で、反転、回転、80% に縮小したものを生成して、各々別のファイルとして train でディレクトリーに保存します。moto ディレクトリーにあった各々 9 個の画像が、各々 225 個の画像になります。
4. これをもとに例えば「python3 learn.py -u 50 -e 100 -m dgkw.model -g」と入力して学習プログラムを実行します。そして「python3 test.py -m dgkw.model」でできたモデルがちゃんと認識するか確認することができます。「-u」や「-e」の値を適切に設定するとリスト 6 のような認識結果も得られます。

リスト 6 水増し画像学習の結果

1 d001.jpg	は「だるま型」です。[[0.999844 0.000005 0.000151 0.000000]]
2 d002.jpg	は「だるま型」です。[[0.999749 0.000000 0.000251 0.000000]]
3 d003.jpg	は「だるま型」です。[[0.719442 0.000005 0.280538 0.000016]]
4 g001.jpg	は「画紙」です。[[0.000000 0.999974 0.000000 0.000026]]
5 g002.jpg	は「画紙」です。[[0.010267 0.687305 0.203553 0.098875]]
6 g003.jpg	は「画紙」です。[[0.000000 0.993213 0.005990 0.000796]]
7 k001.jpg	は「割りピン」です。[[0.000057 0.000892 0.294389 0.704662]]
8 k002.jpg	は「クリップ」です。[[0.232162 0.000000 0.767838 0.000000]]
9 k003.jpg	は「クリップ」です。[[0.000000 0.000000 0.999998 0.000002]]
10 w001.jpg	は「割りピン」です。[[0.000000 0.008512 0.028039 0.963449]]
11 w002.jpg	は「割りピン」です。[[0.000000 0.000098 0.000000 0.999902]]
12 w003.jpg	は「クリップ」です。[[0.000001 0.000002 0.998799 0.001198]]

7.11 YOLO による物体検出

自動車の世界では現在 EV (Electric Vehicle) と自動運転が大きな課題になっています。後者の自動運転で欠かせない技術として高速の物体検出があります。高精度で正確な地図があれば目的地に行くことはできますが、途中の道路上に何があるのかは分かりません。車の運転手は走行中に数秒よそ見をするだけで大事故の危険性が高まります。自動運転車は常に周囲の状況を把握しなければなりません。レーダーを使えば周囲にある物体までの距離がわかります。さらにドップラー効果を利用すれば、その物体が近づいているのか、遠ざかっているのかもわかります。これで衝突は防げるかもしれませんが、レーダーでは信号機が分かりません。赤信号の交差点に突っ込んだら大惨事になります。

画像の中のどこに何があるのかを検出する物体検出は、1970 年代から研究されていましたが、なかなかうまく行きませんでした。黒い目と赤い口から顔を検出するぐらいはなんとかできましたが、様々な姿をした

人の検出は無理でした。ところが2012年に Alex Krizhevsky らが画像認識 (画像分類) のコンテストである ILSVRC2012 で提出した、畳み込みニューラルネットワーク (CNN) を用いたシステムが優勝したことにより、深層学習を利用した物体検出が主流になりました。そして3年後の ILSVRC2015 で優勝した ResNet は人間よりも少ないエラー率を達成しました。

ここで紹介する YOLO は物体検出のシステムで、最初のバージョンが2016年に発表され現在も改良が続けられています。YOLO は “You only look once” から来ており、候補領域をまず抽出し、その中で検出や分類を行うのではなく、一気に抽出から分類まで行う方式である事を示しています。一気にやることにより精度の点で劣るものの、動画をそのまま処理できる速さを実現しました。自動運転での利用を考えると、人が飛び出したのを数秒後に検出したのでは間に合いません。自動運転で使用される物体検出システムは、YOLO を元に更に改良が進められています。

ここでは2020年6月に Glenn Jocher が発表した YOLO ver.5 (<https://github.com/ultralytics/yolov5>) を実際に使ってみます。

YOLOv5 による物体検出

YOLO ver.5 は python3 で書かれた、PyTorch という名前のフレームワークで動きます。そのため YOLOv5 を動かすためには PyTorch を初め多数の関連ライブラリをインストールする必要がありますが、全て mars にインストール済み^{*24}なので、YOLOv5 本体のみインストールすれば使うことができます。

1. 「教材フォルダ」から「yolov5」フォルダを丸ごと mars のデスクトップにコピーします。
2. コピーした「yolov5」フォルダをダブルクリックして開き、「LXTerminal」を起動します。
3. さらに次のコマンドを入力するとサンプルデータの処理を行います。

```
python3 detect.py --source data/images/bus.jpg
```

するとリスト7のようなメッセージが表示されます。14行に「4 persons, 1 bus, 7.9ms」とありますが、画像の中に4人とバスを1題7.9ミリ秒で見つけたと言う意味です。

リスト7 YOLOv5 による画像認識

```
1 miki@mars:~/Desktop/yolov5$ python3 detect.py --source data/images/bus.jpg
2 detect: weights=yolov5s.pt, source=data/images/bus.jpg, data=data/coco128.yaml,
3 imgsz=[640, 640], conf_thres=0.25, iou_thres=0.45, max_det=1000, device=,
4 view_img=False, save_txt=False, save_conf=False, save_crop=False, nosave=False,
5 classes=None, agnostic_nms=False, augment=False, visualize=False, update=False,
6 project=runs/detect, name=exp, exist_ok=False, line_thickness=3,
7 hide_labels=False, hide_conf=False, half=False, dnn=False, vid_stride=1
8 YOLOv5 v6.2-266-g72cad39 Python-3.8.10 torch-1.13.0+cu117 CUDA:0
9 (NVIDIA GeForce GTX 1080 Ti, 11178MiB)
10
11 Fusing layers...
12 YOLOv5s summary: 213 layers, 7225885 parameters, 0 gradients
13 image 1/1 /home/miki/Desktop/yolov5/data/images/bus.jpg: 640x480
14 4 persons, 1 bus, 7.9ms
15 Speed: 0.3ms pre-process, 7.9ms inference, 0.9ms NMS per image at shape
16 (1, 3, 640, 640)
17 Results saved to runs/detect/exp
```

^{*24} 大抵の YOLOv5 の解説では、関連ライブラリのインストールから説明されています。PyTorch などはかなり大きなライブラリなので全て入れると4GB程度のディスク容量を必要とします。

4. リスト 7 の最後の行に「runs/detect/exp」に結果を保存したとあるので、「runs」フォルダ、「detect」フォルダ、「exp」フォルダの順にダブルクリックをして開いていくと、bus.jpg があります。これをダブルクリックすると図 7.2 のような、検出した物体が枠線で囲われているものが表示されます。枠線に「bus 0.85」のように表示されていますが、これは物体が「bus」で確からしさが 0.85 という意味です。

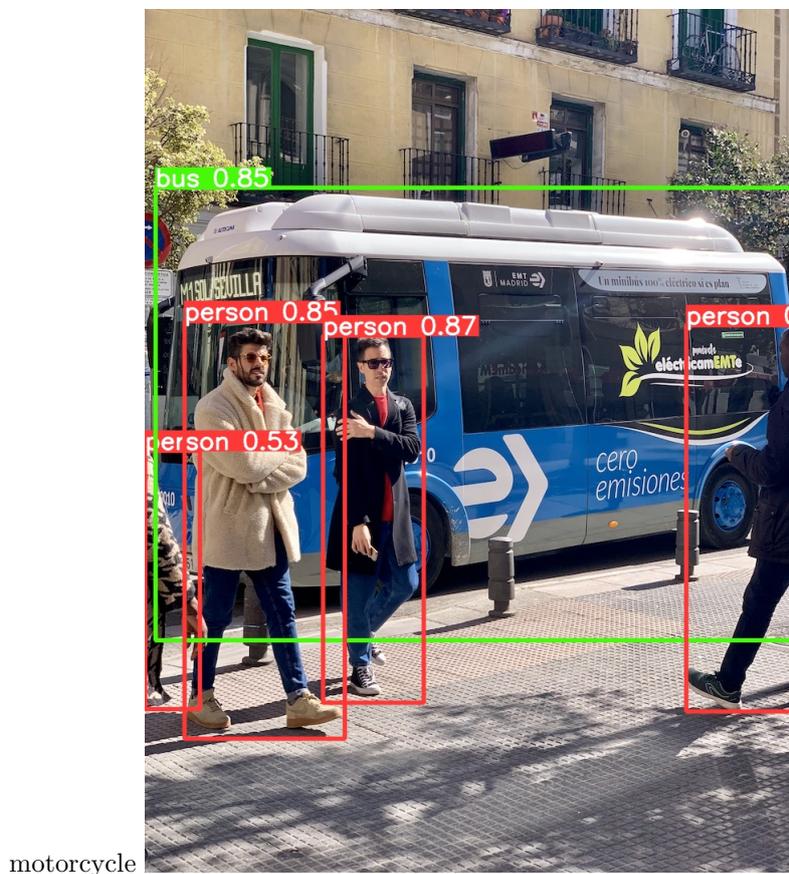


図 7.2 YOLOv5 による画像認識結果

YOLOv5 用学習データの作成

YOLO に自分が検出して欲しい物体を学習させることができます。そのためにはまず学習データを作成する必要があります。4 つの文具を認識させる場合は、画像ごとに何の画像かが分かるようにしていました。YOLO は物体の位置も答えてくれるので、学習データも「何がどこにあるか」が分かるようにする必要があります。このような学習するための情報をアノテーションと呼びます。

YOLO の場合、画像 (~.jpg) ごとにアノテーションの入ったテキストファイル (~.txt) を用意します。アノテーションの形式は 1 行が 1 つの物体に対応しているので、一つの画像に 5 つの物体があれば、5 行になります。1 行には、物体のクラス名、物体の中心の X 座標、物体の中心の Y 座標、幅、高さの 5 つが空白で区切られて並びます。座標や幅、高さは、画素数ではなく、画像全体の幅や高さを 1 としたときの相対的な数値になっています。このような情報を作成するためのツールプログラムがいくつかあり、mars には labelImg^{*25} をインストールしたので、これの使い方を説明します。

1. 「yolov5」を開いたところで「LXTerminal」を起動します。

^{*25} <https://github.com/heartexlabs/labelImg>

2. 次のコマンドを入力して LabelImg を起動します。ただしこれは既にクラス名の入ったファイル (classes.txt) がある場合のコマンドで、そのようなものがない場合は「labelImg」のみで起動します。

```
labelImg data/train/images data/train/labels/classes.txt data/train/labels
```

3. 起動すると図 7.3 のような画面になるので、このように左側のところに「PascalVOC」が出ていたら、「ファイル」メニューの中の「PascalVOC」をクリックして「YOLO」に切り替えます。また「表示」メニューの中の「自動で保存する」にチェックを付けておくと、別の画像に移動する際に保存をしてくれるようになります。中央の細い長方形のところに画像が表示されるので、ウィンドウを最大化しておくといいでしょう。

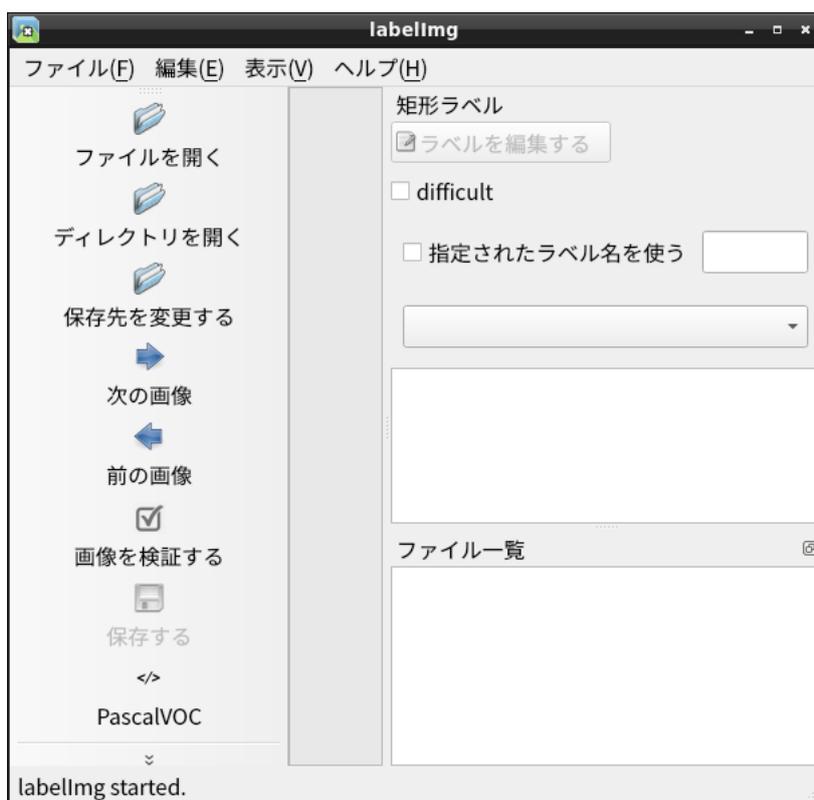


図 7.3 LabelImg の最初の画面 (labelImg 単独で起動した場合)

4. 学習用の画像は「yolov5」の中の「data」の中の「train」の中の「images」に入れておきます。5つのサンプル画像が入っているはずですが、学習用の画像を追加するならば同じところに入れます。この入れる場所は後に出てくる設定ファイルに合っていれば、他の場所でも構いません。train」の中にはもう一つ「labels」があります。ここにはサンプル画像の一つのアノテーション情報の入ったファイルとクラス名の入った「classes.txt」というファイルがあります。分担してアノテーション追加の作業を行う場合は、まず全てのクラス名の入った「classes.txt」を作成して、それをコピーして使うのが良いでしょう。
5. 「labelImg」コマンド単独で起動した場合は、左側にある「ディレクトリを開く」をクリックして、画像があるディレクトリ (images) を開いて行って、画像のファイル名が並んでいるところまで行ったら「Open」をクリックします。
6. 「labelImg」コマンド単独で起動した場合は、左側にある「保存先を変更する」をクリックして、アノテーションを入れるディレクトリ (labels) を指定します。YOLO の場合、画像が入っているディレクトリの隣にこのディレクトリを置きます。

7. 中央に最初の画像が表示されます。そこでアノテーションを付ける物体に対して次の操作を繰り返します。

- (1) **W**を押します。
- (2) マウスポインターの形が変わって、右と下へ黒い細線が伸びます。物体の上と左がちょうど入るところまでマウスポインターを移動します。
- (3) ドラッグしてマウスポインターを物体をちょうど囲う右下へ移動します。
- (4) マウスボタンを離すと、クラスを尋ねてくるので新しいクラスの名前を入力するか、既に存在するクラスをリストの中から選択します。

全ての物体に操作を繰り返すと図 7.4 のようになります。

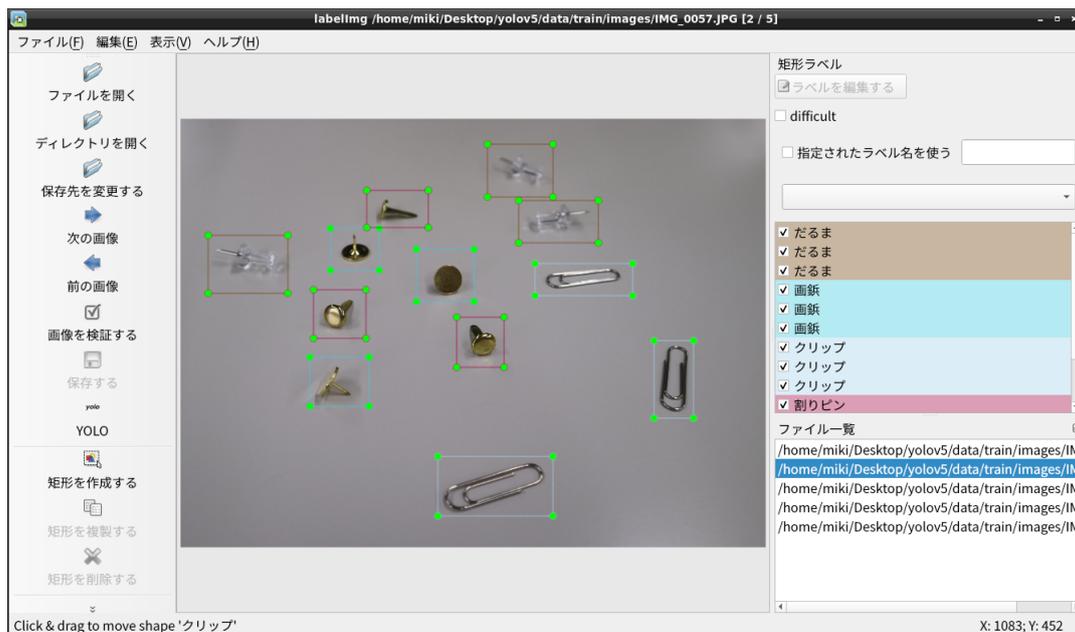


図 7.4 アノテーションの設定が終わったところ

8. まだアノテーションが付いていない画像がある場合は、左側にある「次の画像」をクリックします。

画像の置き場所などが変わると、これまでのアノテーションの結果が表示されなくなることがあります。そのような場合は「ファイル」メニューの「全て初期化する」を選択して、再度設定をやり直すと治るようです。また設定途中で「Attention」のダイアログが出ることがあります。何か設定を変えた場合は「Yes」、そうでない場合は「No」と答えておけば良いでしょう。

学習データによる学習

YOLOv5 で学習させる場合は、まずモデルを選択します。「YOLOv5n」、「YOLOv5s」、「YOLOv5m」、「YOLOv5l」、「YOLOv5x」と5つのモデルがあり、「YOLOv5n」が一番小サイズで、短い時間で認識しますが、精度がその分悪くなります。「YOLOv5x」が一番大きなサイズで、「YOLOv5n」と比べて認識に2倍弱の時間がかかりますが、精度はその分良くなります。他の3つのモデルはその中間の性能になります。特に指定をしない場合は、「YOLOv5s」になります。違うモデルを使用する場合は、後に出てくるコマンドの最後に「-weights yolov5x.pt」のような指定を追加します。

画像がどのディレクトリにあるか、クラスの名前などの設定ファイルを用意します。既に用意したもの(dgkw.yaml)が「data」の中にあるのでこれを利用します。その内容はリスト8のようになっています。

リスト 8 学習のための設定ファイルの内容

```
1 # sample 2022/12/02 by K.Miki
2
3 path: ./ # dataset root dir
4 train: train/images # train images
5 val:   train/images # val images
6 test:  # test images (optional)
7
8 # Classes
9 names:
10  0: 画鋸
11  1: 割りピン
12  2: だるま
13  3: クリップ
```

4 行目で学習用データがあるディレクトリを、5 行目で検証用データがあるディレクトリを指定しています。今回の設定では学習用データをそのまま検証用に使っています。10 行目以降は、「labels」ディレクトリにあった classes.txt の内容にゼロからの番号を付けたものになっています。各業の先頭は同じ数の空白が必要です。この例のように日本語の名前を付けると、各種グラフで文字化けします。

次のコマンドで学習を実行します。

```
python3 train.py --batch 10 --epochs 100 --data data/dgkw.yaml
```

大量のメッセージが表示されます。特に指定をしないと GPU を使用します。「--batch 10」は学習データをまとめる単位で、数字が大きくなるほど認識率が高まりますが、処理時間と GPU のメモリの使用量が増えます。メッセージの中の「CPU_mem」の下に「1.05G」と出ていますが、これがグラフィックボードのメモリーを越えると動きません。mars のグラフィックボードには 12GB のメモリーが付いているので、まだかなり余裕があります。「--epochs 100」は何回学習を行うかを示しています。この数字が大きいくほど学習は進みますが学習時間がかかり、ある程度の数を越えるとあまり結果が変わらなくなってきます。

学習の進み具合は「runs」の中の「train」の中の「exp」（番号付き）の中の画像などを見ることにより分かります。「results.png」がグラフをまとめた画像になっていますが、グラフがまだまだ上がりそうや下がりそうな場合は epochs の数値を大きくすればよいでしょう。

学習結果の確認

次のコマンドで学習結果を用いてテスト画像を認識させることができます。

```
python3 detect.py --source data/train/tests --weights runs/train/exp/weights/best.pt
```

「/exp/」の部分は学習を実行した際の最後に表示される「」「--source」で認識する画像ファイルそのものが、画像ファイルが入っているディレクトリを指定します。「--weights」で認識に使用する学習結果を指定します。同じディレクトリに best.pt と last.pt の 2 つのファイルがあります。best.pt は学習中で一番良かった学習結果で、last.pt は最後の学習結果です。内容が同じものになることもあります。

色々なメッセージが表示されますが、最後の行に「Results saved to runs/detect/exp3」と出たならば、「runs」の中の「detect」の中の「exp3」の中に認識結果が入っています。信頼度の低い結果まで描き込まれるので、一つの物体に検索結果が 3 つも重なっているようなものも出てきます。前にやった一つの画像の一つしか出てこないものの認識であれば、一番確からしいものを認識結果として採用すればよいのですが、こちらは

位置が微妙にずれていますので、そう言うわけには行きません。代わりに「--conf-thres 0.3」のような設定を行の最後に追加して実行すると、信頼度が 0.3 より低い結果は表示されなくなり見やすくなります。

7.12 AI に対抗するには

この節は国立情報学研究所教授の新井紀子著「AI vs. 教科書が読めない子どもたち」(2018 年、東洋経済新報社) の紹介です。

新井氏は 2011 年から「ロボットは東大に入れるか」(東ロボ) という人工知能プロジェクトを開始しました。このプロジェクトは東京大学の入試問題を人工知能に解かせることを目標としたものです。十分な研究費がなく、大学院生も含めると 100 名近い研究者のボランティア活動によって進められたこのプロジェクトは、2013 年のセンター模試の成績は偏差値 45 程度でしたが、2016 年のセンター模試では偏差値が 57.1 まで上がり、センター試験のみの入試であれば、全国に 127 校ある国公立大学のうち、23 大学の 30 学部の 53 学科で合格率 80% 以上という判定が出ました。さらに数学と世界史の 2 教科のみは「東大入試実戦模試」(筆記試験) で偏差値 61.8 と 76.2 という好成績をあげています。ただこのプロジェクトは終了しました。それはこのプロジェクトによって、いくつかの大きな壁が明らかになり、それを越える方法が出てくるまで先へ進めないことがわかったからです。量子コンピュータやビッグデータでも越えられない壁です。新井氏はこれから「AI が神になる?・・・なりません!」、「AI が人類を滅ぼす?・・・滅ぼしません!」、「シンギュラリティが到来する?・・・到来しません!」と述べています。ただ囲碁や将棋で AI にプロが勝てなくなったのと同様に、分野を絞ったセンター模試であれば、高校生の 8 割近くは AI に負ける事が明らかになりました。

これからの世の中ではどんどん AI が使われるようになって、その分の雇用が失われるでしょう。一説によればそれは 50% 程度にもなります。最近の別の研究では 10% 程度とも言われますが、それでも国内だけで数百万人が職を失います。その一方で新しい雇用が生じるのですが、AI のために失業した人が雇用されるわけではありません。かつて産業革命が起こったときにも同様の事が生じましたが、現在その当時の産業革命による失業は残っていません。これは産業革命後の社会に適応した人々が段々生まれて、新しい雇用の対象になったからです。同様に AI による失業も時間が経てば解消される事が予想されますが、それには時間がかかるので現在や近い将来失業する人は救われません。

とりあえず AI によって奪われるような仕事を避けるのが良いでしょうが、それはどのような仕事か。これについては「コミュニケーション能力や理解力が求められる仕事」や「柔軟な判断力が必要な肉体労働」ではないかと言われています。必要なスキルは「高度な読解力と常識、加えて人間らしい柔軟な判断能力」となります。新井氏はこの中から、常識はみんな持っているだろう、としています。みんなが知らない知識は常識にならないからです。柔軟な判断もできるだろうとして、問題は「読解力を基盤とする、コミュニケーション能力や理解力」の有無です。そして新井氏が日本の中高校生を対象に行った調査の結果、残念ながら多くの中高生は教科書の記述を正確に読み取ることができないことが明らかになりました。(うちのような中位の私立大学もボロボロのようです。)そして読解力は高校以下で獲得するのが普通で、それ以降はあまり伸ばす機会がないので多くの大人も同様のレベルと考えられます。

読解力の調査では東ロボプロジェクトで得られた知見が使われています。問題の文章を AI で理解させようとするとき、AI でも割とできるものや、AI ではなかなかできないものがあります。主語と述語、修飾語と被修飾語のペアを見つける「係り受け」解析や「それ」「これ」と言った指示代名詞が何を示しているのか求める「照応」解析は東ロボでも 80% ぐらいまでできました。一方 2 つの異なる文が同じ意味であるかどうか判定する「同義文判定」は長年研究されていますがうまく行きません。AI がなかなか身に付けられない常識が必要なものとして「推論」、「イメージ同定」、「具体例同定」も取り上げています。「推論」は文章に書かれていない常識も用いた意味の理解です。「イメージ同定」は文章と図やグラフを比較して同じ内容のものを見つけます。「具体例同定」は文章にかかっている内容の具体例となるものを見つけます。これらは意味を理解できない AI が苦手とするものです。元となる文章は教科書と新聞記事を利用しています。あまり専門的な文章では難しい用

語が理解の妨げになる恐れがあるからです。また新井氏は教科書の意味が分からないのでは授業にならないので、これぐらいはできるだろうと思っていたようです。

調査した人数は大人も含めて 2018 年までの段階で 2 万人です。その結果は新井氏の本に表やグラフと共に詳しく書かれています。ここでは、その中の表を一つ紹介します。

表 4 学年別のランダム率 (%)

学年	係り受け	照応	同義文判定	推論	イメージ同定	具体例同定 (辞書)	具体例同定 (数学)
小 6	35.8	52.6	95.3	57.3	54.7	56.0	100
中 1	31.6	36.8	78.9	61.8	42.4	68.5	86.4
中 2	26.1	25.4	78.8	54.6	37.3	63.5	82.4
中 3	18.3	15.6	70.2	43.4	31.1	48.7	79.4
高 1	10.8	6.5	57.4	38.6	11.7	47.9	53.1
高 2	10.3	6.2	65.7	37.6	13.4	53.4	57.6

表 4 は調査の結果から回答がランダム、つまりサイコロ振って答えているのと変わらない回答の割合の表です。よって数字が小さいほど理解して答えている人が多いことになります。残念ながら「同義文判定」や「具体例同定」は高校 2 年生でも半分以上の人が間違えています。

外国の調査結果に比べるとこれでもまだそうです。ただ外国は母国語が違う人の割合が高い国が多いので、読解力のテストでは不利な状況にあります。母国語が日本語の人が大多数の日本ならば、本来もっと高い読解力があるはずですが。読解力を伸ばす簡単な方法は見つかりません。新井氏は統計的に様々な要因との関連を調べたそうですが、家庭の貧困が悪い影響を与えているぐらいしか見つからなかったそうです。

この調査に協力的だった埼玉県のパトナ市の小中学校は、読解力の残念な結果を受けて、様々な読解力の向上を試みました。するとこれまで埼玉県で中位レベルの学力だったパトナ市の小中学校が、埼玉県で学力が総合 1 位レベルになったそうです。授業が分からないのは内容が難しいからでなく、単に読解力が足りないためかもしれません。そして社会に出てからも、仕事をするために様々な勉強が必要になります。教科書が読めなければ AI がやって来るまでもなく、仕事ができないことになります。

8. 動画プレゼン

かつて動画の処理はコンピュータにとって鬼門でした。動画のデータ量は文字データや音声データと違って桁違いに大きく、数秒の動画すらメモリーを溢れさせ、数分の動画でディスクがいっぱいになりました。FFT (高速フーリエ変換) アルゴリズムが発見されて、音声データを高速に分解して成分が求められるようになりました。そして成分を適当に間引くことにより、音声データを大幅に圧縮することができるようになり、iPod などの携帯音楽プレーヤーが実現し、圧縮なしの CD (コンパクト・ディスク) はあまり見られなくなりました。それでも動画となると、FFT などを使用してもなかなか追いつけない状況が続いています。どの程度の品質の動画を求めるかにもよりますが、高品質の動画はパソコンレベルでは圧縮などに結構時間がかかります。また大容量のディスクが安価になりましたが、それでも圧縮しない状態の長時間の動画はなかなか保存できません。

一方、文字データ、画像データ、音声データ、動画データなどを比べてみると、他のデータを全て取り込むことができる動画データにはアピール力があります。百聞は一見にしかずとも言います。適切な画像データは文字や音声では伝えきれないものを一瞬で伝えます。動画はさらにその画像が動くわけですから強力ですが、その力を活かすのはなかなか難しいところがあります。お金のかかっている動画と言えば映画があります。数十億円以上の費用がかかった映画が毎年いくつか登場します。しかし数年後にも話題になる映画は、一番費用がかかった映画ではないことも少なくありません。その時代に合った内容の映画でないと人気は出ませんし、普遍的なテーマでなければすぐに忘れ去られてしまいます。

プレゼンテーションは、発表などのことですが、様々な場面で用いられて、何らかの情報を聞き手に伝えるものです。相手がこちらが伝えたい情報をぜひ知りたいたいと言うのであれば、かなり雑なプレゼンテーションでも構わないでしょう。一方相手側がこちらの情報にあまり興味を持っていない場合、なかなか厳しい状況になります。文章を書いても読んでもらえない、話しかけても聞いてもらえない、そういう状況でも自分の会社の新製品について是非知ってもらいたいと言うようなことがよくあります。そこで最強の動画を使うわけですが、しつこくやると嫌われるので TV の CM などは数十秒で終わります。

最初に述べたように動画の処理は現在のパソコンで扱うにしても重たい処理になります。新海誠監督のように 25 分のアニメ映画を 7 ヶ月かけて一人で作ってしまった人もいますが、ここでは適当な動画があればそれも取り込むが、基本は静止画を見せるという PowerPoint によるプレゼンテーションとあまり変わらないものを取り上げます。元々は Windows に付属していた「Windows Live Movie Maker」を利用した動画作成を考えていたのですが、大変残念なことに Windows 8 以降にこれは付属しなくなりました。このソフトは動画編集ソフトとしてはかなり低レベルですが、その分使いやすいという特徴がありました。ここで取り上げる Lightworks^{*26} は著名なソフトではありませんが、Linux だけでなく Windows でも無料で使えるということで選択しました。やはり市販の動画編集ソフトは高機能で良いと思います。有名な市販ソフトは大抵 30 日間の試用が無料^{*27} できるようになっているので、ちょっと使わせてもらって自分の技量やパソコンの能力に合っているかどうか確認しましょう。

動画によるプレゼンテーションの良いところは、

- PowerPoint の発表では発表者が必要ですが、動画の再生には発表者は不要です。よって繰り返し発表するような場合でも負担がありません。
- YouTube などの動画を世界に公開する場があります。YouTube に登録すれば、無料で世界中の人に動画を届けることができます。また見る方もいつでもどこからでも見るができます。

などがあります。

^{*26} <https://lwks.com/>

^{*27} 私の長男は結婚式の披露宴で上映する二人の生い立ち動画を、adobe premiere elements というソフトで作成しましたが、30 日以内に作成して無料で済ませました。結婚式場にこの動画の作成を依頼すると 20 万円ぐらいかかったようです。

8.1 動画作成の手順

レポート作成でも元となる資料の収集は必要になります。動画を作るとなると多種多様なデータを集める必要があります。漠然と集めていると無駄が多いので、全体の構成を考えてから集めましょう。

8.1.1 動画の長さ

全体の構成を考える際に一番の制約となるのは動画の長さです。TVの短いCMは15秒ですし、TikTokは60秒以内です。YouTubeに投稿する動画であればもう少し余裕がありますが、長い動画は途中までしか見てもらえない可能性が高くなります。一発芸で終わりならともかく、大抵は色々伝えたいのだが時間が足りない！という問題で苦勞します。PowerPointではよく細かい字で詰め込んだスライドが見られます。動画で同じことをすると多分誰も読んでくれません。そして読み上げる音声を入れると時間が足りません。とにかく内容を絞り込むしかありません。何も考えずに動画を作成していると、どんどん長くなります。動画を見てくれる人の我慢の限界を考えてまず動画の長さを決めます。

8.1.2 動画の画面の大きさ

PowerPointではデザインにこだわりのある人以外は気にしないところですが、プレゼンテーションを行う際の画面の大きさの設定があります。もちろん絶対的な大きさはプロジェクターとスクリーンの距離によって変わります。PowerPointで設定できるのは縦横比です。「4:3」と「16:9」が選択できるようになっています。昔ながらのTVの画面が「4:3」で、スマホ、大型TVやパソコンの画面のほとんどが「16:9」です。縦横比が合わない場合、切り詰められたり足りない部分が黒い帯になったりします。動画でも通常このどちらかの縦横比を選択することになります。途中で縦横比の変更は可能ですが、デザインの面でやり直しになる可能性もあるので、最初によく考えて決めます。

近年の動画はスマホで再生される機会も多くなっています。その場合は「16:9」で良いのですが、縦向きの使うのが普通のスマホでいつもの向きのまま再生できるようにとすれば「9:16」で作成します。

さらに動画では縦横のドットの数も決めます。ドットの数が多ければ画面が綺麗になりますが、ファイルが大きくなって、通信環境によっては動画の再生に支障が生じることがあります。逆にドットの数が少ないれば画面が荒くなりますが、ファイルが小さくなり、動画の再生や保存の際に有利となります。通常ドット数は、最終的な保存の際に指定する形になるので、いくつか保存して試してみれば良いでしょう。

8.1.3 説明文章

百聞は一見に敵かずとも言いますが、見ただけでは何を言いたいかわからない画像や動画もいっぱいあります。子供が走っている動画を見ただけで、「これは平成25年の秋の幼稚園の運動会のリレーだ」なんてわかる人はいません。最低限の説明を文字で行うとして、それを文字を表示して行うか、音声で入れるかの2つが考えられます。映画で言えば字幕でするか吹き替えでやるかです。ここでは文字で表示する方法を採用します。音声でやる場合の問題は、誰の音声を使うかがまずあります。アナウンサーのように明瞭に話せる人はなかなか居ません。自分でやるのは恥ずかしいとか、録音する際に周囲の雑音が入らないようにするのは意外と難しいこともあります。YouTubeなどでは合成音声を使用している例も見られます。恥ずかしいとか周囲の雑音の混入などの問題はありますが、あまり聞きやすい音声ではない*28こともあって、説明文も同時に表示されています。

文字を表示する際には、一度に全てを表示するのではなく、読むスピードを考えて順番に出します。PowerPointであれば、アニメーションの機能を利用して説明に合わせて順次表示するのが良いとされています。これは一度に全て出してしまうと、見ている側が先読みをしてしまい、せっかくの発表を聞いてもらえなくなるからです。動画の場合は、読み上げ音声なしであれば、先読みされて構わないのですが、全てを一度に

*28 ちゃんとお金を出せば、もう少しまともな合成音声も可能ですが、そういう例を見たことがありません。

出して一度に消すと、文字を読むのが遅い人は着いていけません。順次表示して順次消えていくようにすると、読むのが遅い人もある程度飛ばして見てくれます。とにかく文字数が多いと早く読めません。映画の字幕やTVのニュースの項目の書き方を思い出してできるだけ短くしましょう。

日本人の文字を読む速さは、400～600文字/分と言われていたそうです*29。誰でも間に合う、画像を見てもらう時間も必要と考え、400～600文字/分はきついでしょ。

8.1.4 BGM

説明音声なしになると無音の動画になります。ビデオカメラで撮影した動画の編集による動画であれば、何らかの音声が入っているのでそれを利用すれば良いでしょう。ただいくつかの音声付きの動画を合わせる場合、音量を揃えないとびっくり動画になりやすいので注意します。

画像と文字から作成した動画には元からの音声がありません。無音の動画も寂しいのでBGM (Back Ground Music) を付けることも検討しましょう。ただ音楽にも著作権があります。自分でBGMを作曲できる人はそれほど居ませんので、他の人の作曲した音楽を使うことになります。TikTokではこの著作権の問題をクリアした上で多くの曲を提供しています。またYouTubeは「YouTube オーディオライブラリ*30」というものがあり、ここにある曲は自由に使えるようですが、YouTubeの外では使えません。

Googleで「著作権フリー bgm」をキーワードに検索すると数多くの著作権フリーのBGMを扱ったサイトが出てきます。通常無料で使えるものが多いのですが、使用した動画の最後に誰の曲なのか明記するのが条件とか、商用はだめ、アダルトはだめとかの条件がついていますのでよく確認しましょう。このような曲をなぜ無料で公開しているのか？という疑問も出ると思います。一つは作曲者の宣伝のためと言うのがあろうでしょう。実際に作曲者に有料で別の曲の作成を依頼できるようなサイトもあります。以前このようなサイトを利用した際に困ったのは、曲がありすぎてどれにしたら良いのかわからない、という点です。これから作る動画にふさわしいBGMとはどんな曲でしょうか？また曲を終わりまで聞かないと判断しにくいところも多数の曲から選ぶ際に困る点です。

8.1.5 画像

写真画像にも著作権や肖像権など様々な権利が絡みます。自分の会社のものならば大丈夫とも限りません。椋山女学園では、赤丸に木の「学園章」については申請を出してもなかなか使用許可が出ませんが、緑のSugiyamaの「コミュニケーションマーク」は勝手に使っても構わないようです。なぜそうなっているのかよく分かりませんが、同様の訳のわからないこだわりは、どこの組織でもありそうなので注意しましょう。

無料で使えるイラスト類はGoogleで検索するといくらでも出てきます。ものの説明用の写真などはWikipediaを利用すると説明用文章の元まで同時に手に入るかもしれません。「コンピュータと情報I」のテキストにも出てくる「クリエイティブコモンズ*31」などで探すのも良いかもしれません。

8.2 動画作成の例

ここではLightworksを使用して動画を作成する例を説明します。「現代マネジメント学部の歴史」動画用の素材を用意しましたのでそれを使います。「教材フォルダ」から「movie」フォルダをドラッグしてmarsのデスクトップにコピーしてください。そしてコピーした「movie」フォルダを開いて、図8.1の「Lightworks」フォルダ①を左側にある「ホームフ…」③のアイコンにドラッグして、「ホームフォルダ」にコピーします。同様に「.Lightworks.thereCanBeOnlyOne」ファイル②も同じところにドラッグしてコピーします。Lightworks

*29 「読書速度ハカルくん」(<https://www.sokunousokudoku.net/hakarukun/>)で測定したところ、私は1286文字/分で2/2問正解で平均よりやや速いという判定でした。ちょっと焦って読んだので正解できて良かった。

*30 「オーディオライブラリの音楽と効果音を使用する」<https://support.google.com/youtube/answer/3376882?hl=ja>

*31 <https://search.creativecommons.org/> ただし検索語は英語で入れる必要がある。

の無料版を使用するにはユーザ登録が必要で、90日毎にパスワードの確認を求めてきます。設定ファイルをコピーすることによりあと80日ぐらいは、私の登録情報で使用できるようです。

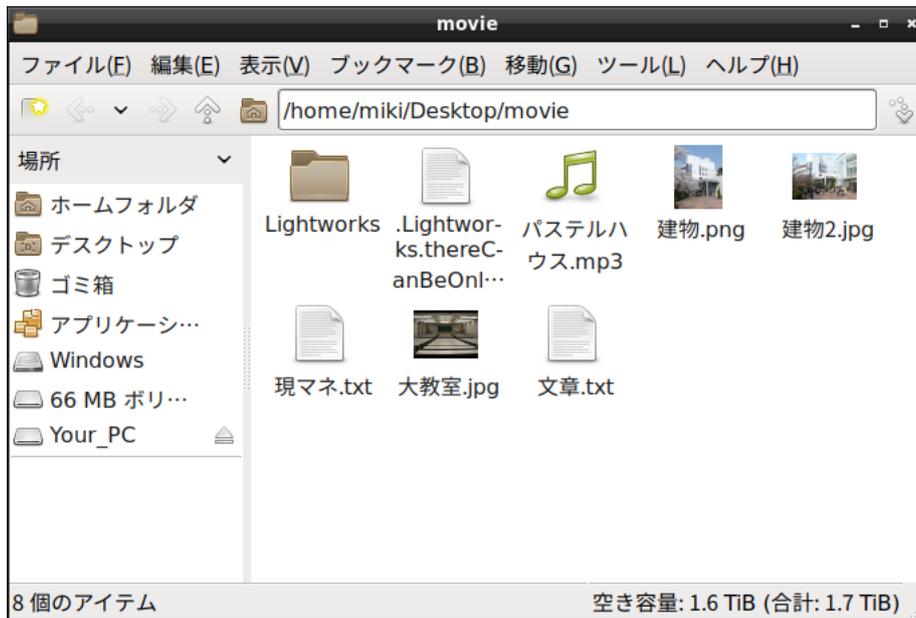


図 8.1 Lightworks の設定ファイルのコピー

「Lightworks」の起動は、メニューの中の「サウンドとビデオ」の中の「Lightworks」をクリックします。すると図 8.2 のような画面が表示されます。(その前に「Upgrade to Lightworks Create」の画面が出たら「今後これを表示しない」を選択してから続行をクリックしてください。)

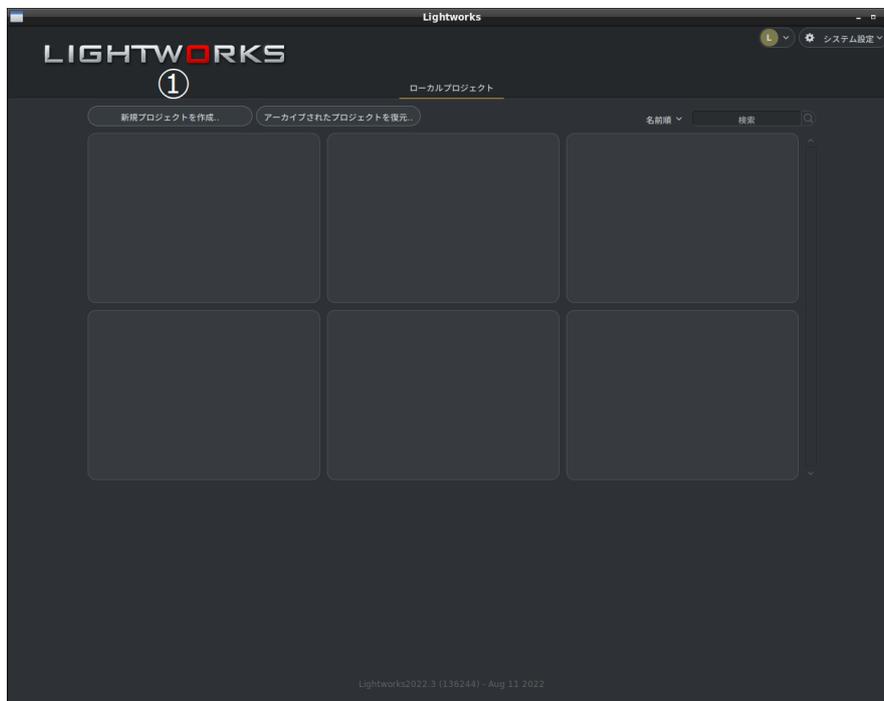


図 8.2 Lightworks の最初の画面

1. 初めての場合や新しい動画を作成する場合は、図 8.2 の①の下にある新規プロジェクトを作成をクリックします。「新規プロジェクトの詳細」というダイアログが出てきますので、名前のところに適当な内

容を入力してから「作成」をクリックします。とりあえず「現代マネジメント学部の歴史」と入れてください。ところが漢字入力に切り替わらないので、コピペで入れてください。(文字入力は「movie」フォルダの中の「文章.txt」を開いてコピペしてください。)

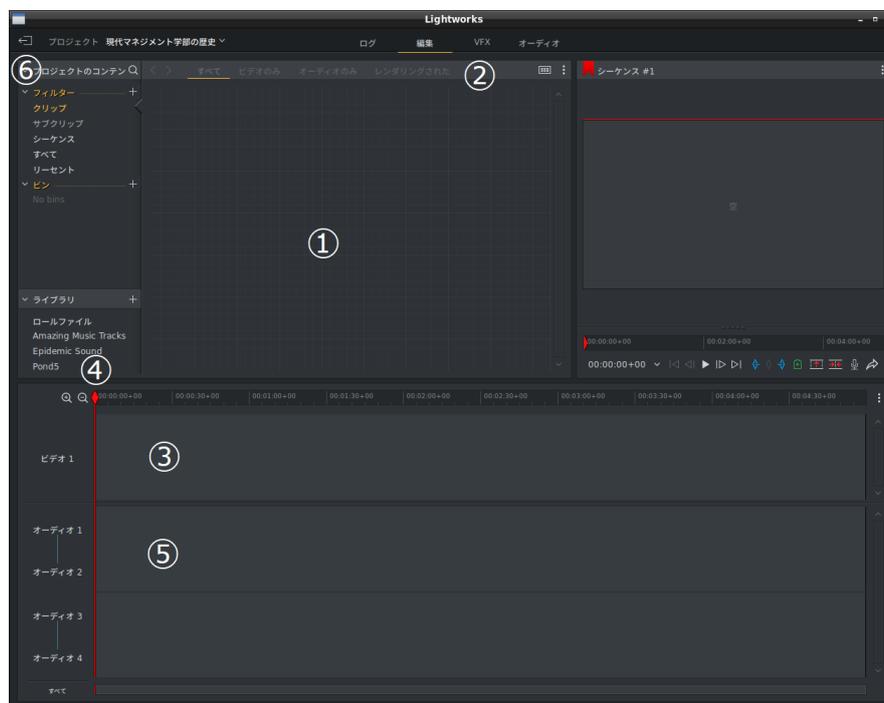


図 8.3 Lightworks の編集画面

2. 図 8.3 のような画面に変わります。①のところに、動画で使用する動画・画像・音声などのファイルをドラッグします。「movie」フォルダにある「建物.png」, 「大教室.jpg」, 「建物 2.jpg」と「パステルハウス.mp3」を入れてください。「パステルハウス.mp3」をドラッグすると「フレームレートを確認してください」と言うダイアログが出るので「インポート」をクリックしてください。
3. ②の上のところに「ログ」, 「編集」, 「VFX」, 「オーディオ」と並んでいますが、ここをクリックすることにより、画面が切り替わります。「編集」の画面で動画、画像、音声などのファイルを③や⑤のビデオトラックや音声トラックに配置したり、切り詰めたりします。「VFX」の画面で配置した動画や画像に様々な効果を追加します。タイトルや説明文章の追加もここでします。「オーディオ」の画面で音声の音量などの設定をします。
4. 作業が終わったら⑥の上をクリックして、図 8.2 へ戻ってから終了します。

8.3 タイトルの挿入

Lightworks ではタイトルや説明文章を、動画や画像に設定を追加するという形で入れることができます。また独立したものとしてタイトルや説明文章を入れることもできます。ここでは長さや位置の調整がしやすい後者の方法の説明をします。

図 8.3 の②の上にある「VFX」をクリックすると画面が切り替わり図 8.4 のような各種効果を設定する画面になります。図 8.4 の①のところの「Titles」をクリックしてから、②の右にある「対象リージョン」を「現在のクリップ (展開)」にし、③の左にある「エフェクトを適用」をクリックすると、④の上に「エフェクト 1」と言うところが出て、緑色になります。また左側に図 8.5 のようなタイトルの設定画面が表示されるので、ここで次のようなタイトルの設定を行います。

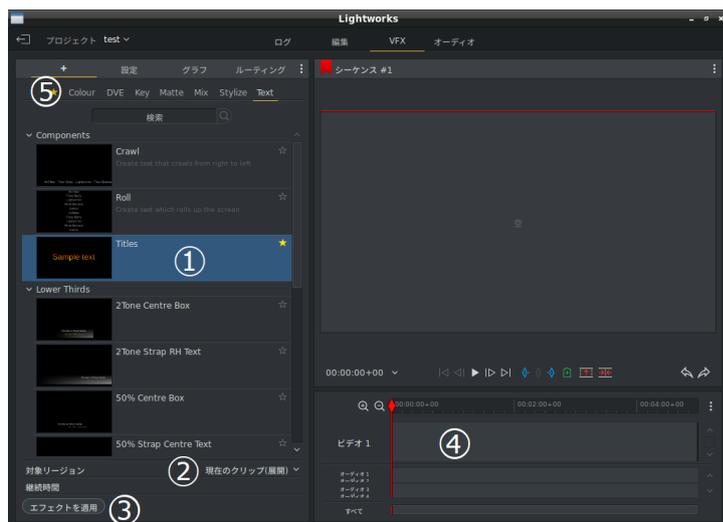


図 8.4 各種効果設定画面

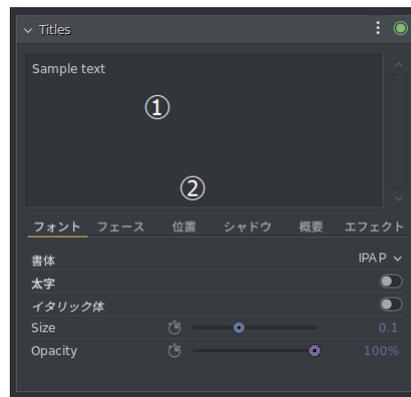


図 8.5 タイトルの設定画面

1. 図 8.5 の①の部分に「Sample text」という文字が入っているので、これを消して「現代マネジメント学部の歴史 A21EA999 梶山花子」を文章.txt からコピペして入れます。
2. 図 8.5 の②の下の「フォント」をクリックすると、書体、文字の大きさ (Size) や文字の不透明度 (Opacity) を変更することができます。mars には日本語フォントをほとんど入れていないので、書体はゴシック体と明朝体ぐらいしかありません。
3. 図 8.5 の②の下の「フェース」をクリックすると、文字の色を変更することができます。「外観」のところで単色、縦方向の色変化、横縞が選択できます。その下の「Colour One」と「Colour Two」の白丸のところをクリックすると色を変更することができます。
4. 図 8.5 の②の下の「位置」をクリックすると、タイトルを表示する位置を変更することができます。「Position X」で横方向、「Position Y」で縦方向、「Leading」で行間の幅、「アンカーポイント」で左寄せ、中央揃え、右寄せが指定できます。
5. 図 8.5 の②の下の「概要」をクリックすると、文字に輪郭線を付けることができます。
6. 図 8.5 の②の下の「エフェクト」をクリックすると、タイトルの最初と終わりに動きを付けることができます。最初 (In) では、ぱっと表示する「No effect」、段々見えるようになる「Fade」、一文字ずつ順番に出てくる「Typewriter」が選択できます。後の 2 つについてはその下で実行時間の設定もできます。最後 (Out) の方は「No effect」と「Fade」しか選択できません。最初と最後に「Fade」の設定を試みましょう。
7. 画面の右下にある図 8.6 の「エフェクト 1」の緑色のバーの右端をドラッグすると、①のように長さを変更することができます。かなり長い時間になっているので、どんどん左へ短くしていきます。②の右の⊕をクリックすると時間軸が拡大されるので、緑色のバーの右端を「00:00:05+00」の辺りまで短くします。
8. 図 8.6 の③の上の三角のボタンをクリックすると動画が再生されます。④の下の赤い縦線が右へ動いていきます。この赤い線のあるところが今表示されているところを示しています。赤線をドラッグするか、③の上に並んでいるボタンで、赤線の位置を変更することができます。

8.4 動画や画像の挿入

図 8.7 の①の上にある「編集」をクリックすると、動画、画像や音声を配置や設定をする画面になります。②にある動画や画像のアイコンを③にある「トラック 1」の適当な場所にドラッグします。画像は表示の長さ

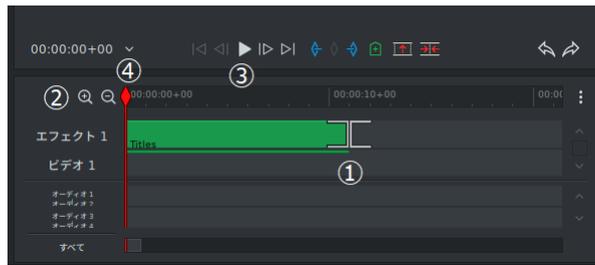


図 8.6 タイトルの表示時間の変更

を調整、動画の場合は不要な部分を削除する必要があります。動画や画像をクリックして選択した上で、図 8.7 の④の下の赤い縦線を必要な部分と不要な部分の境界に動かして、③のキーを押すとそこで動画が分割されますので、不要な部分をクリックして選択して[Delete]キー削除します。

動画や画像の面積が作成中の動画の面積より小さい場合、中央付近に表示されます。また動画や画像の一部だけを見せたいこともあるでしょう。このような場合は、まず図 8.7 の①の斜め上の「VFX」をクリックして画面を切り替え、④の下の赤い縦線を、対象となる動画や画像の上に移動します。するとその上に現在の状態が表示されます。そして対象となる動画や画像の上で右クリックをして、出てきたメニューの「エフェクト」の下の「追加」をポイントすると、サブメニューが表示されるので、その中の「2D DVE」を選択します。「2D DVE」は既に一度使用していると上の方に出てきますが、初めての時は、下の方にある「DVE」の見出しの下にあります。

すると左上の部分が「2D DVE」の設定画面に切り替わるので、「Position X」で水平位置、「Position Y」で垂直位置、「Scale」のところで大きさ、「Crop」の所でトリミングなどの設定ができます。マウスでスライダーをドラッグしますが、かなり大雑把な動きになるので、スライダーの右の数値をクリックして変更したほうが微調整がしやすいでしょう。

5 秒間のタイトルの後に「建物.png」、「大教室.jpg」、「建物 2.jpg」を順番にトラック 1 へドラッグして、長さをそれぞれ 20 秒にしましょう。画像の間隔はなしにして、ぴったりくっつけます。また「建物.png」は左に寄せて、「建物 2.jpg」は右に寄せましょう。

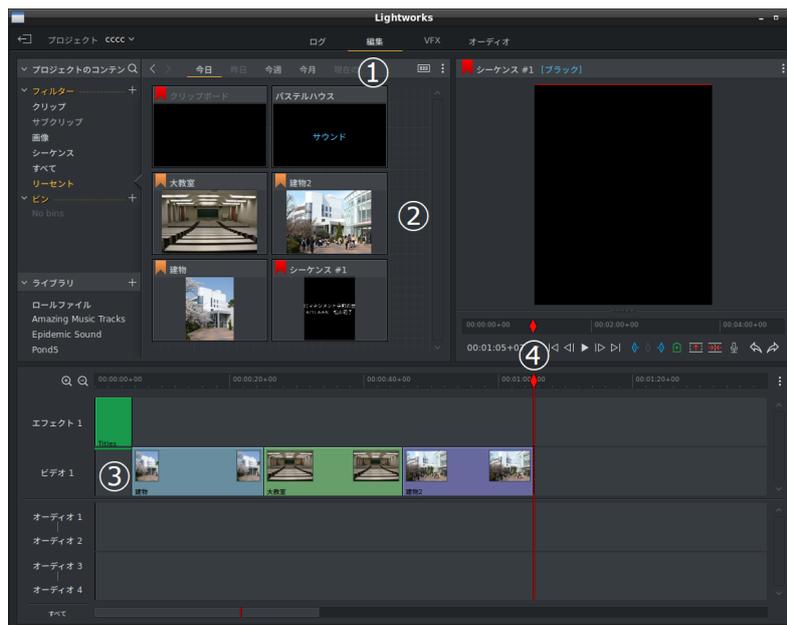


図 8.7 動画や画像の挿入と設定

8.5 動画や画像の切り替えの設定

動画や画像の明るさを最初は暗く、だんだん明るくすること (フェードイン) や、最後を段々暗くすること (フェードアウト) ができます。また動画や画像が、次の動画や画像に次第に切り替わるようにすることができます。まず設定をする場所を選択する必要があります。

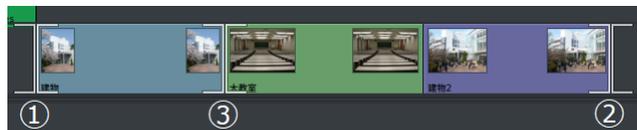


図 8.8 動画や画像の初め、最後、中間の選択

トラック 1 にある動画や画像の最初、最後、2 つの動画や画像の中間のあたりへマウスを持っていくと図 8.8 のように白い線が出ます。フェードインの場合は図 8.8 の①の上のようになった所で右クリックして、出てきたメニューで「トランジション」を選択します。するとサブメニューが出てくるので、「Dissolve」を選択し、「適用」は「始点」を、「継続時間」は適当な秒数とフレーム数を選択します。フェードアウトの場合は、図 8.8 の②の上のようになった所で右クリックして、「Dissolve」を選択し、「適用」は「終点」を選択します。2 つの動画や画像の切り替えの場合は、③の上のようになった所で右クリックして、「適用」は「ここを中心にする」を選択します。切り替えの場合は「Dissolve」以外の「Luma Wipe」(現在のものが暗くなり、次のものが明るくなる)、「Push」(現在のものが次のものに押し出される)、「Squeeze」(現在のものの幅が狭くなり、次のものの幅が広がる)、「Wipe」(現在のものが端から次のものに置き換わる) を選択しても良いでしょう。

フェードインなどの設定を変更する場合は、フェードインの設定がされた部分を右クリックして出てきたメニューで「削除」を選択し、再設定します。

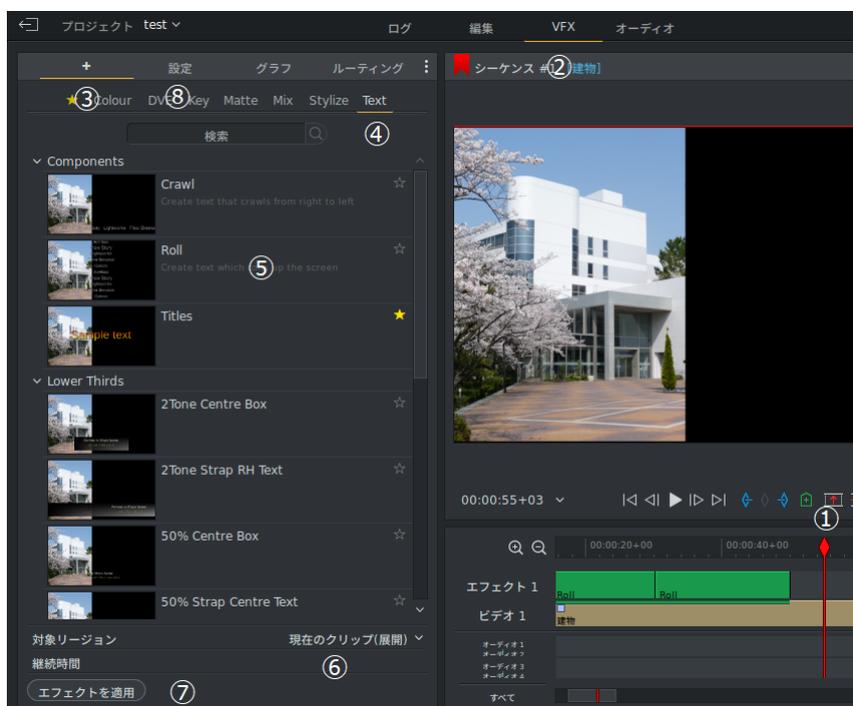


図 8.9 文章の追加

8.6 説明文章の挿入

説明文章を追加するには、図 8.9 の①の赤線の位置を説明文章を入れるところまで動かして、②の上の「VFX」を選択し、③の上の「+」をクリックし、④の上の「Text」をクリックし、⑤のところの「Roll」をクリックして選択します。⑥の上の「対象リージョン」を「現在のクリップ (展開)」にし、⑦の左の「エフェクトを適用」をクリックします。なお、「Roll」の上にある「Crawl」を使うと文字が一行に左から右へ流れるものになります。

文章の追加の際に「エフェクト 2」ができて、そちらに入ってしまった場合は、「エフェクト 1」の方にドラッグして移動します。なお同時に複数の文章を流す場合はそのままでも構いません。

図 8.9 の⑧の上の「設定」をクリックし、⑧の辺りに出てくる「エフェクト 1」を選択して、後はタイトルと同じように内容をコピーして入れます。この時①の赤線の位置が、今設定しようとしている文章 (Roll) を突き刺すようにします。文字の大きさは 0.7 にします。長さを 10 秒程度にします。この長さに合わせて文字が動きますので、再生してみて動きが早すぎるようであれば、長さの数字をより大きくします。

「文章.txt」の説明 1~3 を別々の説明文章にし、それぞれ 3 つの画像の中央付近に出るようにします。エンドタイトルは 5 秒ぐらいのタイトルとして入れましょう。

8.7 音声の設定

図 8.3 の①にある音声ファイルのアイコンを⑤の「オーディオ 1 - オーディオ 2」の所へドラッグします。動画と同様に④の下の子線が必要な部分の最後 (例えばエンドタイトルの数秒後) に動かして、C キーで切り分けて余分な部分を削除します。再生をしてみて、音量が適切でなければ、右クリックして出てきたメニューで「ゲイン (db)」の右にあるスライダーで調整します。

これだけでは音楽がぶつんと切れてしまうので、次のようにフェードアウトの設定をします。図 8.10 の①の上の「オーディオ」をクリックして画面を切り替えます。そして②の横の音声の右上角のところにマウスを近づけると、白い線の三角形が出てくるので左の方へドラッグします。より左の方へドラッグするほどゆっくりフェードアウトするようになります。

左上の角のところで同様のことを行えば、「フェードイン」になります。

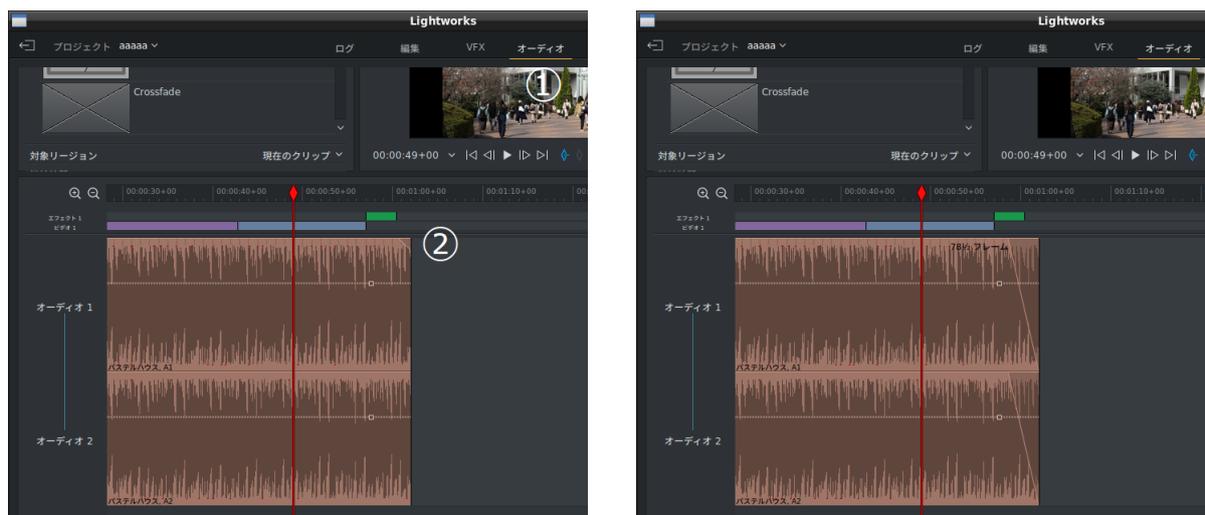


図 8.10 フェードアウトの設定方法

8.8 動画の保存

図 8.3 の②の斜め上の「編集」をクリックして編集画面に切り替えて、③か⑤の空いた所で右クリックして出るメニューで、下の方にある「エクスポート」をポイントするとサブメニューが出るので、「H.264/MP4」を選択します。すると図 8.11 のような設定画面が出るので、必要に応じて設定を変更して「開始」をクリックします。「フレームレート」は毎秒何枚の画像にするかで、数が多いほどなめらかな動きの動画になりますが、ファイルが大きくなります。サイズは動画の解像度で大きいほど、拡大しても荒くならない動画になりますが、やはりファイルが大きくなります。無料版の Lightworks では 720p より高い解像度の動画は作成できません。



図 8.11 出力動画の設定

8.9 動画編集のまとめ

世の中には様々な動画編集ソフトがありますが、基本的な使い方は Lightworks と大きな違いはありません。画像、音声、動画を用意して登録し、それをトラックに順番に並べて一つの動画にまとめます。高価な動画編集ソフトほど操作が簡易で、動画などの切り替えの際などの効果設定が豊富です。ただ豊富な効果設定も多用すると逆効果になることもあるので注意します。

どのような動画編集ソフトを使用しても、最後の動画の生成には時間がかかります。様々な形式の動画や画像を一旦圧縮を解除して形式を揃えた上で再度圧縮し直すからです。よって数分以上の動画を作成するならば、それなりの性能のパソコンを使用しないと時間ばかりかかってしまいます。

元となる動画もコンピュータで作成するとなると、技術だけでなくデザインの才能なども必要になります。そういう方面に興味がある方は良いのですが、それほど才能もないという方はここで紹介したレベルの動画で、内容で勝負した方が良いでしょう。つまりビジュアルに表現するのではなく、伝えたいことを文字や音声で示す方向です。そして伝えたいことが明確な動画は、他の人にもっと見た目を良くしてもらうことは可能です。逆に何を言いたいのかわからない動画は、自分で最初から作り直すしかありません。