

動画教材の自動生成スクリプトの試み

小山裕司¹

概要: 最近、録画型授業、反転型授業、YouTube等の動画 SNS の普及に従い、録画教材作成の機会が急増している。従来の GUI 環境での手作業では、作業の再利用性が低いため、作業時間を要してしまい、教材の品質改善に時間を十分に確保することが難しかった。今回は音声合成、制御構造、変数等の特徴を有するスクリプトで動画教材を表現し、作業効率を高める試みを報告する。

キーワード: スクリプト表現, 動画教材, 変数, 制御構造, 自動処理

An Automatic Scripting for Video Educational Materials

KOYAMA Hiroshi^{†1}

Abstract: Recently, video classrooms, flipped classrooms, and video social networking services such as YouTube have become popular drastically. Thus, opportunities for creating video educational materials have increased rapidly. The conventional method requires a lot of effort because it is manual, not reusable, and inefficient. Therefore, it is difficult to improve the quality of materials themselves. We have developed a prototype scripting system for video educational material to solve these problems. This paper reports several features of this scripting language and its effects.

Keywords: Scripting System, Video Educational Material, Variable, Control Flow, Automation

1. 背景

最近では、録画型授業、反転型授業、YouTube等の動画 SNS の普及に従い、録画教材作成の機会が急増している。東京都立産業技術大学院大学（以下、本学）は90%以上が社会人学生であり、2006年の開学当初から、授業動画の収録及び視聴環境の整備、遠隔型授業等、社会人の修学のための環境改善を継続的に行ってきた。

1.1 ブレンディッドラーニングの取り組み（2020年度）

2014年には、「ブレンディッドラーニング」が概念を実現する取り組みとして、従来の仕組みを拡張し、《録画視聴型の授業（録画型授業）》を開始した。この仕組みでは、週2回の授業を講義（対面あるいは録画視聴）及び演習（対面）の組から構成し、講義の回の対面型授業を欠席した場合でも、次の授業までに録画された授業動画を視聴すれば出席扱いにした（従来の仕組みでは欠席扱い）。これは、従来に増して社会人学生の通学負担を軽減することで修学機会を改善すると同時に、（反転教育類似の）録画視聴の繰り返し及び教材の改善によって学生の理解度を高める効果から高度専門職育成の教育の質の改善に寄与することが目的であった。情報アーキテクチャ専攻では、時間割を工夫する等の取り組みを行い、2014年度の授業科目の約90%にあたる45科目以上がこの仕組みに対応した。また、録画型授業は、通常の授業を単に録画したものが大半を占めるが、専用の動画教材を準備する科目もあった[1]。

1.2 新型コロナウイルス感染症の影響（2020年度）

2020年度は、新型コロナウイルス感染症の流行によって、大半の大学にICTを活用した遠隔型授業、録画型授業が急速に普及し、この体制が現在まで継続している。

本学は、前項に示した「ブレンディッドラーニング」の仕組み及び環境を有し、録画型授業、遠隔型授業が可能であった。この遠隔型授業は専用機材及び回線によって平日夜の秋葉原サテライトキャンパスに対してのみ行われていた。2020年度の専攻改組の際に、学生の利便性向上のため、インターネット上のどこからでも参加できるように改修する計画があり、著者らはこれを転用し、ハイブリッド型授業環境を構築した。この環境の特徴は、映像及び音声に関し、Zoom環境を既存の対面型授業の教室環境に適切に組み込んだことで、教員は従来の通りに授業をすることができ、教員の負担をある程度軽減でき、また感染症の流行が収まり、学生が教室に戻ってきた場合でも対面型授業及び遠隔型授業を比較的に円滑に連携できることにある[2]。

本学も、新型コロナウイルス感染症の流行によって、事実上すべての対面型授業を遠隔型授業として開講したが、学生のインターネット環境等を考慮し、原則すべての授業を録画型授業の受講でも可としたため、同様にすべての授業を録画型授業としても開講した。録画型授業は、通常の対面型授業を録画したものが大半を占めるが、専用の動画教材を準備したり、教材の改善のため録画動画を編集したりする科目もあった。

¹ 東京都立産業技術大学院大学
Advanced Institute of Industrial Technology

1.3 動画教材作成の負担

録画型授業の実行にあたっては、通常の対面型授業を録画するだけであれば、対面型授業を既存の授業収録環境で録画し、公開するだけであるため、教員の負担は視聴確認テストの作成及び評価だけが增加する。

しかし、専用の動画教材を準備したり、教材の改善のため録画動画を編集したりする場合、教員の負担増加は以下に列挙する各種の取り組みによって様々である。

- ① 本学の授業収録の環境を使って、通常通りの授業を録画する。
- ② 自前の録画装置を使って、通常通りの授業を録画する。
- ③ Zoom 環境の録画機能を使って、通常通りの授業を録画する。
- ④ 事前録画の授業のために専用の動画教材を作成する。

2014 年当時、著者は (4) の専用の動画教材を、通常の授業のように PowerPoint で作成したスライド画像を iMovie 等の動画編集環境に読み込み、音声吹き込み、余裕があれば、各種のエフェクトを付けることで準備した (図 1)。



図 1: PowerPoint を読み込んだ iMovie の画面例

90 分の動画教材作成に当初 (2014 年 4 月頃) は約 1 週間 (約 60 時間の試行錯誤) がかかったが、終盤 (2014 年 12 月頃) には最短では約 3 時間に短縮できた。通常の授業でも使う PowerPoint スライド (授業資料) 等の教材作成以外の内訳は以下の通りである。

- 音声の収録 (90 分から 180 分)
90 分の授業であるため、最低でも 90 分を要し、取り直したり、編集を考慮して多めに収録したりすることを考えると、3 時間から 4 時間かかる。
- 動画編集 (60 分から 120 分)
無音部分等を削除したり、文字テロップを付加したり、各種のエフェクトを行ったりする程度に依るが、2 時間から 3 時間を要する。
- できあがった授業動画の確認 (90 分)
90 分の動画の確認には最低でも 90 分を要する。修正が必要であれば、再度収録、編集を要することがある。
- 最後に、最終動画の生成 (動画のエンコーディング)、クラウドストレージへの登録 (アップロード) の時間

がかかる。

当初は課題であった以下の事項は 2022 年現在までにある程度緩和した。

- 当初は動画編集に MacOS 付属の iMovie を使用したが、簡易版であり、修得は容易であったが、機能が単純で幾つかの制約があった。逆に、Apple Final Cut Pro、Adobe Premiere Pro、Blackmagic Design DaVinci Resolve [3]等は概ねあらゆる機能が存在するため、自由度が高いが、活用できるように修得するまでには相当の労力が必要であった。著者は現在 DaVinci Resolve を使い、編集の自由度及び効率があがった。
- 動画教材作成の作業は、試行錯誤の段階から次第に慣れ、無駄を避け、効率的に作業する工夫ができたり、動画の雛型を再利用したりすることができ、作業時間を短縮できた。
- 動画教材作成の処理は PC に負担がかかる。当初は音声収録時に PC の冷却ファンの雑音がかかる程度であったが、PC の性能が上がったため、解決した。特に Apple Silicon 搭載の PC (MacBook Pro) は低発熱であり、また最終動画の生成時間を劇的に短縮できた。

1.4 動画教材作成の課題

現段階での動画教材作成に対する課題を以下に列挙する。

- GUI 環境での手作業
動画教材作成、特に編集作業は概ね GUI 環境での手作業である。GUI 環境での作業は直感的ではあるものの、作業の繰り返し、再利用、自動処理には不向きであり、また PC の負担が大きい。
- 音声収録
音声は、原稿を作成し、雑音を避けた環境で収録する。実際の収録時間だけ要する (1 時間の教材であれば、概ね 1 時間は要する)。雑音、間違いがあれば、編集したり、再収録したりする必要がある。
- 字幕
動画教材に字幕を付加したいことが多い。字幕は原則手動で付加する必要がある。収録した音声に相当する字幕を付ければ動画教材の付加価値があがる。音声から字幕情報を生成してくれるソフトウェアが存在するが、最終的には手作業での調整が必要である。
- 内容の更新
動画教材は年次ごろに内容の修正、改善を行い、継続的に内容を高める必要があるが、一括で内容を変更する変数、マクロ、置換に相当する機能は弱いため、これを効率的に行うことは難しい。また、一連の動画教材も作成するための雛型に相当する機能も弱い。

本稿の第2章では、これらの懸案事項を解決するため動画教材作成のためのスクリプト表現の構想を示す。第3章では、実装の概要を示す。最後の第4章では、試作版から現状の課題及び今後の改善を列挙し、本稿をまとめる。

2. 構想

今回の取り組みでは、現状の動画教材作成の課題をスクリプト表現によって解決することを試みる。動画の表現にスクリプトを活用した事例としては、AviUtl [4]がある。AviUtlは、組み込み目的のスクリプト言語 Lua によって動画中のオブジェクトの動き等を表現することができる。テキストの台詞から動画に字幕及び音声が付与する機能を有する動画編集支援環境としては、ゆっくり MovieMaker4 [5]がある。

今回の動画教材作成のためのスクリプトでは、設定、画像の指定、台詞を時系列スクリプト上に表現し、時系列XMLに変換することで、再利用可（修正可）、自動処理可を実現し、作業効率を高めることを目指す。音声は時系列スクリプト上の台詞から音声合成によって生成する。字幕も時系列XMLに付加する。また、プログラミング言語の変数、制御構造の機能を付与する。

3. 実装

今回の試作版は、動画スクリプトから時系列XML（現状FCPXML, Final Cut Pro XML [6]）ファイルを生成するPython [7]スクリプトへ変換する処理系として実装した。動画スクリプトの例を図2に示す。

```
#source ./sp1/*
#template default
$n = 1
{
  image: s000.png
  quote: "皆さん、こんにちは。 \n 第${n}回の授業（略）"
}
foreach $e (1, 2, 5)
{
  image: se0${e}.png
  quote: "次は演習${e}を行います。 \n スライドの（略）"
  duration: 1m15s
}
```

図 2: 動画スクリプト例

- 1行目： 動画素材（スライド画像等）を収納する場所の指定
- 2行目： 各種設定の指定
- 3-7行目： 変数、画像及び台詞の指定
- 8-13行目： 反復構造、画像の表示時間の明示的指定

動画スクリプト等からFCPXMLファイル等への変換の処理の流れを図3に示す。

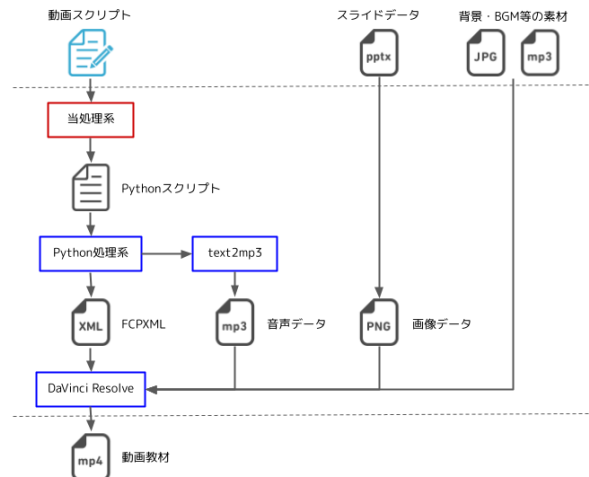


図 3: 変換処理の流れ図

- ① 動画スクリプトのほかに、スクリプト中で仕様する教材のスライド画像をPowerPoint等で準備する。必要であれば、背景、BGM等の素材も準備する。
- ② 今回試作した処理系は、動画スクリプトを一旦Pythonスクリプトに変換し、このPythonスクリプトが時系列XMLを生成する。この際に、スクリプトに書かれた台詞は時系列XML上に字幕として記録され、同時にtext2mp3等の音声合成によって音声データが生成する。これらの変換処理はmakeによって自動一括処理する。時系列XMLファイル例の抜粋を図4に示す。

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE fcpxml>
<fcpxml>
  <resources>
    <format width="1920" id="f0" height="1080" frameDuration="1/30s"/>
    <effect id="title"/>
    <asset format="f1" start="0s" hasVideo="1" id="s000" duration="0s">
      <media-clip src="file:///Users/koyama/footage/sp01/slido/s000.png" kind="original-media"/>
    </asset>
    <asset start="0s" id="a000" hasAudio="1" duration="293/60s">
      <media-clip src="file:///Users/koyama/footage/sp01/audio/a000.mp3" kind="original-media"/>
    </asset>
  </resources>
  <library>
    <event>
      <project name="Timeline SP 001">
        <sequence format="f0" tcFormat="NDF" duration="14033/60s" tcStart="216000/60s">
          <spine>
            <video offset="216000/60s" ref="s000" duration="533/60s">
              <title offset="1s" ref="title" lane="1" duration="413/60s">
                <text>
                  <text-style ref="ts">皆さん、こんにちは。</text-style>
                  <text-style ref="ts">システムプログラミング特論 第1回の授業を始めます。</text-style>
                </text>
                <text-style-def id="ts">
                  <text-style fontSize="50" font="RoundedM+ 1c" fontColor="1 1 1"/>
                </text-style-def>
              </title>
            <asset-clip offset="2s" lane="2" ref="a000" duration="293/60s">
              </video>
            </spine>
          </sequence>
        </project>
      </event>
    </library>
  </fcpxml>
```

図 4: 時系列XMLファイル例の抜粋

- ③ 最後にDaVinci Resolveが時系列XMLファイルを読み込み、動画教材を生成する。時系列XMLファイルを読み込んだDaVinci Resolveの画面を図5に示す。字幕の調整等、必要であればDaVinci ResolveのGUI環境での修正も可能である。



図 5: DaVinci Resolve の画面例

4. 結言

録画型授業では、通常の対面型授業を録画した動画が大半を占めるが、専用の動画教材を準備したり、通常の対面型授業を録画した動画であっても教材の改善のため動画を編集したりすることもある。著者の授業科目では専用の動画教材を準備することで、次に示すように教材の質及び量は概ね上がっている。

- 通常の対面型授業の録画では、無音、準備、操作ミスであったり、同じ内容を繰り返して喋ってしまった部分が生じてしまうが、これらの部分を削除したり、速度エディタ（再生速度 50%等）を活用したりすることで、同じ時間で概ね 1.5 倍から 2 倍の内容を収録できる。しかし、あまり内容が多すぎる場合は理解度が下がる。また、繰り返しは教育上の効果があることもある。
- 授業を適切に準備し、内容及び時間面で計画的に進めることができる。手抜きも減る。

教材の質及び量が向上したこと以外でも録画型授業では学生の理解度が上がっている長所もある。

- 学生の修学コストが下がる（修学時間帯概ね自由、通学時間無し）
- 今までであれば最悪着席していれば、また遅刻であっても出席扱いであったが、視聴確認テストの関係で、集中して授業動画を視聴する必要がある（実際、視聴及び確認テストに 90 分以上要している学生が多い）。
- 内容がある程度理解できている状態で対面型授業での演習を始めることができる。

授業の内容の設計では、録画型授業の長所を活かし、対面型授業では学生の授業参加、ディスカッション、演習、質疑等を適切に組み込み、講義・演習の調整を行い、授業内容の改善をしていくことが重要である。

本稿では、動画教材作成のためのスクリプト表現の特徴及び実装を示した。従来の GUI 環境での手作業では、作業の再利用性が低いため、作業時間を要してしまい、教材の

品質改善に時間を十分に確保することが難しかった。今回試作した動画教材作成のためのスクリプトでは、次に列挙する事項を表現できることから作業効率をある程度高めることができる。

- 設定及び時系列の表現（再利用可、自動処理可）
- 台詞文字列から音声及び字幕の生成（再利用可、修正可）
- 変数、制御構造

また、スクリプトの再利用によって、継続的に動画教材の内容は改善を行うことができ、スクリプトから音声合成が行われるため、作業環境を選ばず、交通機関での移動時にも作業を行うことができることも長所にあげられる。

今回の取り組みでは中核機能を優先し、設計及び実装を行ったが、現状の課題を以下に列挙する。

- 実行環境の整備（クラウド等）
- 音声合成のための外部連携の改良
- スライドの生成（字幕機能の拡張）
- 装飾（背景素材、音楽素材、トランジション、エフェクト等）及び雛型（テンプレート）の充実
- スクリプト表現の仕様の洗練

今後はこれらの改題に対処し、動画教材の効果及び効率の改善をはかる計画である。

謝辞 本発表にあたり、工学院大学附属中学校・高等学校中野由章校長に御支援を頂いた。ここに感謝の意を表する。

参考文献

- [1] 小山 裕司. 録画視聴型の授業の取り組み及び評価, 産業技術大学院大学 紀要. 2015, 第 9 号, pp.13-20.
- [2] 小山 裕司, 大崎 理乃, 木下 修司, 柴田 淳司. 情報系専門職大学院における対面・遠隔ハイブリッド型授業システムの開発, 産業技術大学院大学 紀要. 2020, 第 14 号, pp.193-200.
- [3] “DaVinci Resolve” . <https://www.blackmagicdesign.com/jp/products/davinciresolve/>, (参照 2022/10/30).
- [4] “AviUtl” . <http://spring-fragrance.mints.ne.jp/aviutl/>, (参照 2022/10/30).
- [5] “ゆっくり MovieMaker4” . <https://manjibox.net/yymm4/>, (参照 2022/10/30).
- [6] “FCPXML Reference” . https://developer.apple.com/documentation/professional_video_applications/fcpxml_reference, (参照 2022/10/30).
- [7] “Python” . <https://www.python.org/>, (参照 2022/10/30).