

## SMIL を用いた映像教材作成支援ツールの開発

柳沼良知†, ††

教材のデジタル化、オンライン化が進むにつれて、文字や静止画に比べて表現力のより高い映像を使った教材を容易に作成できるような環境が求められるようになってきている。このため、映像や音声といったメディアの時間的、空間的レイアウトを記述する SMIL を用いた映像教材作成支援ツールの開発を行った。今回開発したツールは、講義映像とスライドを同期した映像教材を作成するツール、スライドとその説明文から音声合成により映像教材を作成するツール、これらの補助として、映像から静止画を取り出し、等間隔、あるいは、色ヒストグラムを用いたシーン検出により、映像の静止画一覧を作成するツールの3つである。これらのツールに対して行った簡単な評価についてもあわせて述べる。

### Development of video contents authoring support tools using SMIL

Yoshitomo YAGINUMA †, ††

As instructional materials are digitized and online contents become popular, contents authoring tools which can easily create streaming video contents, which are usually more attractive than text and image media, are in many cases required. Therefore, developing of video contents authoring support tools using SMIL, which describes temporal and spatial layouts of media such as video and sound, was carried out. The tools are consists of three tools; a tool for synchronizing lecture video and slides, a tool for making video contents from slides and related text using sound synthesis, and a tool for extracting frames form video to make image lists using scene change detection techniques based on time and color histograms. Evaluation of these tools is also discussed.

#### 1. はじめに

近年、計算機の処理能力の向上や記録媒体の大容量化が進み、従来は難しかった映像の扱いが容易になるにつれて、映像をより高度に利用しようとする試みが数多くなされている[1-9]。教育においても、教材のデジタル化、オンライン化が進むにつれて、文字や静止画に比べて表現力のより高い映像を使った教材を容易に作成できるような環境が求められるようになってきた。このため、映像や音声と

いったメディアの時間的、空間的レイアウトを記述する SMIL を用いた映像教材作成支援ツールの開発を行った。今回開発したツールは、講義映像とスライドを同期した映像教材を作成するツール、スライドとその説明文から音声合成により説明音声付きの映像教材を作成するツール、これらの補助として、映像から静止画を取り出し、等間隔、あるいは、色ヒストグラムを用いたシーン検出により、映像の静止画一覧を作成するツールの3つである。以下、その概要について述べる。

†メディア教育開発センター

National Institute of Multimedia Education

††総合研究大学院大学

Department of Cyber Society and Culture, The  
Graduated University for Advanced Studies

#### 2. 講義映像とスライドの同期による映像教材の作成

講義映像とスライドの同期による映像教材

作成ツールは、時間的に変わるスライドの中に、講義映像を組み込んで Web ブラウザ上で表示できる映像教材を作成するツールである。このツールは、スライド (JPEG 等の画像ファイル) を再生し、時間情報を記録するためのツールである Easy Streaming NIME Presenter と、その結果と講義映像を同期し、Web ブラウザ上で表示できる形式に変換する Easy Streaming NIME Layouter からなる。

図 1 に、Easy Streaming NIME Presenter のインタフェースを示す。左上の「ファイルがあるディレクトリを指定」ボタンを押してスライド画像のファイルがあるディレクトリを指定する。(スライド自体は、市販のプレゼンテーションツールで作成する。これらのツールは通常、スライドを連番を付けた画像ファイルとして出力可能であるため、これらの画像ファイルを読み込み処理を行う。) 画像は、ファイル名の中の数字が小さい順に画面下側の領域に読み込まれる。「記録開始」ボタンを押すことで、スライドを再生したタイミングが記録される。「次へ」、「前へ」ボタンでスライドを移動できるほか、下側のアイコンをクリックすることで、そのスライドへ飛ぶことができる。「記録終了」ボタンを押すことで、図 2 のような、どのスライド (画像ファイル) がいつ再生されたかを記録したテキストファイルが作成される。

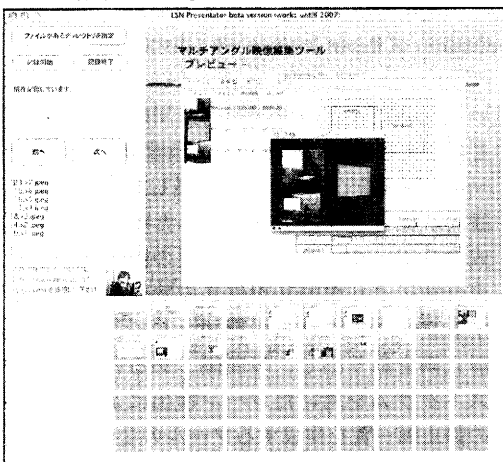


図 1 スライド再生ツール(Easy Streaming NIME Presenter)

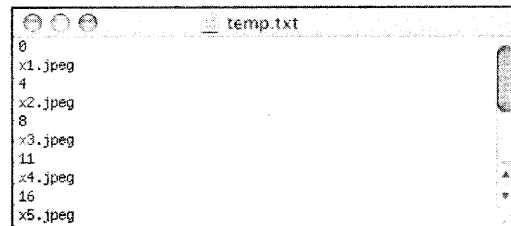


図 2 Easy Streaming NIME Presenter の出力ファイル

スライドの再生時間の情報を記録した Easy Streaming NIME Presenter の出力ファイルをもとに、スライドと講義映像を同期し、Web ブラウザ上で表示できる形式に変換するツールが Easy Streaming NIME Layouter である。図 3 にそのインタフェースを示す。

「ディレクトリを指定」ボタンで、Easy Streaming NIME Presenter で指定したディレクトリを指定する。このディレクトリには、スライドの時間情報を記述したファイル (temp.txt) とそれにより作成したスライドのムービーファイル (x.mov) が存在する。また、あらかじめ、講義映像はストリーミング形式に変換して、このディレクトリに y.mov という名前で保存しておく。映像フォーマットとしては Apple<sup>®</sup>の QuickTime<sup>®</sup>[10]形式に対応している。

「ここが 2 枚目」ボタンは、講義映像を人が見て、スライドを 2 枚目に切り替えるタイミングでクリックし、スライドと講義映像の同期を行うボタンである。

「ファイルを書き出し」ボタンを押すことで、図 4 に示すような、講義映像とスライドを同期した SMIL(Synchronized Multimedia Integration Language)[11]形式のファイルを書き出す。SMIL は、XML ベースで映像のレイアウトやその再生順序等を記述する言語である。SMIL ファイル自体はテキストファイルであるためデータサイズが小さく、実際の映像データを直接編集するわけではないので再生する画像の劣化もない。また、同時に、この SMIL ファイルから、QuickTime 形式のムービーとそのムービーを Web ブラウザ上で再生す

るための HTML ファイルを作成する。SMIL ファイルを QuickTime 形式のムービーに（形式的に）変換するのは、ユーザの環境により、SMIL ファイルを扱うための Plug-in の設定が異なっており、作成した SMIL ファイルを正しく再生できない場合があるためである。ただし、作成するムービーは SMIL ファイル同様、どのファイルをどういったレイアウトでいつ再生するかのレファレンス情報のみをもったファイルである。図 5 に Web ブラウザ上で再生した例を示す。この時、表示しているムービー全体が秒何フレームかの大きなムービーとして保存されているわけではなく、小さな講義映像のみが秒何フレームかのムービーとして保存されており、背景のスライドは静止画を切り替えのタイミングで読み込んでいることから、スライドはクオリティが高いまま、全体として小さなデータサイズで映像教材を作成できている。現在レイアウトとしては、スライド映像の右下の部分に講義映像を埋め込んだ形式を教材の形式としている。

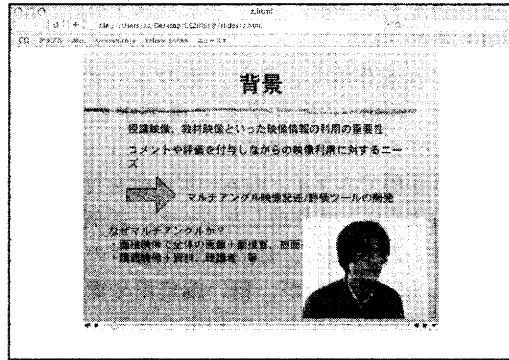


図 5 Web ブラウザでの再生

### 3. 音声合成によるスライドと説明文からの映像教材の作成

SMIL を用いた映像教材作成ツールのもう 1 つの例として、スライドとその説明文から音声合成により説明音声付きの映像教材を作成するツール (Easy Streaming NIME Slide2Movie) の試作を行った。音声合成エンジンは、Mac OS X に付属のものを用いているため、合成できる音声は現在のところ英語のみである。

このツールを用いて映像教材を作成する場合は、まず、図 6 の様な、どのスライドと説明文から映像教材を作成するかを記述したテキストファイルを用意する。このファイルの一番頭には、何枚のスライド (JPEG 等の画像ファイル) を使うかの枚数が書いてあり、その後に、スライドの枚数分だけ、スライド画像のファイル名、スライドのタイトル、音声合成するテキストが並んでいる。

このツールのインターフェースは至ってシンプルであり、立ち上げると、「ファイルを指定」というボタンが 1 つついたウィンドウが立ち上がるので、このボタンを押して、先ほどのテキストファイルを指定する。すると、

- 1) tempdir というディレクトリを作り、そこに画像ファイルをコピーする。
- 2) 説明文を Mac OS X の音声合成エンジンを用いて音声ファイルに変換し、tempdir に保存する。
- 3) それぞれの画像を「説明文の長さ+2 秒」

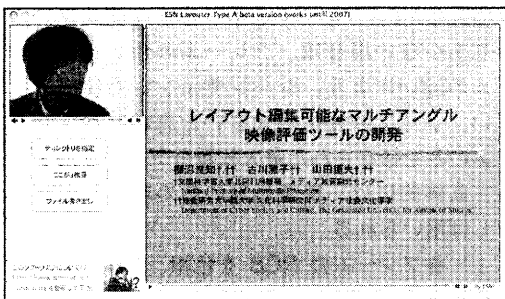


図 3 レイアウト編集ツール(Easy Streaming NIME Layouter)

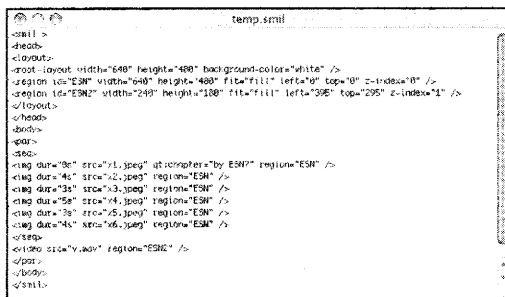


図 4 Easy Streaming NIME Layouter の出力 SMIL ファイル

ずつ再生する SMIL ファイルを作成する。

4) その SMIL ファイルをもとにしてムービーファイルを作成し、また、そのムービーファイルを Web ブラウザ上で再生するための HTML ファイルを作成する。

という手順でスライドのムービー化を行う。

このようにして作成したムービーの例を図 7 に示す。現状では、合成された音声の品質が必ずしも高くないが、粗い編集用や簡易的な教材の作成ツールとして利用できると思われる。なお、各スライドのタイトルとして入力された文字列は、図 7 の右下のようにメニューとして見ることができ、このタイトルを選ぶことで、そのスライドに飛ぶことができる。

```
xi.jpeg
title
The title is Multi-angle video annotation tool which has layout editing function.
x2.jpeg
background
In order to evaluate and utilize instructional movies, video annotation tool is sometimes requi
x3.jpeg
multi-angle video annotation tool
This video annotation package consists of two tools: a multi-angle video editing tool (Easy Str
xi.jpeg
multi-angle video editing tool
The multi-angle video editing tool (Easy Streaming NIME Multi-angle Editor) is a Java program t
x5.jpeg
video annotation tool
The video annotation tool (Easy Streaming NIME Annotator) is a tool for annotating multi-angle
```

図 6 読み込みファイルの形式

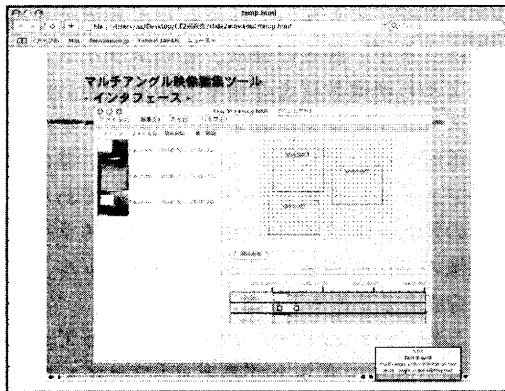


図 7 ムービー作成結果

#### 4. 映像把握のための映像一覧の作成

SMIL を用いた映像教材作成ツールの補助として、映像から静止画を取り出し、等間隔、あるいは、色ヒストグラムを用いたシーン検出により、映像の静止画一覧を作成するツール (Easy Streaming NIME Frame Extractor) の作成を行った。このツールにより映像から

必要なフレームを取り出し、前述の映像教材作成ツールの編集素材として利用することができる。対応しているフォーマットは QuickTime 形式のムービーである。

インタフェースは、図 8 によってなっている。「ムービーを読み込む」ボタンを押すことで、ムービーを読み込める。読み込んだムービーは左側のムービープレーヤに表示される。「このフレームを保存する」ボタンを押すことで、現在再生中のフレームを temp.jpeg という名前で保存できる。

映像一覧を作成する際には、処理するフレームの間隔と取り出すフレームの最大数を指定する。取り出すフレームの最大数を十分大きくした場合 (例えば、999999 等)、このツールは単純に指定された時間間隔で映像のフレームを取り出し、JPEG 形式で保存する。また、それらの一覧表示を行うための HTML ファイルも同時に出力される。取り出されたフレームのサイズはもとの映像のサイズと同じである。取り出すフレーム数が、(映像全体の長さ) / (フレームを取り出す時間間隔) よりも少ない場合は、フレームの色ヒストグラムの違いにより、シーンチェンジの検出を行い、シーンの変化が大きいと思われるフレームのみを指定された最大枚数分だけ取り出す。

具体的には、映像処理の前処理としてしばしば用いられる、

- 1) 各フレームの RGB を色相、明度、彩度に変換する。
- 2) 色相について、N 段階 (ここでは N=16) のヒストグラムを作成する。
- 3) 連続する 2 枚のフレームの色相のヒストグラムの違いを

$$\sum_{n=0}^{N-1} (H_2(n) - H_1(n))^2 / H_1(n)$$

で計算し、この値が大きいフレームを取り出す最大枚数分選び出す。(H<sub>1</sub>(n)、H<sub>2</sub>(n)は、各ヒストグラムの値である。)

というアルゴリズムでシーンチェンジの検出を行っている。

このツールを使った映像一覧の作成結果を

図 9、図 10 に示す。図 9 が、145 秒の映像に対して 2 秒ごとにフレームを取り出した結果であり、図 10 が色ヒストグラムの違いによるシーンチェンジ検出を行い、30 枚のフレームを取り出した結果である。図 9 では、同様のフレームが繰り返し出てくる部分が多いのに対し、図 10 では、そのような繰り返しが少なくなっており、良好なフレームの抽出ができていくことが分かる。

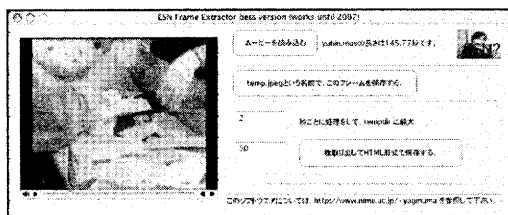


図 8 映像一覧作成ツール(Easy Streaming NIME Frame Extractor)

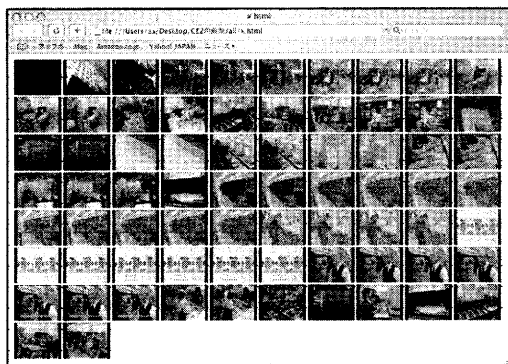


図 9 映像一覧作成結果 (時間等間隔)

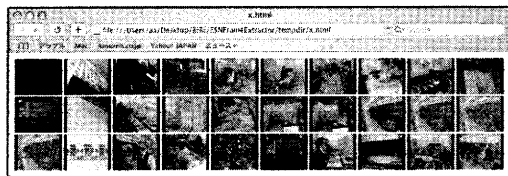


図 10 映像一覧作成結果 (色ヒストグラムによるシーンチェンジ検出)

## 5. 評価

今回開発したツールの簡単な評価として、評価者 1 名 (大学院生。ユーザビリティ評価

の経験あり) による各ツールのインスペクション評価を行った。それぞれの評価に対しては、重要度を A、B、C の 3 段階で表し (A が最も重要度が高い)、Easy Streaming NIME Presenter について 13、Easy Streaming NIME Layouter について 10、Easy Streaming NIME Slide2Movie について 2、Easy Streaming NIME FrameExtractor について 2、全体に対して 1 の計 28 の評価コメントを得た。主なものとしては、Easy Streaming NIME Presenter については、「スライドの全画面表示ができない(A)」、「前へ」「次へ」のキーボードショートカットがない(B)」等の問題点が上げられた。Easy Streaming NIME Layouter については、「スライド再生のタイミングを後から修正したい(A)」、「スライド再生のタイミングを視覚的に一覧したい(A)」等の意見が出された。Easy Streaming NIME Slide2Movie については、「作成の元になるテキストデータファイルの作成が容易になると良い(B)」等の意見が出された。Easy Streaming NIME FrameExtractor については、「全枚数を書き出す方法が分かりにくい(C)」等の意見が出された。これらの点については、今後改善をはかっていく予定である。

## 6. まとめ

本稿では、SMIL を用いた映像教材作成支援ツールとして開発を行った、講義映像とスライドを同期した映像教材を作成するツール、スライドとその説明文から音声合成により映像教材を作成するツール、これらの補助として、映像から静止画を取り出し、等間隔、あるいは、色ヒストグラムを用いたシーン検出により、映像の静止画一覧を作成するツールの概要について述べた。また、これらのツールに対する評価の結果についてもあわせて述べた。今後は、共同研究等によりこれらのツールの実証実験を行い、機能やインターフェースの改良を行っていく予定である。

※Apple、QuickTime 等、本文中の会社名や製品名は、該当する各社の商標または登録商標です。

## 7. 参考文献

- [1] 柳沼良知,古川雅子,山田恒夫,"レイアウト編集可能なマルチアングル映像評価ツールの開発",情報処理学会研究報告,2003-CE-71, pp.21-26,2003.10
- [2]古川雅子,柳沼良知,山田恒夫,"映像評価ツールを用いた言語・非言語行動観察",日本教育工学研究報告集,JET03-4, pp.69-74,2003.7
- [3] 張文利,柳沼良知,坂内正夫,"関連メディアを用いた映像の記述・利用システム",画像電子学会研究会予稿,01-01-01, pp.1-8,2001.6
- [4] 柳沼良知,坂内正夫,"DP マッチングを用いたドラマ映像・音声・シナリオ文書の対応付け手法の一提案",電子情報通信学会論文誌,Vol.J79-D-II, No.5, pp.747-755,1996.5
- [5] 柳沼良知,和泉直樹,坂内正夫,"同期されたシナリオ文書を用いた映像編集方式の一提案",電子情報通信学会論文誌,Vol.J79-D-II, No.4, pp.547-558,1996.4
- [6] 孟洋,佐藤真一,"映像コーパスに基づく映像フィルタリングシステム",画像電子学会研究会予稿,01-01-01, pp.9-14,2001.6
- [7] 杉山善明,有木康雄,"多重部分空間法に基づくテレビスポーツニュース映像の自動分類",電子情報通信学会論文誌,Vol.J81-D-II, No.9, pp.2112-2119,1998.9
- [8] 長坂晃朗,宮武孝文,上田博唯,"カットの時系列コーディングに基づく映像シーンの実時間識別法",電子情報通信学会論文誌,Vol.J79-D-II, No.4, pp.531-537,1996.4
- [9] 佐藤隆,坂内正夫,"ライブハイパメディアにおける映像情報の獲得",電子情報通信学会論文誌,Vol.J79-D-II, No.4, pp.559-567,1996.4
- [10] <http://www.apple.co.jp/quicktime/>
- [11] <http://www.w3.org/AudioVideo/>