

メディアファイルのアクセシビリティ 向上のための字幕作成ツール

青森公立大学 経営経済学部 神山 博

kamiyama@nebuta.ac.jp

PC を利用したメディアに代替コンテンツが提供されることが少ない一因は、同期が容易なツールがないことにある。本研究では PC で利用されるメディア用のキャプションニングツールを試作した。

キーワード：字幕、クローズドキャプション、音声解説、SMI, SAMI, SMIL, OC, CC, AD

1. はじめに マルチメディアの代替コンテンツ

近年、映像や音声は文字の代替として利用され、マルチメディアによる配信が容易に可能となってきた。結果として、遠隔教育を初めとしてあらゆる場面でメディアファイルを利用する機会が増えているが、多くの場合音声の代替としての文字情報を付与しないために、これまでの文字ベースのコンテンツと比べて、聴覚障害を持つ利用者に対するアクセシビリティが著しく低下してしまっている。

TV 番組においては 2007 年までに字幕化率 100% を目指すという総務省指針があるが、PC を利用したメディアについては、情報アクセス権に強い影響を与えるにも関わらず、アクセシビリティについての配慮がされているとは言い難い。そこで PC 用のメディア形式のキャプションニングツールを試作し、評価した結果を報告する。

2. 代替コンテンツの種類と方法

マルチメディアでは情報提供者が複数の媒体を併用して情報伝達を行うための相補性と、利用者が対話的に利用することのできる双方向性を併せ持つ。アクセシビリティに関しては前者の相補性を利用した代替コンテンツの提供が行われる。映像の代替としては音声による説明(AD)をサウンドトラックに挿入する事で、映像コンテンツの状況説明の代替コンテンツが提供される。AD に関しては国内でも 30 年ほど前から副音声に対応した受像器が販売され広く普及しているために、一般の人でも AD を視聴することができる。AD の著作権は制作者自身が保有し、また映像とは独立した副音声トラックに収容するの

で権利処理も比較的容易と言える。

一方音声の代替では、字幕表示領域に字幕を表示する(Captioning)ことになるが、ハードウェアや著作権の制限等の違いによって、原映像に重ねる形のオープンキャプション(OC)や、専用のキャプショントラックに記録するクローズドキャプション(CC)などの異なる方法がとられる。国内で販売されるアナログ TV 受像器には通常、字幕トラックを受信する機能がついていないので、OC が使われることが多い。

OC は字幕が原映像に重ねられて常に画面に表示されるが、もとの映像トラックの内容を変更することになるので、権利処理上の問題があるだけでなく、字幕を他の言語を変更する事が困難であるという欠点がある。一方 CC では個々の利用者が自ら字幕表示をオン・オフすることができるので、映像を変更する際の権利処理の問題が発生しないばかりか、字幕の編集や言語の変更についても翻訳・翻案に関わる権利処理のみとなり、技術的にも容易という利点がある。

代替コンテンツを提供する際に重要なのは、それぞれのコンテンツ同士を同期させることであるが、同期マルチメディアの規格としては、SMIL (Synchronized Multimedia Integration Language) や、SAMI (Microsoft Synchronized Accessible Media Interchange)などが代表的であるが、いずれも容易な作成と汎用的なアクセスに重点が置かれて設計され、仕様は無償で公開されており、多くの人が利用可能な Windows Media や Quick Time、Internet Explorerなどでサポートされている。このような代

替コンテンツ同期技術によって、映像・音声・字幕、音声説明などを同期することができる。

本研究では Windows Media Player(以下 WMP)でサポートされている SAMI 形式の同期マルチメディアを制作するためのツールを開発した。

3. SAMI 形式の代替コンテンツ制作手順

SAMI では、映像・音声ファイルと同じファイル名を持った SAMI ファイルを、同一のディレクトリにおくことで同期マルチメディアを実現するが、SAMI の仕様では HTML と同様のタグ構造を持っており、図 1 のようなタグを WMP 内蔵の SAMI のレンダリングエンジンが解析することで音声と同期させながら、ウィンドウ内の字幕表示領域に字幕を表示する。

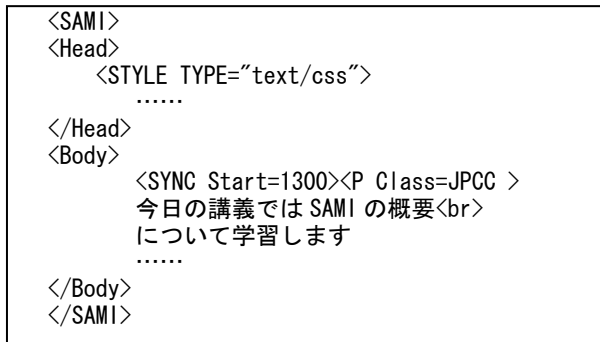


図 1 SAMI のタグ構造

<SYNC>タグでは字幕の表示時刻をミリ秒の単位で指定する。これにより音声との同期を取ることができるが、現時点で汎用の使いやすい制作ツールがない。図 2 に同期マルチメディアを制作する流れをしめす。まず音声トラックに収録された音声から文字おこしを行い、音声テキストを作成する。

文字おこし技術としては近年発達した音声認識システムによる方法の他、ステノワードや高速入力者による文字おこし作業を行う方法があり、後者は人材の確保が必要となるが、標準的なテキストエディターを用いて作成できる。しかし実際には人の話速は 1 分間に 200~400 文字程度であり、一方文字入力速度は、熟練した入力者でも 200 文字程度なので、文字入力しながらも容易に映像・音声ファイルの再生・停止・巻き戻し操作が可能となるような仕組みが必要となる。また複数の話者が登場するようなコンテンツでは、話者によって表示文字をワンタッチで変更できるような仕組みが必要となる。

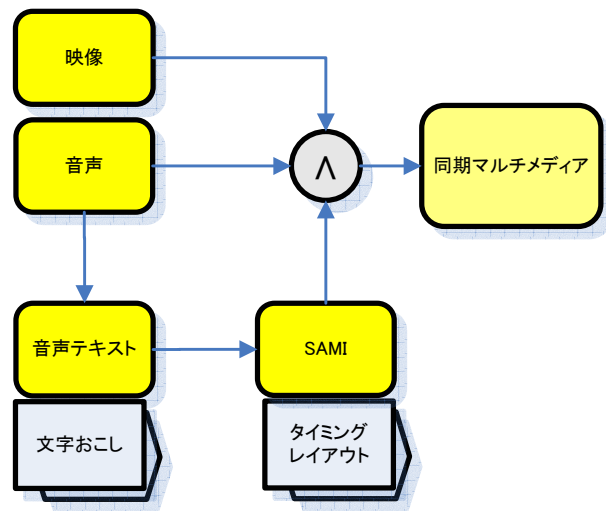


図 2 同期マルチメディアファイル制作の手順

こうしてできた音声テキストに、タイミングや書式情報等を付加して SAMI ドキュメントを作成し、最後に映像・音声ファイルと同じディレクトリに置くことで WMP が再生可能な同期マルチメディアとなるが、ここでもタイミングの微調整や書式の微調整などを容易にできる仕組みが必要となってくる。

これらを可能とするオーサリングツールを試作し、ドラマ、ニュース、ドキュメントなどのコンテンツを用いてツールの有効性について検証をおこなった。詳細については講演で報告する。

4. 今後の課題と展望

メディア再生制御や SYNC 挿入が平易となったことで、同期マルチメディアに必要な SAMI ドキュメントが容易に制作できるようになり、概ね当初の目的を果たすことができた。

現時点では文字書式情報はスタイルシートを用意して同一のタグを出力するようにしているが、今後は書式を自由に設定できるように機能拡張をする予定である。また利用者がメディア再生時に、自分で表示を変更できるようなツールがあると、利用時の利便性が著しく高まることが予想される。これについても今後検討していきたい。

参考文献

- 聴覚障害者の情報バリアフリーと著作権の壁 山田桂子 月刊社会教育 Vol. 44, No. 5(2000) pp. 61-68.
 聴覚障害者の情報保障の意味 高岡正 月刊社会教育 Vol. 44, No. 5 (2000) pp. 52-60.