

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 22 日現在

機関番号：14303

研究種目：基盤研究(C) (一般)

研究期間：2014～2016

課題番号：26350325

研究課題名(和文) 手書き板書方式ビデオ教材共同開発のためのシステム基盤構築

研究課題名(英文) System infrastructure for collaborative authoring of hand-written based educational video materials

研究代表者

永井 孝幸 (Nagai, Takayuki)

京都工芸繊維大学・情報科学センター・准教授

研究者番号：00341074

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：本研究では手書き板書方式(いわゆるカーン・アカデミー方式)の教材を対象に、ペンストローク再生と合成音声を組み合わせたビデオ教材のオープンな開発基盤について研究を行った。教材開発基盤の中心となるペンストローク記録再生ツールについては、JavaScriptを用いることでOS環境に依存しないWebアプリケーションとして実装した。作成した手書き板書教材はクラウドストレージに保存することでインターネット上ですぐに共有することができる。音声合成サーバについてもHTTP経由で受け取った台詞データを音声ファイルとして送り返す汎用的な実装を行った。

研究成果の概要(英文)：In this research, we have studied infrastructure for open collaborative authoring of educational video materials combining pen-stroke replay and synthesized speech. The key component for open collaborative authoring of educational material is a tool for pen stroke recording and playback. We implemented the tool as a Web application independent of the OS environment by using JavaScript. The created teaching material can be shared immediately on the Internet by storing it in cloud storage. For speech synthesis, we developed a general-purpose implementation that receives dialogue data through HTTP and send back synthesized voice audio.

研究分野：教育学習支援情報システム

キーワード：ビデオ教材 合成音声 ストロークデータ ペン入力 手書き入力 手書き板書方式 タブレット

1. 研究開始当初の背景

現在、オンライン教育は大量の履修者に対して学習履歴にもとづいたデータ分析を行い、各学習者に合わせた教育を提供する段階に進みつつある。この段階で鍵となるのは学習者に応じた教育コンテンツを作り出す技術であり、大量の学習者に対応するためには教育コンテンツのオープンな開発基盤が不可欠である。

OpenCourseWare, iTunes U による教育コンテンツの公開を始めとする教育のオープン化の流れは MOOCs(Massive Open Online Courses) へと結実し、大量の学習者に対してビデオ教材とオンラインクイズ・仮想実験室等を組み合わせることで学習コースそのものをオンラインで無償提供することが可能となった。しかしながら MOOCs では受講生の落第率が非常に高い(一説によれば 90%程度)ことが知られており、真に効果的な教育手段となるには一律のコース内容を提供するのではなく、個々の学習者の属性を考慮した教育内容を提供することが課題と言える。ここで個々の学習者に応じた教材、中でも作成に手間を要するビデオ教材の開発が課題となる。

Coursera, edX, Udacity 等のトップ大学を源流とする MOOCs では世界最高水準の教育を無償で提供するという理念が発端となっており、各大学の講師本人が前面に出た放送品質のビデオ教材が用いられている。これらの教材は高品質ではあるが講師本人・MOOCs 提供元の助けを借りずに内容を変更することは困難であり、一部の学習者に適した教材にとどまらざるを得ない。一方、これと別の流れとして、サルマン・カーン氏の個人的な動機(親族の勉強の手伝い)で始めた活動が世界規模にまで広がったカーン・アカデミーの事例がある。氏の作成したビデオ教材は特段に手間のかかったものではない。氏がペンタブレット上に算数の問題等の説明を直接筆記していき、氏のナレーションとともに筆記の過程が表示されるという素朴なものである。ここで注目すべきは、この教材ビデオにおいてカーン氏自身の姿は全く映っておらず、「教室における手書き板書場面から教員を消し去った状態」であるということである。結果として、この属人性の少ない素朴な手書き板書方式ビデオが反転授業などの形で世界中で利用され、ボランティアによる翻訳も行われるなど Wikipedia のような持続的改善コミュニティが形成されつつある。

講師本人が前面に出たビデオ教材の場合、多人数での共同開発は不可能であり、大量の学習者に対して各学習者に合わせた教材を作り出すには教材の属人性が少ないことが決定的に重要である。そこで、本研究では属人性の少ない手書き板書方式ビデオ教材に着目し、手書き板書方式ビデオ教材のための共同開発基盤を構築することで、学習者に応じたビデオ教材の持続的な改善を可能とする。

2. 研究の目的

手書き板書方式ビデオ教材の共同開発基盤を構築することが本研究の目標である。そのために(1)板書ペンストロック情報の編集・多言語化機能を備えた手書き板書方式ビデオ教材作成システムの開発、(2)合成音声を用いたナレーションの差し替え機能を備えた教材作成システムの開発、(3)教材共同開発のためのオンラインコミュニティシステムの構築を行う。

ビデオ教材を属人性の低い状態に保つために手書き板書のペンストロック情報を編集単位とし、ナレーションには合成音声を用いる。多くのビデオ教材は編集後の完成品として流通しているため学習者に合わせたカスタマイズ(多言語対応、説明事例の追加、記法の修正等)が困難である。特に、カーン・アカデミー方式のビデオは手書き部分の他言語化が困難であることが再利用の障壁となっている。提案方式はペンストロック情報・音声ともに差し替え可能とし、配信形式に変換する前の素材をオンラインで共同編集することでこの問題を解決する。自然な合成音声については感情音声合成の分野で研究が急速に進みつつあり、当面不自然さのある音声であっても、台詞情報を保持しておけば後に再計算することで不自然さを解消できる。

3. 研究の方法

(1) 市販タブレット機材を用いたペンストロック記録・再生方式の検討

市販のタブレット機材を用いたペンストロックの記録・再生方式について要素技術の開発を行う。現在、手書き板書方式のビデオ教材はデスクトップ PC とペンタブレットを組み合わせられて作成されることが多い。しかしタブレット端末が急速に普及しつつあることから、教材の共同開発を前提とした場合はタブレット端末による教材作成システムを実現することが重要である。

開発の第一段階として、広く普及しているタブレット端末(アップル社製 iPad, Google 製 Nexus10, Windows Surface 等)と市販のスタイラスペンを組み合わせたペンストロック(ペン先座標の時系列データ)の収録・編集システムを構築する。

(2) 市販タブレット機材を用いた手書き板書方式ビデオ教材作成システムの開発

開発の第二段階として、映像素材(講義スライド素材・Web ページ・アプリケーション操作映像・カメラ映像等)に対してペンストロックによる説明を記入することができるように上記システムの機能を拡張し、手書き板書方式での教材作成システムとして利用可能な状態

にする。

ここでタブレット端末の場合、映像素材の取り込み手段が問題となる。デスクトップPCでは映像取り込み用のオプションボード類を接続することで外部映像信号を経由して幅広い素材を取り込むことができる。一方、現状のタブレット端末には映像信号取り込み用のインタフェースが用意されておらず、映像信号取り込み用の外部機器も市販されていない。そこで一般的なタブレット端末に装備されている USB・無線 LAN 等のインタフェースを用いた映像素材取り込み方法を実装する。

これに加えて合成音声を用いたナレーションの差し替え機能の実現に取り組む。講義内容に合わせた発話辞書の改良、多言語対応におけるペンストロークと音声の再生タイミング調整が課題である。コンテンツの共同開発のためには開発したシステム全体がフリーであることが重要なため、日本語の音声合成エンジンには OpenJTalk, 英語の音声合成エンジンには eSpeak, MARY 等のオープンソースライセンスに基づいたソフトを第一候補として用いる。

4. 研究成果

(1) 市販タブレット機材を用いたペンストローク記録・再生方式の検討

タブレット端末の候補として iPad, Android, Windows タブレットから画面サイズ・タッチパネル実現方式の異なる複数機種を選定し、手書き教材の作成に耐えるか検証を行った。その結果、静電容量方式のタッチパネルでは十分なペン入力精度が得られず、デジタイザ方式のタッチパネルが不可欠との結論を得た。A4 サイズ相当の大きさで教材を作成するには、画面サイズは 12 インチ以上の大きさが理想的であるが、10 インチ以上あれば実用的な利用が可能であると結論した。

次に、ペンデジタイザから得られたペンストローク情報を元に、文字ストローク切り出し処理を行う手法について検討を行った。教材で作成する文書は日本語で一行ずつ規則正しく記述されるものと仮定し、(1) 行の高さの推定、(2) 行の高さに基づく文字幅の推定、(3) 推定文字幅に基づく近接ストロークのクラスタリングの順で文字切り出しを行う手法を考案した(図 1)。

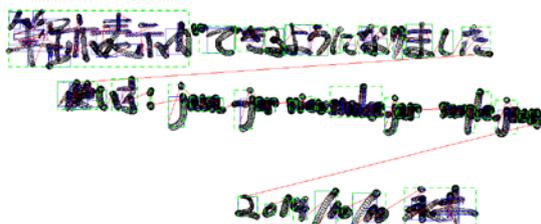


図 1 ペンストロークの自動分割結果

ここで、ストロークのクラスタリングは、幅制約付き区間集合分割問題へ帰着し、貪欲

法に基づくアルゴリズムで解くことにより求めた。幅制約付き区間集合分割問題は、入力として (閉) 区間の集合 $I = \{I_1, \dots, I_n\}$, 任意の $I_i, I_j \in I$ に対する重み $w(I_i, I_j)$, 及び自然数 H を受け取り, I の分割 $P = \{P_1, \dots, P_m\}$ で各 P_k の要素の和集合の幅が H 以下であるものの中で, 次式で定義される重み $w(P)$ を最大化するものを求める問題である。

$$w(P) = \sum_{k=1, \dots, m} \sum_{I_i, I_j \in P_k} w(I_i, I_j)$$

このとき、ストロークのクラスタリングは、次の手順により幅制約付き区間集合分割問題へ帰着することができる。

- 水平方向に平行な軸への射影を用いて、ストローク集合に対応する区間集合を求める。
- 得られた区間集合から、交差関係に基づいて特に結びつきの強い 2 つの区間をその和集合で置き換えることを繰り返すことで新しい区間集合を得る。

得られた区間集合、2 つの区間に対するそれらの近さを表す重み関数、及び文字幅の推定値を入力とする幅制約付き区間集合分割問題を解くことで、ストロークのクラスタリングを得る。

タブレットへの外部映像入力方式についても検討を行い、(1) HDMI 入力に対応した UVC キャプチャデバイスを用いる方式、(2) 小型 Linux サーバと HDMI 信号入力ブリッジを組合わせて簡易キャプチャデバイスとして用いる方式の 2 つの方式が有望であるとの結論を得た。図 2 は UVC キャプチャデバイスを用いて Android 端末上で HDMI 信号をキャプチャしている様子である。Android の基本機能としては UVC キャプチャデバイスに対応していないが、UVC キャプチャデバイスと通信する Android アプリケーションを作成することで、汎用の Android タブレット端末で HDMI 信号のキャプチャが可能となる。

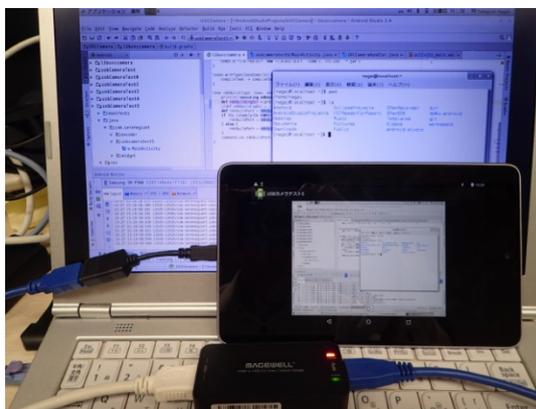


図 2 Android 端末上で UVC キャプチャデバイスを用いて HDMI 信号をキャプチャ

図 3 はハイビジョン信号キャプチャボードを追加した小型 Linux ボード (Leopardboard 368) を用いて試作した HDMI

信号キャプチャデバイスである。このデバイスを用いてフル HD ハンディカム (1080i), iPad2 (720p) の HDMI 出力を 20fps 以上のフレームレートで安定して長時間 (12 時間以上) 連続して RTSP ストリームとして配信し、動画キャプチャできることを実証した。

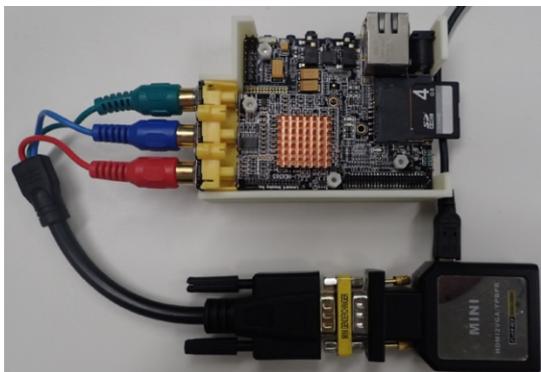


図 3 小型 Linux サーバで試作した HDMI 信号入力対応 RTSP ストリーミングサーバ

本研究で想定している手書き板書方式教材の作成が普及価格帯の市販機材で行えることが確認できたことは、本研究の実用化に向けた重要な一歩である。

(2) 市販タブレット機材を用いた手書き板書方式ビデオ教材作成システムの開発

ペンストローク記録再生ツールの開発にあたり実装方式について複数方式の検討を行った。研究開始当初はペンタブレットに最適化した Android アプリでの実装を進めていたが、後に iOS や Windows 環境でもタッチペンを搭載した携帯型端末が多数登場した。教材の共同開発環境としてはできるだけ多くの端末を利用できることが重要であるため、OS 環境に依存しない Web アプリでの実装方式を採用することとした。

最新の Web ブラウザではペン入力データを直接受信できることを利用し、ペン入力デー

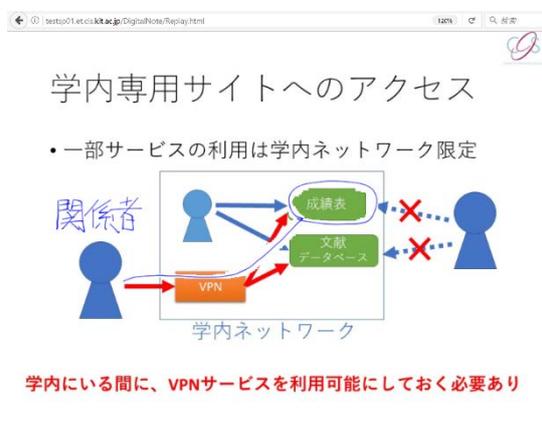


図 4 Web ブラウザ上で実装した手書き板書方式教材作成ツールの画面

タの記録・再生を行う Web アプリケーションを JavaScript で記述した (図 4). ペン入力データをクラウドストレージに保存する機能も実装しており、作成した手書き板書教材はインターネット上ですぐに共有することができる。

音声合成サーバの構築においては、所属研究機関の変更ならびに熊本地震の影響のために実用に向けて音声データを調整する人員を割当てる見通しが立たなくなったことから、生成音声の再配布が可能な安価な商用ソフトを用いてサーバの構築を行った。この音声合成サーバは HTTP 経由で受け取った台詞データを音声ファイルとして送り返す汎用的なもので、HTML ページに埋め込んだ JavaScript からも利用可能である。

講義の 1 トピックの説明に相当する 7 ページのスライドを題材に、実際の説明内容 (説明時間約 7 分 10 秒, 約 2,600 文字) を用いて音声合成サーバでのナレーション音声作成実験を行った。実験に用いた仮想サーバ (仮想 CPUx2, 物理 CPU: Core i7 3.4GHz) 上で音声合成に要した時間は 1 秒程度であり、実用上支障ないことを確認した。一方で説明内容の書き起こしには 2 時間程度の時間を要した。このことから、台詞情報をオープンコンテンツとして流通させることが教材作成効率の飛躍的な向につながると考えられる。

得られた知見として、合成音声によるナレーションについては、人間による吹き込み作業と比べて内容の修正が容易である一方、専門用語の読みについて発話内容を修正する必要があり、実用化にあたっては専門用語の読みに関するデータベースの整備が課題となることが分かった。

本研究の最終年度には音声マークアップ言語の標準規格 SSML に準拠した商用の音声合成サービスが登場し、高品質な合成音声を安価で作成できる環境が整った。この点で合成音声の利用を前提とした本研究の方向性は合っていたと考えられるが、商用サービスであっても合成音声は講師の発話と比べて抑揚や間の取り方が平板であり、学習者にとって重要な非言語情報をいかに盛り込むかが残された課題となった。

市販の教材作成ソフトとしてペン入力に対応したものも存在するが、いずれも映像を生成することに主眼が置かれておりペンストローク情報自体の保存や共有は想定されていない。開発した音声合成サーバとペンストローク記録再生ツールを組み合わせることでインターネット上の協同作業として教材の継続的な改善を行うことが可能となり、Wikipedia のように協同作業としての教材作成が一般化することが期待される。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕（計 0 件）

〔学会発表〕（計 0 件）

〔図書〕（計 0 件）

〔産業財産権〕

○出願状況（計 0 件）

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

○取得状況（計 0 件）

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕

ホームページ等

<https://www.et.cis.kit.ac.jp>

6. 研究組織

(1) 研究代表者

永井 孝幸 (NAGAI, Takayuki)
京都工芸繊維大学・情報科学センター・准
教授
研究者番号：00341074

(2) 研究分担者

名古屋 孝幸 (NAGOYA, Takayuki)
公立鳥取環境大学・人間形成教育センター・
准教授
研究者番号：90349796