

科学研究費助成事業 研究成果報告書

平成 26 年 5 月 27 日現在

機関番号：17401

研究種目：若手研究(B)

研究期間：2011～2013

課題番号：23700976

研究課題名(和文) 高解像度教育コンテンツ映像の活用に向けた映像加工システム基盤の構築

研究課題名(英文) Development of video processing framework for high-definition educational video

研究代表者

永井 孝幸 (Nagai, Takayuki)

熊本大学・総合情報基盤センター・准教授

研究者番号：00341074

交付決定額(研究期間全体)：(直接経費) 3,100,000 円、(間接経費) 930,000 円

研究成果の概要(和文)：本研究の目的は大量の高解像度教育コンテンツ映像を学習素材として活用する際に必要となる「映像の著作権保護」と「プライバシー保護」を実現するための映像加工システム基盤を構築することである。人物検出・匿名化処理の高速化に不可欠な画像処理の高速数値演算ボード対応、オープンソースを用いたビデオ加工基盤全体の構築によるライセンス費用の無償化、ユーザ属性に基づいたコンテンツへのアクセス制御を実現するための認証基盤の開発といったシステム基盤の構築と合わせ、実際の講義室環境に適した講義スライド検出・顔検出技法の改良、ならびにタブレット端末を用いたビデオ収録・加工に必要な要素技術の開発を行った。

研究成果の概要(英文)：To solve copyright and privacy issues in using high-definition educational videos, we developed video processing infrastructure as follows. First, we applied GPU technology to speed up face detection and video transcoding. Second, we implemented the entire video processing system by open source software to eliminate license fee. Third, we developed user authentication infrastructure with Shibboleth and Groupware to realize fine-grained access control to educational materials. We also improved slide detection and face detection methods that match with real classroom conditions. Moreover, we developed high-definition video streaming device and video masking tool to use tablet devices in video capture and post production.

研究分野：総合領域

科研費の分科・細目：科学教育・教育工学

キーワード：講義ビデオ 著作権保護 顔検出 GPU 認証基盤 タブレット

1. 研究開始当初の背景

現在,OpenCourseWare や iTunes Uをはじめとして高等教育機関における教育コンテンツの蓄積と活用が世界的に加速しており,良質の教育コンテンツをオープンコンテンツとしてインターネット上で提供することが高等教育機関の重要な機能となりつつある.映像コンテンツについては収録をはじめとして作成に多くの人手・設備・費用が必要であり,大規模な取り組みは予算・人員に余裕のある一部の大学に限られていた.しかしながら,北米・ヨーロッパの大学を主要メンバーとする OpenCast プロジェクトにおいて商用システム並の機能を備えたオープンソースの講義ビデオ収録システムの開発が始まるなど,映像収録・配信のための技術面・設備面でのハードルは下がりつつある.

この結果,近い将来には多くの講義が日常的に収録されるようになる予想される.しかし,収録した映像を学習素材として活用するには関係者の権利保護のために依然として多くの人手が必要であり,撮影量の増加に対応するためには権利保護作業の負担を低減するための新たな映像加工システム基盤の構築が必要である.「フェアユース」規定の元に教育目的でのビデオ公衆配信が広く認められている英米法体系の国々では講義ビデオの蓄積・活用が急速に進む一方で,現行の日本の著作権法では教育目的のビデオ公衆配信は教育上の例外とは認められていない.ビデオ素材を活用した教育の推進において権利保護作業の負担が大きな足枷となっており,本研究は映像加工システム基盤を構築することでこの問題を解決しようとするものである.

2. 研究の目的

本研究の目的は大量の高解像度教育コンテンツ映像を学習素材として活用する際に必要となる「映像の著作権保護」と「プライバシー保護」を実現するための映像加工システム基盤を構築することである.これにより映像の権利者・教授者・学習者の権利を侵害することなく,高解像度映像を活用した教育活動の大規模な展開が可能になる.

本研究では高解像度教育コンテンツ映像を学習素材として活用する際に必要となる「映像の著作権保護」と「プライバシーの保護」を可能にする映像加工システム基盤を構築することを目指す.具体的には権利保護作業の大幅な省力化が見込まれる以下の2点を課題とする.

- (1) プライバシー保護のための人物検出・匿名化システム基盤の開発

大量の教育コンテンツ映像の中から個人の特定につながる顔や歩行場面を検出し,該当箇所加工を施すことで人物の匿名化を行うシステム基盤を構築する.標準価格帯のPCと顔認識ソフトウェアを用いた予備的な実験では,フルハイビジョン解像度の映像から顔検出を行うのに1フレームあたり3~4秒かかることが分かっている.素朴な実装では顔の検出だけで撮影時間の100倍の時間がかかってしまうため,近年大幅に価格性能費の向上した高性能GPU等の高速計算技術を画像処理に活用することで計算時間の大幅な短縮を実現することが鍵となる.

- (2) 著作権保護のための映像分析・加工システム基盤の開発

教育コンテンツ映像に映っている書籍・写真・図表・ビデオといった著作物を抽出し,利用許諾の確認,および,利用許諾の得られていない著作物を取り除く処理を効率よく行うための映像分析・加工システム基盤を構築する.

ある著作物の利用許諾が得られているかどうかは映像からだけでは判断できないため,人間による総合的な判断が必要である.そのため,映像から抽出された潜在的な著作物を閲覧・分別するための操作性の高い編集用インターフェースを開発することが鍵となる.

3. 研究の方法

- (1) 人物検出・匿名化処理の高速化

先行研究で用いられている技法を参考にしながら,機能要件を満たす人物検出・匿名化処理をシステムとして実装する.一般に顔検出処理には大量の計算が必要であり,日常的に収録される大量の高解像度映像に対してプライバシー保護の処理を施すためには計算時間の大幅な短縮が不可欠である.そこで,CPUに対して数十倍の演算性能を持つ高性能GPU(3次元画像処理に用いられる高性能数値演算ボード)を画像処理に用いることでシステムの実現に必要な計算速度の向上を実現する.高性能GPUによる計算だけで必要な計算速度が得られない場合は,教育コンテンツ映像に特化した手法を組み合わせることで計算速度の向上を図る.

- (2) オープンソースによる高解像度ビデオ加工基盤の開発

高解像度教育コンテンツに対して日常的に分析・加工を行うためには大量の計算機資源が必要であり,GPUを用いた画像処理の高速化に加えてランニングコストの低下が課

題となる．ビデオ加工基盤を可能な限りオープンソースソフトウェアを用いた実装とすることで，高額なライセンス費用が発生しないシステムとすることを旨とする．

(3) ユーザ属性にもとづくコンテンツアクセス制御基盤の開発

日常的に収録された教育コンテンツ映像を利用する際，プライバシー保護のためには教員・受講者・その他学生等，利用者のユーザ属性に応じてコンテンツのアクセス制御を実施する必要がある．そこで，ユーザの所属グループに応じたアクセス制御を実現するためのシステム基盤を構築する．

(4) 講義スライド・顔検出精度の向上

既存の講義スライド検出・顔検出の手法では「講師 PC の映像信号を直接収録する」「教卓付近の講師映像のみ収録する」等といった仮定が設けられており，本研究で想定する「固定ハイビジョンカメラによる講義室全景映像」に出現する講義スライド・顔画像においては性能が低下する．そこで，実環境において収録された固定ハイビジョンカメラ映像に対して有効に機能するよう講義スライド検出・顔検出の手法を改良する．

(5) タブレット端末用ビデオキャプチャデバイスの開発

講義ビデオにおいて配信すべきで無い範囲を指定するには，収録後の編集段階で指定する方式だけでなく収録時に指定する方式も考えられる．タブレットデバイスにビデオ信号キャプチャ用のデバイスを組み合わせることで，収録時にビデオ映像を見ながら対話的に配信範囲を指定することが可能となるが，タブレットに接続可能な映像信号入力用のデバイスが存在しない．そこで，USB カメラの開発に利用される小型の Linux コンピュータを用い，ハイビジョン映像信号を入力とする汎用キャプチャデバイスを開発する．

(6) タブレットデバイスを用いたビデオマスキングツールの開発

映像コンテンツに対してマスキングを施すのは手間のかかる作業であり，通常のビデオ編集ソフトではフレーム毎にマスク領域を指定する必要がある．この手間を省力化するため，タブレットデバイス上でペン入力により直接的にマスク領域を指定するツールを開発する．

4. 研究成果

(1) 人物検出・匿名化処理の高速化

講義ビデオにおける人物検出・匿名化処理を高速化するため，講義ビデオ加工インフラの GPU (高速数値演算ボード) 対応を行った．高解像度ビデオのデコード処理に GPU を適用することで，90 分の講義ビデオを加工する際に CPU のみの処理では 3 時間かかっていた処理を 2 時間に短縮することに成功した．更に顔検出処理を高速に行うため，画像処理ライブラリ OpenCV の GPU 対応版を用いて顔検出処理の実装を行った．これにより，「90 分の講義ビデオ (1920x1080) を 960x540 に縮小し，1/3 秒ごとに顔検出を行う」という現実的な条件の下，顔検出に要する時間が CPU のみでの処理では 40 分程度要していたものが 20 分程度に短縮された．顔検出処理を含めて 90 分講義ビデオ 1 本あたりの処理時間が 3 時間未満となり，実用的な処理時間を達成している．

既存の講義ビデオ収録システムでは権利保護に必要な顔検出等の高度な画像処理が組み込まれていないが，本成果により画像処理を組み込んでも実用的な時間でビデオ収録を行うことが可能となり，権利保護機能を備えた収録システムが普及することが期待される．

(2) オープンソースによる高解像度ビデオ加工基盤の開発

ビデオ加工基盤の全学的な利用にはサーバを大量に用いることが必要であり，高額なライセンス費用の発生しないフリーな実装であることは決定的に重要である．ビデオ素材のデコード・分析，ストリーミング配信形式への変換処理を ffmpeg, OpenCV を始めとするオープンソースソフトウェアを用いて実装し，自作ツールと組み合わせることでシステム基盤を Linux へ全面的に移行することで GPU (高速数値演算ボード) のドライバ・CUDA ライブラリを除いて完全にフリーな実装となった．



Figure 1 講義室全景映像から合成した講師追跡映像とプロジェクタ映像の Matterhorn を用いたマルチトラック配信の様子

更にオープンソースのビデオ配信システム Matterhorn との連携ツールを開発し、固定ハイビジョンカメラで撮影された講義室全景映像から後処理によって生成された講師追跡映像とプロジェクター映像を同時配信することを可能とした(Figure 1)。

既存の講義ビデオ配信システムでは講師映像とプロジェクター映像の同期配信を行うには収録時にマルチトラック収録を行う必要があり、講師 PC の映像信号キャプチャに要する機材も含めると収録機材のコストが高価となる問題があった。これに対し、今回開発したシステムでは講師 PC の映像信号キャプチャを行うこと無く講義室全景映像から後処理によってプロジェクター映像を生成するため、マルチトラック配信の実現に必要なシステムの構築コストを大きく下げることが可能となる。

(3) ユーザ属性にもとづくコンテンツアクセス制御基盤の開発

ユーザ属性に応じたコンテンツへのアクセス制御を実現する認証基盤技術である Shibboleth を用いたシングルサインオン環境を構築するとともに、オープンソースのコンテンツ管理システムである「Plone4」を用いて粒度の細かいアクセス制御機能・出版ワークフロー機能を備えたコンテンツ管理システムの構築を行った。

オープンソースのグループ管理ミドルウェア Grouper を用いて人事データベース・学務情報データベースの情報を集約し、LDAP 上にグループ情報を同期させるシステム基盤を構築した。これにより、熊本大学の全教職員・学生に対して所属と履修科目に応じたユーザグループを常時更新できることを確認した。このユーザグループ情報を Shibboleth に受け渡すことで、コンテンツへのアクセスを科目の担当教員・履修学生に限定することが可能となる。

更に Shibboleth から受け渡されたロール情報に基づいてユーザ ID を選択することのできる Shibboleth-CAS ゲートウェイを開発し、複数ロールを持つユーザがロールに応じて CAS 対応システムを利用することを可能とした(Figure 2)。

細粒度アクセス制御・出版ワークフロー機能を備えたコンテンツ管理システムはプライバシー保護と講義ビデオの活用を両立させるために不可欠な要素技術であり、本成果は実用化に向けた重要な一歩である。

(4) 講義スライド・顔検出精度の向上

講師 PC のビデオ信号を直接収録する方式では、画像変化量に対して単純閾値を用いてスライド切替えを検出するのが一般的である。しかし本研究で前提とする講義室全景を収録した講義ビデオの場合はプロジェクタ



Figure 2 CAS 認証時にロール情報に基づいてユーザ ID を選択

スクリーン自体の揺らぎも画像変化量に含まれるなど外乱要因が多いため、スライド検出率が大幅に低下する。そこで、講義ビデオ中のプロジェクタスクリーン領域を小ブロックに分割し、各ブロックの変化量の出現確率に基づいてスライド変化を検出する手法を考案し実際の講義ビデオに適用した。講義スタイル毎のスライド検出の見落とし率については分析中であるが、検出された場面を一覧することで講義ビデオ全体を見直すのに比べて大幅に少ない時間で講義中に利用されたスライドを把握することが可能になった。

収録した数百講義のビデオに対して人物検出のための顔検出を行った結果、講義場面によって誤検出の割合が非常に高くなることが判明した。分析の結果、プロジェクタスクリーンや学生の衣服などが誤検出されることが多いことを確認している。講義ビデオに含まれる顔の検出精度を向上させるため、実際の講義ビデオから抽出した顔画像を用いて顔検出辞書の改良を行った。OpenCV 画像処理ライブラリに付属している標準の顔辞書を用いて得られた検出画像を初期学習用画像データとし、Haar-like, HOG, LBP の各画像特徴量を用いて顔辞書を算出した。

実際の講義ビデオから抽出した評価用顔画像を用いた顔検出性能の比較の結果、一般に顔検出に向いていると言われる Haar-like 特徴量を用いた顔辞書よりも、HOG 特徴量を用いた顔辞書のほうが高い検出率を示すことが確認できた。

顔認識に関する既存の研究では、整った撮影条件の下で撮影された顔画像を学習用画像に用いて顔辞書を構成している。同等の学習用画像を集めることが困難な講義ビデオに対し、より高い検出性能を持った顔辞書を構成できたことは、実際の講義ビデオに対して顔検出を適用する際の重要な成果である。

(5) タブレット端末用ビデオキャプチャデバイスの開発

代表的なタブレット端末には Android タブ

レット, iPad, Windows タブレットがあるが, いずれも外部映像信号を直接取り込むためのインターフェースを備えていない。しかし, いずれの機器も高速ネットワークインターフェースを備えていることから, 汎用のビデオ信号取り込み方法として「ハイビジョン映像信号を RTSP 形式に変換してネットワーク配信し, タブレット端末で受信する」方式を採用することとした。従来, この方式での映像配信デバイスは放送用の高価な製品しか存在しなかったが, 教育コンテンツでは画質への要求がある程度限定されることから要求性能を絞り込み, デジタルカメラ・ネットワークカメラ開発用の小型 Linux ボード (Figure 3) を用いてハイビジョン信号対応の RTSP 配信用デバイスを安価に開発することに成功した。

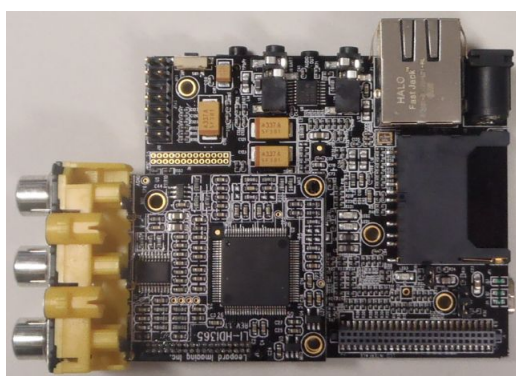


Figure 3 RTSP 配信用デバイスの開発に用いた小型 Linux サーバ

開発したデバイスをオープンソースのビデオ収録端末 Galicaster と組み合わせ, タブレットとフルハイビジョンカメラの HDMI 出力を 24 時間以上継続してマルチトラック収録することに成功しており, ビデオ信号取り込みインターフェースを備えていない情報端末でのビデオ収録の実用性を確認できたと言える。

タブレット端末をハイビジョン映像の収録に用いる取り組みは事例がなく, 本成果により急速に普及しているタブレット端末を安価に映像収録機材として用いることが可能となることから, その波及効果は大きい。

(6) タブレットデバイスを用いたビデオマスクングツールの開発

権利保護のためのマスクング作業を省力化するため, タブレットデバイスを用いたビデオマスクングツールの開発を行った。マスクング領域の指定に「タブレット用ブラシを用いて色を塗る」という直感的な操作 (Figure 4) を採用することで, 1 フレーム辺り数秒でマスクング作業を行えることを確認した。

講義映像から抽出されたキーフレームに

対してマスクング領域を指定することで, 権利保護のためのマスクング作業が大幅に省力化されることが期待される。

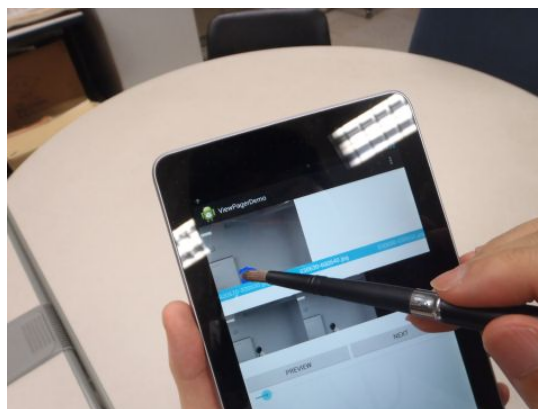


Figure 4 タブレット用ブラシを用いたマスクング領域指定の様子

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 3 件)

永井孝幸, 杉谷賢一, 河津秀利, 中野裕司, "学認対応認証基盤とユーザ ID 体系移行用 CAS ゲートウェイの構築", 第 11 回 CLE 研究発表会 情報処理学会研究報告 Vol.2013-CLE-11 No.20 (2013-12), 査読無

Nagai Takayuki, TOYOTA Toshiyuki, NAGOYA Takayuki, IMAI Masakazu, NISHIZAWA, Koki, "Implementation of high-definition lecture recording system for daily use", Proceedings of Global Engineering Education Conference (EDUCON), 2013 IEEE, pp.520-525, 査読有
<http://dx.doi.org/10.1109/EduCon.2013.6530155>

永井孝幸, 豊田寿行, 名古屋孝幸, 西澤弘毅, 今井正和, "講義ビデオの活用に向けた講義音声の発話特徴分析", 第 4 回 CLE 研究発表会 情報処理学会研究報告 Vol.2011-CLE-4 No.4 (2011-05), pp.1-8, 査読無

[学会発表](計 5 件)

永井孝幸, 杉谷賢一, 河津秀利, 中野裕司, "学認対応認証基盤とユーザ ID 体系移行用 CAS ゲートウェイの構築", 情報処理学会 教育学習支援情報システム研究会, 2013 年 12 月 15 日, 琉球大学

Nagai Takayuki, TOYOTA Toshiyuki, NAGOYA Takayuki, IMAI Masakazu, NISHIZAWA Koki, "Implementation of high-definition lecture recording

system for daily use", IEEE Global Engineering Education Conference (EDUCON), 2013/03/13, Technische Universitat Berlin, Berlin, Germany
Yasushi Kodama, Katsusuke Shigeta, Takayuki Nagai, "About Internationalization of Matterhorn and Introduction to Japanese University Community", Matterhorn Unconference, 2013/01/30, University of California San Diego

永井孝幸, "映像配信を利用した教育情報システムの最新事情", 法政大学情報メディア国際シンポジウム, 2013年03月08日, 法政大学市ヶ谷キャンパス

永井孝幸, 豊田寿行, 名古屋孝幸, 西澤弘毅, 今井正和, "講義ビデオの活用に向けた講義音声の発話特徴分析", 情報処理学会 教育学習支援情報システム研究会, 2011年5月13日, 長岡技術科学大学

〔その他〕

ホームページ等

<http://www.cc.kumamoto-u.ac.jp/nagai/research>

6. 研究組織

(1) 研究代表者

永井 孝幸 (NAGAI, Takayuki)

熊本大学・総合情報基盤センター・准教授
研究者番号：00341074