

実書籍に対する仮想情報付加による学習支援の提案

相 楽 恭 宏[†] 高 井 那 美^{††} 高 井 昌 彰^{†††}

本稿では、教科書などの実書籍に対して仮想的な情報の付加を複合現実感技術を用いて実現し、新しい形の学習支援システムを提案する。本システムでは、ビデオカメラの付いたヘッドマウントディスプレイを装着したユーザを想定し、ユーザのビデオカメラから取り込んだ書籍を含む映像を用いて、画像処理を行なうことで複合現実画像を生成する。撮影映像内の書籍には、回転、スケーリング、および変換を統合した平面射影変換が加わる。そのため、書籍特有の情報を用いることで同変換画像に対して画像認識・情報付加を簡単な計算により解決し、学習支援機能の実時間性を確保する。現在の開発中のプロトタイプシステムの状況について述べ、本システムの可能性を検討する。

Learning Assistance with Real Books by Adding Virtual Information

TAKAHIRO SAGARA,[†] NAMI K. TAKAI^{††} and YOSHIAKI TAKAI^{†††}

This paper presents a learning assistance system with real books using augmented reality. We suppose a user to be wearing a Head Mount Display with a video camera. We use the images from the video camera to create mixed reality images for the HMD. A book in the image is converted by projection conversion including rotating, scaling and flaping. We propose a simple calculation to add virtual information using visual characteristic of pages in the book.

1. はじめに

現実世界での学習において、我々の情報取得行動は書籍に頼ることが多い。これは、

- 書籍が人間にとってなじみの深いメディアであり、使用法を教わる必要がない、
- 高速ブラウジングが可能である、
- 全体の中で、どこを読んでいるのかが明確である、などの書籍メディアの利点¹⁾に起因する。

一方、eラーニングの研究が進められ、その学習効率や双方向性などの有用性が注目されている。しかし、eラーニングで用いられるマルチメディアコンテンツの利用や、ネットワークを通じたコミュニケーションなどの機能は、書籍媒体単体では実現困難な機能である。

また、テキストやグラフを用いて何らかの情報を得るとき、注釈やアンダーライン、メモを加えることによって、重要な点の明確化や後日再び参照する際に素

早く注目すべき点の判断、作業の継続をスムーズに行えるなど、情報の価値を高めることができる²⁾。

そこで、本研究では実書籍に対して複合現実を用いることにより、書籍に対して仮想的な情報を付加できるシステムを構築し、実書籍を用いた新しい形の学習支援方法を提案する(図1)。

また、中原ら⁵⁾は、日本や韓国などのアジア地域ではVOD(Video On Demand)や学習を仲間やグループで行う方を好む傾向があると指摘し、eラーニングが個人学習であるという問題点を挙げている。

本論文では、書籍に付加した仮想情報をネットワークを通して共有することで、学習者に他学習者の存在を認識させ、協調学習を可能にする基盤を提供することを考える。

2. システムの概要

2.1 システム構成

本システムの利用者は、装着型のビデオカメラ、ディスプレイ(HMD)を身に付ける。ビデオカメラでは、利用者が実際に書籍を見ている光景が撮影され、計算機に映像が送られる。ディスプレイには、撮影された映像を重ねて、計算機上で生成された注釈情報等が表示される。

今回構築したプロトタイプシステムでは、注釈情報

[†] 北海道大学 大学院 情報科学研究科
Graduate School of Information Science and Technology, Hokkaido University

^{††} 北海道情報大学 経営情報学部
Hokkaido Infomation University

^{†††} 北海道大学 情報基盤センター
Information Initiative Center, Hokkaido University

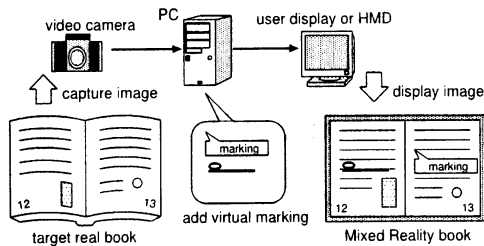


図 1 システム概要
Fig. 1 system concept

入力インターフェイス部分はシステム範囲外とし、書籍の認識と情報の付加方法に焦点を絞る。なお、可視化空間内への注釈情報の付加インターフェイスは園田らによる空中での手書き文字入力システムの提案⁶⁾や、久木元らによる PDA を用いた注釈付与が可能な可視化システムの実装⁷⁾などがなされている。

本システムではまず、利用者が見ている光景の中から画像処理を行い書籍を認識する。次に、認識した書籍画像から書籍の特徴をパラメータとして保持する。このパラメータは必要に応じて毎フレーム更新される。そのパラメータを用いて仮想的な書籍を内部モデルとして生成し、そのモデルに対して付加情報を与えることで複合映像を出力する。

2.2 システムの実装環境

プロトタイプシステムにおいて、ビデオカメラは固定カメラ (DCR-VX2100) を用いて撮影した。映像は 320*240 のフレームサイズ、深さ値を 24bit で記録し、デスクトップ型 PC (CPU Pentium4, OS: Linux, Capture Card: BT878) 上で処理した。

また、撮影する書籍は中学生の英語の教科書^{*}に絞り実験を進めた。

3. 書籍認識

書籍に情報を付加するために、入力画像から書籍紙面を認識する必要がある。

書籍部分のみを抽出するために、はじめに 2 値化処理を施す。入力画像は書籍を映したものであるという状況に限定されるため、出力画像の 2 値画素の割合状態から閾値を動的に決定する P タイル法を用い、入力画像の HSV 値を用いて 2 値化処理を行う。この方法により、比較的照明条件などに左右されない出力結果を得ることができる。実際に 2 値化処理が施された画像が図 2 である。

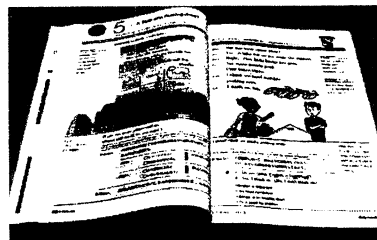


図 2 2 値化処理
Fig. 2 binarization

次に、2 値化画像のノイズの除去を行う。ノイズ除去は 2 つの段階に別かれる。最初に、2 値化画像に対して 8 近傍縮小、8 近傍膨張を順に行う開口処理を施す。この処理により、書籍輪郭線上のヒゲと、小さなノイズを除去する。次に、画像内で画素連結領域が最大となる領域を書籍領域と決定し、それ以外の領域をノイズと判断して除去する。画素連結領域の測定には輪郭線追跡法を用いる。画像をラスタ走査して最初に検出した図形画素から半時計回りに探索を行い、連結領域の境界線を抽出する。領域の横方向サイズ W_x と縦方向サイズ W_y を求め、 $W_x \times W_y$ を特徴量として用いて最大領域を決定する。

輪郭線追跡の前に開口処理を加えることにより、計算量の削減と、対象書籍の正確な輪郭線を得ることができる。また、ノイズの除去に輪郭線追跡を用いることで、ノイズの除去と同時に書籍の歪み具合をパラメータとして取得することができる。

4. パラメータの決定

抽出した書籍領域情報から、情報付加のために必要なパラメータを決定する。

書籍は四角形に近似できる限定された形状のため、書籍四端点の位置情報は、単純に書籍領域内画素で x, y 座標値の組合せ

- $(-x - y)$
- $(x - y)$
- $(-x + y)$
- $(x + y)$

が最大となる画素で特定される。

書籍上部の歪みは、ノイズ除去時に抽出した書籍の輪郭線情報を用いて算出する。

図 3 に従い、書籍上部左端点座標を (x_1, y_1) 、右端点座標を (x_2, y_2) とすると、書籍上部の歪み具合を保持する配列 $d[i]$ ($i: 0 \leq i \leq x_2 - x_1$) は x 軸と書籍

^{*} "New Horizon English Course2", 東京書籍株式会社

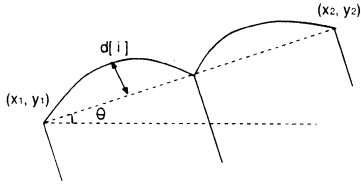


図3 歪み配列の決定

Fig. 3 decision of distortion array

がなす角度 θ を用いて以下の式により決定する。

$$d[i] = (y'_i - \frac{y_2 - y_1}{x_2 - x_1} x_i) \cos \theta \quad (1)$$

なお、 θ は次の式により求めることができる。

$$\cos \theta = \frac{x_2 - x_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} \quad (2)$$

また、書籍の見開きの中心部分は書籍上部、下部の輪郭線の変化量から決定することができる。

5. 情報の重ね合わせ

決定したパラメータを用いて、付加情報と入力画像から、出力画像を生成する。プロトタイプシステムでは付加する情報は事前に作成されたラスタデータとして与えられる。

入力画像内の書籍は、回転・拡大縮小・あおりなどの影響を受けているため、この画像に対して直接情報の付加を行うことは、付加位置の対応付けの点などから困難である。そのため、入力画像内の書籍イメージと、歪みのない四角形である内部書籍イメージとの対応をとり、内部イメージに対して情報の付加を行うことで出力画像を生成する。

内部イメージから出力画像への変換は2段階の処理からなる。

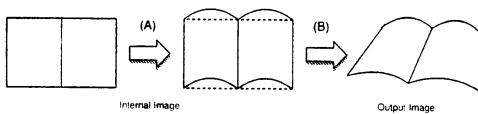


図4 2段階の画像変換

Fig. 4 two-step approach of image conversion

はじめに、正規化された内部イメージに対して、書籍の歪みに従った変化を与える(図4(A))。これは、先の章で取得した歪み配列 $d[i]$ を用いて、内部イメージの y 座標値を増減させることでおこなう。内部イメージの x 座標値が i である y 座標値を y_i とすると、以下の式

$$y_i = y_i + d[i] \quad (0 \leq i \leq i_{max}) \quad (3)$$

を用いて y 座標値をずらす。

次に、 y 座標値をずらした内部イメージに平面射影変換を施す(図4(B))。内部イメージ上の点を $m_1 = [x_1, y_1, 1]^T$ 、出力画像上の点を $m_2 = [x_2, y_2, 1]^T$ とすると、2枚の画像間の平面射影変換行列 H の関係は次のように表される¹⁰⁾。

$$\lambda m_2 = H m_1 \quad (4)$$

$$\lambda \begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix} = \begin{pmatrix} h_0 & h_1 & h_2 \\ h_3 & h_4 & h_5 \\ h_6 & h_7 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} \quad (5)$$

$$\begin{cases} x_2 = \frac{h_0 x_1 + h_1 y_1 + h_2}{h_6 x_1 + h_7 y_1 + 1} \\ y_2 = \frac{h_3 x_1 + h_4 y_1 + h_5}{h_6 x_1 + h_7 y_1 + 1} \end{cases} \quad (6)$$

上の式で得た8個の未知係数が平面射影変換のパラメータとなる。

先に取得した撮影画像の書籍端点座標 (x_2, y_2) の組4個と、内部イメージのコーナ座標 (x_1, y_1) の組4個から、式(6)中の未知係数8個を推定する。

式(6)を変形して

$$\begin{cases} h_0 x_1 + h_1 y_1 + h_2 - x_1 x_2 h_6 - x_2 h_7 = x_2 \\ h_3 x_1 + h_4 y_1 + h_5 - h_6 x_1 y_2 - h_7 y_1 y_2 = y_2 \end{cases} \quad (7)$$

を得る。式(7)に4つの端点の対応座標の組 $(x_1, y_1), (x_2, y_2)$ を代入することで式が8個得られ、この8元連立1次方程式を解いて8パラメータを得る。

以上により求めた8パラメータを式(6)に代入して、 (x_2, y_2) を計算する。

変換結果の座標値は整数値ではないので、補間する必要がある。補完の方法には最近傍法、バイリニア法、バイキュービック法等があり、後者ほど処理時間がかかるが精度がよい。本システムでは処理時間と精度の兼ね合いから、バイリニア法を用いて補完を行なう。

以上の2段階の変換手順を踏むことで、より少ない計算量で付加情報を持つ出力画像(図5(b))を得ることができる。

図5は付加情報を明確に把握できるように、文字列ではなく帯状の模様を付加した例を示した。また、1段階目の変換を行わず、直接平面射影変換のみ行なった場合、出力結果は図5(a)のようになる。実際に文字列、アンダーラインの付加を行い、変換を行ったもの

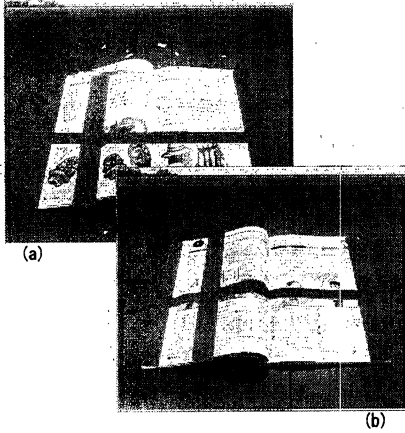


図 5 情報付加変換

Fig. 5 conversion to add information

が図 6 である。

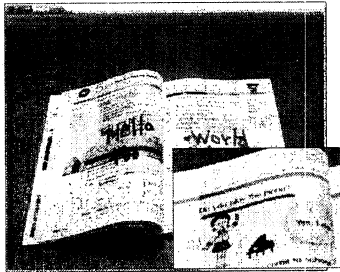


図 6 書籍への情報付加例

Fig. 6 example of adding information

6. 実時間性の検証

本システムの実時間性を検証するために、情報付加を行う際の画像処理のスループットと、実際の画像入力から表示までの遅延時間を 10 分間測定した。なお、ここでの情報付加は、画像認識、特徴抽出、パラメータ決定、出力画像生成、表示のプロセスの繰り返しを指し、書籍紙面の識別は含まれない。

結果は、表画像表示のスループットは平均毎秒 18.3 フレームであり、処理の遅延時間は 55msec であった。

表 1 実行時間

Table 1 lag time & throughput

画像処理のスループット	18.3 frame/sec
入力から表示までの遅延時間	55 msec

あくまで体感的なものであるが、実際に使用しても画像の表示に関して違和感を感じることなく使用可能であった。

7. ページの識別

紙面に付加する仮想情報を決定するために、現在のページを識別する必要がある。画像ベースの文書検索の手法としては、テキストを単語ごとに分割して各単語の文字数の並びを特徴量として表現する Hull の方法⁸⁾や、射影変換の不変量である単語重心の複比を用いて特徴量を計算する中居らの方法⁹⁾などがある。

本研究では、現在のところ対象書籍を中学生の教科書に絞っていることを考慮し、挿絵とページ内の文章が占める領域を特徴量として用い、識別を行う(図 7)。

識別のための領域判定は、挿絵・写真領域判定と文字領域判定とに別けて考える。

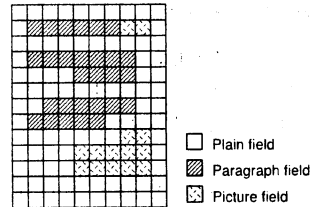


図 7 領域判定イメージ

Fig. 7 field determination

はじめに挿絵・写真領域を判定する。判定には色情報として彩度、明度を用いた。平面射影変換により正規化したイメージをグリッド状に区切り、各格子領域内の彩度、明度の平均値を計算する。最大領域となる色の値を白紙面の値とし、その差が閾値 θ 以上の格子を挿絵・写真領域と判定する。

次に文章領域の判定を行う。文章領域判定の前処理として、2 値化画像に対してガウシアンフィルタをかけ、その後再び 2 値化処理を行う。先に挿絵・写真領域と判定された格子以外で、2 値化結果値 0 が半数を占める格子を文字領域とする。前処理を行うことにより、図 8 のように文字部分が拡大されるため、文章領域判定の精度が向上される。

以上により判定した領域を、あらかじめ登録してある複数の紙面の領域情報と比較し、差が最小となるものを対象紙面と決定する。

なお、現在の判定方法では判定・検索にかかる時間を考慮していないため、実用化のためには、登録ペー

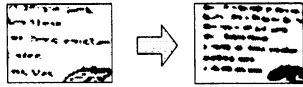


図 8 文字領域識別のための前処理
Fig. 8 preparation

ジが大規模になった際の高速な判定方法¹¹⁾を考える必要がある。

8. 付加情報の共有

実世界においても他人のノートや教科書への書き込みはしばしば有用である³⁾。また、梅村ら⁴⁾は e ラーニングにおける集団学習について、学習アチーブメントの向上などの有効性を示し、集団学習の場を

- 協調学習が行なえる、
- 学習中に他学習者の存在が認識できる、

と定義した。

本システムでは、同一書籍に対する付加情報を P2P ネットワークにより共有することにより、自律的な学習グループの形成を図る。これにより、学習者は他の学習者の存在を認識することが可能になり、他学習者の付加情報を参考に学習を進めることができる。これにより共有可能になる情報は書籍のコンテンツではなくユーザによる著作物であるため、知らないユーザ間で情報を共有する際に、誤って書籍の著作権を侵害してしまうといった危険性を回避できる。

付加情報の共有のみではなく、チャットの機能などを実装することにより、円滑な集団学習を支援することも期待できる。

9. まとめと今後の課題

実書籍を用いた複合現実による学習支援の可能性を検討し、プロトタイプシステムの開発を進めることで方向性を示した。今後は、ネットワーク内ユーザ間での付加情報共有の設計・実装を行うことでプロトタイプとしてのシステムを完成させる。また、現在の情報付加の位置精度やロバスト性などを向上させ、評価実験を行う。実際にシステムを利用した複数ユーザによる学習支援実験を行い、既存の e ラーニングシステムとの差別化を明確にすることも今後の課題である。

参 考 文 献

- 1) 岡田健一, 松下 温, “本メディアを超えて:Book-Window” 情報処理学会論文誌, Vol.35, No.3, pp.468-477(1994).
- 2) 伊藤清美, 柳沢昌義, 赤堀侃司, “Web 教材への書

き込みを可能とするシステムの開発と評価” 第 19 回日本教育工学会大会講演論文集, Vol.2, pp.753-754(2003).

- 3) 三浦克宜, 斎藤 一, 斎藤健司, 前田 隆, “電子教材に対する注釈付けと学習支援機能について”, 情報処理北海道シンポジウム 2003, pp.118-121, April 2003.
- 4) 梅村 透, 赤倉貴子, “Virtual な時間共有により集団学習の場を提供する e-learning システムの開発と評価”, FIT2005 情報科学技術レターズ, pp.215-218, Aug 2005.
- 5) 中原 敦編著, 北村士郎, 荒木淳子, 松田岳士, 浦嶋憲明, 小松秀徳, “ここからはじまる人材育成-ワークプレイスラーニング・デザイン入門-,” 中央経済社, 東京, pp.104-105, 2004
- 6) 園田智也, 村岡洋一, “空中での手書き文字入力システム” 信学論 (D-II), vol.J86-D-II, no.7, pp.1015-1025, July 2003.
- 7) 久木元伸如, 江原康生, “VR 環境における PDA を用いたパラメータ操作と注釈付与が可能な可視化システムの実装と評価” 信学論 (D-II), 電子情報通信学会論文誌, vol.J88-D-II, no.3, pp.596-604, March 2005.
- 8) J.J.Hull: “Document Image Matching and Retrieval with Multiple Distortion-Invariant Descriptors”, Document Analysis Systems, pp.379-396(1995)
- 9) 中井友弘, 黄瀬浩一, 岩村雅一, 松本 啓之亮, “複比とハッシュに基づく文書画像検索”, 電子情報通信学会技術研究報告, PRMU2004-249(2005-3)
- 10) K.Kanatani, Geometric Computation for Machine Vision, The Oxford Engineering Science, no.37, 1993.
- 11) 中井友弘, 黄瀬浩一, 岩村雅一, “デジタルカメラによる文書画像検索 - 1万ページから 0.1 秒で検索する”, FIT2005 情報科学技術レターズ, pp.215-218, Aug 2005.
- 12) 片山 淳, 中村高雄, 山室雅司, 曾根原 登 “電子素化し読み取りのための i アプリ高速コーナ検出アルゴリズム” 電子情報通信学会論文誌, Vol.J88, D-II No.6 June 2005