

## バージョン管理を用いた漢訳仏典翻刻支援システム

村川 猛彦<sup>†1</sup> 福岡 整<sup>†1</sup> 丁 敏<sup>†2</sup>  
中川 優<sup>†1</sup> 三宅 徹誠<sup>†3,\*1</sup> 落合 俊典<sup>†3</sup>

經典画像と既存のテキストファイルを対照表示させてテキスト部を編集可能とし、經典画像と同一内容のテキストファイルを効率良く作成することを目指して構築した、漢訳仏典翻刻支援システムについて述べる。本システムの特徴として、主にソフトウェア開発で用いられているバージョン管理ソフトウェア Subversion を導入し、テキストデータの履歴管理を行っている点、また編集を複数の作業者が同時に実施するための機能だけでなく、専門家がその差分表示を見比べて妥当なテキストを選定もしくは作成するための機能も実装した点があげられる。この分業体制を通じて、各参加者の小さな労力により、質の高い經典テキストデータの作成が期待できる。經典画像として金剛寺一切經撮影画像、既存テキストとして大正新脩大藏經テキストデータを用いて評価実験を行い、有効性を確認した。

## Transcription Support System for Chinese Translated Buddhist Sutras Using Version Control

TAKEHIKO MURAKAWA,<sup>†1</sup> HITOSHI FUKUOKA,<sup>†1</sup>  
MIN DING,<sup>†2</sup> MASARU NAKAGAWA,<sup>†1</sup>  
TETSUJO MIYAKE<sup>†3,\*1</sup> and TOSHINORI OCHIAI<sup>†3</sup>

We developed the support system of transcription for Chinese translated Buddhist sutras so that users can smoothly compose the text file. In this system, the version management software Subversion is introduced for maintaining the history of text files. In addition, we implemented not only the interface for rewriting by operators who look at a pair of shot image and editable text file, but that for choice or editing by an expert who can read the differential lines. Based on the division of labor, the system will be expected to produce text files which coincide with the shot images while the participants use their energy modestly. We conducted the experiments using shot images of Kongoji Issaikyo and the text data derived from Taisho Tripitaka to confirm the efficacy.

### 1. はじめに

大阪府河内長野市にある、天野山金剛寺が所蔵する一切經（金剛寺一切經）は、国内古写經の内容理解において重要な經典となっている。これまで、その經卷単位の存欠状況を調査するとともに、高精細デジタルカメラを用いた悉皆撮影を実施してきた<sup>1),2)</sup>。筆者らは、その撮影画像を研究等に適切に利用できるための手法の確立を目指している。

經典画像は視認性に優れているが、そのままでは検索ができない。大正新脩大藏經テキストデータを用いた全文検索および文書内検索が効果的であるが、画像とテキストの内容上の不一致（異同）により、該当ファイルや該当箇所をうまく取得・発見できないことがある。

検索の精度を向上させるためには、經典画像に合ったテキストデータを作成すること、すなわち翻刻が不可欠である。膨大な数の經典が存在することと、異同が見られるとはいえ大部分の經典に対しておおよそ対応するテキストデータが利用可能なことから、画像の各文字を順にテキスト化するという、単純な形態の翻刻ではなく、經典画像と既存のテキストファイルを対照表示させ、テキストデータを編集可能とすることで、經典画像と同一内容のテキストファイルを効率良く作成できると考えられる。

本論文では、上記の着想をもとに構築した漢訳仏典翻刻支援システム<sup>3)-6)</sup>について述べる。本システムの特徴として、以下の点があげられる。(1) バージョン管理ソフトウェア Subversion を用いて、複数人による同時作業により異同を効率良く発見できる。また、同ソフトウェアの履歴管理機能により、翻刻作業の前後のテキストファイルを比較し、その差分を取得できる。(2) テキストファイルの編集を複数人が別々に実施するための「翻刻支援機能」と、専門家がその差分表示を見比べて妥当なテキストを選定もしくは作成するための「採用支援機能」を提供している。この分業環境を通じて、各利用者の小さな労力により、質の高い（編集漏れの少ない）經典テキストデータが作成できる。

本論文の構成は以下のとおりである。2章では準備として、漢訳仏典とその分類について

<sup>†1</sup> 和歌山大学

Wakayama University

<sup>†2</sup> 日本電気株式会社

NEC Corporation

<sup>†3</sup> 国際仏教学大学院大学

International College for Postgraduate Buddhist Studies

\*1 現在、元興寺文化財研究所

Presently with Gangoji Institute for Research of Cultural Property

述べるとともに、Subversion の機能等を紹介する。3 章では漢訳仏典翻刻支援システムについてその概要を記したのち、サーバのファイル構成、翻刻支援機能、採用支援機能に分けて説明する。本システムの有効性を確認するために実施した評価実験について 4 章で述べ、5 章では実験結果、Subversion の使用および関連研究をもとに考察を行う。

## 2. 準備

### 2.1 漢訳仏典とその分類

本研究が対象とする漢訳仏典は、奈良平安古写経・鎌倉写経（古写経）と大正新脩大蔵経（大正蔵）の 2 種類である。これらは流通経路および電子化方法が異なっている<sup>7)</sup>。

中国で漢訳された仏典が、中国への修行僧による書写もしくは輸入により日本に入ると、その経典がさらに書写され国内で流布していった。仏教学、国文学等において特に重要なのは、奈良・平安・鎌倉時代に書写されたものである。金剛寺一切経は、現存する平安鎌倉写経の 1 つである。近年、その存否状況の調査と、デジタルカメラを用いた経典撮影（電子画像化）を座主の協力のもとで実施してきた。金剛寺一切経の経典撮影画像は 3 千巻を超え、その中のいくつかの経典はテキスト化が試みられている（たとえば文献 8）が、網羅的なテキスト作成には至っていない。

なお本論文では、『放光般若経』巻第一および般若訳『大方広仏華嚴経』巻第三を、システム動作説明用および予備実験・本実験用の経典画像として用いている。これらは、奈良写経の転写本である可能性が高く、古い形態を保持していると思われること、また将来的に、金剛寺以外の古写経との比較をする際にも有用と考えられること等から選定した。

仏教学において漢訳仏典を引用する際に広く用いられる大正新脩大蔵経は、高麗版を中心とした刊本一切経を底本とし、20 世紀前半に編纂された。刊本一切経の作製時期は 10 世紀以降とされている。大正新脩大蔵経については、大蔵経テキストデータベース研究会 (SAT)<sup>\*1</sup>、および中華電子仏典協会 (CBETA)<sup>\*2</sup>により、そのテキスト化が行われ、約 9 千巻にも及ぶそのテキストデータは無料で検索および入手が可能である<sup>9)</sup>。

CBETA と SAT のテキストデータについて、同一経典のテキストをいくつか比較したところ、字体の違いは相当数あり、SAT の方がおおむね、現代日本語に近い字形を使用している。いくつかの例を表 1 に示す。CBETA で外字扱いされていた箇所が多くが SAT では

\*1 <http://21dzk.l.u-tokyo.ac.jp/SAT/>

\*2 <http://www.cbeta.org/>

表 1 CBETA と SAT の間で字形が異なる例

Table 1 Example of Chinese characters found in CBETA and SAT text files.

CBETA		SAT	
眾	(773E)	衆	(8846)
說	(8AAA)	説	(8AAC)
刹	(524E)	刹	(5239)
歩	(6B65)	步	(6B69)
為	(70BA)	爲	(7232)
眞	(771F)	真	(771E)

括弧内は Unicode のコードポイントで、「U+」は省略した。

1 文字に割り当てられていたのに対し、その逆の事例は見られなかった。したがって本研究では、SAT のテキストデータを使用している。

### 2.2 バージョン管理ソフトウェア Subversion

テキスト管理のため、本研究では Subversion を使用している。これは集中型バージョン管理ソフトウェアとして知られており、主にソフトウェア開発において用いられている。この節では、本論文で使用している、Subversion に関する用語や概念を説明する。

Subversion は一般に、1 台のサーバに複数のクライアントが非同期的に接続する、クライアント/サーバ型で運用される。サーバで管理している領域を「リポジトリ」と呼び、管理するディレクトリおよびファイルは、「リポジトリ URL」で指定される。「作業コピー」は、リポジトリ URL に結び付けられた（結び付ける行為を「チェックアウト」という）、クライアント側で変更可能なファイル領域を指す。なお、作業コピーの実体はディレクトリもしくはファイルであるが、リポジトリに関しては、必ずしもリポジトリ URL に対応するディレクトリもしくはファイルの実体が存在するとは限らない（具体例は 3.2 節で述べる）。

リポジトリと作業コピーとの間の情報交換には次の 2 種類がある。すなわち、「アップデート」はリポジトリから作業コピーへ最新のファイルを送る作業であり、「コミット」は反対に、作業コピーで行った修正をリポジトリへ送って反映させる作業をいう。リポジトリの変更は「リビジョン」と呼ばれる通し番号で管理され、1 回のコミットにより、リビジョンは 1 増える。

Subversion はしばしば SVN と略され、小文字の svn は、各種操作のコマンド名である。同一ファイルの任意の 2 つの時点（リビジョン）の相違内容のことを「差分」と呼ぶ。svn コマンドに適切な引数を与えて実行することで、行単位の差分を取得できる。

1 つのリポジトリに複数の作業コピーが結び付けられ、運用していると、短時間に同一



図 1 異同の例

Fig. 1 Example of passage which is not coincident with the text.

ファイルを複数のクライアントがコミットしたときに、「衝突」と呼ばれる状態になり、後のコミットは失敗する。これは Subversion が楽観的な制御をしているためつねに起こりうる。通常の利用において、衝突が起こった場合には当事者間で対応を決め、解消をすることになるが、本システムではそのような対処が困難なため、異なる方法をとっている。

Subversion を導入するメリットは以下のとおりである。まず、これまで述べてきたように、複数の作業者の共同作業をサポートする。そして、修正内容だけでなく、コミットをした者や時刻等の情報、「コミットログ」と呼ばれる任意の変更記録情報を同時に残すことができる。さらに、修正内容を任意の時点でサーバに送ることができるため、一種のバックアップツールとして使用できる。

### 2.3 翻刻テキストの必要性

経典画像から対応するテキストファイルの特定を支援するため、筆者らはこれまで、CBETA の提供するテキストデータをもとに、横方向検索やワイルドカード検索といった付加機能を設けた全文検索システムを試作した<sup>10),11)</sup>。また、CBETA や SAT の Web サイトにおいても、テキストデータが整備され、全文検索だけでなく大正新脩大蔵経の分類をもとにしたファイル検索も可能となっている。

しかし金剛寺一切経と大正新脩大蔵経は、漢訳仏典という共通点を持ちつつも、書写時や木版に至るまでに記載ミスや意図的な変更があり<sup>1),7)</sup>、大小の相違点、すなわち「異同」が存在する。たとえば図 1 は「諸優婆夷塞」と読めるのに対して、これに対応する（前後が一致している）大正新脩大蔵経テキストは「諸優婆塞」である。画像側はテキスト側に比べ、「夷」という文字が余分にあるため、そのまま検索しても、テキストファイルを取得でき

ないという問題が生じる。漢訳仏典において「優婆塞」「優婆夷」という用語がよく見られること等から、画像内の「夷」は書写ミスによるものと推測できる。

大正新脩大蔵経の経典テキストデータを、対応する古写経経典画像に合うよう修正すれば、その修正テキストデータを用いて経典画像の検索が可能となるだけでなく、修正前後のテキストデータの異同を計算機により効率良く発見でき、またデータ分析を通じて、古写経と刊本との違いを知る手がかりになると考えられる。

本システムを用いて古写経経典画像に合ったテキストデータを作る行為を、本論文では「翻刻」と呼んでいる。筆者らの過去のシステムでは、「校訂」と記述していたが、これは既存テキストを改訂 (revise) して目的に合ったファイルを作ることにちなんだものである。しかし「校訂」という言葉は通常、異本との照合や語学的な検討 (critics) を経たうえて、表現を決定したり変更したりするといった意味で使われ<sup>12)</sup>、本システムはそのような行為を直接的に支援するものではない。別のいい方をすると、個別具体的な経典に文字単位で対応するテキストの作成を支援するのであって、(漢訳仏典としての) 原文の復元を目指すことは想定していない。むしろ、たとえば『原本通りの内容を一字一句の増減・変更も無く伝えるように出版すること』<sup>13)</sup> として定義される「翻刻」のほうが適切であると考え、この語を使用することとした。

## 3. 漢訳仏典翻刻支援システム

### 3.1 システム概要

本システムは大きく 2 つの機能からなる。テキストファイルを、必ずしも専門家でない作業者が経典画像とともに閲覧し、相違箇所があればテキストの中身を書き換える「翻刻支援機能」と、作業結果を専門家が閲覧し、修正された行について採否を決め、翻刻テキストを作る「採用支援機能」である。加えて、両機能を円滑に運用するためにテキストファイルを管理・変換する処理も実装している。本システムを用いた古写経翻刻の流れを図 2 に示す。なお、前処理において経典画像におおよそ対応するテキストを発見し、翻刻支援機能でできるように改行箇所を画像と対応づける作業は、人手により行う必要がある。

以下では、翻刻支援機能を使う者を「作業者」、採用支援機能を使う者を「専門家」と呼び分けている。また作業者と専門家を合わせて「利用者」と呼ぶ。システムの利用者という意味で、「ユーザ」を用いることもある。実際には、専門知識を持つ者が翻刻支援機能を使っても差し支えない。ただし、翻刻支援機能では、書き換えた理由を記載する欄をあえて設けておらず、照合および文字修正に集中するよう配慮している。

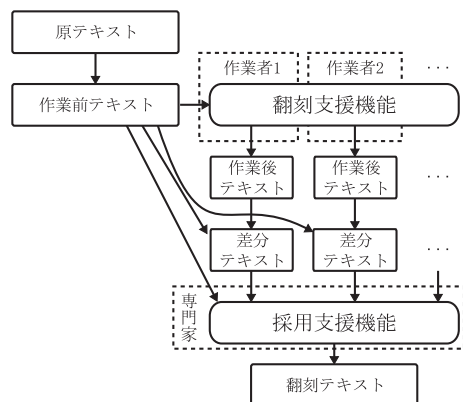


図 2 本システムを用いた古写経翻刻の流れ  
Fig.2 Workflow of transcription using the system.

複数人の作業者が同一の経典を編集できるようにする必要性について、実例をあげながら述べる。般若訳『大方広華嚴經』巻第三の中に、「觀察如来不可思議菩薩衆海觀察如来不可思議三昧神通種種變現觀察如来不可思議入一切諸世界海觀察如来不可思議一切法幻智境界觀察如来不可思議普現三世一切諸佛悉皆平等觀察如来不可思議无量无边諸語言道開悟一切諸法門海」という記述があり、下線を施した「觀察如来不可思議」は 6 回出現している。その間の字句は、「一切」が複数回出現するものの、いずれも異なる。他にも、連続する行で何文字かが同じものや、1 行を挟んで行頭の何文字かが同じものも見られる。頻出する語には「菩薩」「三昧」「一切」等がある。これらの頻出字句を知るとは、翻刻をより円滑に実施するのに有用であるが、それにとられるあまりその間の文字や頻出字句自体の要修正箇所を見落としやすい。1 人で全体を何度も通して読めば、見落とし数を減らすことはできるが、1 度読んだ際の記憶が悪影響を及ぼすこともありうる。それに対して、複数人の作業者が、各自 1 回ずつ読んで修正する形態をとり、採用支援機能によって誰か 1 人でも修正した行を専門家が見直すという作業環境があれば、各利用者の小さな労力により、編集漏れの少ない翻刻テキストが得られると考えた。なお、この作業環境において専門家は 1 人とし、複数の専門家を擁する場合は 5.1 節で検討する。

システムの構築にあたり、サーバには Debian Linux を用いた。サーバアプリケーションを開発する際によく用いられる LAMP (Linux: オペレーティングシステム, Apache: Web サーバ, MySQL: データベース管理システム (DBMS), PHP: アプリケーション開

発言語・処理系)のうち、MySQL の部分を Subversion に置き換えた構成といえる。

### 3.2 ファイル管理

サーバの持つファイル構成について検討する。本システムでは、リポジトリだけでなく、各利用者の作業コピーもサーバ側で保持することとする。ユーザ認証を入れ、認証ユーザごとに異なる作業コピーを保持する。これにより、利用者は作業コピーの管理を意識することなく、使い慣れた Web ブラウザがあれば作業ができる。

SVN サーバと Web サーバは同一の計算機とする。通常、クライアントが SVN サーバのリポジトリに接続する方法は、リポジトリ URL の先頭部分(「http://」、「https://」、「file://」、「svn+ssh://」等)により決まるが、本システムでは「file://」を用いている。リポジトリおよび作業コピーの所有者(Linux のユーザ名)は、Web サーバプロセスの所有者と同一とし、SVN のクライアント・サーバ間の認証は特に行っていない。

2.2 節で述べた衝突の可能性について、本システムでは、事前に分岐することにより解決を試みた。まず、新たに編集対象となった経典テキストファイルについて、リポジトリ URL を割り当てる。このリポジトリ URL は、どのユーザの作業コピーにも対応づけられない。ある利用者が翻刻支援機能もしくは採用支援機能を用いて、そのテキストファイルを編集したいときに、Subversion の「分岐」と呼ばれる操作を用いて、もとのリポジトリ URL から、ユーザ名を含むリポジトリ URL として複製を行い、さらに新規に作業コピーを確保し、この複製されたリポジトリ URL を用いてチェックアウトする。以後の編集において、同一の経典テキストファイルであっても利用者は別々のリポジトリに結び付けられた作業コピーを編集するため、衝突が起こらない。また差分は、分岐した時点のリビジョンと最新状態との間で求められれば十分である。

『放光般若経』巻第一を編集する場合のリポジトリ URL およびファイル名は次のとおりとした。下線部は、サーバに実在するディレクトリ名またはファイル名である。

リポジトリ URL (対応ユーザなし):

```
file:///opt/Butten/repository/main/T08n0221_001/body.txt
```

リポジトリ URL (ユーザ user01):

```
file:///opt/Butten/repository/user/user01/T08n0221_001/body.txt
```

作業コピー (ユーザ user01):

```
/opt/Butten/workingcopy/user01/T08n0221_001/body.txt
```

### 3.3 翻刻支援機能

本節で述べる「翻刻支援機能」とは、経典画像とテキストを対照表示させ、不一致の見ら



図 3 翻刻支援機能の画面例

Fig. 3 Screenshot of transcription support interface.

れる箇所についてそのテキスト部を修正できるようにしたものである。この機能は、複数人の作業者が同じまたは異なる時間帯に利用することを想定している。Web アプリケーションとしての翻刻支援機能は、ログイン画面、經典選択画面、編集画面等から構成される。はじめにユーザは、ログイン画面でユーザ名とパスワードを入力する。認証に成功すると、經典選択画面が表示され、翻刻作業を行う經典を選択すると、經典画像と經典テキストが対照表示された、編集画面となる(図3)。

この画面は、Scroll Viewer<sup>14)</sup> という名称で筆者らが開発した画像・テキスト対照表示アプリケーションに、テキスト編集機能を追加したものである。ここで対照表示アプリケーションについて簡単に述べる。画像部、縦書きテキスト部のいずれにおいても、マウスでドラッグすると任意方向に移動でき、操作していない画像部またはテキスト部も連動して表示領域が変わる。内部的には、画像を 128 × 128 ピクセルに断片化し、Ajax の仕組みを利用して描画に必要な画像のみを送付するようにしている。またテキスト部は、經典あたりのバイト数がそれほど大きくないので、編集画面表示にあたりすべてを Web クライアントが受け取っている。連動させるための、縦書きの各行に関する座標情報(画像処理においては X 座標すなわち列方向)をあらかじめ經典画像ごとに作成し、サーバで保持している。

翻刻支援機能とするにあたり、編集可能な箇所に の印を付け、その行番号を表示させている。經典画像は通常、数百行からなり、ドラッグ操作のみで末尾へ行くのは困難であるため、直前や直後の画像表示領域、また經典の先頭(画像の右端)や巻末(左端)に容易に移

動できるように、画像の上にボタンを配置している。

作業者は、この対照表示アプリケーションを利用して、(1) 目視で画像・テキスト間の異同を探索する。(2) 異同を見つけたら、テキスト下にあるボタンをクリックすることで、編集のためのテキストボックスが表示される。(3) このテキストボックス内の文字列を適切な内容に変更し、右横のボタンを押すと、その内容に変更される。サーバ内では 1 回の修正につき 1 回のコミットを行っている。変更した行は、色が変わるようになっており、作業者がどの行を修正したかを、容易に確認できる。この(1)–(3)の繰返しで、順次、翻刻を行っていく。

ここで縦書きと横書きについて述べる。經典画像は縦書きであるので、それと対となって参照できるように、対照表示アプリケーションにおいては縦書き表示としている。しかしテキスト編集用のテキストボックスについても、縦書き表示をさせたところ、カーソルの移動や文字の入力・削除が直感的でなく、実際、システムの動作確認時に作業者が大きく混乱した。Web ブラウザ上での縦書きテキスト編集が有用な事例を知ることができず、結局、テキストボックスについては横書きを採用した。なお、マウスで画像部もしくは縦書きテキスト部をドラッグすると、それに連動してテキストボックスにおいても表示内容が対応する行番号のものに変わるようにしている。作業者は好みや状況に応じて、經典画像・縦書きテキスト・横書きテキストの 1 つ以上を参照して、經典を読み進めればよい。

### 3.4 採用支援機能

本節で述べる「採用支援機能」とは、複数の作業者が翻刻支援機能を利用して作成した、作業後テキストに対して、専門家が閲覧し、適切な修正を取捨選択するものである。この機能は、經典データごとに、1 人の専門家が利用することを想定している。

翻刻支援機能を用いた作業を終えると、作業者ごとに作業後テキストが作られているので、作業前テキストとの間で文字単位の差分を求め、差分テキストとする。Subversion における svn diff コマンドや、Linux の diff コマンドによる差分では、行単位でしか情報が得られず、その行のどこを修正したのかという文字単位の取得には対応していない。そこで、自前で文字単位の差分を算出している。初期には、作業前テキストおよび作業後テキストの文字間に改行を挿入してから、diff コマンドを実行し、その出力をもとに求めていた<sup>6)</sup>が、その後、作業前テキストおよび作業後テキストの対応する行ごとに、編集距離に基づき文字単位で差分を求める処理を PHP で記述して使用している。ただし通常の編集距離で採用される、挿入・削除・置換に加えて、前後の 2 文字を逆に配置する「交換」についても、距離

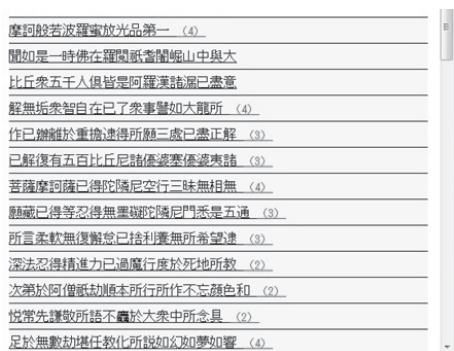


図4 採用支援機能の画面例(初期画面)

Fig. 4 Screenshot of initial decision support interface.



図5 採用支援機能の画面例(行編集)

Fig. 5 Screenshot of decision support interface for editing.

1の違いとして処理対象としている<sup>\*1</sup>。異同箇所には前述の4種類で異なる色付けを施しており、各作業者がどこを修正したかを採用支援機能で見やすくしている。すべての作業者の修正後テキストに対して、色指定付きの差分テキストを生成すると、採用支援機能が使用可能となる。

画面例について述べる。まずユーザ名とパスワードで認証を行い、経典を指定すると、図4のように表示される。各行は編集前テキストであり、行末の数字は、その行を変更した作業者の人数を表している。

1回の採用作業は以下のとおりである。(1) 任意の行をクリックすると、その行に関する、各作業者の作業後テキスト、編集可能なテキストボックス、および備考記入用のテキストボックスが、アコーディオンメニューとして動的に引き出されて表示される(図5)。(2) このとき、以下のいずれかの操作を行う。(a) 誰かの作業者の内容を採用したければ、その者の右にある「採用する」のボタンを押す。(b) 各作業者の修正を取捨選択したい場合、編集可能なテキストボックスを用いて書き換え、「編集して採用する」のボタンを押す。(c) どの作業者の修正も取り入れず、編集前テキストの内容が適切だと判断したときは、単に「編集して採用する」のボタンを押す。(a)-(c)のいずれにおいても、備考記入用のテキストボックスには、その理由等を入力できる。また必要なら翻刻支援機能にもアクセスして、経典画

像を参照する。(3) 「採用する」または「編集して採用する」のボタンを押すと、その修正がサーバに送られ、(a)、(b)のいずれかであればコミットも行われる。

1回の採用作業により、画面上では、採用した内容に変更されるとともに、変更箇所の色付けがなされ、行末の数字が灰色になる。これにより、専門家がどこまで採用を行ったかを知ることができる。行末の文字の表示色を変えるための情報は、リポジトリや作業コピーとは別に、採用支援機能のユーザおよび経典の組合せごとに管理されている。また、「採用する」のボタンを押したときであっても、システムでは修正内容のみを管理し、誰の修正を採用したかについては管理しない。

## 4. 評価実験

### 4.1 実験概要

本システムを利用し、適切な作業後テキストおよび翻刻テキストが作成できるかを確かめるために、評価実験を行った。

実験協力者は、作業者として経典に関する知識のない大学生5人と、専門家として古写経を専門とする研究者1人である。

実験手順は以下のとおりである。まず各作業者は、別々のユーザとして翻刻支援機能にログインし、画像に合致するよう編集作業を行った。作業者ごとにリポジトリを分岐しているため、同一時間での作業による衝突の心配はなく、作業者には自由な時間に翻刻作業してもらった。5人の作業者が、翻刻作業を終えると、次に、専門家が採用支援機能を利用し、作

\*1 本システムでは非専門家による文字照合に基づく翻刻作業を想定したので、書写後に経典に付与される符号類については検討の対象外としている。しかし、交換に対応するものとして、顛倒符と呼ばれる符号が知られており、金剛寺一切経の経典においても付与されている事例がある。

業前テキストと得られた5つの差分テキストを参照しながら、適切な修正を見つけ採用していった。

『放光般若経』巻第一を用いて実施した実験を「予備実験」と呼び、4.2節で述べる。その結果に基づき、翻刻や採用の作業をしやすいよう整備し、般若訳『大方広華嚴経』巻第三を用いて実施した実験を「本実験」と呼び、4.3節で説明する。本実験では、翻刻テキストを得ただけでなく、専門知識を有する研究者が別途、翻刻作業を行い作業後テキストを作成し、それらの比較を試みている。

#### 4.2 予備実験およびその結果

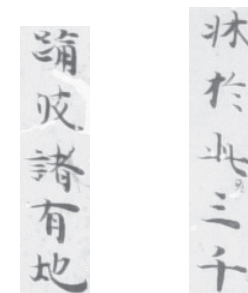
『放光般若経』巻第一を用いた予備実験について述べる。画像は Scroll Viewer 構築時に作成しており<sup>14)</sup>、これを使用した。原テキストは SAT から取得し<sup>\*1</sup>、不要な文字等を取り除いたうえで、經典画像と行が合うように手作業で改行位置の変更を行い、作業前テキストとした。作業前テキストは446行、7,543字の分量がある。

作業時間や範囲については特に指示しなかった。その結果、各作業者はテキスト全体を読み翻刻にあたったが、1人あたりの作業時間は5時間程度となった。また異体字等についての注意をしていなかったこともあり、字体変換をする修正が目立った。専門家が先頭から順に、1人でも修正した箇所を見ていくのはかなりの労となり、採用支援機能の使用は先頭から70行までで打ち切った。

作業者の修正および専門家の採用について、いくつか例を示す。図6左の文字列は、作業前テキストでは「踊没諸有地」となっていた。これに対し、作業者の1人が「踊肢諸有地」に置換したが、専門家はこれを「踊波諸有地」と判断し、書き換えた。また図6右はある行の行頭であり、作業前テキストでは「床於此三千」となっていた。作業者のうち4人がこれを「狀於此三千」としたが、専門家は、「床」の異字である「牀」に変更している。

表2は、各作業者および専門家の修正行ごとに、先頭から70行までの内容に関してその修正を分類したものである。3.4節で述べた挿入・削除・置換・交換のほか、1つの行で複数文字の修正をしていれば別に数えた。

各作業者の修正に対する採用状況は表3のとおりである。表の2列目以降の意味について述べる。「採用行数」は、作業者が修正し、かつ専門家がそれと同じものを採用した行数をいう。なお、3.4節で述べたように、採用支援機能においてどのユーザの修正を採用したかについては管理していないので、行ごとに一致しているか否かを比較し、計数した。「採



(a) 踊波諸有地 (b) 牀於此三千

図6 テキスト修正の文字例 (予備実験)

Fig. 6 Example of passages in preliminary experiment.

表2 利用者ごとの修正の分類 (予備実験)

Table 2 Types of editing in preliminary experiment.

作業者名	挿入	削除	置換	交換	複数	合計
user01	3	0	8	2	2	15
user02	3	0	27	1	19	50
user03	1	0	0	2	1	4
user04	2	0	17	1	10	30
user05	1	0	13	0	9	23
専門家	2	0	6	0	1	9

いずれも行単位の集計である。「複数」は、1つの行での複数箇所の修正を意味する。

表3 採用状況 (予備実験)

Table 3 Adoptions in preliminary experiment.

作業者名	採用行数 (%)	採用支援行数 (%)	見逃し数 (%)
user01	2 (13.3)	5 (33.3)	4 (44.4)
user02	1 (2.0)	8 (16.0)	1 (11.1)
user03	2 (50.0)	2 (50.0)	7 (77.8)
user04	2 (6.7)	7 (23.3)	2 (22.2)
user05	2 (8.7)	5 (21.7)	4 (44.4)

用支援行数」は、作業者が修正し、かつ専門家も修正した行数をいう。その内容が一致しているかは問わない。値は必ず、採用行数以上となる。この尺度を導入したのは、「ここは修正すべきではないか」と作業者が判断したものについて、その修正内容が適切でなかったと

\*1 [http://21dzk.l.u-tokyo.ac.jp/SAT/T0221\\_08,0001a01:0221\\_08,0007b07.html](http://21dzk.l.u-tokyo.ac.jp/SAT/T0221_08,0001a01:0221_08,0007b07.html)

表 4 修正人数ごとの行数 (予備実験)  
Table 4 Numbers of persons and edited lines in preliminary experiment.

作業人数	修正行数	採用支援行数 (%)
0	17	—
1	16	2 (12.5)
2	17	1 (5.9)
3	11	2 (18.2)
4	6	3 (50.0)
5	3	1 (33.3)
合計	70	9

しても、専門家の採用決定に影響を与えたことを表すためである。最後に「見逃し数」は、専門家が修正したが作業者は修正をしなかった行数であり、専門家の採用行数 (予備実験では 9) から採用支援行数を引くことで求めている。採用行数および採用支援行数のパーセンテージを求める際の分母は、作業者の修正行数であり、見逃し数においては専門家の採用行数である点に注意を要する。

作業人数の採用に対する影響を見るための算出を試みた。表 4 は、ちょうど  $n$  人 ( $0 \leq n \leq 5$ ) の作業者が修正した行数ごとに、上で定義した採用支援行数を求めたものである。修正した作業者が多い行ほど、そのまま採用されるか、修正のうえで採用される割合が高くなるのが期待されるが、そのようにはなっていない。

表 2 を見ると、修正行数は user02 が 50 行なのに対し、user03 が 4 行にとどまっている等、ばらつきがみられる。このようになった原因の 1 つとして、どのようなときにテキストを書き換え、どのようなときには書き換える必要がないかの指針が定まっていなかった点が指摘できる。この点の解決を試み、作業員および専門家の負担を減らしてシステムが利用できるよう、次節で述べる本実験で配慮した。

#### 4.3 本実験およびその結果

般若訳『大方広仏華嚴經』巻第三を用いた本実験について述べる。画像は Scroll Viewer 構築時に作成しており、これを使用した。原テキストは SAT から取得し<sup>\*1</sup>、作業前テキストとするよう前処理を行った。

作業員は 5 人で、うち 4 人 (user01–user04) は前節の実験と同一ユーザであり、user05 は交代している。専門家は前節と同じである。利用者の負担低減させつつより適切なテキス

\*1 [http://21dzk.l.u-tokyo.ac.jp/SAT/T0293\\_,10,0670c03:0293\\_,10,0675c27.html](http://21dzk.l.u-tokyo.ac.jp/SAT/T0293_,10,0670c03:0293_,10,0675c27.html)

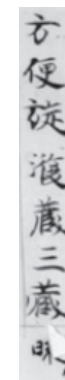


図 7 テキスト修正の文字例 (本実験)

Fig. 7 Example of passage in main experiment.

トを作成するための試みをいくつか導入している。作業範囲は經典画像の先頭から 200 行とした (作業前テキストは 3,312 文字となった)。ばらつきを抑えるために、作業員とは別に、システム構築に関わった者があらかじめ翻刻支援機能を使用し、字形が違うが同一視できる (したがって修正の不要な) 漢字画像と文字のリストを作り、作業員に参照させた。また各作業員には、本システムの全体像を示したうえで、見逃しがあってもよく、他の人が補ってくれると考えて気楽に作業すること、迷った場合は修正をしないことを伝えた。翻刻支援機能において、画像部とテキスト部の表示領域を広げ、上下方向の移動をしなくても、1 行全体が表示されるようにした。

作業員の修正および専門家の採用について、例を示す。図 7 に対応する翻刻前テキストは「方便施洩藏三蔵味」である。この行は 5 人中 2 人の作業員が変更しており、1 人は「方便施洩藏三蔵味」、もう 1 人は「方便施洩藏三蔵味」としている。ここで、図 7 末尾の「味」は經典の本文の外にあるとともに、カタカナ書きの「カ」が添えられている。これは行の最後の文字の「蔵」は「味」ではないかとする意図で、書写後に書き加えられたものと推測される。この場合、翻刻においては「味」を採用せず、専門家はテキストボックスで編集して「方便施洩藏三蔵」とした。

本実験で得た、作業員ごとの修正の分類に、各作業時間を加えたものを表 5 に示す。なお、作業時間には休憩時間は含まれていない。前節と同様にして算出した、採用状況、および修正人数ごとの行数を、それぞれ表 6、表 7 に示す。この実験でも、user02 の修正行数



表 5 利用者ごとの修正の分類 (本実験)  
Table 5 Types of editing in main experiment.

作業者名	挿入	削除	置換	交換	複数	合計	作業時間
user01	6	2	7	0	1	16	52 分
user02	5	2	46	0	7	60	1 時間 39 分
user03	3	0	7	0	0	10	55 分
user04	4	2	6	0	1	13	35 分
user05	5	0	12	0	0	17	53 分
専門家	4	0	14	0	3	21	1 時間 17 分

表 6 採用状況 (本実験)  
Table 6 Adoptions in main experiment.

作業者名	採用行数 (%)	採用支援行数 (%)	見逃し数 (%)
user01	10 (62.5)	13 (81.3)	8 (38.1)
user02	14 (23.3)	21 (35.0)	0 (0.0)
user03	8 (80.0)	9 (90.0)	12 (57.1)
user04	9 (69.2)	10 (76.9)	11 (52.4)
user05	9 (52.9)	12 (70.5)	9 (42.9)

表 7 修正人数ごとの行数 (本実験)  
Table 7 Numbers of persons and edited lines in main experiment.

作業者数	修正行数	採用支援行数 (%)
0	138	—
1	39	4 (10.3)
2	6	3 (50.0)
3	7	5 (71.4)
4	6	5 (83.3)
5	4	4(100.0)
合計	200	21

は多い。しかし、それ以外のユーザの修正行数等はばらつきが抑えられている。

さらに、5人の作業者のうち何人だけが実施していればどのくらい採用されるかを算出した(図8)。具体的には、5人の作業者の各部分集合(0人を除く)について、その集合による作業結果から、専門家が1人以上修正した行を見て採用もしくは編集して採用をしたと仮定すると、採用支援行数がどれだけあるかを推計し、作業者数ごとに平均値を求めている。各作業者は独立して翻刻支援機能を用い編集している点に注意すると、採用支援の行数は、想定作業者数に対して単調に増加するが、比例関係ではなく、人数の増加にともない伸

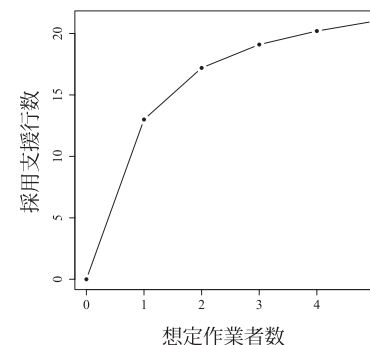


図 8 想定作業者数ごとの状況 (本実験)

Fig. 8 Number of persons and estimated adopted lines in main experiment.

表 8 専門知識を持つ作業者の修正の分類  
Table 8 Types of editing by a worker with expertise.

挿入	削除	置換	交換	複数	合計	作業時間
5	2	21	0	1	29	1 時間 34 分

表 9 専門知識を持つ作業者の推定採用状況  
Table 9 Estimated adoptions for a worker with expertise.

採用行数 (%)	採用支援行数 (%)	見逃し数 (%)
14 (48.3)	18 (62.1)	3 (14.3)

びが緩やかになると推測できる。図8からも、その傾向を確認することができる。

本実験で得た翻刻テキストを、専門知識を持つ者が翻刻支援機能を使った場合と比較することとした。具体的には、採用支援機能を使用した専門家とは別に、古写経を専門とする研究者1人の協力を得て使用してもらった。以下、このユーザを「専門知識を持つ作業者」と呼ぶ。事前に、5人の作業者に参照させたときと同様にシステム全体像および漢字画像と文字のリストを提示したが、比較対象を考慮し、他の人が補うものではないことを伝えた。

専門知識を持つ作業者による、修正の分類および作業状況を、表8および表9に示す。表9において採用行数、採用支援行数および見逃し数を求めるにあたり、前述の翻刻テキストと、専門知識を持つ作業者による作業後テキストとを比較した。表6では、各作業者の作業後テキストは翻刻テキストの作成に影響を及ぼしているのに対して、ここでの比較ではそのような依存関係がない。この違いのため、表9の各値は推定値としての算出となる。

見逃しの3カ所はいずれも、専門家が置換したもので、そのうち1つは、SATのテキストでは外字としていた文字であり、残りの2つは字形（筆使い）からでは修正すべきか判断の迷う文字であった。専門知識を持つ作業者が修正して、専門家がそのとおりにしなかったのは、複数の修正のうち専門知識を持つ作業者が1カ所しか修正していなかった行のほか、ここでも、判断の迷う文字であった。

## 5. 考察

### 5.1 評価実験の検討

本実験の結果より、作業者が多くなるほど、見逃しが少なくなることを確認した。また予備実験および本実験において、多くの作業者が行った修正をそのまま採用した事例だけでなく、多くの作業者が行った修正と異なる文字に置き換えて採用とした事例も見られた。これらは、作業者の修正内容が、専門家の編集作業に役立っていることを意味し、複数人の作業者と1人の専門家が連携して、本システムで經典画像に合ったテキストを作ることができたといえる。

表2と表5を比較すると、予備実験の修正行数については、5人中4人の作業者の修正行数が専門家よりも多いが、本実験ではこれが5人中1人になっている。この違いは、次のように説明できる。すなわち、予備実験では専門家は各作業後テキストのフィルタ役を担っていた。ただし、採用行数ならびに採用支援行数が低かったことを考慮すると、手間のかかるフィルタリング作業であったと推測できる。一方、本実験では、作業者個別に見ると十分でない作業内容に対して、集約して専門家の採用（翻刻テキストの決定）に影響を与えることができている。また表5と表8を見ると、専門知識を持つ作業者については、修正行数、作業時間とも、本実験の各作業者の平均を超えている。これは、他の作業者が補うことをしないという指示が影響し、慎重に翻刻作業をしたためと考えられる。図8にあてはめると<sup>\*1</sup>、専門知識を持つ作業者1人は、そうでない作業者2人分の働きをしていることを意味し、その一方で、専門知識は持たなくても人数を増やせば、質の高いテキスト作成ができることを示唆している。

翻刻の作業者数について検討する。参加者が多いほうが、より多くの要修正箇所を発見できる。しかし実施の時間や、採用機能使用時の一覧表示での見やすさを考慮すると、無制限

\*1 ただし、前述のとおり図8と表9は推定値であるほか、すべての行を任意に書き換えれば採用支援行数は専門家の修正行数と等しくなるため、採用支援行数のみを作業者の質を見るための尺度とするのは適切ではなく、修正行数や採用行数を併用する必要がある。

にするわけにはいかない。本実験では、1つの經典に5人が（別々に）翻刻するのは多すぎるという想定で翻刻支援機能を使ってもらい、専門家のチェックを経て作られた翻刻テキストを基準として、5人のうちの $n$ 人（ $1 \leq n \leq 5$ ）で実施したと仮定したときに、専門家のチェックを経たとして作られる翻刻テキストについて、どのくらい採用されるかを推計し、図8に示すとおり、想定作業者数ごとの採用支援行数を求めている。いい換えると、それらの表の採用支援行数は絶対的なものではなく相対的なものであり、「5人の作業者なら見逃しをしない」ではなく、「5人の作業者ならここまで修正できた。ではそれより少ない人数だとどうか?」という観点で、算出や検討を実施している。

実際には、作業者や翻刻対象經典、また同一經典であっても作業者と専門家の実施状況（計算機環境、体調等）次第で、採用支援行数や見逃し率の値は変わりうる。したがって、さらに実験を重ねることで、作業者数と採用支援行数等の関係を明らかにし、作業前テキスト（「正解」が作られていないテキスト）に対する、作業効率の見積りに役立てる必要があると考えている。

相対比較すなわち「枠内での比較」は、次のような発展も可能とさせる。すなわち、特定の作業者の組合せの場合に、その人数の平均よりも採用支援行数が多いか少ないときに、その原因を差分に立ち返って検討するといった、より細かい分析も期待できる。

一方、採用支援機能の作業者についても、1人の専門家のみで妥当なテキストを作成できるのかという疑問も生じる。これについては、採用支援機能の結果も、他の専門家による採用支援機能の入力とすればよい。具体的には、採用支援機能において、經典を選択したのち、その經典に対して翻刻支援機能もしくは採用支援機能を用いて編集したユーザの一覧を表示し選択できるよう、システムを修正することで、実現できる。そのような再帰的または階層的な採用支援機能の活用を通じて、編集ミスの箇所が減り、テキストの質がさらに向上すると思われる。

図9は、翻刻作業の形態とそのバージョン管理方式のうち3種類を表した概念図である。右方向を時間の経過とし、黒丸が修正（Subversionを用いた場合、コミット）のなされた時点である。(b)および(c)において最下段の線は、対応ユーザなしのリポジトリURLを、それ以外の水平方向の線は、何らかのユーザに結び付けられたリポジトリURLを表す。分岐も1回のコミットがなされる。また複数人が同時に使用して編集する場合でも、コミットの特長から、黒丸は同じ時刻に並ばない。なお(a)の運用形態も、翻刻支援機能を用いてユーザ認証のための情報を複数人で共有すれば、本システムで可能である。

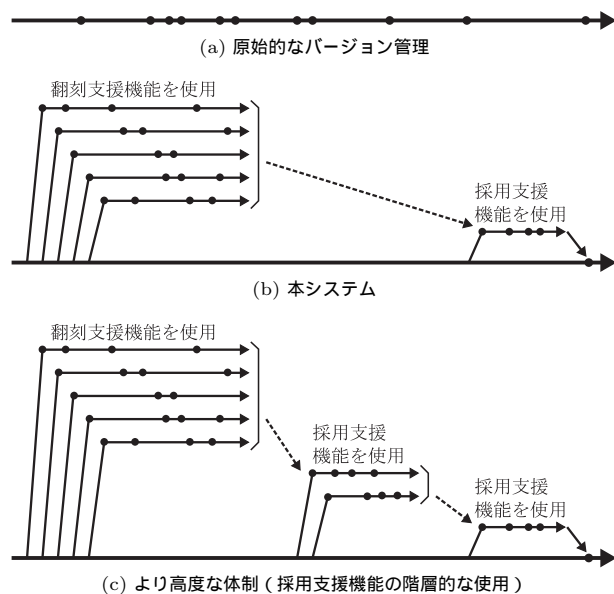


図9 翻刻作業とバージョン管理

Fig.9 Transcription environments and version managements.

## 5.2 Subversion の代替

本論文で述べたシステムでは、テキストデータの管理に Subversion を用いている。これを他のバージョン管理ソフトウェア、もしくは DBMS に置き換えることができるか検討する。

テキストデータの管理に使用した操作のうち主要なものは、分岐、コミット（変更箇所のほか、時刻、ログ、およびどのユーザがコミットしたかの情報を含む）、任意の時点のテキストの取得である。差分テキストは、3.4 節で述べたとおり svn diff コマンドではなく、翻刻支援機能使用前後のテキスト、具体的には同一のリポジトリ URL でリビジョンのみ異なる 2 つのテキストファイルを取得して、別途算出している。したがって、これらの操作が行えればよく、たとえば Subversion 以前に普及していた CVS（Concurrent Versions System）や、後発で分散型バージョン管理ソフトウェアの Git 等に置き換えることは、決して難しい。

バージョン管理ソフトウェアではなく DBMS によるテキスト管理については、次のように考えている。上記の操作をデータベースに対して行うよう、システムを修正することは原

理的に可能である。しかしその実装に先立ち、テーブル、カラム等のスキーマ設計が不可欠である。設計において、1 回のコミットにおける修正情報と、修正されたテキスト（作業後テキストおよび翻刻テキストを含む）を保持するよう、冗長性を廃しつつ効率の良いアクセスをするためのデータ構造を決定するのは容易ではない。またシステムにおいて何らかの機能拡張を実施する際、スキーマの見直しが必要となることもある。たとえば、コミットを取り消す機能を追加する場合、Subversion では直接的な取り消しはできないが、直前のコミットを逆方向に適用することで、前の状態に戻ることができる。DBMS での管理も、不可能ではないが、アプリケーションとスキーマの両方を修正しなければならないのは、運用において弱点といわざるをえない。

以上より、挿入・削除・置換・交換が考えられるテキスト間において効率良くその異同を管理するには、関係データベースをはじめとする DBMS ではなく、テキストや差分に特化した、バージョン管理ソフトウェアを使用するのが適切であると考えている。

## 5.3 関連研究

漢字からなる古典資料のテキスト化について、関連する研究を 2 種類あげ、本研究との違いを述べる。

まず、墨で書かれた古代史料という点で本研究と共通点のある、木簡について、その文字認識が行われている。その中でも、文脈処理を行いた欠損文字列の推定<sup>15)</sup>、欠損をとまなう文字に対してグレーゾーンを用いた文字パターン検索方法<sup>16)</sup>は、注目に値する。ただしその場合、木簡を分析して得られる語彙情報、具体的には年代や地名、あるいは単漢字との対照は行えるが、既存テキストと対照するのは、多くの木簡からテキストを翻刻しているのでない限り、困難である。また木簡は記述の複製を主目的としていないため、結局 1 点 1 点についてその内容の特定を行わなければならない。それに対し国内古写経については、すでに述べてきたように漢訳仏典として大正新脩大蔵經のテキストデータが利用可能となっており、これと照合することで、効率良く翻刻が行えるほか、多くの経典に対する差分テキストや翻刻テキストを分析（テキストマイニング）して、人手で誤って認識されやすい文字や、一切経間の特徴的な異同を発見することが期待できる。なお、そのような既存テキストが存在しないような史料に対しても、光学式文字読取装置（OCR）等を用いて文字認識を行い、それを作業前テキストとすることで、本システムの適用が可能である。

次に、漢訳仏典のテキスト化に関しては、SAT のデータベースや検索システムと合わせて、その管理方法が報告されている<sup>9),17),18)</sup>。そこでは、大正新脩大蔵經に限定することなく、サンスクリット語やチベット語等の文献と連携させた、インド学仏教学のための検索・

閲覧・編集のシステムが進められている。大正新脩大蔵経は最も入手や引用のしやすい漢訳仏典であるが、筆者らは、金剛寺一切経だけでなく、七寺一切経ほか国内古写経のテキスト比較が容易に行えるようなシステムを目指している。その際、史料もしくは原文の間の関連付け方法が、SAT と本研究とで異なる。すなわち、SAT の活動は多言語の既存テキストにおける字句間の対応付けを支援するのに対し、筆者らは、経巻レベルでは対応付けられているテキストにおいて不一致箇所を発見し保持することへの支援を主眼としている。

## 6. おわりに

本論文では、DBMS の管理ではなくバージョン管理ソフトウェア Subversion を用いて構築した漢訳仏典翻刻支援システムについて述べ、評価実験によりその有効性を確認した。翻刻支援機能および採用支援機能とに分け、複数人の作業者と 1 人の専門家の分業体制により、修正漏れの少ないテキスト翻刻ができ、字句レベルの分析や比較といった経典テキストの活用が期待できる。

今後の課題としては、実験を通じて利用者より要望のあった、画像およびテキストの拡大縮小機能、また採用支援機能から直接、対応する経典画像を参照する機能等の実装があげられる。本システムを他の経典に適用して、翻刻作業をより円滑に進めるための方策を見出すこと、SAT テキスト・翻刻テキスト・校訂テキストといった同一経巻の異なるテキストに対応した横断型の全文検索システムを構築すること等にも取り組みたい。

謝辞 国際仏教学大学院大学戦略的研究基盤形成支援事業研究員上杉智英氏には、本実験において協力を得た。また金剛寺御当局より御高配を賜った。ここに謹んで謝意を表する。

本研究は、文部科学省学術フロンティア推進事業「奈良平安古写経研究拠点の形成」の研究成果である。

## 参 考 文 献

- 1) 落合俊典：金剛寺一切経の基盤的研究と新出仏典の研究，平成 12-15 年度科学研究費補助金基盤研究 (A) (1) 研究成果報告書，課題番号 12301001 (2004)。
- 2) 落合俊典：金剛寺一切経の総合的研究と金剛寺聖教の基礎的研究，平成 16-18 年度科学研究費補助金基盤研究 (A) 研究成果報告書，課題番号 15202002 第 1 分冊 (2007)。
- 3) 丁 敏，村川猛彦，福岡 整，中川 優：Subversion を用いた仏典テキスト校訂支援システム，人文科学とコンピュータシンポジウム論文集，情報処理学会シンポジウムシリーズ，Vol.2008, No.15, pp.61-66 (2008)。
- 4) 福岡 整，村川猛彦，野田大地，中川 優：Subversion を用いた仏典テキスト校訂支

援システムの評価，人文科学とコンピュータシンポジウム論文集，情報処理学会シンポジウムシリーズ，Vol.2009, No.16, pp.61-66 (2009)。

- 5) Murakawa, T., Fukuoka, H., Noda, D. and Nakagawa, M.: Transcription Support System using Subversion, *Proc. 6th International Conference on Web Information Systems and Technologies (WEBIST 2010)*, pp.150-155 (2010)。
- 6) 福岡 整，村川猛彦，中川 優：古写経翻刻のための採用支援機能の構築，人文科学とコンピュータシンポジウム論文集，情報処理学会シンポジウムシリーズ，Vol.2010, No.15, pp.299-304 (2010)。
- 7) 落合俊典：日本古写経データベースと漢訳仏典研究，文部科学省私立大学学術研究高度化推進事業国際仏教学大学院大学学術フロンティア「奈良平安古写経研究拠点の形成」2009 年度公開シンポジウム講演資料集，pp.21-25 (2009)。
- 8) 国際仏教学大学院大学学術フロンティア実行委員会：集経礼懺儀巻下，Vol.4, 日本古写経善本叢刊 (2010)。
- 9) 下田正弘，永崎研宣：大蔵経と人文系データベース，情報処理学会研究報告人文科学とコンピュータ，2009-CH-82, pp.1-6 (2009)。
- 10) 田中猛彦，仁野洋平，中川 優：仏典データベースのためのテキスト処理について，情報処理学会研究報告人文科学とコンピュータ，2006-CH-69, pp.33-40 (2006)。
- 11) 村川猛彦，丁 敏，中川 優：仏典全文検索システムの構築と評価，人文科学とコンピュータシンポジウム論文集，情報処理学会シンポジウムシリーズ，Vol.2007, No.15, pp.221-228 (2007)。
- 12) 秋山陽一郎：校訂とはいかなる行為か？，漢字文献情報処理研究，Vol.6, pp.36-43 (2005)。
- 13) 新明解国語事典第六版：三省堂，p.1388 (2005)。
- 14) 村川猛彦，中川 優：経典デジタル画像閲覧のための情報処理技術，金剛寺一切経の総合的研究と金剛寺聖教の基礎的研究研究成果報告書 (平成 16-18 年度科学研究費補助金基盤研究 (A)：課題番号 15202002)，Vol.1, pp.35-52 (2007)。
- 15) 西嶋佳津，齋藤 恵，末代誠仁，中川正樹，馬場 基，渡邊晃宏：木簡解読支援のための文脈処理の提案と実装，人文科学とコンピュータシンポジウム論文集，情報処理学会シンポジウムシリーズ，Vol.2005, No.21, pp.241-246 (2005)。
- 16) 末代誠仁，齋藤 恵，戸根康隆，石川正敏，中川正樹，馬場 基，渡邊晃宏：古代木簡解読支援のための文字パターン検索，情報処理学会論文誌，Vol.50, No.4, pp.1444-1455 (2009)。
- 17) 永崎研宣，下田正弘：大正新脩大蔵経テキストデータベース構築のためのコラボレーションシステムの開発，情報処理学会研究報告人文科学とコンピュータ，2006-CH-70, pp.33-40 (2006)。
- 18) 永崎研宣，下田正弘：東洋古典文献研究におけるデジタルテキストの適切な記述方法について—インド学仏教学のための学術知識基盤の構築に向けて，人文科学とコンピュータシンポジウム論文集，情報処理学会シンポジウムシリーズ，Vol.2010, No.15,

(平成 22 年 12 月 20 日受付)

(平成 23 年 4 月 8 日採録)

(担当編集委員 相良 毅)



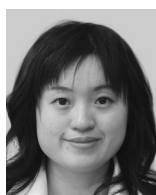
村川 猛彦 (正会員)

昭和 46 年生。平成 10 年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了。博士 (工学)。奈良先端科学技術大学院大学情報科学研究科助手, 和歌山大学システム工学部助手を経て, 平成 15 年より和歌山大学システム工学部講師, 現在に至る。Web アプリケーション, デジタルアーカイブに関する研究に従事。



福岡 整

平成 21 年和歌山大学大学院システム工学研究科博士前期課程入学, 現在, 同在学中。データベースシステムの研究に関心を持つ。



丁 敏

平成 21 年和歌山大学大学院システム工学研究科博士前期課程修了。同年より, 日本電気株式会社公共・医療ソリューション開発本部勤務。



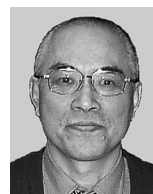
中川 優 (正会員)

昭和 45 年大阪大学基礎工学部卒業, 昭和 47 年同大学院修士課程修了。同年日本電信電話公社武蔵野通研, 平成 6 年近畿大学生物理工学部教授, 平成 9 年和歌山大学システム工学部教授。工学博士 (大阪大学)。IEEE, 情報知識学会, 人工知能学会各会員。



三宅 徹誠

昭和 45 年生。平成 18 年京都大学大学院文学研究科文献文化学専攻仏教学専修博士課程単位取得満期退学。修士 (文学)。国際仏教学大学院大学附置国際仏教学研究所以非常勤研究員, および同大学附置日本古写経研究所非常勤研究員を経て, 平成 23 年より元興寺文化財研究所嘱託研究員, 現在に至る。日本印度学仏教学会, 日本仏教学会, 仏教史学会各会員。



落合 俊典

昭和 23 年生。佛教大学大学院文学研究科仏教学専攻博士課程単位取得満期退学。修士 (文学)。華頂短期大学教授を経て, 国際仏教学大学院大学教授, 現在に至る。七寺古逸経典研究にはじまり日本古写経の研究に従事。