手書き情報の利活用のための
認識と共有に関する研究

提出先　　大阪大学大学院情報科学研究科

提出年月　2016 年 1 月

池 田 尚 司

研究業績

A. 論文

1. Hisashi Ikeda, Naohiro Furukawa, Masashi Koga, Hiroshi Sako, and Hiromichi Fujisawa, “Context-Free Grammar-Based Language Model for String Recognition”, *International Journal of Computer Processing of Oriental Languages*, Vol. 15, No. 2, pp. 149-163, 2002.

2. Hisashi Ikeda and Yukio Ohsawa, “Visualization of Insight Process in Concept Creation Focusing Handwriting Features”, *International Journal of Knowledge and Systems Science*, Vol. 4. No. 1, pp.18-31, 2013.

3. 池田尚司，古川直広，薦田憲久, “手書きによる情報共有に基づく設備保全作業支援システムの開発とプラント保全業務への適用,” 電気学会電子・情報・システム部門論文誌, Vol. 135, No. 6, pp. 580-588, 2015.

B. 国際講演

1. Hiromitsu Nishimura, Hisashi Ikeda, and Yasuaki Nakano, “A Segmentation Method for Touching Handwritten Japanese Characters”, in *Proceedings of IAPR Workshop on Document Analysis Systems (DAS98)*, pp. 130-139, 1998.

2. Hisashi Ikeda, Yukio Ogawa, Masashi Koga, Hiromitsu Nishimura, Hiroshi Sako, and Hiromichi Fujisawa, "A Recognition Method for Touching Japanese Handwritten Characters", in *Proceedings of 5th International Conference on Document Analysis and Recognition (ICDAR99)*, pp.641-644, 1999.

3. Hisashi Ikeda, Naohiro Furukawa, Masashi Koga, Hiroshi Sako, and Hiromichi Fujisawa, "A Context-Free Grammar-Based Language Model for Document

Understanding", in *Proceedings of IAPR Workshop on Document Analysis Systems (DAS2000)*, pp. 135-146, 2000.

4. Hisashi Ikeda, Naohiro Furukawa, Katsumi Furukawa, and Hiromichi Fujisawa, "Toward a Personalized Digital Library for Providing "Information JIT"," in *Proceedings of IAPR Workshop on Document Analysis Systems (DAS2004)*, pp. 47-50, 2004.

5. Naohiro Furukawa, Hisashi Ikeda, Yosuke Kato, and Hiroshi Sako, "D-Pen: A Digital Pen System for Public and Businesses Enterprises", in *Proceedings of 9th International Workshop on Frontiers in Handwriting Recognition (IWFHR-9)*, pp. 269-274, 2004.

6. Hisashi Ikeda, "Human Memory Expansion by Personal Handwriting for Realizing Information JIT", in *Proceedings of WM2005 Professional Knowledge Management, Experiences and Visions*, pp. 650 – 651, 2005.

7. Hisashi Ikeda, Kosuke Konishi, and Naohiro Furukawa, "iJITinOffice: Desktop Environment Enabling Integration of Paper and Electronic Documents", in *Adjunct Proceedings of 19th annual ACM Symposium on User Interface Software and Technology (UIST2006)*, pp. 93-94, 2006.

8. Hisashi Ikeda, Naohiro Furukawa, and Kosuke Konishi, "iJITinLab: Information Handling Environment Enabling Integration of Paper and Electronic Documents", in *Proceedings of CSCW 2006 Workshop Collaborating over Paper and Digital Documents (CoPADD 2006)*, pp. 25-28, 2006.

9. Naohiro Furukawa, Junko Tokuno, and Hisashi Ikeda, "Online Character Segmentation Method for Unconstrained Handwriting String using Off-stroke Features", in *Proceedings of 10th International Workshop on Frontiers in Handwriting Recognition (IWFHR-10)*, pp. 361-366, 2006.

10. Kosuke Konishi, Naohiro Furukawa, and Hisashi Ikeda, “Data Model and Architecture of a Paper-Digital Document Management System”, in *Proceedings of ACM Symposium on Document Engineering 2007 (DocEng 2007)*, pp.29-31, 2007.

11. Sandip Rakshit , Subhadip Basu , and Hisashi Ikeda, “Recognition of Handwritten Textual Annotations using Tesseract Open Source OCR Engine for information Just In Time (iJIT)”, in *Proceedings of International Conference on Information Technology and Business Intelligence*, pp. 117-125, 2009.

12. Hisashi Ikeda and Yukio Ohsawa, “Toward Visualization of Insight Process in Concept Creation Focusing Handwriting Features”, in *Proceedings of 7th International Conference on Communication and Information Technology (CIT13)*, pp.136-143, 2013.

C. 国内講演

1. 池田尚司，小川祐紀雄，古賀昌史，西村広光，酒匂裕，藤澤浩道，“手書き接触漢字切出しに関する検討”，1998 年度電子情報通信学会ソサエティ大会，D12-14，p. 236，1998.

2. 古川直広，加藤陽介，池田尚司，酒匂裕，“レイアウト駆動型文字列認識方式”， 電子情報通信学会技術研究報告 パターン認識・メディア理解，Vol. 103，No. 295，pp. 67-72, 2003.

3. 古川直広，加藤陽介，池田尚司，酒匂裕，“再帰遷移ネットワーク型文字列照合方式の高速化”，情報処理学会第 66 回全国大会, pp. 115-116, 2004.

4. 古川直広，徳野淳子，池田尚司，“自由文読取のための文字切出し方式の開発”，情報科学技術フォーラム一般講演論文集, Vol. 4, No. 3, pp. 39-40, 2005.

5. 小西康介，古川直広，池田尚司，“文書とのインタラクションを考慮した文書情報表現方式の検討”，情報科学技術フォーラム一般講演論文集，Vol. 5，No. 4，pp. 421-422，2006

6. 古川直広，池田尚司，小西康介，“紙とペンによる情報アクセス方式の開発”，情報科学技術フォーラム一般講演論文集, Vol. 5, No. 4, pp. 423-426, 2006.

7. 池田尚司，小西康介，古川直広，“文書へのアノテーションを活用する文書管理システムの開発”，情報科学技術フォーラム一般講演論文集, Vol. 5, No. 4, pp. 427-428, 2006.

8. 古川直広，池田尚司，小西康介，“デジタルペンを用いた研究ノートの開発”，インタラクション 2007, pp. 59-60, 2007.

9. 池田尚司，古川直広，藤澤浩道，“紙・電子文書統合管理による手書き情報の活用方式”，2009 年度 HCG シンポジウム，CD-ROM, 2009.

10. 川口英夫，川上憲人，有馬秀晃，池田尚司，坂入実，“書字の時間構造を用いたメンタルヘルスの可視化”，第 38 回可視化情報シンポジウム， Vol. 30，No. 1，pp. 159-160，2010.

11. 池田尚司，額賀信尾，“非構造化データを扱う情報処理基盤の実現を支えるメディア処理技術”，2012 年人工知能学会全国大会，CD-ROM, 2012.

12. 池田尚司，額賀信尾，三好利昇，柳瀬利彦，“ビッグデータ解析と社会イノベーションに向けた取り組み”， 2014 年人工知能学会全国大会，
https://kaigi.org/jsai/webprogram/2014/pdf/568.pdf，2014.

13. 池田尚司，古川直広，薦田憲久，“手書きによる情報共有に基づく設備保全作業支援システムの開発”，第 59 回電気学会情報システム研究会， I S-14-18, pp.29-33, 2014.

D. 解説等

1. 池田尚司，額賀信尾，小林義行，神田直之，渡邉裕樹，平山淳一，“非構造化データ利活用のためのメディア処理技術”，人工知能学会誌，Vol. 28, No. 1, pp. 114-121, 2012.

E. 書籍等

1. Hiromichi Fujisawa, Hisashi Ikeda, Naohiro Furukawa, Kosuke Konishi, and Shoichi Nakagami, “Information Just-in-Time: Going Beyond the Myth of Paperlessness”, in (Bidyut Baran Chaudhuri and Swapan Kumar Parui Eds.) Advances in Digital Document *Processing and Retrieval*, pp. 35-49, World Scientific, 2008.

内容梗概

本論文は，筆者が 1996 年から現在まで(株)日立製作所中央研究所，および，大阪大学大学院情報科学研究科マルチメディア工学専攻にて取り組んだ手書き情報利活用に関する研究成果をまとめたものである。

近年 PC やスマートフォンの普及によって，オフィスワークのみならず，プラントの保全などの現場業務もその多くが電子情報を介して行われるようになった。その一方で紙と手書きによる情報の閲覧や入力も未だに多くの場面で用いられている。情報の閲覧や表現，共有において，紙には筆記しやすいというメリットがある一方で，電子文書にはアクセスの場所，人数の制約がなく，再利用もしやすいというメリットがある。業務の生産性や効率向上を考えるには，電子か紙かという二者択一ではなく，それぞれのメリットに応じて紙と電子媒体を使い分けつつ，媒体に関わらず自由な情報の活用を可能とする環境を実現することが重要である。本論文では，人間が主に紙文書上に筆記した手書きの情報を取り込んで計算機上で扱えるような情報に変換，加工し，これを計算機上，あるいは再印刷した紙文書上で参照できるようにするとともに，この情報を用いて業務を効率的に行えるようにすることを手書きの利活用と定義し，その実現のために以下の三つの課題に着目する。

(1) 文字どうしが接触した手書き文字列からの文字切り出し方式

(2) 知識照合による文字列認識精度向上とそのための知識獲得方式

(3) 紙上の既存情報に付加される筆記(手書きアノテーション)の解釈方式と手書きによる情報共有の保全業務の現場への適用

上記の(1)と(2)は手書き文字列の認識に関わる課題であり，(3)は手書きアノテーションの解釈と再利用に関わる課題である。(1)に対して，計算機に取り込んだ手書き文字列の認識タスクにおいて，接触した文字を含む文字列を高い精度で認識するための接触箇所を含む連結成分の強制切断による文字候補パタンの生成と文字列認識方法について提案する。(2)に対して，認識対象の文字列の集合を知識(言語情報)として用意しておき，文字識別の誤りなどを訂正，補完することで文字列の認識精度を向上させる知識処理において，認識対象の文字列を文法として定義する統語的言語モデルを導入し，これを用いた探索的文字列認識方法と，統語的言語モデル構築法について提案する。(3)に対して，紙文書に引かれた線や追記された手書きの意味づけを行い，業務の現場において共有する手法を提案する。これらによって，人間が主に紙文書上に筆記した手書きの情報を計算機上で扱えるような情報に変換，加工し，これを計算機上，あるいは再印刷された紙文書上で参照できるよう

にするとともに，さらにこの情報を用いた業務を効率的に行えるようにすることが本論文の貢献である。本論文は全 5 章で構成される。

第 1 章では，日常の業務における手書きの重要性を示し，手書きの利活用を容易にすることの意義を述べる。また，手書きの利活用システムの全体像を示し，本論文で取り上げる三つの課題を位置付ける。

第 2 章では，日本語における文字の接触の仕方を分析し，その結果に基づき，文字領域のストローク成分の形状により接触箇所の推定を行い，連結部分を強制的に切断することで文字候補パタンを生成する手法について検討する。また，日本語の文字は「偏」や「旁」などの複数の要素から構成されるため，接触箇所の切断だけでは正しい文字候補パタンを得ることができない場合がある。これに対して，隣接する文字候補パタンをマージすることで，正しい文字候補パタンを生成する手法を提案する。これを地名表記文字列が手書きされた帳票サンプル画像を用いて評価し，提案手法により接触文字列の認識精度が向上するとともに，処理時間の増加も実用上問題のないレベルにとどまることが確認できた。

第 3 章では，地名表記や組織名などの固有名詞などにおいて，正式な表記に加えて書き手の違いや筆記領域の面積の制限から生じる省略など様々な異表記を再帰的遷移ネットワークによって記述する統語的言語モデルを提案する。この言語モデルを用いて，文字列認識結果の候補列を入力とする再帰的遷移ネットワークの探索による文字列照合方式を提案する。この際，文字切り出しや文字識別の誤りによって，文字候補パタンが欠落したり，余分な文字候補パタンが含まれてしまい，グラフ探索の途中で入力の文字候補パタンと言語モデルとの対応が取れなくなってしまう。こうした場合に対してもグラフの探索処理を進めるアドバンテージ遷移を導入することにより，文字列認識精度の向上を実現する。また再帰的遷移ネットワークと等価な表現力をもつ文脈自由言語によって認識対象の文字列の集合を表現し，これを再帰的遷移ネットワークに変換することで，人間にとって可視性が高く，構築と保守の容易な統語的言語モデル構築手段を実現する。この時，規則を用いて「東恋ヶ窪」が「東恋が窪」となるような異表記を自動的に加えることを可能にする。最後に異表記の少ない地域と，異表記のバリエーションがきわめて大きい地域を対象に，手書きの地名表記を含む帳票サンプル画像を用いて，提案手法を評価する。人手で異表記を追加した従来のトライ構造の言語モデルに対して，異表記を半自動的に追加した再帰的遷移ネットワークによる言語モデルでは特に異表記のバリエーションが大きい地域程，実際に書かれる地名表記のバリエーションに対する表現力が高く，また言語モデルが必要とするメモリ量も増大しにくいことが分かった。グラフ探索型文字列照合におけるアドバンテージ遷移の効果とともに，言語モデルの半自動構成手法における異表記の追加精度につ

いて，人手による継続的な改善作業を上回ることが確認できた。

第 4 章では，業務における手書き情報の共有について，手書きのアノテーションを再利用可能なように解釈し，電子文書と共に保持する紙･電子文書統合管理方式について述べる。実業務の例として設備保全業務を取り上げ，紙と手書きが多用されている業務における情報共有の課題について分析したところ，現場や事務所などでは緊急性の高いものから低いものまで多種多様の情報が報告されているものの，障害の予兆などのその時点では緊急性は低いが将来重要になる可能性の高い事項が時間の経過とともに他の情報に埋もれて共有されにくいことが分かった。この問題に対して，現場で筆記した手書きの意味をその紙に印刷されていた電子文書の情報を用いて解釈し，関連する項目に関する報告の履歴を自動的にまとめて保持することで情報の再利用と共有を容易にする，紙・電子文書統合管理方式について提案する。さらに，デジタルペンを用いて提案した紙・電子文書統合管理方式を実装し，プラント等の設備保全業務，特に現場での情報入力の効率向上と保全担当者間の情報共有の確実化を目指して，設備保全作業支援システムを開発した。このシステムを筆者が所属する研究所内の廃水処理プラント保全業務に適用したところ，代替策を講じて稼働を維持したり，次の障害発生まで様子を見ると言った場合に対して，運転日誌上に印刷して現場保全担当者の目に触れるようにしておくことで効果があることが保全担当者へのインタビューにより確認できた。

第 5 章では，手書き利活用を実現するための機能として，文字列認識における文字切り出しと知識照合，アノテーション解釈とアノテーション統合紙文書化が実現できることを述べ，今後の研究の方向性に関して議論する。

目次

第1章

序論

1.1. 研究の背景

手書きは人間にとって重要な情報の表現手段として広く使われてきた。紀元前に古代エジプトから広まったパピルスや，ヨーロッパを中心に普及した羊皮紙に文字が筆記され，物語や宗教文書として流布している[1]。日本でまだ紙が普及していない奈良時代以前にも，木簡に文字を筆記することで行政関連の通達や報告が行われており[2]，この後，平安時代以降，現代に至るまで，行政文書だけではなく，文学作品や手紙など，多くの情報が紙に手書きされ，文書として保存されている。

1970 年代以降，ワードプロセッサ，その後 PC (Personal computer)が普及し，多くの文書がキータイプで作成されるようになり，手書きの重要性が減少するのではと言われた。これに対して Sellen らは，文書の精読時に手書きにより注釈を加えるといった，読む行為と書く行為を同時に行える点を紙の利点として挙げており，思考・判断，協調作業の手段，およびこうした作業の記録手段としての手書きの役割はなくならないと主張している[3]。21 世紀を迎えた現在，行政文書や企業で使用される公式の文書はほとんどが電子文書として作成されるようになったが，まだ手書きが使われている場面がある。一つは自身の考えを整理したり，相手に考えや情報を正確に伝達するための手書きである。これは主に情報を表現し，読み手に伝えるという目的で行われている。もう一つは教育の現場において思考能力を高めるために手書きを行う場合である[4]。本論文では前者の手書きを対象とする。

前者の手書きの例として，医療や保全などの現場や，自治体や金融機関，さらには交通機関や運送業者の窓口などで，情報の伝達や各種の手続きを正確に行うために筆記される，住所や氏名，会社名，製品名，患者や設備の状態などを含む手書きがある。特に医療や保全，営業などの現場業務において，複数の人が協調して作業する際の情報の共有手段として手書きが重用されている。ある人が作業中に気付いたことを紙に書き留め，これを他の作業者にメモ書きとして渡すことで，口頭で伝えるよりも確実に情報を共有することができる。他の例として，会議などで自分の意見を正確に相手に伝えるために，説明をしながらホワイトボード等に手書きする文字や図がある。さらに，講演などを聞きながら，気付いたことをノートに手書きすることで理解が進むことが多い。一人で熟考する際も，手元の紙に考えの断片を書き出しながら，これを線で結ぶなど構造化することで，考えが整理，

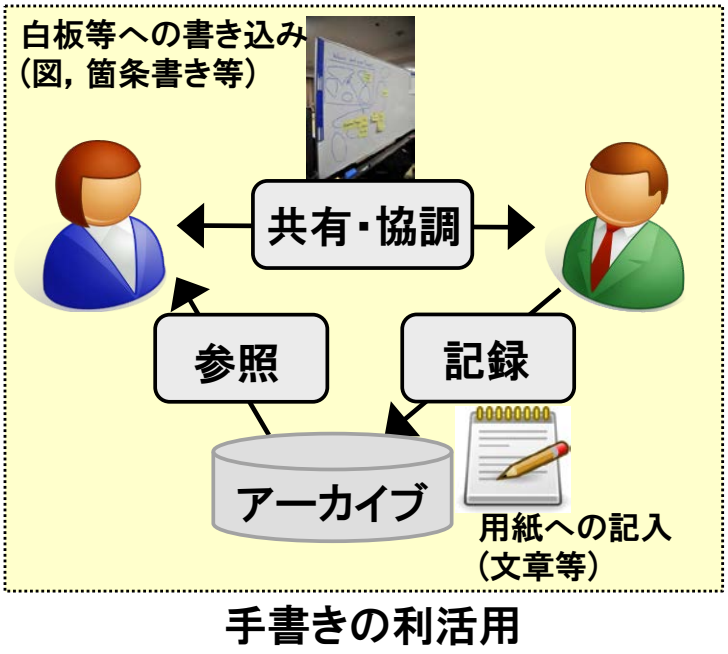


図 1.1 手書きの果たす役割と本研究における位置付け

具体化されていく。

図 1.1 に示すように，手書きは人間が情報を記録，参照するとともに，これを共有することにより協調の手段となっている。但し，これらはいずれも紙に書かれており，再利用が困難であった。ICT 技術によって，記録した情報の再利用を容易にするとともに，情報の共有を効率化することが期待されている。

図 1.1 に示すような手書きの利活用を ICT 技術により実現するためには様々な課題を解決する必要があるが, 本研究では特に以下に示す手書きの認識に関する課題に着目する。手書きの認識とは筆記者の意図を計算機により自動的に解釈するタスクであり，その認識された意図に応じて再利用や共有サービスが実現されるため，手書きの利活用に向けた最も重要な課題の一つと言える。

手書きされた情報を再利用可能な形態で保存するために，OCR（Optical Character Recognition）によって文字を計算機上で処理しやすいテキストコードに変換する手段が提案されてきた。これまでに，金融機関の入出金帳票の金額や電話番号といった数字について OCR による手書き文字の数値データ化が実用化されている。一般的に紙文書をスキャナ等により画像データに変換し, そこから文字行領域を抽出する。文字行画像に対して，筆跡成分の連続性に基づき一文字毎の文字領域画像に分割し，これに対して文字識別処理

を行うことで，文字毎の認識結果を得る。しかし，筆記するスペースの制限や連続した動作で筆記されるため，手書きされた文字は互いに接触してしまうことが多い。この場合，文字の境界をまたいで筆跡成分が連結してしまうため，文字行画像から一文字ごとの文字領域画像を正しく抽出することができない。このように，文字の接触による切り出し失敗は文字識別の誤りと並び，文字列認識における誤読や不読の大きな原因となっている。文字どうしが接触している場合であっても，正しく文字領域を切出し，認識結果を得ることが第一の課題である。

さらに，申請書に記載される住所や商品名，会社名等，あるいは保全記録票に記入される機器の状態など，帳票には数値以外の文字が記入される場合が多い。漢字やひらがな，カタカナを対象に含めると，認識対象の文字種が八千以上になり，ここに人による筆記のくせや筆記スペースの制約などによる省略などの異表記が加わるため，認識対象の文字の多様性はさらに増大する。これにより文字列の認識精度は低下してしまう。これまでに認識対象の単語の集合を知識として用意しておき，文字認識結果と照合することで誤認識を修正し，単語認識精度を向上させる手法が提案されてきた。ここで用いる認識対象の文字列に対する知識のことを言語情報と呼ぶ。しかし，複数の単語からなる文字列の認識については，単語に関する知識との照合だけでは十分な精度を得ることが困難である。これに対して，認識対象の文字列を表現する言語情報が提案されてきたが，異表記を網羅した言語知識を一般的な情報システムのリソース上で実行可能な形態で実現することは困難になってきた。手書き文字列の認識精度向上に十分であり，メモリ量を抑えた言語情報の導入とそれを用いた高速な文字列認識の実現，ならびに言語情報の構築の簡易化が第二の課題である。

記録される手書きは文字だけではなく，下線や図など文字以外である場合も多い。病棟や手術室における患者の状態の管理，店舗や倉庫における在庫管理，プラントや工場における運用，保全業務といったオフィスを離れた現場作業では，チーム内の他の作業者に短時間で正確に情報を伝達する必要がある。このため，全ての情報を文字だけで表現するよりも，印刷された文章や図面など既に存在する情報に対して，重要だと思う部分に下線を引いたり，関連があると思われる文を引き出し線で結んだり，あるいは，文書中のある部分から引き出し線を書いて余白に注釈を加えることが多い。これは筆記者の意図やその文章の理解内容を正確に，かつ，効率的に伝達することができるからである。本論文ではこうした紙上の既存情報に付加される筆記を手書きアノテーションと呼ぶ。計算機を介してこうした情報を他の作業者に伝えることで協調作業を支援したり，アーカイブして再利用するためには，単に下線の存在を認識したり，余白に書かれた文字列を認識するだけでは

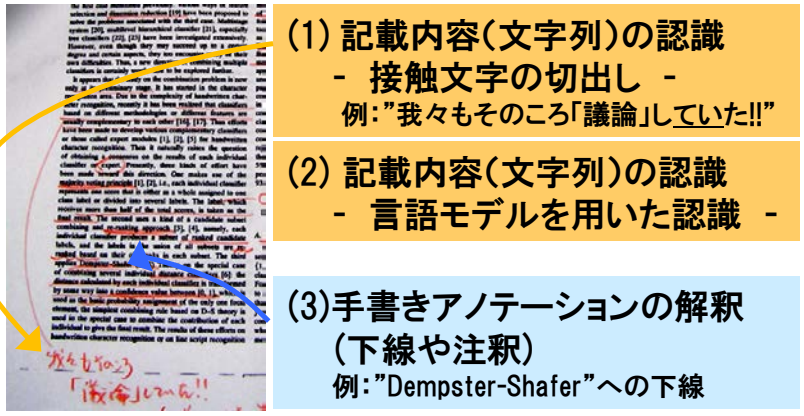


図 1.2 本研究の課題

不十分である。紙上に印刷されていたどの情報に対して下線が引かれたのか，どの部分に対してコメントがなされたのかを含めて解釈する必要がある。すなわち，計算機による手書きアノテーションの解釈により協調作業，あるいは手書きアーカイブの再利用の支援を実現することが第三の課題である。

本研究の目的は，手書き情報の利活用の手段として，図 1.2 に示す(1)記載内容，すなわち文字列の認識，特に文字が互いに接触している場合であっても認識可能な手法，(2)認識対象の文字列に対する知識を言語情報として利用した文字列の認識精度向上手法，および，(3)手書きアノテーションの解釈による手書き情報の共有の容易化手法，に関する提案を行うことである。

1.2. 関連研究

1.2.1. 手書き接触文字の切出し

文字列の認識は，(a)入力となる文字行の画像から文字パタンの候補を切出し(文字切出し)，(b)文字パタンの候補に対して識別を行い(文字識別)，(c) 得られた文字識別の候補から，言語情報を用いて文字列の認識結果を選択する(知識照合)という 3 つのステップによって実現されるのが一般的である[5]。基本的にこれらの 3 つのステップは順に実行されるため，文字パタンの候補を正しく切出すことは，最終的な文字認識列の認識処理に大きく影響を与える。紙文書をスキャナ等で画像情報として採取し，これに対してカラー処理やノイズ除去等の前処理を行うことで，文字列の筆跡部分を黒，背景部分を白とする二値画

像が得られる。これが文字列の認識処理に対する入力となる。これを文字行画像に分割し，文字行画像から文字パタンの候補を切出す。文字と文字は空白によって区切られているため，入力の二値画像から黒画素が連なった領域である連結成分を抽出し，これを組み合わせることで文字パタンの候補を生成することができる[6]。

ここで，文字どうしが接触している場合を考える。ペンデバイスを用いるなど手書き情報が時系列の筆跡データとして取得できる場合は，文字の接触がある場合でも，時間情報を用いることで接触しているストロークは分離可能である[7]。しかし手書き情報を静止画像として取得する場合は，複数の文字の構成要素が同じ連結成分に含まれてしまうため，文字を切出すことができない。これに対して，Baird[8]，Lu[9]や Ohta[10]らは手書きの英ブロック体や印刷英字に対して，文字行画像の各画素を行方向(すなわち，横書きの場合は縦方向に，縦書きの場合は横方向)に射影をとり，射影された画素の数が相対的に少ない箇所を文字の境界の候補とすることで接触文字を分離する手法を提案している。筆記体の英文字列では必ずしも行方向の射影で文字の境界が検出できず, Paul らは筆跡画像の輪郭解析により文字境界候補を推定する手法を提案している[11]。また，Shridhar らは横書きの文字行画像を行方向に走査し黒画素の最上部と最下部を算出し，その差分が小さい範囲を文字の境界候補とし，その領域に対して輪郭解析をすることで，文字の境界候補を検出する手法を提案した[12]。英筆記体の場合は，基本的に接触箇所は 1 か所であることが多いのに対して，日本語，特に漢字の場合は複数個所で接触する可能性があり，この手法でも文字の境界候補を精度よく検出することは困難である。

認識対象の文字列集合から出現する可能性のある文字列を限定した上で文字識別を適用し，認識対象の文字列として認識されなかった部分を文字の接触が起こっている可能性があるとして，行方向に切断位置をスライドさせながら文字認識処理を適用し，認識が成功した位置が文字の境界であるとする方法が仲林らによって提案されている[13]。認識対象の文字列が関東地方の市町村名 14,000 と比較的少数に限定されている条件で有効に機能している。しかし，認識対象の文字列数が多い場合は，接触を含むパタンや誤切断したパタンに対して誤った文字がマッチしてしまう可能性が高くなるという問題がある。

文字を切出すというアプローチとは別に，Gilloux らは HMM (Hidden Markov Model) を用いて，文字を切り出さずに英筆記体の単語全体を認識する Holistic アプローチを提案した[14]。しかし，英文字が 26 (大文字，小文字を区別しても 52)に限定されているのに対して，日本語の漢字，ひらがなは 4 千字種以上になり，これらの HMM を学習するためには，大量の学習サンプルが必要になるという問題がある。

文字間の境界を文字識別処理によって推定する，あるいは文字間の境界を考慮せず単語

として認識するアプローチでは，文字パタンの候補を精度よく生成することは困難であることが明らかになってきたため，Casey らは連結成分の組み合わせで文字パタンの候補を複数生成し，これを文字識別処理で検証することにより正しい文字の境界候補を絞り込む，Recognition-based Segmentation と呼ぶ手法を提案した[15]。この方法では文字パタン候補を漏れなく生成することが重要であり，Fujisawa らによって，文字パタン候補の相対的な大きさなどに基づき生成する文字パタン候補の数を削減することで，文字パタン候補の可能性がある連結成分の組み合わせをできるだけ多く出力する，Over-segmentation と呼ぶ手法が提案された[16]。

文字どうしの接触を含む日本語の手書き文字列を認識するには，Over-segmentation による Recognition-based segmentation の手法において，接触している文字画像からこれを強制的に切断し，文字パタン候補を生成する手法が必要となる。

1.2.2. 言語情報を用いた文字列認識

手書きされた英文字列「mn」が「mn」なのか「nm」なのかを，文字パタンだけから判断するのは難しい。また，偏とつくりがそれぞれ別の文字となりうる漢字を多くもつ日本語においては，手書きの「柿」が「柿」なのか「木，市」なのかを判断することは困難な場合がある。さらに，個々の文字パタンの識別においても，英筆記体の「e」と「l」や，手書きの仮名文字「り」と「リ」を計算機が正しく識別することは困難である。このように，前項で述べた(a)の文字切出しや(b)の文字識別においては，入力の画像情報のみから文字パタンや文字識別結果を一意に得ることは現実的ではない。従って，文字切出しや文字識別の結果の候補を複数保持しておき，これに対して(c)の知識照合において，正しい候補を選択したり，あるいは正しい識別候補がない場合は，言語情報をもとに識別誤りを訂正し，候補を補完し，文字列の認識結果を得るという多重仮説生成型の認識手法が提案されてきた[6][15][16]。

このような知識照合に用いられる言語情報のうち，認識対象の文字列のインスタンスを単純に列挙するのではなく，構造化して保持したり，あるいは，何らかの数式や規則により認識対象の文字列を表現したものを言語モデルと呼ぶ。言語モデルは，記述型と統計型の 2 種類に分類することができる[5]。記述型とは，認識対象の文字列を何らかの方法で列挙する方式である。単語認識に対しては，対象となる単語のリストを言語情報として用いて，知識処理を実現する場合が多い。Kimura らは英筆記体で書かれた地名単語の認識に[17]，Srihari[18]や Mahadevan[19]，Chen[20]は，英語の都市名や通り名の認識にこの

形態の言語情報を適用し，Weinman らは標識上の英単語の認識に適用している[21]。

しかし，単純に列挙する方法では，言語知識が大規模になると知識照合の処理時間が問題になる。照合処理を高速化するために，対象文字列中で用いられる語彙の集合を順序付きの木構造であるトライ構造[22]で実装した言語モデルが提案されている。Zhu らは，日本語の病名に関する単語の認識をトライ構造の言語情報を用いることで行っている[23]。このように，記述型の言語情報の利用は，認識対象の単語集合が明確な場合等，認識対象の文字列が漏れなく収集できる時は，認識精度の向上に有効である。

しかし，複数の単語が並ぶ文の認識を行おうとすると，言語知識として認識対象の文字列を列挙することは困難になる。単語認識を通り名，都市名を表す単語ごとに順次実行することで，英語の地名表記を認識する手法が Malburg[24]や Dengel[25]らによって報告されているが，これは英語の単語が分かち書きをされていること，都市名と通り名の連接で地名表記が構成されるという前提条件を用いたことによる。一方，日本語や中国語の文は単語が分かち書きされていないため，文字行の画像から単語の境界を同定することは容易ではない。単語の境界の同定が比較的容易な印刷活字の日本語に対しては，トライ構造の言語情報を用いて精度向上を行う手法が伊東らによって提案されているが[26]，手書きの日本語や中国語の文の認識に Malburg の手法が適用できるかは，計算量やメモリリソースの点で明らかになっていない。

認識対象を列挙しきれないという問題に対して，大量のサンプルから算出した文字や単語の頻度情報に基づき，単語どうしの隣接しやすさを確率として表わす，統計型の言語情報が提案されている。文字列認識の場合は，着目する単語の一つ前から n 個前までの単語の並びが与えられた際に，着目する単語の現れやすさを表す n-gram[27]を用いることが一般的である。Marti ら[28]，あるいは Vinciarelli ら[29]は，HMM（隠れマルコフモデル）による単語認識に，n-gram による言語モデルを適用し，英文文字列認識精度向上に効果があることを示している。また Wuthrich らは，n-gram による言語モデルを用いて，中世の英文書の文認識を行っている[30]。さらに，現在普及しているオープンソースの OCR ソフトである OCRopus においても，n-gram に基づく言語モデルを用いて手書きの英文の認識が実装されている[31]。日本語の文に対しては，永田が単語長の確率と単語表記の確率をもとに統計的言語モデルを構築し，認識精度の向上に効果があることを報告している[32]。統計モデルに基づく言語情報では，認識対象の全てのインスタンスを列挙することなく，出現頻度に基づく「認識対象らしさ」を評価することができ，これにより平均的に文字列認識精度を向上させることが可能である。一方で，統計モデルの作成には，大量のサンプルデータが必要である。収集したサンプルデータにおいて頻度が小さい表記につ

いては，言語モデル内の統計値も小さくなり，知識照合される可能性が低くなるという問題がある。

このように，記述型の言語情報は，そこに含まれる文字列に対しては，確実に知識照合の効果が得られるのに対して，認識対象として列挙できなければ知識照合が機能しないという問題がある。統計型の言語モデルは，認識対象を網羅的に収集できない場合でも，サンプルを大量に収集することで，知識照合の効果をある程度得ることができるが，収集したサンプルの中に含まれない対象については,知識照合の効果を得にくいという問題ある。

紙に手書きされた複数の単語から成る日本語文字列の認識においては，その認識対象は必ずしもすべてを列挙できるものではなく，また言語モデル中の統計情報によって認識率に大きく差がつくものであってもならない。文字の接触などもあり単語の境界の同定が困難な手書き日本語文字列に対しては，トライ型に代表される記述型の言語モデルを計算量やリソースの観点で適用可能にすることが必要である。

1.2.3. 手書きアノテーションの解釈と共有

複数の人間による正確，かつ効率的な情報共有のための手書きアノテーションの解釈に関する研究は，まずタブレット等の電子機器の登場にともない，そこに表示される web ページや電子文書を対象に，アノテーションを電子文書にどう対応付けるかという観点で始まった。

Schilit，Golovchinksy, Price は，文書の精読を支援し，その結果を電子化することで再利用可能にする，XLibris と呼ぶシステムを実現している[33][34][35]。タブレット上に表示された電子テキストに対して，手書きにより修正やコメントを書き加えることが可能になっている。伊東らは,複数の人間がタブレット上で研究論文へのアノテーションを行い，これを分析することで関連文献の推薦や同様の興味を持つ研究者を抽出するといった協調作業支援を可能にしている[36]。Bargeron, Brush, Shilman は，単にディスプレイ上に筆記したストロークが電子化されるだけではなく，表示された電子テキストとの論理的対応関係を保持して管理できるシステムを実現している。これにより，別のディスプレイに表示される場合など，描画領域サイズや改行位置などの表示環境が変わっても，電子テキストと修正やコメントとの位置的対応関係を保って手書きデータを表示できるようにしている[37][38][39]。大賀らはタブレット PC 上で入力した手書きアノテーションを文字認識によりテキスト化し，ストロークデータとともにディスプレイに表示された電子テキストにリンク付けする手法を提案している[40][41]。また，石井，鈴木らはタブレット上でアノ

テーションが付けられた文字列をクエリとして検索アプリケーションを起動する手法を提案している[42][43][44]。

筆記データを電子的に取得できるペン型デバイスが利用可能になり，これを用いて紙上に筆記したアノテーションを解釈することで，紙文書上の情報と電子文書上の情報を対応付ける研究が行われている[45]。Guimbretièreは，Anoto社のデジタルペン[1]を用いて紙に筆記したアノテーションを電子化し，もとの電子文書に重ねてディスプレイ上に表示するための手法を提案している[46]。アプリケーションの種類を限定しているものの，紙上での校正作業を電子文書に反映させることを可能にしている。Liaoらは，これを応用して，紙を用いた立体工作を例に，接合面等を紙上にアノテーションとして筆記することで，計算機上で三次元形状を生成するシステムを開発している[47]。

デジタルペンを用いた紙上のアノテーションの解釈に関する研究では，紙上の手書きアノテーションを元の電子文書と関連付けて電子化することは実現されており，紙上の情報を計算機上で処理，閲覧することは可能である。しかし，特にオフィスを離れて行う現場作業においては，携帯性や耐久性の観点からタブレットのような電子機器を用いることができない場合も多く，そこでは紙を介して情報の共有を行う必要がある。

Yeh[48]や Weibel[49]らは生物学者などの屋外での観察調査の際に用いるフィールドノートにデジタルペンを適用している。現場で撮影した写真とノートへの手書き情報を紐づけて計算機上に保存するなど，ノート上の情報の計算機上での再利用性には優れた手法を提案している。しかし，例えば計算機上に取り込んだ情報を印刷した場合に，印刷した文書と計算機上の文書との紐付けを行うことができないため，印刷した文書に再度手書き情報を追記したとしても，その追記事項は計算機上の文書に反映されることはない。さらに，基本的には現場の情報を計算機に取り込み，これを計算機上で活用することに焦点を当てており，現場へ動的に情報提供を行うことについては考慮されていない。

このように，現場作業においては，一旦計算機上に入力した手書きアノテーションに関する情報を，再度紙媒体を介してユーザに提示するとともに，その紙上で筆記したアノテーションを再び解釈して計算機上に取り込むという，紙媒体と電子媒体との間の情報のサイクルを実現する必要があるが，まだ実現されていない。

1.3. 研究の方針

前節までの議論を踏まえ，本論文で想定する手書きの利活用を「人間が主に紙文書上に

[1] http://www.anoto.com

筆記した手書きの情報を計算機上で扱えるような情報に変換，加工し，これを計算機上，あるいは再印刷された紙文書上で参照できるようにする，さらにこの情報を用いた業務を効率的に行えるようにすること」と定義する[50][51]。これを具体化したシステムの例を図 1.3 に示す。

紙文書上に筆記された手書きには, 文字列や図や絵, さらには下線や引き出し線など様々な種類がある。これらを手書きアノテーションと総称する。図 1.3 のアノテーション解釈機能において，当該紙文書にすでに印刷されていた電子文書のコンテンツ等の，関連する情報と手書きアノテーションとの対応関係を解析し，それぞれの手書き情報が持つ意味を算出し，これを含めてアノテーションデータとして保持する。この時，手書きの文字列に対しては，文字列認識によってテキストコード列に変換してアノテーションデータとして保持する [2]。意味付けされたアノテーションデータを手掛かりに筆記データと元の紙上に印刷されていた電子文書データを検索し，計算機上で参照，編集を可能としたり，筆記データと電子データとを紙文書として再印刷することで人間が参照，さらに手書きを行うことを可能とする。本論文にて述べる研究内容は，図 1.3 の吹き出し中に示すように位置づけられる。

(1) 文字どうしが接触した手書き文字列からの文字切り出し方式

文字どうしの接触を含む日本語の手書き文字列を高精度に認識するために，連結成分の組み合わせにより文字の可能性のあるパタンをできるだけもれなく生成し，その文字パタンの候補に対して文字識別処理を行うことで妥当な文字パタンの候補のみを残そうとする over-segmentation による Recognition-based segmentation の手法を拡張する。文字の接触を含むと想定される連結成分を抽出し，接触箇所を強制的に切断することで, 接触した文字どうしを分離する。日本語，特に漢字は「偏」や「旁」といった複数の要素によって構成されており，連結成分を切断するだけでは，必ずしも正しい文字パタンの候補を得ることができるとは限らない。これに対して，連結成分の切断だけではなく，隣接連結成分との統合を合わせて行うことで文字パタンの候補をより精度よく生成できるようにする。

[2] 認識した手書き文字列を電子文書中に反映することも考えられる。

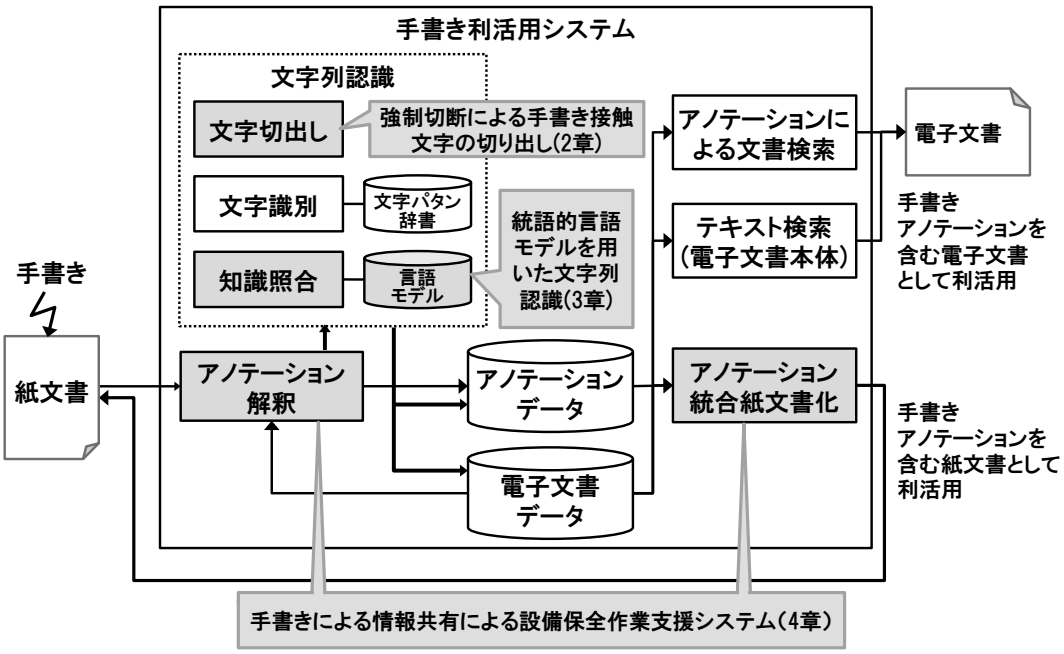


図 1.3 手書き利活用システムにおける本研究の位置づけ

(2) 知識照合による文字列認識精度向上とそのための知識獲得方式

認識できなかった文字や認識を誤った文字を補完，修正することで，文字列認識精度を向上させることを目的として，認識対象の文字列に関する知識を統語的な言語モデルとして保持し，これを用いて認識対象文字列の照合処理を行う手法を提案する。例えば地名表記や人名，会社名などの固有名詞は多くの場合表記にゆれが生じ，多くの異表記が派生することになる。トライ型の言語モデルで異表記を表現しようとしてもそのサイズが大きく実用的な時間で認識処理を行うことができない。これに対して，多層型のグラフ構造による言語モデルとこれを用いた文字列照合処理手法について提案する。この時，文字列認識の失敗や誤りにより生じる，文字列認識結果の不完全を考慮した照合アルゴリズムにより高精度な文字列認識を実現する。さらに，提案する多層型グラフと文脈自由文法の表現力の等価性を利用し，認識対象文字列の集合を文脈自由文法により記述し，これを言語モデルに変換する手段を実現する。そこでは，あらかじめ意図した汎用性のある異表記に関するルールに基づき，異表記も含めた認識対象文字列を表現する言語モデルの半自動的な生成を可能にすることで，実際の応用場面においても容易な運用を可能にする。

(3) アノテーション解釈方式と手書きによる情報共有の保全業務の現場への適用

筆記された文字列を認識しテキストコード化する以外の手書きの利活用の方法として，紙文書に引かれた線や追記された手書きの意味づけを行い，これを業務の現場において共有する手法を提案する。紙面上での手書きの位置と，印刷された情報との位置関係から，印刷されたどの情報に対する手書き(注釈)なのかを推定し，手書き情報の意味を補足する。

解釈した手書きアノテーションをデータベースにて管理することで，必要なアノテーションを抽出して計算機上で参照したり，これを再度紙文書として印刷し手書きアノテーションを共有する手法を提案する。さらにこれを保全業務の現場に適用し，複数の担当者が他者の書いた手書きアノテーションを共有することにより，業務を効率化できることを評価する。

1.4. 論文の構成

本論文では，第 2 章以降を以下の通り構成する。

第 2 章では，文献[52][53][54]に基づき，接触した文字を含む文字列を認識するための文字行画像の強制切断による文字パタンの候補(文字候補パタンと呼ぶ)の生成と，文字列認識方法について述べる。まず文字の接触の仕方の分析に基づき，文字領域成分の画像の形状により接触箇所の推定を行い，この部分の画素を強制的に削除することで分離された文字候補パタンを生成する手法について述べる。既存の文字切り出し手法によって生成された文字候補パタンの並びを切り出し仮説グラフとして表現する。提案手法によって切断された接触文字の候補パタンを切り出し仮説ネットワークに追加するだけでなく，ネットワーク上で隣接する文字候補パタンをマージすることで，正しい文字候補パタンを生成する手法を提案する。次にこれを地名表記文字列が手書きされた帳票サンプル画像を用いて評価し，提案手法の有効性を確認する。

第 3 章では，文献[55][56][57]に基づき，異表記を含む言語モデルを用いた探索的文字列認識を実現するための統語的言語モデルとその構築法について述べる。まず，異表記を含む認識対象の文字列を効率的に表現できる多層型のグラフ型言語モデルについて述べる。次に，この言語モデルを用いた手書きの文字列を探索的照合手法について提案する。さらに，このグラフ型言語モデルと表現力において等価な文脈自由文法で記述された認識対象の文字列集合から，グラフ型の言語モデルを半自動的に構築する方法について提案する。表記のゆれが大きい地名表記を例に，提案する統語的言語モデルとこれを用いた文字列照

合文字列認識精度を評価し，提案手法の有効性を確認する。

第 4 章では，文献[58][59]に基づき，紙･電子文書統合管理方式と実業務への適用を議論する。実業務の例として設備保全業務を取り上げ，その課題について整理する。次に，紙・電子文書統合管理方式について提案し，これを用いた手書き情報の共有による設備保全作業支援方式について述べる。これを実際のプラント保全業務に適用し，手書き情報の効果について実験的に評価し，その効果を確認する。

第 5 章では，結論として本研究で得られた成果をまとめたのち，今後に残された課題について議論する。

第2章

強制切断による手書き接触漢字列認識

2.1. 緒言

手書き帳票の読み取り処理は一般的に次の手順で文字列の認識結果を得る[5]。(1)入力画像から文字行を抽出し，(2)文字行から個々の文字パタンの候補を切出し，(3)個々の文字パタンの候補に対して文字識別処理を行い，(4)得られた文字識別結果に対して予め用意した読取り対象の単語辞書などの事前知識を用いて，識別処理の誤りを訂正する。上記(2)の文字切出し処理においては，連結成分と呼ぶ前景画素の連続した領域を単位として，これらを組み合わせて一つの文字パタンの候補を生成する手法が一般的である[5]。例えば,「切」という漢字が書かれた画像からは,「七」と「刀」に相当するという連結成分が抽出される。これらの連結成分を相対的な位置関係や大きさなどに基づき「文字らしさ」を考慮しながら組み合わせることで,「七」,「刀」,「切」に相当する画像が得られる。本論文ではこの画像のことを文字候補パタンと呼ぶ。連結成分から文字候補パタンを生成する際，正しい文字候補パタンが漏れないように，連結成分の組合せを十分に多く生成する。しかし隣接する文字どうしが互いに接触している場合は複数の文字の構成要素が単一の連結成分に含まれることになり，この手法では文字を正しく切出すことが出来ない。第 1 章で述べたように，本論文で対象とする，医療や保全現場，あるいは公共機関窓口のように，筆記にかけられる時間や紙上のスペースに制限がある状況で記入された手書き文字列は文字どうしが接触することが多い。このため接触した文字を切出すことが出来なければ，文字識別以降の処理が正しく行われず，文字列の読み取り率を低下させてしまう。

本章では，手書き文字列において複数の文字が接触している場合に，一つの連結成分中で文字が接触していると推定される部分を強制的に切断して，正しい文字候補パタンを得る方式を提案する。アルファベットや数字の場合は文字の接触箇所を含む連結成分は大きくなる傾向にあるため，行方向に相対的に長い連結成分に接触箇所が含まれると仮定し，これを切断することで接触した文字を分離することができる[9]。一方,「偏」や「旁」等の相対的に小さい複数の構成要素からなる漢字の場合，接触箇所を含む連結成分は必ずしも大きくなるとは限らない。ここで小さい連結成分に含まれる接触箇所を切断しようと閾値となる連結成分の行方向の長さを短くすると，多数の切断が行われ文字候補パタン数が増加してしまい，文字識別処理以降の処理時間が増大してしまうという問題が生じる。こ

のように日本語文字列から接触箇所を含む連結成分だけをもれなく取り出すことは難しい。

そこで，(1)連結成分を単位として生成した文字候補パタンに対してその大きさに基づき接触可能性を判定し，(2)接触箇所を含む可能性があると判断された文字候補パタンに対してのみ，これを構成する連結成分の形状に基づき接触の有無を推定し，接触箇所を含むと推定された連結成分を切断し，切断された連結成分を含む文字候補パタンを生成するという二段階のアプローチを提案する。これにより，連結成分の不必要な切断を抑えながら，接触箇所を含む連結成分が相対的に小さい場合でも漏れなく切出すことを可能にする。地名表記文字列を対象に提案手法を評価し，有効性を検証する。

2.2. 手書き文字の接触と切出し処理のアプローチ

2.2.1. 文字切出し処理の概要

自由に書かれた手書き文字は文字ごとの大きさが一様でなく，また同じ種類の文字であってもその変形が大きいため，文字行から個々の文字を切出すことは困難である。

手書き文字列における文字切出しにおける有効な方法に，文字候補パタンに基づく切出し（Pattern-Oriented Segmentation）[16]と呼ばれる方法がある。文字候補パタンに基づく切出しでは，連結成分を組合せることにより正しい文字パタンが含まれるように文字候補パタンを生成する。文字行画像の先頭から末尾まで続く文字候補パタンの並びが文字切出しの候補である。この候補のことを文字切出し仮説と呼ぶ。これらの文字切出し仮説をグラフによって表現する。本論文ではこのグラフを「切出し仮説グラフ（Segmentation Hypothesis Graph）」と呼ぶ。図 2.1 に示すような漢字文字列に対する切出し仮説グラフは図 2.2 のようになる。図 2.2 においてエッジが文字候補パタンを表し，ノードがパタンの境界を示す。グラフの左端のノードから右端のノードに至るパスがそれぞれの切出し仮説を表す。ここで生成された文字候補パタンにはそれぞれ文字識別処理が行われ，その文字候補パタンの「文字らしさ」が評価される。そしてもっとも「文字らしい」と評価された文字候補パタンの列が文字切出し結果として選択される[60]。

しかし，連結成分を最小単位としているため，文字どうしが接触している場合は，これらを分離することは出来ない。そこで接触箇所を見つけて，連結成分を切断し正しい文字候補パタンを生成する必要がある。接触したアルファベットやアラビア数字の文字列は，行方向に長大な連結成分を見つけ，この連結成分を切断することにより切出すことが出来る。長大な連結成分における切断箇所は行方向に垂直な周辺分布形状を分析したり[61][62][63]，輪郭を抽出しストロークの交差している場所を見つけることにより[64][65]

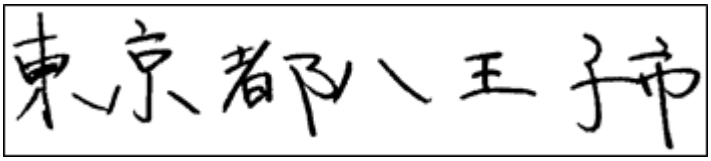


図 2.1　漢字文字列の例

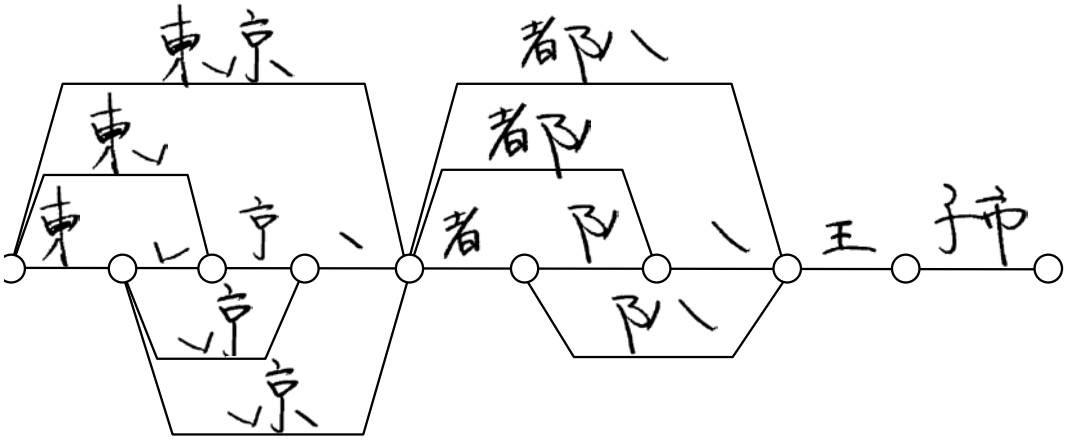


図 2.2　切出し仮説グラフの例

推定できる。また，Nishiwaki らはニューラルネットワークを用いて接触文字の切り出しを行う手法を提案した[66]。

一方，日本語の文字(特に漢字)は，「偏」や「旁」，あるいは「点」といった複数の構成要素からなるため，英数字と同様の方法では，接触文字を同程度の精度で切出すことは困難である。図 2.3 は図 2.1 の文字列の接触箇所を明示したものである。「東京」における二つのはらいのように，接触箇所を含む連結成分は，他の接触を含まない連結成分と比較してその大きさが小さい場合がある。連結成分の大きさで切断の是非を判断し，かつ小さな連結成分までも切断しようとすると，文字行中の多くの連結成分に対して切断処理を行うことなり，文字候補パタン数が増えてしまい，文字切出しの精度が低下するとともに，処理時間が増大してしまう。さらに，漢字においては，接触箇所を含む連結成分を切断して生成された文字候補パタンは正しい文字候補パタンの一部分である可能性がある。正しい文字候補パタンを得るためには,連結成分の切断後,隣接した文字候補パタンを統合し,

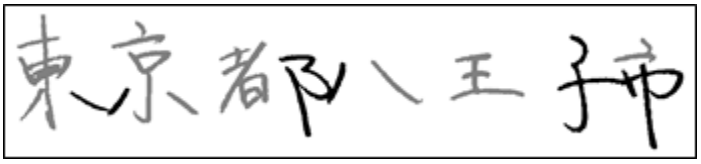


図 2.3　図 2.1 の文字列の接触箇所(濃色が接触箇所を含む連結成分)

新たな文字候補パタンを生成しなければならない。例えば,「川越」の「川」の右端のストロークと「越」が接触している場合，接触箇所を切断した後,「川」の右端のストロークと残りの 2 本のストロークを統合しなければ「川」に対する文字候補パタンが生成されない。この問題は，特に「川」が大きく書かれている場合に生じやすい。

2.2.2.　接触文字切出しのアプローチ

前節で述べたような問題点を解決するために，大きな連結成分を接触文字とみなして切断するのではなく，連結成分を組合せて作成した文字候補パタンに対して接触文字かどうかの判定を行ない，そのパタンの中の連結成分を切断することにより，接触文字の切出しを行う手法を提案する。

提案する接触文字切出し手法は以下の 5 つのステップからなる(図 2.4 参照)。

(1) 連結成分を組み合わせて生成された文字候補パタンから接触を含むパタンを同定

(2) 接触したと推定された文字候補パタンにおいて，接触箇所を含む連結成分を同定

(3) 接触箇所を含む連結成分を切断・分離

(4) 切断・分離により生成された連結成分から新たな文字候補パタンを生成（これを接触を含むと判断されなくなるまで再帰的に繰り返す）

(5) 新たに生成した文字候補パタンとその隣接した文字候補パタンとを統合し新たな文字パタン候補を生成

このように，文字候補パタンを接触文字と判断すれば「切断」し，切断した文字候補パタンを隣接する文字候補パタンと「統合」することにより，接触文字の切出しを行う。この手法を「切断統合法（cut-and-merge method)」と呼ぶ。

連結成分の切断に対しては，文字の接触のしかたを分析した結果に基づいて複数の手法を用意する。接触箇所を含む連結成分にこれら手法を適用して切断処理を行い，必要であれば複数文字候補パタンを生成する。生成される文字候補パタンの数を抑制するしくみを

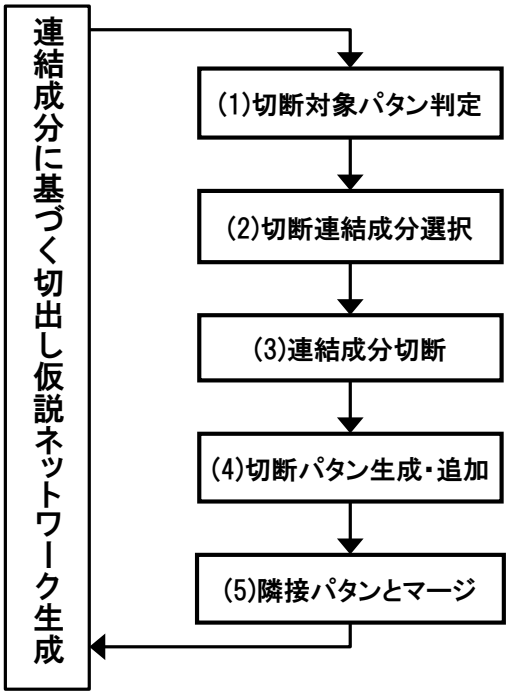


図 2.4　接触文字切出しに対するアプローチ

用意しつつ，複数の候補を生成することにより切出し精度の向上を図る。

接触文字の切出し処理は，文字切り出し部において，文字行毎に連結成分を組合せて文字候補パタンの並びを生成することで，切り出し仮説グラフを構築するプリセグメンテーションプロセスの後半の処理として起動される。接触文字の切出し処理が行われた後に文字識別による文字候補パタンの評価と最適な切出し仮説の選択が行われる。

2.3.　手書き接触漢字切出しアルゴリズム

2.3.1.　接触文字候補パタンの同定と接触箇所の推定

漢字が接触している場合，接触箇所を含む連結成分は必ずしも接触箇所を含まない連結成分に比べて大きくなるとは限らない。例えば，図 2.3 には三つの接触箇所がある。最後の「子市」の接触は，接触箇所を含む連結成分の行方向の幅は，一文字のそれよりも長く，接触文字であると推定することができる。しかし前二つの「東京」と「都八」については，接触箇所を含む連結成分の幅は一文字の幅と比較すると，必ずしも大きいとは言えず，その幅，あるいは面積といった連結成分の大きさに関する情報からは，この連結成分が接触文字を含んでいるかどうか判断できない。

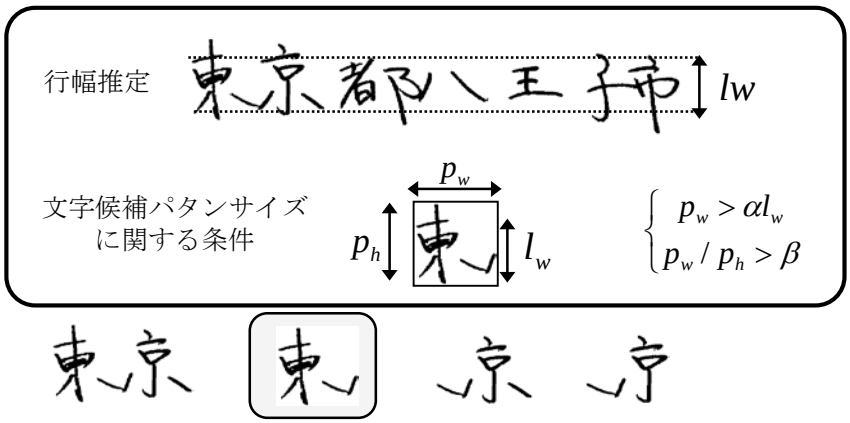


図 2.5 接触箇所を含む文字パタンの選択

そこで，連結成分を文字らしくなるよう組合せて生成した文字候補パタンを用いて，そのサイズによって接触を含むかどうかの判定を行う。これは，接触箇所を含む連結成分が小さくても，それを含んだ文字候補パタンは複数の文字からなるため一文字分の幅より大きくなり，その文字候補パタンのサイズから接触しているパタンを含むかどうかを検出できると考えたためである。

接触文字候補パタンを選択した後に，そのパタンを構成する連結成分の中から，接触箇所を含む連結成分を抽出する。このとき，ある接触箇所を含む連結成分は複数の文字候補パタンに含まれることが多い。そこで切断処理の重複を避け，処理時間の増大を防ぐため，同じ連結成分が切断対象として抽出された文字パタン候補が複数存在する場合は，そのうちの一つを残して切断処理の対象から除く。

接触を含む文字パタン候補の同定は，切出し仮説グラフを用いて以下の 2 ステップで行う。

(a) 切出し仮説グラフ上のすべての文字候補パタン，すなわちエッジに対して，その幅とアスペクト比が以下の条件を満たすかどうか調べる。

文字候補パタンが $p_w > \alpha\, l_w$ かつ $p_w/p_h > \beta$ を満たすならば接触箇所を含むと見なす。

ここで，p_w, p_h はそれぞれ文字候補パタンの幅と高さであり，l_w は行幅の推定値である。図 2.3 に示した例は横書きであるが，本アルゴリズムは縦書きの場合にも

適用可能であるので，ここでは行高ではなく行幅と呼ぶ。これは行画像を複数の領域に分割し，その領域内の黒画素の上端と下端の値のメディアンから求める。また，α, β は定数であり，その値は 2.4 節の評価実験で用いたデータとは異なる地域の手書き地名文字列のデータを用いた実験により，$\alpha = 1.1$, $\beta = 1.4$ とした。

この条件を満たす文字候補パタンが複数存在する場合は，その幅が最も広い（行方向に最も長い）連結成分を選択する。

(b) 切出し仮説グラフ中の文字候補パタンを調べる際に，異なる文字候補パタンから同じ連結成分が接触箇所を含むとして抽出された場合は，そのパタンの幅が最も狭い（行方向に短い）文字候補パタンのみを残し，あとのパタンは選択しない。

これは切断後の隣接パタンとの統合処理において最も多様な文字候補パタンを生成できるようにするためである。このステップによって接触した文字候補パタンを抽出することにより，接触箇所を含む連結成分が小さい場合でも接触した文字候補パタンとして抽出することができる。

図 2.5 の例では，文字列「東京」において「東」の右側の「はらい」と「京」の左側の「はらい」が接触している。切出し仮説グラフ中の文字候補パタンのうち，サイズに関する条件を満たすものが抽出される(ステップ(a))。これに対して後述の手法により接触箇所を含むと推定される連結成分が同定される。この例の「東京」に対しては，接触箇所があると推定された連結成分を含む文字候補パタンが 4 つ生成される。このうち，最も行方向に短い左から 2 番目の文字候補パタンが強制切断を行う文字候補パタンとして選択される(ステップ(b))。

2.3.2. 接触箇所を含む連結成分の強制切断

接触タイプ

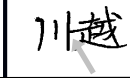
実際に接触した漢字を含む文字列の画像を調査した結果，縦書きの場合は，図 2.6(a)に示すように，上下二文字において，行頭側の文字の一部が下方に伸びて行末側の文字に接触する場合が多い。横書きの場合は，図 2.6(b)に示すように，「川越」の行末側の「越」を書く際に，左に行きすぎて行頭側の「川」の一部に接触するような場合の他，図 2.6(c)のように文字間隔が狭く全体が密着している場合もある。そこで，主に漢字などの日本語文字に関して接触する二つのストロークの接し方によって，表 2.1 のように縦書き二種類，横書き三種類に文字の接触を分類した。

タイプ V1 では，縦と横のストロークが交わって接触が生じており，ストロークの水平

(a)縦書き接触文字　　(b)横書き接触文字(1)　　(c)横書き接触文字(2)

図 2.6　文字の接触例

表 2.1　文字の接触の分類

文字方向	接触タイプ		例	各文字方向における割合
縦書き	V1	⊥		83.3%（80/96）
	V2	Others		16.7%（16/96）
横書き	H1	⊣		36.7%（40/109）
	H2	⊢		44.0%（48/109）
	H3	Others		19.3%（21/109）

方向の幅が急激に変化する。タイプ H1 と H2 も縦と横のストロークが交わって接触が生じており，ストロークの垂直方向の幅が急激に変化する。

一方タイプタイプ V2 と H3 では同じ方向のストロークが重なるように接していたり，あるいは接触箇所がループ状になっていたりする場合である。このとき接触箇所を含む二つのストロークの幅の変化は顕著ではなかったり，接触箇所の近傍に三本以上のストロークが存在したりする。

埼玉県川越市付近の地名が書かれた帳票画像中の縦書き 96，横書き 109，合計 205 の漢

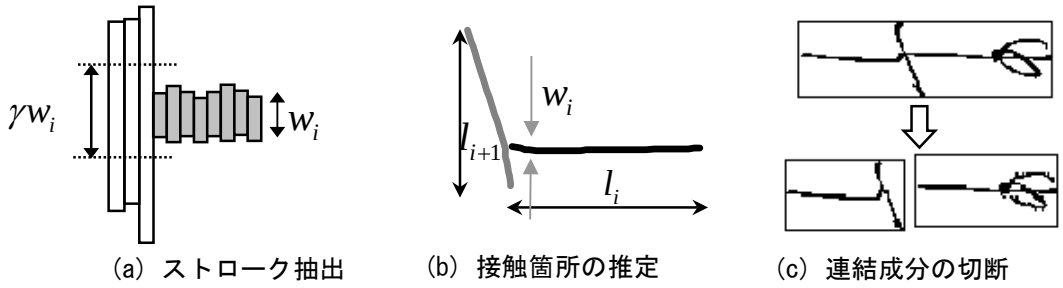


図 2.7　ストローク幅の変化による切断処理

字接触を調べると，表 2.1 の右の列に示すように，縦書き，横書きとも接触の 80%以上が縦と横のストロークが交わるタイプ V1，タイプ H1,タイプ H2 であることがわかった。そこで，接触箇所でストロークの幅が大きく変化するタイプ V1，H1，H2 については，ストロークの幅の変化を捉えることにより連結成分の切断を狙う。上記タイプに加えて，タイプ V2,H3 については，接触箇所でのストローク幅の変化が小さいが，連結成分の行方向の周辺分布をとるとその鞍部になっていることが多い。そこで，連結成分の行方向の周辺分布の形状を解析することにより連結成分の切断を狙う。これら二つの手法を併用することによって，表 2.1 に挙げた接触を含む文字を切出す。

次に二つの連結成分の切断方法について述べる。

ストローク幅の解析による連結成分の切断

この方法では，連結成分に含まれるストロークを抽出した後，ストロークに接触箇所が存在するかどうか判定し，存在する場合は切断をおこなう。この方法は，表 2.1 の V1，H1，H2 タイプのような T 字型の接触はある程度の長さを持つ 2 本のストロークが互いに垂直に接しているという考えに基づきストロークに接触箇所が存在するか判断する。

まず，接触箇所を含む連結成分を文字行方向に垂直な方向に幅が 1 画素の短冊状の領域に分割する。それぞれの短冊状の領域を行頭から行末方向へ走査し，隣り合う短冊の高さを比較することにより，その変化の大きさがしきい値以内であれば，同じストロークを構成するものと判定する。すべての短冊状の領域についてこの比較を行い連結成分に含まれるストロークを抽出する。

w_i をこれまでに走査したストロークiの幅（短冊状の領域の高さの平均）とする時，次に走査する短冊状の領域の高さが γw_i よりも低く，かつ w_i/γ よりも高ければ，その短冊状の領域は当該ストロークの一部とし，そうでなければ別のストロークと判断する。実

験の結果, $\gamma = 2$ とした。図 2.7(a)では, 右から左に短冊状の領域を走査することにより, 灰色の短冊(横長のストローク)と白色の短冊(縦長のストローク)の 2 本のストロークであると判定する。

このようにして連結成分内のストロークを抽出した後，それぞれのストロークに対して以下の三つの条件を満たすかどうかを調べる。これらの条件を満たしたストロークに接触個所が存在すると判断する。

条件 1：ストローク i はその長さがしきい値以上の行方向に平行なストロークである

条件 2：ストローク i に垂直に交わるストローク i+1 の長さがしきい値以上である

条件 3：ストローク i はストローク i+1 よりも長い

連結成分の切断は接触個所を含む二つのストロークのうち，ストローク i（行方向に平行なストローク）のストローク i+1 との境界に最も近い短冊状の領域を一つ削除することにより行う。この短冊状の領域の削除の結果連結成分の数が増えなければ，この連結成分はループを構成していると考えられる。ストローク幅の変化により切断する本手法では，このような場合は扱わず，切断を無効とする。すべてのストロークについて上記の判定処理を行った結果，複数の接触個所の候補が見つかった場合，連結成分あたりの切断個所の数に上限を設け，文字候補パタンの増加による処理時間の増大を押さえる。2.4 節の評価実験で用いたデータとは異なる地域の手書き地名文字列のデータを用いた実験により上限値は 1 と設定した。

切断候補の順位付けには，接触個所を含む二つのストロークの長さの積を用いる。これは接触個所を含み行と平行なストロークは複数の文字にまたがるため長くなる傾向にあるためであり，さらにもう一方のストロークに関しても「はね」などを抽出しないよう長いストロークを優先させるようにするためである（図 2.7(b)参照)。図 2.7(c)に実際の切断の例を示す。

周辺分布形状の分析による連結成分の切断

連結成分を切断する二つ目の方法は，連結成分の黒画素の行方向に垂直な周辺分布を利用する手法であり，上述のストローク幅を用いる手法と相補的に働き，接触文字の切出し精度を向上させる。

一般に文字の接触は，最初の文字の「はらい」が偶然次の文字に接して生じることが多く, 接触個所のストロークの幅は細くなる傾向がある。毛筆の続け字などでも同様である。また文字の間隔が狭いために接触してしまった場合も接触個所，すなわち文字と文字の境界では，黒画素の周辺分布の値が小さくなりやすい。従って，接触個所を含む連結成分の

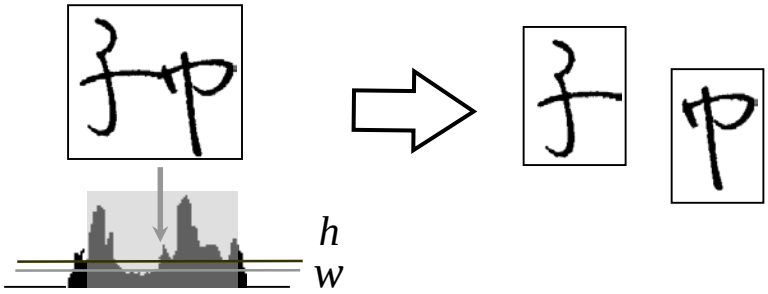
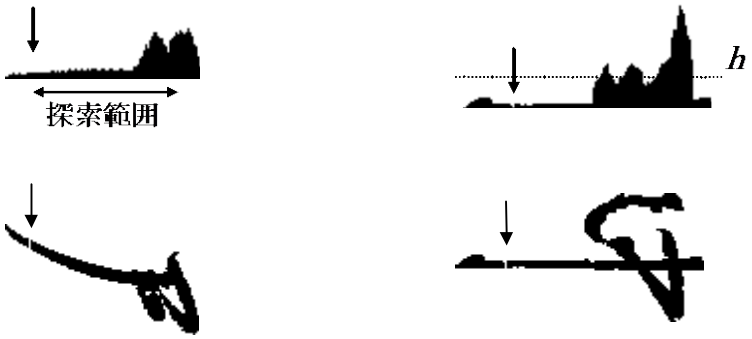
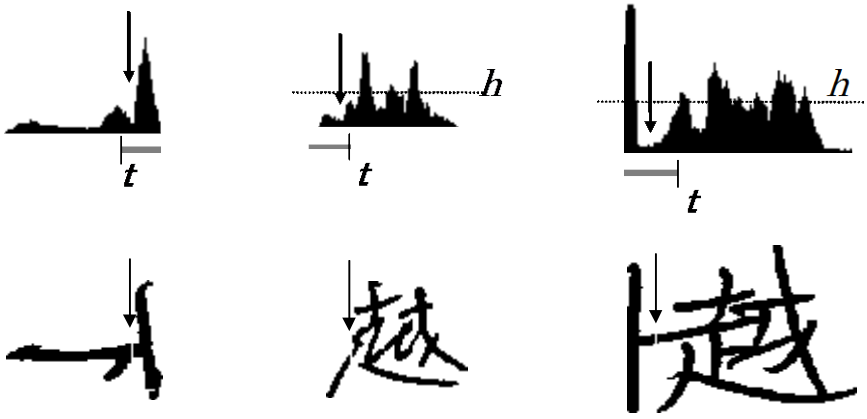


図 2.8　周辺分布形状の利用による切断処理



(a)「戸」の一部　　(b)「中」

図 2.9 不適切な切断箇所候補の例



(a)「下」　　(b)「越」　　(c)「川」の一部と「越」の接触

図 2.10　投影ヒストグラムの形状により適切/不適切と判定された切断箇所の例

周辺分布の極小値を見つけることにより接触部分を識別することが可能である。この特徴を利用した切断処理の流れを述べる。

連結成分の黒画素の行方向に垂直な周辺分布を計算し，その両端部を除いた領域を接触個所の探索範囲とする。これは，連結成分の書き始めや書き終わりの部分は筆圧の変化などでストローク幅が極端に細くなる場合が多く，これらを接触個所と誤るのを防ぐためである。

第一に判定処理のしきい値を算出する。上述の手法により得られたストローク幅を基に，「二」や「三」のように行に並行な方向に複数のストロークが存在し，行に垂直な方向に周辺分布を取ると単一のストローク幅以上の値を示す領域におけるストローク幅の最小値，および単一のストローク幅の最大値を推定する。ただし，ストローク幅がストローク解析によって得られなかった場合は，周辺分布の両端を除いた接触個所探索領域での最小値を単一ストローク幅の最大値とすることとする。ストローク幅の最小値を L_w としたとき，ストロークが複数存在する場合の最小値 h ，単一のストローク幅の最大値 w は，

$$h = \lambda L_w \qquad \text{(式 1)}$$

$$w = \mu L_w \qquad \text{(式 2)}$$

で表わすことができる。但し，$\lambda > \mu$ とする。実験の結果，$\lambda = 4, \mu = 2$ とした。

第二に，周辺分布形状が谷から山，もしくは山から谷への変化点を抽出し，接触個所の推定を行う。図 2.8 に示すように，周辺分布の値が w より小さく，かつ少なくともその両側に h より大きな値のある部分の極小値を連結成分の切断箇所とする。両側に h より大きな値を持つことを条件としたのは図 2.9 に示すようなストロークの末端の膨らみを接触個所と誤るのを防ぐためである。極小値が連続する場合は，文字方向を考慮して，縦書きの場合は最も行末に近い側，横書きの場合は行頭側と行末側の二つを採用する。

このようにして求められた複数の接触箇所の候補を周辺分布の値の小さい順に順位付けする。式 1, 式 2 の定数 λ, μ の値を変化させることにより接触箇所候補の数を調整できる。切断処理では正解を含むよう過剰に切断し，以下に述べる検定処理により不適切な接触個所候補を除去する。

2.3.3. 接触箇所候補の検定処理

強制切断による文字候補パタンの増加による処理時間の増加，誤認識の抑制のために，切断箇所の位置による検定処理を行い，切断箇所となりえないものについては，切出し仮説生成をやめる。推定された切断箇所に対して，(1)これを含む連結成分中の位置，(2)これ

を含む文字候補パタン中の位置，(3)これを含む文字候補パタンと隣接する文字候補パタンの位置，による 3 つの処理を順次行う。

(1)切断対象の連結成分における位置情報利用

接触する二文字において，行頭側の文字の一部が行末側に伸びて行末側の文字に接触する場合が多いので，一般に，切断箇所は文字候補パタンにおいて極端に行頭側に位置しない。従って，図 2.10 の(a)(b)に示すように，接触対象の連結成分を文字候補パタンにおいて切断箇所が閾値以上行頭側に位置する場合は，これらを取り除くこととする。閾値の値は，文字候補パタンの行方向の長さを基準として決める。文字候補パタンの文字行方向の長さを P_w としたとき，閾値 t は，

$$t = c\,P_w \text{ (但し } c \text{は定数，} 0 < c < 1) \qquad \text{(式 3)}$$

とする。ただし，横書きの場合，文字全体が密着していることにより接触している場合もあるので，図 2.10 (b)のように，行頭側に h（複数のストロークが重なった場合の周辺分布の最小値の推定値）以上の値がないという条件を共に満たす場合に，切断箇所から取り除くこととし，図 2.10 (c)のように，行頭側に h 以上の値がある場合は切断する。

(2)切断対象の文字候補パタン内の位置関係利用

接触を含むと判断した文字候補パタンに対して，その連結成分から行方向の長さ(幅)が最大の成分を選択し切断処理を行う。この時，接触を含むと判断された文字候補パタン内での切断対象とならなかった連結成分を取り出し，これと接触を含む（切断対象の）連結成分との間の位置関係から切断の是非を検証する。

具体的には，切断の対象とならなかった連結成分と，それ以外の成分との行方向への投影をとり，切断対象の成分の投影画像が，もう一方の投影画像に完全に包含されており，かつ，もう一方の成分の投影画像が行方向に連続であれば，切断対象の連結成分の選択を誤ったとして，切断処理を行わない。例えば「辻」という文字の「十」を構成する連結成分が切断対象とされたときに，その連結成分を含む文字候補パタンの他の連結成分，つまり，「しんにょう」を構成する連結成分は「十」の連結成分を行方向に包含している。このとき「十」の切断は行わない。

(3)切断対象の文字候補パタンと隣接する文字候補パタン内連結成分との位置関係利用

接触を含むパタンに対して，隣接するパタンを取りだし，接触パタンにおける切断位置と，隣接するパタンにおける，切断箇所からもっとも近い連結成分までの距離を求め，こ

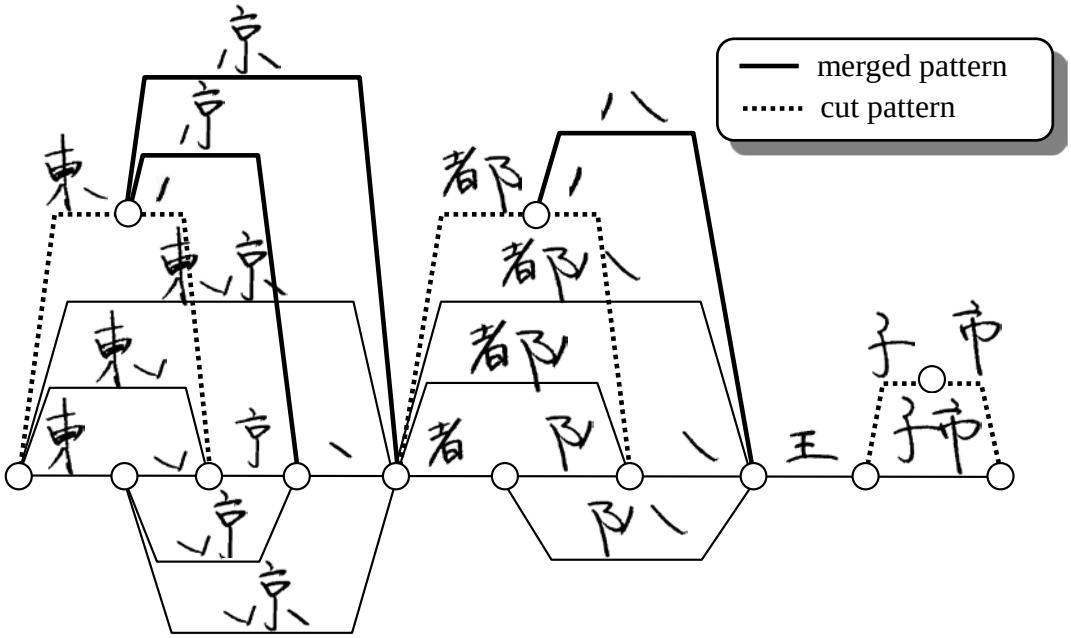


図 2.11 接触文字切り出しの結果を反映した切出し仮説ネットワーク

れがしきい値以上離れていれば，誤切断であるとして，切断結果を採用しない。これは，はねなどにより 90 度方向の異なる直線成分ができた場合，これを切断してしまうのを防ぐことを目的とする。実験の結果，上記しきい値は文字幅の推定値の 1/7 とした。例えば，横書きの「御池通」「池」と「通」が十分離れている場合に，「池」の最後のはねの部分を接触箇所として切断することを回避する。

2.3.4. 新しい文字候補パタンの生成

切断によって生成された二つ以上の連結成分と，接触文字候補パタン中の他の連結成分から新たに文字候補パタンを生成する。前処理部で切断をするように判定されたパタンのうち最小のパタンについて，切断後の連結成分をそれぞれ含む二つのパタンを新たに切出し仮説グラフに追加する。

切断の対象外となった連結成分は，元のパタンにおける切断連結成分との相対的な位置関係でいずれかのパタンに振り分ける。もし切断された連結成分が，接触個所を含まない他の連結成分に行方向に包含されていれば，この切断候補を棄却する。こうして生成された文字候補パタンは，図 2.11 の点線で示すように，切出し仮説グラフに新しいエッジとして追加される。

ここまで述べた，接触文字候補パタンの抽出，接触を含む連結成分の切断，文字候補パ

タンの生成のプロセスは，切出し仮説グラフ上で未探査のノードが存在しなくなるまで繰り返す。すなわち，切断処理によって新たに生成された文字候補パタンに対しても接触文字候補パタンかどうかの判定処理を行う。これにより 3 文字以上が接触している文字候補パタンに対しても切断処理が再帰的に働くことになり，一文字毎に正しく切出すことが可能になる。

もし最初に文字候補パタンを生成した際に，接触した文字が別の文字パタン候補に分断されてしまった場合，文字パタン候補を切断するだけでは，正しい文字パタン候補を得ることができない。そこで，最後に切断によって生成された文字候補パタンとそのパタンに隣接する文字候補パタンを統合して新たに文字候補パタンの生成を試みる。

文字候補パタンの幅や隣接するパタンとの間隔などの周辺特徴量[5]が文字候補として適当であれば，統合された文字パタンが図 2.11 の太い実線のように切出し仮説グラフに追加される。ここでの文字候補パタンに対する判定は，最初に切出し仮説グラフを生成する際のしきい値を用いて行う。

表 2.2　接触文字切り出し処理の概要

処理流れ	内容	
接触文字候補パタン抽出	アスペクト比判定	
	行方向黒画素連続判定	
	切断連結成分判定	
	切断連結成分重複判定	
	最適切断パタン判定	
連結成分切断	ストローク幅解析による切断	ストローク（直線成分）の抽出
		ストローク幅変化点（切断線）の抽出
	周辺分布形状解析による切断	投影計算
		投影形状変化点（切断線）の抽出
切断パタン生成	連結成分ソート	
	切断パタン重複判定	
	切断パタンのパタンテーブルへの追加	
切断パタンと隣接パタンのマージ	アスペクト比によるマージ判定	
	マージパタンのパタンテーブルへの追加	

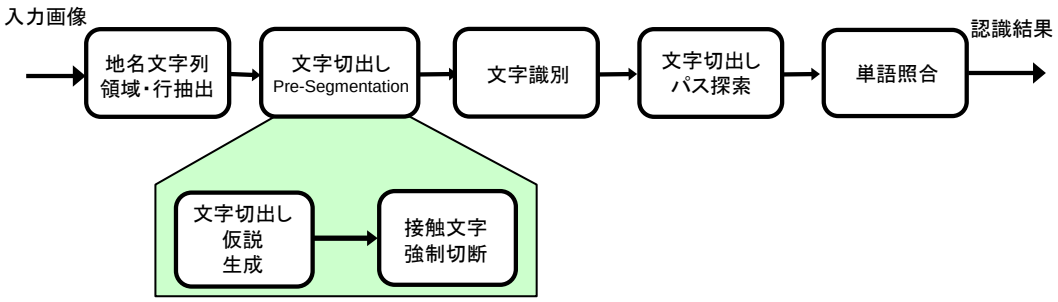


図 2.12 手書き地名文字列読取り処理の流れ

以上述べてきた接触文字切出し手法の各ステップにおける処理の内容を表 2.2 にまとめる。提案する手法により，プリセグメンテーションプロセスの一部として，接触した文字に対する文字候補パタンを生成し，切出し仮説グラフに追加することができる。この手法は，縦書き，横書きの両方に適用可能である。両者の間の違いは，縦書きは表 2.1 にも示したように，行頭側の文字から伸びたストロークが行末側の文字に接している場合を対象にするのに対して，横書きは行末側の文字から伸びたストロークが行頭側の文字に接している場合も対象とする点にある。従って，接触個所を含むストロークの位置関係が逆（縦→横に加えて横→縦）になっている場合も扱うように変更することで，両方の書き方に適用することができる。

2.4. 評価実験の結果と考察

提案する接触文字の切出し手法の有効性を評価するため，接触箇所の切断精度と測るとともに、本手法の適用を想定する地名文字列認識プログラムを用いて地名文字列読取り精度の向上を調べた。

地名文字列認識プログラムの構成は図 2.12 のようになっている[12]。宛名領域・宛名行抽出において，郵便物画像上から抽出された宛名文字行に対して，文字切出し部において連結成分を組合せて文字候補パタンを生成し，このパタン系列の候補を文字切出し仮説として複数生成する。文字候補パタンに対して文字識別処理を行い，そこから得られた類似度，あるいはパタンのサイズ，間隔等から計算されるペナルティを用いて文字候補パタンに確信度付けし，次の文字切出しパス探索部において，最適な文字切出し仮説を得る。最

後に，単語照合部において宛名単語列との照合処理を行い，文字識別の誤りの訂正を行って，宛名認識結果を得る。

単語照合部では，文字識別の誤り訂正処理を含んでいるが，郵便番号の読取り結果による地名読取り結果の補完は行っていない。評価を行ったマシンは日立ワークステーション VJ240 (CPU:PA-8200, SPECint95=17.3，SPECfp95=25.4)である。

埼玉県川越市内の地名表記が書かれた手書き帳票画像 6,634 枚を用いて評価した。採取した帳票画像の全体をデータセット A とし, そのうち地名表記の漢字の部分, すなわち「川越市藤間 1-2-3」における「川越市藤間」の部分（これを町域とよぶ）に接触を含むすべてのサンプル 939 枚を取出し，これをデータセット B とする。さらに接触を含むデータセット B のうち，165 枚（接触箇所は 213 箇所）をランダムに抽出し，データセット C とする（表 2.3 参照)。

表 2.3　評価実験条件

評価マシン	Hitachi WS VJ240 (CPU:PA-8200 236MHz, SPECint95=17.3, SPECfp95=25.4)	
評価データ	データセット A	川越市内の手書き住所画像 6,634 枚
	データセット B	データセット A のうち地名漢字部分に接触を含む 939 枚
	データセット C	データセット B の一部である 165 枚

表 2.4　接触文字の切り出し精度

	Total	切出し成功		切出し失敗	
縦書	95	88	92.6%	7	7.4%
横書	140	104	74.3%	36	25.7%
全体	235	192	81.7%	43	18.3%

評価の項目は以下の三点である。

(1) 接触箇所の切断精度

(2) 連結成分の切断手法の併用に関する有効性

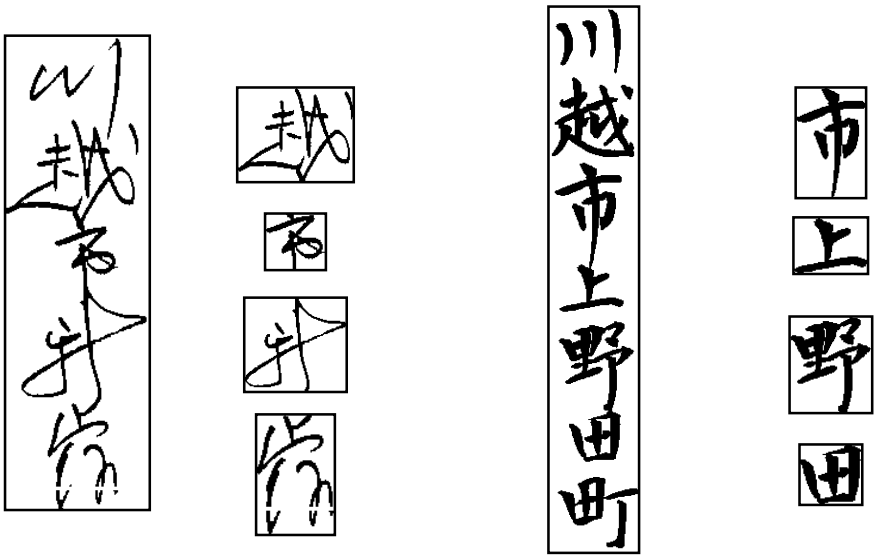
(3) 接触文字切出し処理による地名読取り精度の向上

(a) 「藤間」

(b) 「川越」

(c) 「郵便局」

(d) 「松原」



(e) 「越市新宿」　　(f) 「市上野田」

図 2.13　接触文字切出しの成功例

最初の評価項目として，データセット C を用いて，その接触した文字が正しく切出せているかどうかを調べた。ここで，接触文字を正しく切出せるとは，連結成分を切断して追加した文字候補パタンを目視した結果正しい位置で切断されていること，あるいは切断に

より追加した文字候補パタンに対して，地名文字列読取りで用いられている文字識別モジュールが正解候補を出力したことと定義する。その結果，表 2.4 に示すように，縦書きでは，95 の接触箇所のうち 88 箇所（92.6%）が切出しに成功し，横書きでは 140 の接触箇所のうち，104 箇所（74.3%）の切出しに成功した。全体では 235 の接触箇所のうち，192 箇所（81.7%）の切出しに成功した。正しく切出せた接触文字の例を図 2.13 に示す。(d)のように接触箇所を含む連結成分が相対的に小さい場合にも，文字識別ができる程度に切出せている。

接触を含む文字を正しく切出すことが出来なかった例として，図 2.14 に示すような 4 つの場合がある。

(1) 斜め同士のストロークが接触している
(2) 複数箇所で接触している
(3) 接触個所を含むストロークの検出に失敗した
(4) 行全体にわたって文字が密に書かれており個々のストロークの長さがしきい値を満たさなかった

縦書きに比べて横書きの切断精度が低いのは，(1)が横書きに顕著に見られる特徴であることが影響している。この場合各文字パタンの幅も小さくなっていることが多く，接触文字の判定の際接触を含んでいても，文字候補パタンの幅のしきい値を越えることが出来ず判定に失敗してしまう。文字列の読取りに失敗した場合に，文字候補パタンの幅に対するしきい値を動的に小さくするなどして，接触文字パタンの判定処理を行うなどの方法が考えられる。

図 2.14 の(b)における「字新」や(d)の「新田」のような複数箇所での接触に対しては，今回提案するストローク幅変化点による切断手法では切断することが出来ない。周辺分布形状変化による切断手法では一本のストローク幅の推定値をパラメタとして用いているため，このような場合に対応することは困難である。

斜めどうしのストロークが接触している場合については，今回提案するストローク幅抽出手法では，隣接する短冊状の領域のずれがしきい値以下であることを同一のストロークとみなす条件としているため，斜めのストロークの抽出精度は低い。これに対しては，輪郭抽出によりストロークを抽出するか，あるいは提案手法において，隣接する短冊状の領域のずれる方向を記憶しておくことにより，ずれ幅のしきい値を緩くすることで，斜め方向のストロークの抽出制度を高める等の手法が考えられる。

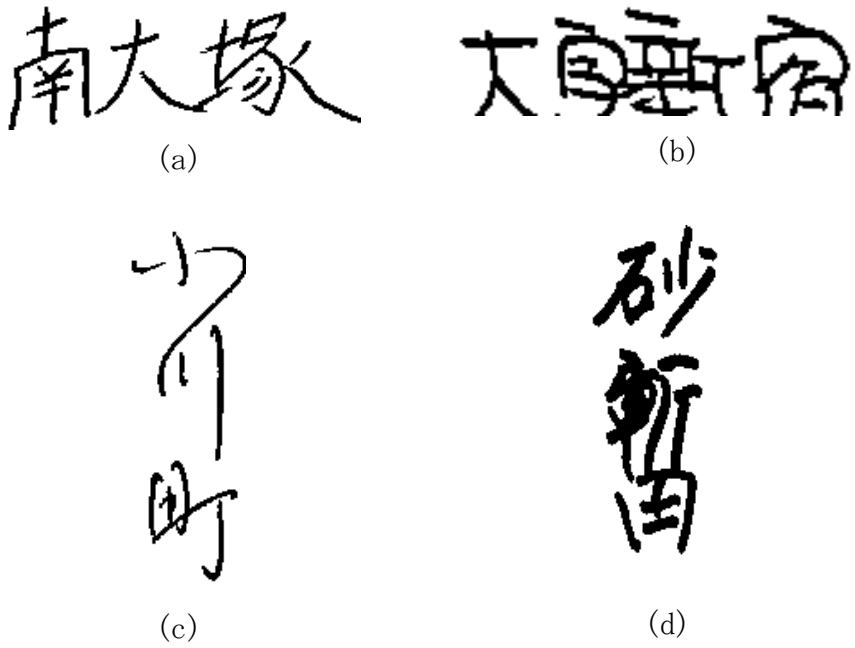


図 2.14　接触文字切出しの失敗例

表 2.5　切断手法と地名文字列（町域）認識精度

		既存手法[16]		提案手法(強制切断による接触文字切出しあり)					
				ストローク幅変化		周辺分布変化		両方	
データセットA	Total	6366		6366		6366		6366	
	Accept	4948	77.7%	5084	79.9%	5194	81.6%	5204	81.7%
	Correct	4812	75.6%	4948	77.7%	5056	79.4%	5070	79.6%
	Error	136	2.7%	136	2.7%	138	2.7%	134	2.6%
	Reject	1418	22.3%	1282	20.1%	1172	18.4%	1162	18.3%
データセットB	Total	897		897		897		897	
	Accept	380	42.4%	503	56.1%	607	67.7%	615	68.6%
	Correct	349	38.9%	474	52.8%	582	64.9%	591	65.9%
	Error	31	8.2%	29	5.8%	25	4.1%	24	3.9%
	Reject	517	57.6%	394	43.9%	290	32.3%	282	31.4%

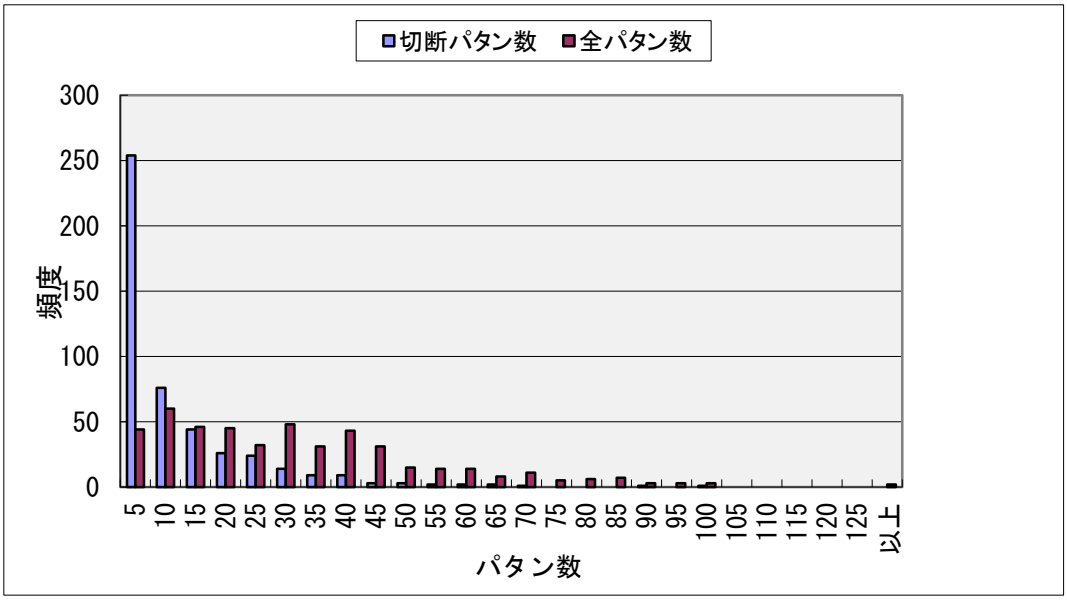


図 2.15　接触文字切出し処理と文字候補パタン数の分布

第二番目の評価として，接触箇所を含む連結成分を切断する手法として，ストローク幅の変化により連結成分を切断する手法と，連結成分の周辺分布形状により切断する手法を組合せたことに関する効果と，連結成分の切断とその後の隣接パタンとの統合処理を組合せたことによる効果を調べた。ストローク幅変化と周辺分布形状による連結成分の切断手法につき，一つの連結成分についてそれぞれ最大 2 箇所の接触箇所候補を出力するものとし二つの手法を組合せた場合は，連結成分当たり最大 4 箇所の接触箇所候補を出力するようにした。データセット A と B について宛名読取りプログラムによる町域文字列の認識率で評価した結果を既存手法[16]による認識精度と比較して表 2.5 に示す。

接触と非接触の混在する手書帳票画像サンプル（データセット A）に対しては，既存手法(接触文字切出しを行わない場合)のアクセプト率 77.7%と比べて，ストローク幅による切断手法，周辺分布変化による切断手法，その両方を用いた手法では，それぞれ 79.9% (2.2pt)，81.6% (3.9pt)，81.7% (4.0pt)と向上している。周辺分布形状による切断と両者の併用による手法とでは 0.1pt しか向上が見られないが，二つの切断手法の併用による効果は，正解率が 79.4%から 79.6%へ 0.2pt 向上し，誤読率が 2.7%から 2.6%へ 0.1pt 低下したことに現れている。

さらに接触文字を最低一箇所含むサンプル（データセット B）について評価すると，接触文字切出しを行わない場合のアクセプト率 42.4%と比べて，ストローク幅による切断手

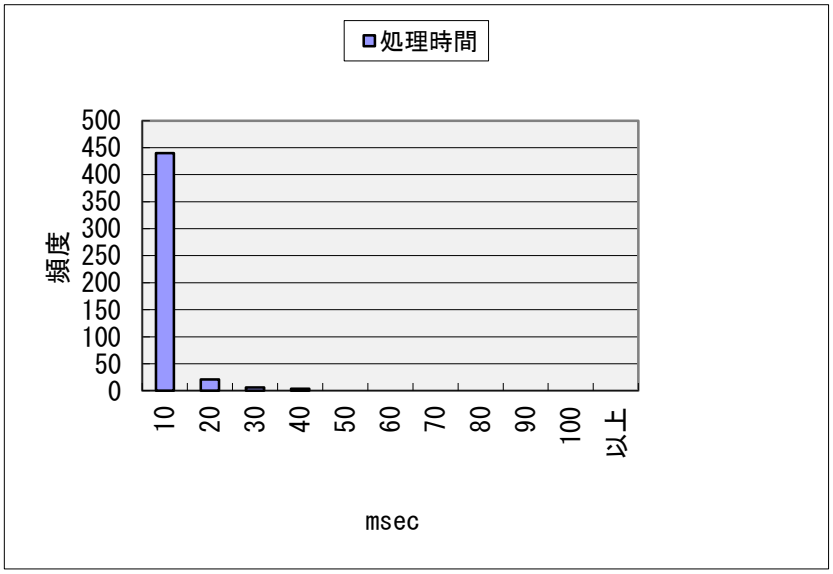


図 2.16　接触文字切出し処理時間の分布

法，周辺分布変化による切断手法，その両方を用いた手法では，それぞれ 56.1% (13.7pt)，67.7% (25.3pt)，68.6% (26.2pt)と向上している。ここで接触文字切出しを行わない場合でもアクセプト率がゼロではないのは，図 2.12 に示すように，地名表記読取り処理の最後で単語照合処理が行われており，文字識別の誤り訂正機能が働いているためである。

第三番目の評価項目として，接触文字切出し処理により新たに増加した文字候補パタンの数を調べた。一行あたりの文字候補パタン数の平均が 30.0 であり，そのうち，接触文字切出し処理で生成したパタン数の平均が 9.2 である。その分布は図 2.15 のとおりである。また接触文字切出し処理の処理時間の平均は一行あたり 5.0msec であり，その分布は図 2.16 のようになる。文字候補パタン一つあたりの文字識別の処理時間は約 10msec であるため，接触文字切出し処理を行うことにより，表 2.5 に示すような精度向上を得るかわりに宛名読取り処理に要する処理時間は一文字行あたり約 100msec 増加する。実際は，接触文字切出しにより上位の文字行候補をアクセプトすることにより，下位の文字行候補の処理をしなくてすむために，宛名読取り処理全体の処理時間の増加は速くなることが期待できる。

2.5.　結論

本研究では，接触した手書き漢字の切出し手法に関して検討をし，これを開発した。提案する方式の特徴は以下の四点である。

(1) 連結成分を組合せて生成された文字候補パタンに対して接触文字か否かを判定
このため，接触個所を含む連結成分が小さい場合であっても，接触個所を見つけることが可能である。漢字の場合は一文字が偏や旁，あるいは点等の小さい部品からやなる場合が多く，隣接する文字間で，その部品同士が接触しても，接触個所を含む連結成分は他に比べて大きくなるとは限らない。そのため，本特徴は接触漢字の切出しには有効に働く。

(2) 連結成分の切断処理と，切断により生成された文字候補パタンと隣接文字候補パタンとの統合処理からなる，「切断統合法（Cut-and-Merge method)」の導入
連結成分の切断により生成された文字候補パタンと既存の文字候補パタンを統合して新たに文字パタン候補を生成することにより，接触文字が複数の文字パタンに分離してしまった場合でも，接触文字に対する候補パタンを生成可能となる。

(3) 連結成分の切断に複数の手法を用いて多様な接触の仕方に対応
今回実装した方式では，連結成分の切断は，ストローク幅の変化を捉えて切断する手法と，連結成分の周辺分布の形状の変化を捉えて切断する手法の二つを併用した。これにより「はらい」が隣の文字に接触する場合や文字同士が密着したような場合についても，それぞれの手法で切出せるようになった。

(4) 接触個所の切断処理を再帰的に適用
切断によって生成した文字候補パタンが，接触判定処理により接触を含むと見なされる限り，連結成分の切断処理を繰り返し適用する。すなわちある文字候補パタンに対して再帰的に切断処理が行われることになる。これにより，複数の文字が接触しているような場合でも，それぞれの文字候補パタンを切出すことができる。

今後の課題として，接触漢字切出し精度の向上が挙げられる。手書き接触漢字の切出し精度は，接触していない文字の切出しに比べて劣っており，これを向上させることが必要である。しかし，ただ切断する連結成分の数を増やし，生成される文字候補パタンの数を増やしていくだけでは，文字列読取り処理全体の処理時間が増大してしまう問題がある。そこで，接触個所ではない連結成分の切断の増加を極力抑えながら，接触個所を含む連結成分の切断精度を高めていくことが課題である。具体的な手法としては，以下が考えられる。

(a) 連結成分の切断の際に接触個所と接触とは無関係のストロークとの位置関係を利用した検定処理の導入
ストローク幅変化による切断を行う際，接触個所を含む二本のストロークにのみ着目しているが，その他のストロークを行方向に投影すると，接触個所を包含し

ているような場合は接触個所として不適切であると考えられる。このような検定処理を加えることにより，切断処理に関するしきい値を緩めても，不適切な切断候補の数を抑えることができる。

(b) 文字候補パタン全体での周辺分布形状の利用

連結成分だけの周辺分布ではその形状の変化が小さくても，文字候補パタン全体の周辺分布では，接触個所近傍の周辺分布形状の変化が大きくなり，これを検出できる可能性がある。提案手法で接触個所を切断できなかった場合は，周辺分布形状を調べる対象を文字候補パタン全体に広げれば良い。

第3章

統語的言語モデルを用いたグラフ探索型文字列認識

3.1. 緒言

ブロック体で書かれたアルファベットの単語列の認識には，単語間の空白の情報を利用することができる[67]。本章では，日本語のように単語に分かち書きされず単語間の境界があいまいなため，単語認識の連接では十分な認識精度を得ることが困難な言語における文字列の認識精度向上のための，一般的な情報システムのリソースで実行可能な言語モデルとその利用法について提案する。

業務において手書きされる文字列は，地名，氏名，会社名，製品や商品の名前，金額，日時，症状や設備の状況などの記述を含むことが多い。これらの記述のうち，金額や日時を除くと複数の単語の組合せによって構成されることがほとんどである。加えて，特に地名や会社名，製品や商品名，症状や設備の状況などは，省略や漢字の使用の有無など複数の表記が用いられることも共通している。複数の単語からなる文字列として，本章では地名表記を取り上げる。これは，複数の単語の組合せで表現される，複数の表記が許容されるという特徴を持つ文字列の中で，地名表記は一つの表記中に含まれる単語数が多く，表記の多様性に富むからである。地名表記とは「東京都国分寺市東恋ヶ窪 1 丁目 280 番地」のように特定の場所を示す文字列のことであるが，本章では「東京都国分寺市東恋ヶ窪」といった都道府県名，市町村名，街区名の部分のみを対象とする。提案する認識手法は番地名を含む地名表記文字列に適用することも可能である。地名表記の認識とは，記述された文字列を逐一認識することではなく，その文字列が表わす場所を特定することである。すなわち，入力された文字列が「国分寺市東恋ヶ窪」であっても，「国分寺市東恋が窪」であっても，さらに「コクブンジシヒガシコイガクボ」であっても，郵便番号「185-0014」で表わされる場所であると同定する必要がある [3]。これは地名表記に限定された観点ではなく，例えば，保全業務において「脱水層のスクレーパー」や「脱水スクレーパー」を同じ部品のことだと解釈しなければならない場合があることから，文字列の認識とは文字を逐一認識することではなく，その文字列が表わす対象に紐づけることであるとすることで一般性を失わない。

[3]日本の地名表記の場合，都道府県名，市町村名，街区名によって特定される場所は，7 桁の郵便番号によって特定される場所とほぼ一致する。

これまでに，文字毎に用意された枠内に手書きされた日本語の地名表記に対して，丸川らは，文字毎の複数認識結果を並べて構成する文字候補ラティスからオートマトンを生成し，単語辞書中の文字を入力し状態遷移をさせることで，最終的な認識結果を得る手法を提案している[68][69]。単語を列挙した言語情報を用いながら，認識誤りの修正機能を持つ地名文字列認識を実現したが，枠のない書式に書かれた文字列に対しては文字切出しの曖昧性が生じるため，オートマトンが複雑になり処理時間の増大が問題となる。また，冒頭の「東京都国分寺市東恋ヶ窪」の例のような異表記を網羅することは考慮されていない。Koga らは，都道府県，市町村，街区など複数の単語の並びからなる地名表記をトライ構造によって表現し，文字切出し候補を入力としてトライ構造の言語モデルを探索することで，文字列認識結果を得る手法を提案している[70]。この手法でも異表記を網羅することは考慮されていない。仮に異表記を含めたとしても，地名表記を木構造で表現しているため，街区部分は同じ表記であっても市町村部分で異表記がある場合は，街区部分を含めて重複してトライ上に表記データを保持する必要があり，メモリ容量の増大を招くという問題がある。このように，異表記を含めて，認識対象の文字列をすべて列挙する既存の記述型の言語情報として表現することは困難である。

一般に異表記は標準の表記に比べて，筆記される頻度が低いことが多い。従って，統計的な言語モデル[71][72][73][74]を用いたとしても，発生頻度の低い表記は認識できない可能性がある。1.2.2 節で述べたように，複数の単語から成る日本語の地名表記の認識には，インスタンスを列挙することなく，かつ認識精度がサンプルの頻度に依存しない言語モデルとこれを利用した知識照合手法が必要である。さらに実用上の観点からは，その言語モデルが容易に作成できるようにする必要がある。

本章では，複数の単語から構成され，かつ多数の異表記を含む文字列である地名表記を認識するために，インスタンスを列挙するのではなく，文法規則に基づき認識対象の文字列を表現する統語的言語モデルを用いた文字列認識手法提案する。まず，異表記を含む場合でもメモリ容量を抑えることができるように，再帰的なグラフ構造による言語モデル(以下，グラフ型言語モデルと呼ぶ)を検討する。次に，異表記を含む場合でも高精度かつ実用的な時間で処理可能な，再帰的なグラフの探索による文字列照合手法について提案する。さらに再帰的なグラフと文脈自由文法の表現力の等価性を用いて，文脈自由文法によって記述した対象文字列に対してルールベースで異表記を追加し，これをグラフ構造に変換することでグラフ型言語モデルを構築する手段を提案する。これによってグラフ型言語モデルの構築が簡易化可能となる。

以下，3.2 節では，地名表記の異表記について述べ，再帰的なグラフ構造による統語的

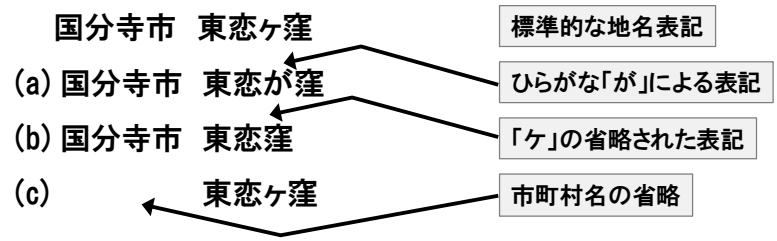


図 3.1　地名表記の異表記の例

言語モデルについて述べる。3.3 節では，提案する言語モデルを用いて，文字認識結果候補列から認識結果を選択する，グラフ探索型文字列認識手法について述べる。3.4 節では，実際のアプリケーションへの適用を見据えて，提案する言語モデルの実装方法について述べる。3.5 節では，提案するグラフ型言語モデルにおける対象文字列の表現力と，これを用いた文字列認識精度と処理時間について評価を行う。これにより，提案する統語型言語モデルを用いた文字列認識手法の有効性を示し，3.6 節で本章の結論を述べる。

3.2.　グラフ型言語モデル

3.2.1.　地名表記における異表記

地名表記の認識において，認識対象となる地名表記の多様性は図 3.1 に示すように，以下の 3 つに分類される。

(a)　文字レベル：地名表記中のひらがな，カタカナの違い

(b)　単語レベル：“St.”や“L.A.”といった省略や略号

(c)　句レベル：都道府県名，市名の省略や別名など

文字レベルや単語レベルの異表記は，都道府県名や市町村名，街区名など地名表記のどの階層でも起こりうる。地名表記の異表記の総数は，これらの組み合わせによって非常に多数となる。さらに，ある単語がひらがなで記述されていても，その単語が表記中に再度現れる際にはカタカナで書かれていることもあるように，異表記の生じ方には首尾一貫性はない。何らかのルールによって文字の使い方を検証することも困難である。また句レベルの異表記については，さらに考慮すべき表記が存在する。京都のような通りの名前の組み合わせによっても住所が表現される都市においては，例えば，「京都市下京区高倉」という地名は「京都市下京区烏丸通り錦」や「京都市下京区烏丸通り四条」といった別名で表

現されることが多い[75]。しかも，通りの名前の順番が入れ替わることも多く，さらに，通りの名前に省略表記などの異表記が存在する。複数の通りの名前を用いることで単語長も増え，全体の異表記の総数も大きく増え，一つの地名に対する異表記の数も数千に上る場合がある。

3.2.2. 再帰的遷移グラフによる言語モデル

地名表記は都道府県，市町村，街区という階層構造をとっている。京都府や京都市は一つしか存在しないが，南区という地域は全国に複数存在する。さらに錦町という街区はさらに多くの地域に存在する。従って，地名表記をトライ構造にて実装しようとすると，葉に近いところで同じ部分木が多数でき，本質的に冗長性が高まるという性質がある。さらに 3.2.1 項で述べた異表記を表現しようとすると，同じ部分木に異表記が追加され，さらに冗長性が高まってしまい，製品として実装可能なメモリ量をはるかに超えてしまうという問題がある。

こういった冗長性をなくすために, 都道府県, 市町村といった地名表記に表出する単位, あるいは，そこで生じる異表記毎にグラフで表現し，これらグラフの多層構造で地名表記を表現する言語モデルを提案する。この言語モデルは，エッジが文字，あるいは文字列に対応し，ノードが文字と文字の間を表す無閉路有向グラフ(DAG（Directed Acyclic Graph))[76]の集合によって表される。提案するグラフを構成するエッジは 3 種類に分けられる。1 つ目は,「プッシュダウンエッジ」と呼ぶもので，別のグラフへのポインタをそのラベルとして持つ。二番目は，文字列を表す「ノーマルエッジ」，三番目は null 遷移に対する「ヌルエッジ」であり，これは入力がない場合の遷移を許容する。認識対象の文字列全体は,「幹ネットワーク(trunk network)」と呼ばれるグラフによって表現される。都道府県や市町村，あるいは異表記などを表現する別のグラフを生成しておき，プッシュダウンエッジによって別のグラフへの参照が記述される。このように，認識対象の文字列はグラフを用いて再帰的に表現することができる。こうしたグラフは，再帰的遷移ネットワーク(RTN: Recursive Transition Network) [77]と呼ばれており，文脈自由文法と同等の記述能力をもつことが知られている。

図 3.2 に「京都府京都市下京区高倉町」の地名表記を表すグラフ型言語モデルの例を示す。図中でエッジのラベルが「ε」となっているのがヌルエッジ, エッジのラベルが「<>」で囲まれているのがプッシュダウンエッジである。図中には 1 つの主グラフ(幹ネットワーク)と都道府県名(京都府)，区名(下京区)，通りの組合せ(烏丸通四条)，通り名(烏丸通)に対

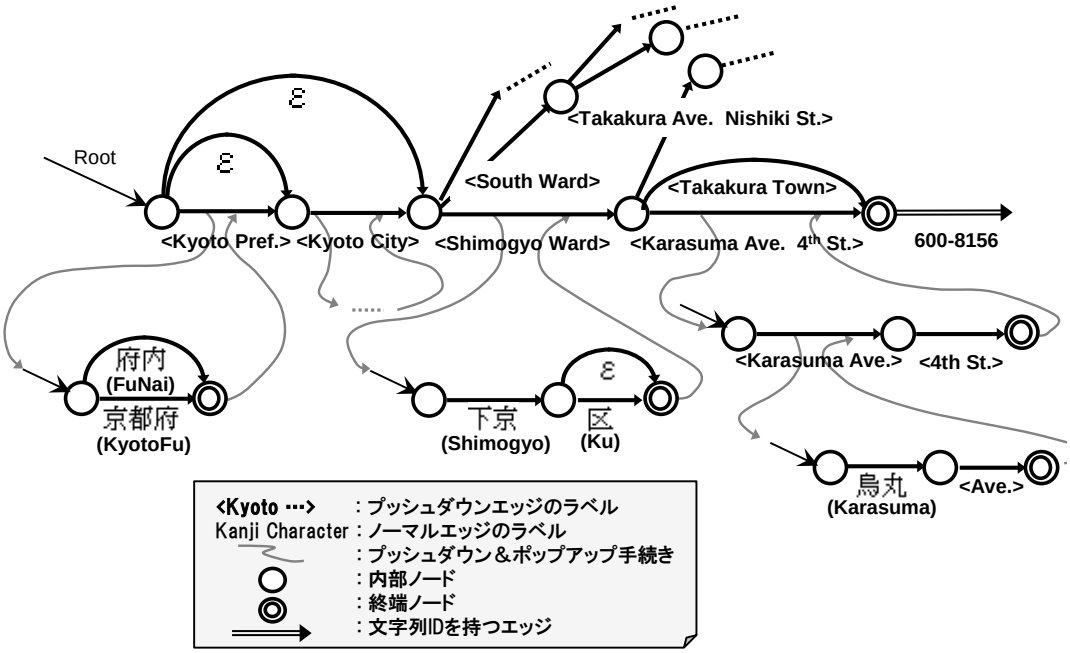


図 3.2　言語モデルのグラフ表現の例

応する 4 つの部分グラフが記されている。「下京区」に対応する部分グラフは，「下京区」あるいは「下京」という文字列を表現している。また「烏丸通四条」を表す部分グラフは「烏丸通」を表す別の部分グラフと「四条通」を表す別の部分グラフを参照している。さらに，「烏丸通」を表す部分グラフは，「烏丸」という文字列と通り名に対応する部分グラフによって「烏丸通」が表現されることを示している。主グラフは，これら 4 つの部分グラフを参照しながら，読み取り対象の文字列全体，つまり地名表記を表現する。それぞれのグラフはルートノードと終端ノードを持つ。グラフの探索はルートノードから始まり，終端ノードのいずれかに到達した際に成功する。主グラフの終端ノードから派生するエッジは，そのグラフで表現される文字列に対する ID を表す。本章で取り上げる地名表記に対する言語モデルの場合，郵便番号を ID として用いる。

トライ構造は決定性有限オートマトンで記述できることが知られている。提案する再帰型遷移ネットワークの方が決定性有限オートマトンよりも表現力が高いので，提案するグラフ型言語モデルでは既存の言語モデルで表現していた文字列を全て表現できることが保証される。

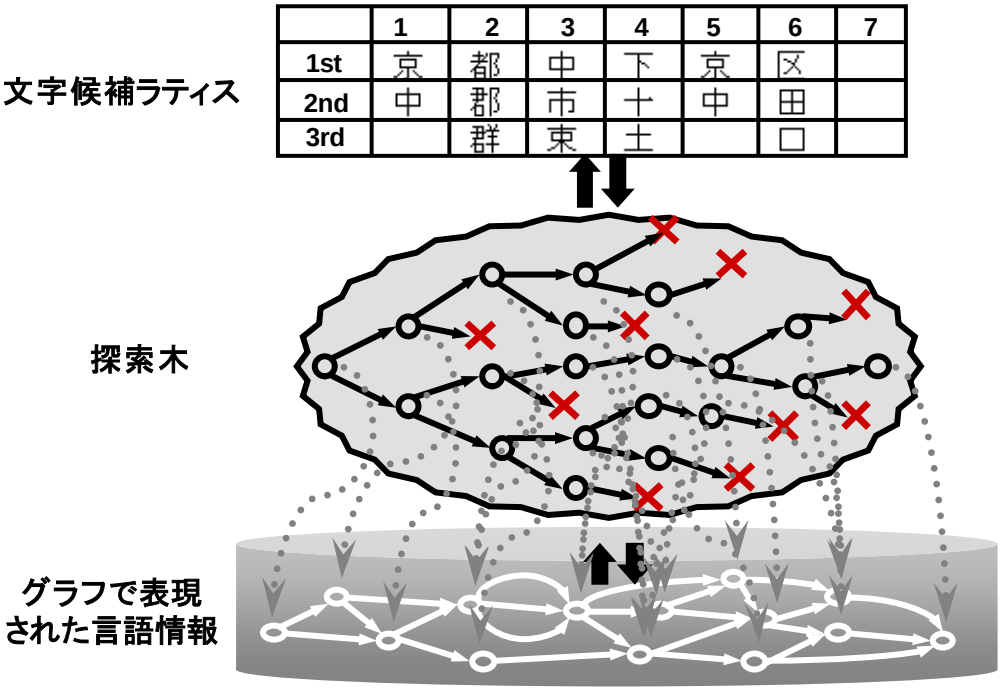


図 3.3 グラフ探索の様子

3.3. グラフ探索型文字列認識

3.3.1. 文字列候補に対するグラフ探索

本節では, 言語モデルを用いて文字列認識を行うための文字列照合方式について述べる。文字列照合とは，言語モデル中に記述された表記の中から，もっとも確からしい文字候補の系列を選ぶという問題を解くことである。確からしいとは，個々の文字の識別結果に基づき計算されるスコアによって，文字列認識結果の候補と言語モデル中の文字列の一致度が最も高いと判断されることを意味する。

自然言語の構文解析においては，言語モデルへの入力は文字の並びである。一方で，文字列認識における言語モデルへの入力は，文字切出しの仮説上の各文字候補に対して複数の文字識別結果候補を保持する「文字候補ラティス」と呼ぶ束構造のデータとなる。図 3.3 上部が文字候補ラティスの例であり, それぞれの文字ごとに候補がランク付けされている。そして, 3.2.2 節で説明したグラフ型の言語モデルを探索することによって最も確からしい文字列候補を見つける。グラフ探索処理は，「幹ネットワーク(trunk network)」と呼ばれるグラフの根ノードから開始される。探索処理がプッシュダウンエッジに到達すると，その終端ノードがプッシュダウンスタックに積まれ，探索プロセスはそのプッシュダウンエ

ッジによって参照される部分グラフの根ノードに移動し，探索処理を開始する。部分グラフの探索が，その終端ノードに到達すると，プッシュダウンスタックを参照し，もとのグラフのプッシュダウンエッジが接続する終端ノードにポップアップする。このようなプッシュダウンとポップアップを繰り返した後，幹ネットワークの終端ノードに到達したら，これまでたどってきたエッジのラベルの並びから成る文字列がこのグラフによって受容される。幹ネットワークの終端ノードから派生するエッジ上の数値が，受理された文字列の ID である。部分グラフの探索に失敗した場合は，プッシュダウンスタックに積まれたプッシュダウンエッジの終端ノードの情報を棄却するとともに，そこまでの探索を打ち切る。グラフ探索の結果は，探索木と呼ばれる木構造に保管される。なお，探索木のノードをセルと呼ぶ。そしてそれは，図 3.3 で示すように，グラフの対応するノードへのポインタを持つ。

グラフ探索は動的計画法に基づき，文字候補ラティスの列毎に最初の文字から順に行われる。文字候補ラティスの列が処理されている時，セルと呼ばれる情報が当該列中の各文字候補に対して生成される。そして，その文字候補とグラフのエッジ上の文字の照合結果に従って，探索木が伸びていく。図 3.3 の例では，まずルートノードが生成されたあと，一文字目の候補である「京」と「中」に対応するノードがルートノードの子ノードとして生成される。また，探索木の枝は探索スコアによって枝刈りされる。探索スコアとはそれまでの文字パタン候補に対する文字識別結果の類似度の累計値である。

文字候補ラティスの最後の列の照合が終わると，文字列候補スコアに従って，もっとも確からしい文字列候補が選択される。文字列$\boldsymbol{x} = x_1, x_2, \ldots, x_n$に対して，その長さを$n = |\boldsymbol{x}|$とすると，文字列$\boldsymbol{x}$の文字列候補スコアは，文字候補$x_i$の文字候補スコア$S_c(x_i)$を用いて，

$$S_T(\boldsymbol{x}) = \rho \sum_{i=1}^{\mathrm{n}} w_i S_c(x_i)$$

と定義される。但し，ρは正規化パラメタ，w_iは重みベクトルの i番目の要素である。文字候補スコア$S_c(x_i)$は文字候補x_iに対する文字識別結果の候補順位$r(x_i)$を用いて，

$$S_c(x_i) = \lambda^{r(x_i)-1}, \qquad (0 < \lambda < 1)$$

とする。候補順位の指数で表現するのは，文字識別の上位候補を下位候補よりも重視するためである。これは，本論文では統計的文字識別手法[78]を導入しているが，学習により得られた文字毎のモデルとの統計的距離に基づき識別結果を算出するという手法の性質上，下位候補の順位の信頼性は上位候補に比べると低くなることによる。これらのパラメタの値は 3.5 節で評価に用いるデータとは別の地域の地名表記データを用いて行った実験結果より

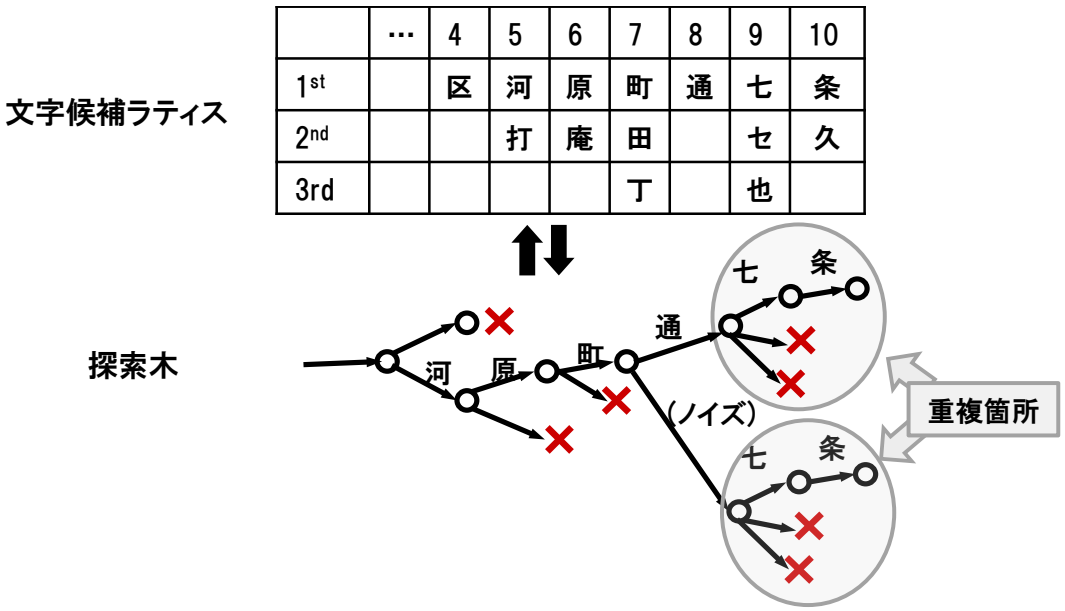


図 3.4　グラフ探索の冗長性の例

$$\lambda = 0.75, \quad w = \left({}^{0.3}/_{n_a}, \ldots, {}^{0.3}/_{n_a}, {}^{0.7}/_{n_b}, \ldots, {}^{0.7}/_{n_b}\right), (n_a + n_b = n)$$

と設定した。但しn_aは都道府県や市区町村名を表す部分文字列の長さ，n_bは街区名などを表す部分文字列の長さである。

3.3.2.　アドバンテージ遷移を伴う探索

上記のグラフ検索アルゴリズムは, 文字候補ラティス中の全ての文字と完全に一致した場合のみ，文字列候補を受理する。しかし，文字候補ラティスの内容は必ずしも正しいとは限らない。文字列探索処理の頑強性を高めるため，誤った入力文字候補に対する謝り訂正機能を設ける必要がある。この入力となる文字候補ラティスの不完全性は以下の原因によって生じる。

(1)　文字候補ラティスの列の中に正しい文字候補が存在しない(誤認識),

(2)　ノイズパタン，あるいは文字切出し誤りによって生じた余剰の候補(湧き出し),

(3)　文字候補ラティスの列の欠損

本論文では，(1)と(2)を扱う。これらの不完全性を受容するために，「アドバンテージ遷移」を用いた探索を行う。アドバンテージ遷移を用いた探索では，グラフ探索において文字候補ラティス中の文字候補とグラフの照合に失敗しても，探索プロセスがグラフの次のノードに遷移させる。

アドバンテージ遷移によって生成されたセルには，その文字列候補スコアにペナルティが加算される。但し，1 回のアドバンテージ遷移が 4 文字以上の文字列に対して行われたときのみ，ペナルティが加算されるものとする。また，1 回のグラフ探索におけるアドバンテージ遷移の回数に対象となる文字列の長さに比例した上限を設ける。

しかし,アドバンテージ遷移を用いることで,探索木のサイズが大きくなる問題がある。探索木のサイズは文字列認識に要する処理時間に大きく影響を与えるため，探索木のサイズが大きくなりすぎることを防ぐ必要がある。

3.3.3. 探索木の枝刈り

グラフ探索に要する処理時間は探索木のサイズに比例して長くなる。実験によって，アドバンテージ遷移を伴うグラフ探索は，これを伴わないグラフ探索に比べて，探索木のサイズが 6 倍になることが分かっている。探索木のサイズの増大は，アドバンテージ遷移によって異なる文字列と照合された後，文字候補ラティス上の同じパスを探索することで生じる。例えば，図 3.4 の文字候補ラティスにおいて，一位候補を順に選択した「河原町通七条」に対する探索木と,「通」の文字候補をノイズ(湧き出し)として扱うアドバンテージ遷移により「河原町七条」に対する探索木が生成される。そして，この後の「七条」に対するグラフ探索が重複してしまう。このような場合に,「七条」に対するグラフ探索を 1 回だけ行うようにすることで，グラフ探索の冗長性を低減する必要がある。

グラフ探索が冗長になっているか否かは，文字候補ラティスの同じ列において生成されたセルの情報を比較することで，同定することができる。任意の 2 つのセルがグラフの同じノードを指している場合，そのセルから生成される部分探索木は同一である。従って，冗長な探索木は，その 2 つのセルが当該ノードに至るまでの文字列候補に対する文字列候補スコアの低い方のセルを削除することで枝刈りをすることができる。このようなグラフ探索アルゴリズムを図 3.5 に示す。

```
SearchAlgorithm(M[n][m]) {                          // Input : Character candidate lattice
    Path = ϕ;                                       // Path on searching
    Goal = ϕ;                                       // Found Path
    for i = 1 to n do {                             // Repeat for lattice columns
        PlantSeed(Path);                            // Root of the search tree set
        Transition(Path, M[i]);                     // Transition on the candidates in the column i
        CheckGoal(Path, Goal);
        CheckOverlap(Path);
    }
    return Goal;                                    // Output the searched path
}

Transition(Path, M[n]) {
    Next = ϕ;                                       // Next Path
    foreach p = v_1, v_2, ..., v_k ∈ Path do {      // v_k is a node of the graph
        Flag = FALSE;
        foreach e = (v_k, w) ∈ E do                 // for edges starting from node v_k
            for j = 1 to m do                       // for character candidates in column of lattice
                if σ(e) == M[j] then{
                    add p_w to Next;                // State Transition
                    flag = TRUE;
                }
            if (flag == FALSE) and (penalty(p) < k) then
        AdvantageTransition(Next, M[m], p);
    }
    Path = Next;                                    // Update Path
}

AdvantageTransition(Next, M[m], p = v_1, v_2, ..., v_k) {
    foreach e = (v_k, w) ∈ E do {                   // for edges starting from node v_k
        add p_w to Next;
        foreach f = (w, z) ∈ E do{
            for i = 1 to m do
                if σ(e) == M[i] then
                    add p_z to Next;                // in case of missing lattice-column
        }
    add p_{v_k} to Next;
}

CheckOverlap(Path = {s_1, s_2, ..., s_n}){
    for i = 1 to n − 1 do
        for j = i + 1 to n do{
            p = s_i;                                // p = p_1, ..., p_k
            q = s_j;                                // q = q_1, ..., q_l
            if p_k == q_l then                      // Overlap detected
                if penalty(p) < penalty(q) then
                    delete q;
                else
                    delete q;
}
```

図 3.5　グラフ探索アルゴリズム

```
{
      <St.> ::= 通[り];                                  //「り」が省略可能であることを表す
      …
      <4th   St.> ::= 四条[<St>];                        //「通り」が省略可能であることを表す
      <Karasuma Ave.> ::= 烏丸[<St>];                    //「通り」が省略可能であることを表す
      <Karasuma Ave. 4th St.> ::= < Karasuma Ave.> <4th.St.>;
   begin (600-8156)
      <N600-8156> ::= <Kyoto Pref.> < Kyoto City><Shimogyo Ward>
                                    ( <Takakura Ave. Nishiki St.> |
                                      <Takakura Town> |
                                      <Karasuma Ave. 4th St.> );
                                                  // 上記3行は〈高倉通錦通〉の表記
                                                  // 〈高倉町〉の表記
                                                  // 〈烏丸通四条通〉の表記
                                                  // のいずれかで記述されることを表す
   end
}
//「京都府京都市下京区高倉」の記述例
```

図 3.6 SPDL による地名表記を表現する文脈自由言語の例

3.4. グラフ型言語モデルの半自動生成

3.4.1. 文脈自由文法による言語モデル

グラフ型言語モデルを用いた文字列認識を住所認識のような実際のアプリケーションに適用するには, 言語モデルを簡易に作成できることが必要である。グラフ構造のモデルは, テキストで記述されたモデルに比べて，人間による可読性に劣るため，グラフ構造のモデルを直接作成するのは困難である。前述の通り，提案するグラフ型言語モデルは再帰的遷移ネットワークであるため，その表現力は文脈自由文法と等価である。そこで，文脈自由言語で記述された言語モデルを人間にとって可読性の高いテキストで表現し，これをグラフ型の言語モデルに変換することで言語モデルの作成を容易化する。

本項では, バッカス・ナウア記法(Backus-Naur-Form)を拡張した言語 SPDL (Syntactic Phrase Description Language)による, 文脈自由言語での地名表記の記述について述べる。

図 3.6 に SPDL で記述した地名表記を表す文脈自由言語の例を示す。単語とその異表記が統語的カテゴリとして定義され，認識対象の文字列はこれらのカテゴリの並びとして定義される。「begin」と「end」に挟まれた式が「京都府京都市下京区高倉」の地名表記を定義する最上位の式である。その左辺の「<N600-8156>」は「京都府京都市下京区高倉」が示す地名に対する ID である。縦線“|”は選択を，カッコ“[]”はその中の文字列がオプションであることを意味する。“< >”は統語的カテゴリの名前を表し，“()”は評価の順

番を陽に表現する。

一旦，各単語とその異表記が統語的カテゴリとして定義されると，それらを他の文字列で表現する際に，その統語的カテゴリのラベルを参照することより，新しい文字列に対して異表記を含んだ形で表現できる。図 3.6 に示す地名表記の場合，「り」の有無を受容するように統語カテゴリ<St>::=通[り]を定義することで,「四条通」という表記に対して,「四条通り」という表記も異表記として含めることができるようになる。これによって「京都府京都市下京区高倉」に対する地名表記が文脈自由言語で記述される。

3.4.2. グラフ型言語モデルの生成

異表記を含むグラフ型言語モデルを効率的に生成するために，上述したような異表記の追加を文脈自由言語上で行った上で，これをグラフ構造に変換することが有効である。この際の異表記追加を可能な限り自動化し，言語モデル作成における人間の負荷を軽減するために，図 3.7 に示すような，文字列の集合から文脈自由文法の言語モデルへの変換と，文脈自由文法の言語モデルからグラフ型言語モデルへの変換の二段階からなる統語的言語モデルの獲得のための方法を提案する。第一段階の SPDL テキストの生成のために以下の 4 つの情報を用いる。

(1) 異表記を含まない地名表記のリスト(文字列の集合)
(2) 異表記に関するヒューリスティクス(異表記追加規則)
(3) 人手により追加される異表記のサンプル
(4) 地名表記認識器によってリジェクトされる異表記

異表記を含まない地名表記のリストはまず SPDL テキストに変換される。次に予め用意された規則(ヒューリスティクス)に基づき異表記が追加される。この規則によって正しい異表記が生成されない場合，人手により SPDL テキストに異表記を追加することもできる。文字列認識器は文字列の認識の際，言語モデルに含まれていない文字列は認識することができず，リジェクトしてしまう。その文字列が本来認識すべきものである場合は，これを人手により追加するとともに，事例ベースとして再利用するために蓄積される。このように，言語モデルは修正され，再構築されていく。

第二段階において，言語モデルのグラフ表現が生成される。SPDL テキストが入力となるため，これをグラフ表現に変換することは，文脈自由言語を解析するコンパイラによって可能となる。

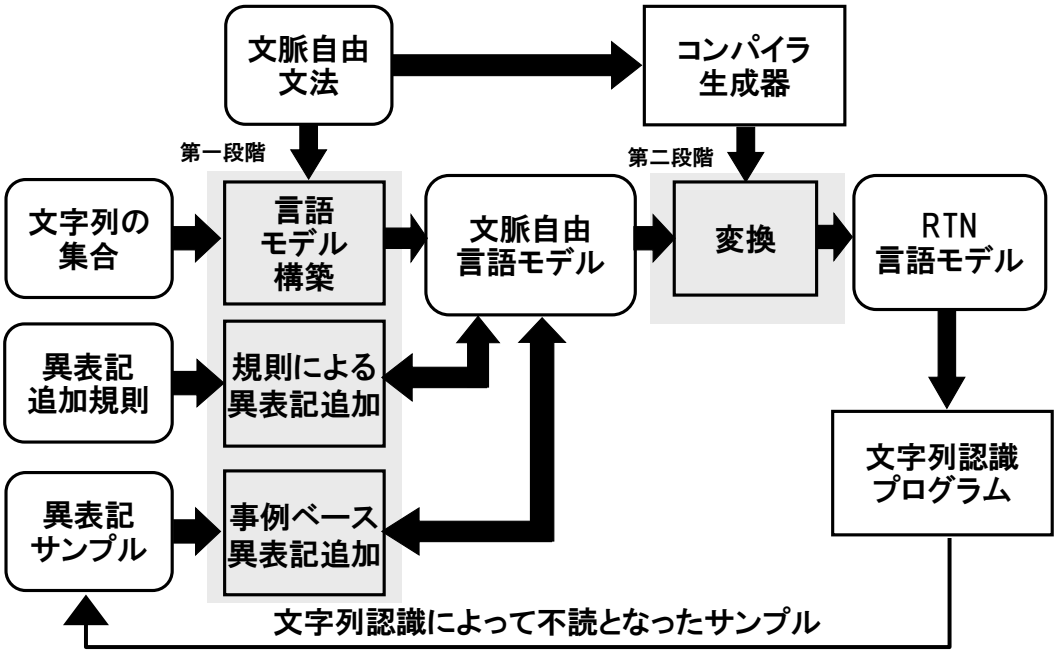


図 3.7 言語モデル生成の手順

(1) SPDL で記述された言語モデルの生成

言語モデルの生成は 3 つのステップからなる。第一は，異表記のない地名表記のリストから SPDL を生成することである(図 3.5(a)参照)。地名表記は「市」のような特定のキーワードによって，単語に分割され，各々の単語は非終端記号となる。そして，地名表記は非終端記号の系列として記述される(図 3.5(b)参照)。

第二に，地名表記の異表記をルールによって自動的に追加する。ルールの条件部が満たされた場合, ルールの帰結部が当該表記の記述として追加, あるいは, 置換される(図 3.5(c)参照)。定義文の右辺のある文字列を統語的カテゴリのラベルで置換することで，異表記が SPDL に追加される。例では「ケ」が「ヶ」や「が」を受容するよう「<Ga>」に置換されている。

第三に, ルールによって生成することのできない異表記が SPDL 文を入力することで追加される。図 3.5(d)の矢印の上部にある「東町」に関する定義文が該当する。あるタイミング，例えば，SPDL テキストから再帰的遷移ネットワークに変換される際に，入力した式の右辺(この例では東町)は，元の式の右の項と記号「|」によって結合される。そして，その行の先頭に，この行がアドホックに追記されたことを示すタグが挿入される。

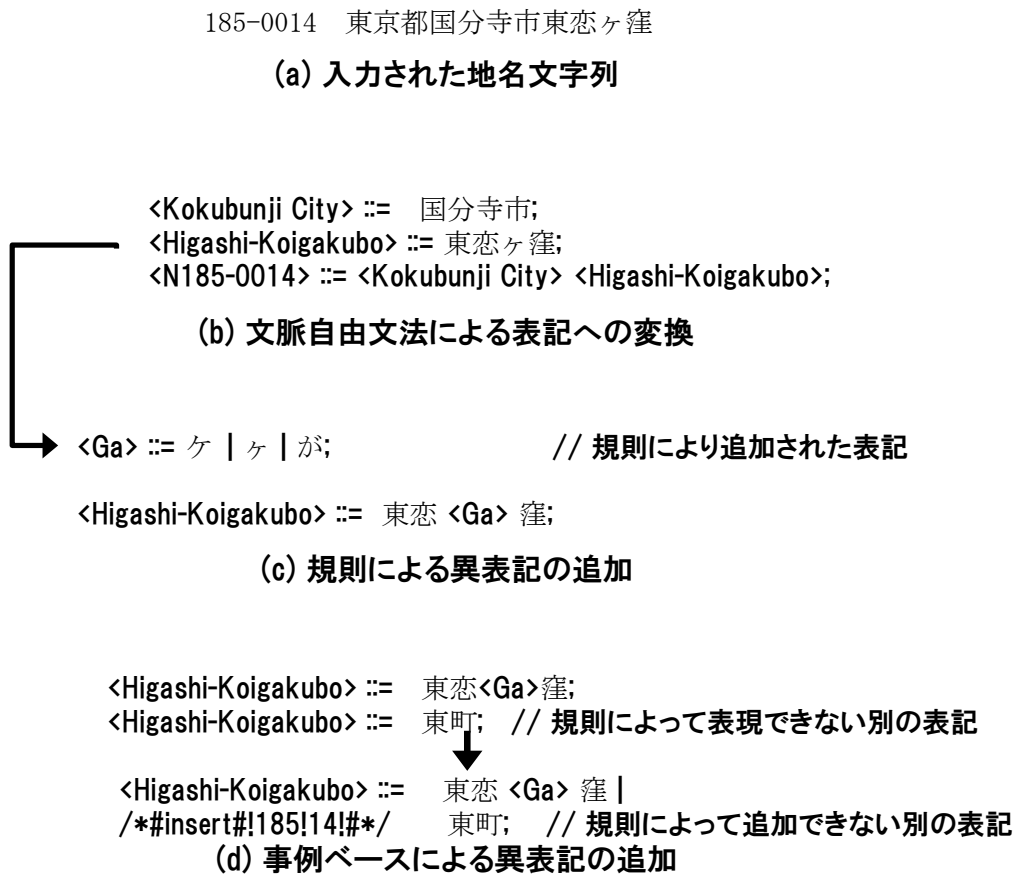


図 3.8　SPDL テキストの生成例

(2) グラフ表現への変換

次に SPDL テキストはグラフ表現に変換される。上述したように，提案するグラフ型言語モデルの文字列の表現力は,SPDL によって記述される文脈自由言語と等価であるので,文脈自由言語からグラフ型言語モデルへの変換プログラムは yacc や lex[79]といったコンパイラ生成器によって自動で生成することができる。このとき，3.3 節で述べたグラフ探索処理を高速化するために，文脈自由言語から再帰的遷移ネットワークを構築する際に,グラフの根から近いノードからの分岐数がグラフの葉に近いノードのそれよりも小さくなるように最適化する。例えば，図 3.5(d)における「東恋ヶ窪」を定義する「東恋<Ga>窪 | 東町」に対して,「|」の両側の表記に対してエッジを作成して，これを同じノードから派生させるのではなく,「東」に対するエッジを作成して，その後段に「恋<Ga>窪」と「町」に対するエッジを派生させる。

表 3.1 グラフ構造とトライ構造による言語モデルの比較

言語モデル			京都市外(地域 A)	京都市内(地域 B)
異表記追加せず	表記数		681	8,400
	提案手法(グラフ型)	ノード数	580	6,753
		エッジ数	581	5,667
	文献[70](トライ型)	ノード数	2,211	36,498
		エッジ数	2,212	36,499
異表記追加後	表記数		8.010	2,861,056
	提案手法(グラフ型)	ノード数	1,179	16,995
		エッジ数	1,283	29,193
	文献[70](トライ型)	ノード数	26,216	9,417,171
		エッジ数	26,217	9,417,172

3.5. 評価

3.5.1. グラフ型言語モデルの表現力とメモリ効率

グラフ型言語モデルのメリットは，必要とするメモリサイズの小ささと，これによって表現可能な地名表記の網羅率の高さである。まずグラフ構造で表現された言語モデルのメモリサイズの効率性を評価する。一般的な地名表記(地域 A)と京都市内の地名表記(地域 B)に対する言語モデルの 2 つを対象として評価を行った。地域 A では，郵便番号によって区別される地名が 178 箇所存在し，その地名表記の平均文字列長は 8.8 文字である。これに対して地域 B では，516 箇所の地名が存在し，その平均文字列長は 16.8 文字である。これらの 2 地域に対して異表記なしの場合と異表記を含む場合の 2 種類の言語モデルをそれぞれ構築した。但し前者についても都道府県名と市町村名の省略は許容するモデルとした。もう一方は，提案手法によって生成された異表記を追加したモデルである。比較のために，同じ表記を網羅するトライ構造の言語モデルも用意した。グラフ構造とトライ構造のそれぞれの言語モデルについて，ノードとエッジの数を表 3.1 に示す。異表記を含むグラフ構造の言語モデルのノードとエッジの数は，含まないモデルに対して，地域 A は 2 倍，地域 B は 5 倍に増加した。その一方で，トライ構造のモデルについては，地域 A，B に対して，それぞれ，12 倍，258 倍に増加した。このとき，地域 B のモデルに含まれる地名表記の

数は，異表記を追加することによって，340 倍に増加しているのに対して，グラフのノードとエッジの数は高々5 倍にしか増えていないことが分かる。グラフのノードとエッジの数の増加が抑えられたのは，グラフ構造によって，多くの地名表記に共通して表れる部分文字列を効率よく表現できているためである。

文字列読取りシステムの実装にあたり，言語モデルに割り当てられたメモリサイズに対して，異表記を追加しない場合は，グラフ型，トライ型の両方の言語モデルともに保持できたが，異表記を追加した場合，トライ型については，割り当てられたメモリ量を超過して実装することができなかった。

3.5.2. 文字列照合の処理速度とグラフ探索型地名表記文字列の認識精度

文字列照合処理は入力である文字候補ラティスから言語モデル中に含まれる文字列を探索する処理であるので，言語モデルが同じ文字列集合を表現している限り，言語モデルの構造にかかわらず認識精度は同じである。ここでは言語モデルの構造が異なる場合の文字列照合を含む文字列照合の処理速度を評価する。

前述のとおり京都市内の地名表記(地域 B)に対して，異表記を追加したトライ型の言語モデルはメモリ量の制限から計算機上で処理できなかったため，同地域における異表記を追加しない言語モデルを使用する。トライの探索は表記数に比例して処理時間が増加するので，追加すべき異表記数について増加する処理時間を推定して提案手法と比較する。その際の地名表記あたりの平均の文字列照合の処理時間は 1.0msec であった。地域 B は異表記を追加することで表記数が 340 倍になるので，処理時間も約 340msec(= 340 × 1.0)になると推定される。これに対して，異表記を含むグラフ型言語モデルを用いたグラフ探索による文字列照合の処理時間は，地名表記あたり平均で 38.1msec であった。この結果，グラフ探索型の文字列照合は，トライ探索型の文字列照合よりも高速に実行できる見込みを得た。

次に言語モデルへの異表記の追加に関する効果を調べる。京都市内とそれ以外という二か所の地名表記に対する文字列認識精度の比較を表 3.2 に示す。正読とは正しい認識結果を出力した場合，誤読とは誤った認識結果を出力した場合，不読とは信頼度が閾値に満たない等によりどの地名表記にも該当しないと判断する場合を表す。地名表記文字列を言語モデルへの異表記の追加によって，文字列認識の正解率において 8.5pt から 43.3pt の向上が見られ，誤認識率において 0.2pt から 16.9pt の低減が確認できる。

表 3.2　グラフ型言語モデルの異表記の有無による地名認識精度の比較

対象地域	結果	異表記なし		異表記追加	
地域 A(簡単) 総計 1,781	正読	1,357	76.2%	1,508	84.7%
	誤読	25	1.4%	22	1.2%
	不読	399	22.4%	251	14.1%
地域 B(複雑) 総計 1,854	正読	607	32.7%	1,395	76.0%
	誤読	360	19.4%	45	2.5%
	不読	887	47.8%	385	20.8%

次に，アドバンテージ遷移の効果と冗長な探索の枝刈りの効果を評価する。上記の 2 地域の地名表記の認識精度をアドバンテージ遷移の有無により測定し，その結果を表 3.3 に示す。表中の条件の「候補なし」と「湧き出し」は，アドバンテージ遷移の対象として想定する，文字候補ラティスの列の中に正しい文字候補が存在しない場合と，ノイズパタンや文字切出し誤りによって余剰の候補が生じた場合のことである。それぞれの場合にアドバンテージ遷移の実行の有無を切り替え，文字列認識精度を測定した。京都市内の地名表記(地域 B)に対しては，アドバンテージ遷移による誤認識した文字の訂正の効果が顕著であった。その一方で，京都市以外の一般的な地名表記(地域 A)に対しては，効果が確認できなかった。京都市内の地名表記にのみ効果が顕著であった理由として，以下が考えられる。

(1) 一般的な地名表記に比べて京都市内の地名表記が長く，1 文字あたりの情報量が少ないためアドバンテージ遷移により読み飛ばしを行っても，他の部分で補完できる余地が大きい。

(2) 京都市内の地名は通り名と町名を用いて表記されることが多く，冗長性をもつため，一方が認識できずアドバンテージ遷移によりスキップされても，他方を認識することで地名表記の認識が成功する可能性がある。

文字列照合の処理時間の観点からの冗長なグラフ探索の枝刈りの効果についても評価した。一般的な地名表記に対して，平均処理時間は 45.5%減少し，京都市内の地名表記に対しては 48.1%減少することが分かった。

表 3.3　アドバンテージ遷移の効果

	アドバンテージ遷移実行条件		正読		誤読		不読	
	候補なし	湧き出し						
地域 A 1,781 サンプル	なし	なし	1,507	84.6%	19	1.1%	255	14.3%
	なし	あり	1,507	84.6%	20	1.1%	254	14.3%
	あり	なし	1,509	84.7%	21	1.2%	251	14.1%
	あり	あり	1,508	84.7%	22	1.2%	251	14.1%
地域 B 1,854 サンプル	なし	なし	1,247	68.0%	20	1.1%	556	30.3%
	なし	あり	1,283	69.9%	26	1.4%	513	28.0%
	あり	なし	1,348	73.5%	62	3.4%	411	22.4%
	あり	あり	1,395	76.0%	45	2.5%	381	20.8%

3.5.3. グラフ型言語モデル生成手法

言語モデルにおける地名表記の網羅率を評価するために，そのモデルが表現することのできる表記数を調べた。人手によって書かれた実際の帳票から収集した地域 B の地名表記 6,578 個に対して，提案するグラフ型言語モデルによる網羅率を調べた。その結果を表 3.4 に示す。提案手法によって異表記を言語モデルに追加することによって，収集した地名表記に対する言語モデルの包含率は 34pt 向上していることが分かった。また，2 年間をかけて人手で異表記を追記してきた同じ地域に対するトライ型の言語モデルには 32 万の表記が含まれていた。これは提案手法によって生成された言語モデルに含まれる表記数(2,861,056)の約 1/9 である。

表 3.4 提案する言語モデルの表現能力

言語モデル	言語モデルが包含する表記数	実サンプル中の表記の包含数	包含率
異表記なし	8,400	4,249	64.6%
異表記追加	2,861,056	6,488	98.6%

3.6. 結論

本章では，文字列認識と解釈のために，言語情報の表現，獲得，そして利用の方法について提案した。再帰的遷移ネットワークを用いた言語モデルとすることにより，認識対象の文字列表記の多様性を言語モデルの中に少ないメモリ容量で効率的に含めることができる。さらに文字列照合の処理時間を試算した結果，提案するグラフ型の言語モデルを用いた文字列照合は，従来のトライ探索型の文字列照合よりも高速に実行できる可能性を確認した。

さらにグラフ型言語モデルを用いた文字列照合において，文字識別結果の誤りやノイズパタンの混入といった不完全な入力に対してもこれを許容してグラフ探索を行う，アドバンテージ遷移を提案し，文字列認識率向上に寄与することを確認した。文字切出しの誤りや文字認識の誤りによって生じた文字候補パタンを許容することで，頑強な文字列認識・解釈が可能となる。

再帰的遷移ネットワークと文脈自由文法の表現力の等価性に着目し，認識対象の文字列集合を文脈自由言語で表現し，これを再帰的遷移ネットワーク構造に変換することで，人間にとっての可読性を高め，言語モデルの構築や保守を容易にした。文字列表記の異表記は，予め定義したルールと，人手による操作によって，言語モデルに追加される。人手によって追加された異表記は，収集，保存され，次の言語モデル生成の際に，事例ベースとして利用される。地名読取りのような実際のアプリケーションにおいては，認識のための言語モデルの生成コストは重要なファクタである。

本章で提案したグラフ型の言語モデルとその利用法は，人名や企業名といった文字列の認識においても適用可能である。

今後の課題として，文字列認識システムのメモリ容量の制限から評価できなかった，異表記を含んだ言語モデルに関して，従来のトライ型の言語モデルとの処理時間，認識精度の精緻な評価を行うことがあげられる。また，使用できる計算リソースの増大に伴い，HMM 等の確率的遷移を伴うモデルで認識対象文字列を表現することも可能になると思われる。こうした検討についても地名表記のデータの収集と共に進めていく必要がある。

さらに，今回提案した言語モデルで導入した再帰的遷移ネットワークは，文脈自由文法に加え，プッシュダウンオートマトンと等価な表現能力を持つことが知られている。異表記を含む文字列集合の記述とその照合について，オートマトン理論の観点での解釈を与えることも今後の課題としたい。

第4章

手書きによる情報共有に基づく設備保全作業支援システム

4.1. 緒言

本章では，設備保全業務において作業者が協調作業をする際に重要な役割を果たす，手書きによる情報共有を取り上げる。作業者が筆記する手書きアノテーションの解釈を行い，担当者間の情報供給を支援する方法を提案するとともに，これを設備保全作業支援システムとして実装し，実際の保全業務に適用した結果を評価する。

水道，電力，ガスや化学工場などのプラントでは安定的な稼働を維持するための保全業務が重要である[81]。保全とは，エネルギーや製品の産出といったその設備の目的を遂行するために，設備の稼働状況を把握しつつ，障害が発生，あるいは発生しそうな箇所があれば，代替策の実行や修理を行うことにより，設備の安定した運用を支える業務である。

ここで課題となるのが，設備の状態に関する保全担当者間での情報共有である。設備の安定運用のためには，日々変化する設備の状態に関する情報をもとに，担当者がその場で適切な対応を取らなければならない。プラントの保全業務は複数の担当者によって行われることが多く，設備の状態に関する最新の情報を把握しておくことが重要である。

多くの設備保全の現場では，事務所でのミーティングなどで，担当者が現場で取得した設備状況に関する情報を口頭で申し送ることにより，他の保全担当者との情報の共有が行われている。その際に，ホワイトボード等に現場で手書きした情報を再度書き出すこともあるが，ここで共有すべき情報が失われてしまうことがある。さらに，事務所で申し送りされた情報は現場で参照することができないため，保全担当者が現場に赴いた際に，申し送りが必要な情報が想起されないリスクがある。タブレット機器を携帯すれば，こうした情報を随時提供できるが，重量や汚れの問題から，導入可能なプラントは限定的である。このように保全現場で情報が共有されない可能性があるという問題がある。

情報共有の実現は，設備の状態に関する情報の取得，取得した情報の解釈，保全担当者への情報の提供，の 3 つの要素からなる。紙とペンという現場の保全担当者が慣れた道具を使った業務形態を変えずに情報共有を実現するために，本論文では，デジタルペンと紙を用いた設備状態に関する情報の取得と現場での担当者への情報提供を行う保全業務支援方法について検討する。そして，保全担当者による手書きの文章や図による申し送り事項の意味を計算機が解釈，保持するための手書きアノテーション解釈技術を提案する。提案

方式では申し送り事項として共有すべき情報を他の保全担当者が携行する紙へ印刷し，現場での保全担当者間の情報共有を確実にする。さらに，提案方式を応用した設備保全支援システムについて述べる。

以下，4.2 節において，設備保全業務とその課題について整理したうえで，これを解決する設備保全支援システムを実現するアプローチについて述べる。4.3 節で提案するデジタルペンを用いた設備保全支援システムについて述べる。4.4 節で，開発したシステムを実際の設備保全業務に適用し，評価した結果を示し，4.5 節で結論を述べる。

4.2. 設備保全業務とその課題

4.2.1. 設備保全業務の現状

一般的なプラントなどにおいては，複数の処理が連続して行われるため，各処理に対応した機器が並び，その規模も数十メートルから数百メートル四方と広域にわたることが多い。そこでの設備保全とは，各機器が設置された現場を巡回して，これらの点検を行い，障害があればその対策や修理を行う業務のことである。

本章では，設備保全業務の例として筆者の所属する研究所における廃水処理プラントの保全業務を取り上げる。この廃水処理プラントは，研究室で排出されるフッ酸排水に各種の薬剤を添加し凝集沈澱，ろ過・吸着の各処理を行い，下水道に放流可能な水とする設備である。図 4.1 に示すように，各種の薬剤を排水と反応させたり，中和させたりするための水槽や撹拌するための翼，処理水を水槽に注入・排出するためのポンプ，弁などから構成される。通常は朝にプラントを稼働し，夕方に停止される。設備保全の担当者は，プラントの始動と停止作業を行う他，始動時，停止時とその中間を加えた 1 日 3 回，プラントの各設備の稼働状況を事務所から離れた(徒歩数分)現場に出向いて確認する。稼働状況の確認とは，各部位の目視確認の他，各水槽に投入されている薬剤の残量，ポンプの圧力，水槽の PH の値の記録等である。これを運転日誌と呼ばれる，日毎に用意する A4 一枚の帳票に記入する。こうした運転日誌は自治体の条例によって長期間の紙での保存が義務づけられている。他に，「薬剤の残量が少なくなったため追加した」といった保守作業の記録や，水槽への排水の注入用の弁の動作不良などを発見した際には，修理・交換が必要との記述を申し送りとして記入する。これらは運転日誌の特記事項欄に文章や図を用いて記入される。

点検の記録，特に薬剤の残量やポンプの作動時間のような，処理した排水の量に直接関わる部分については，毎日夕方に保全作業員の事務所に設置された PC に入力，管理され

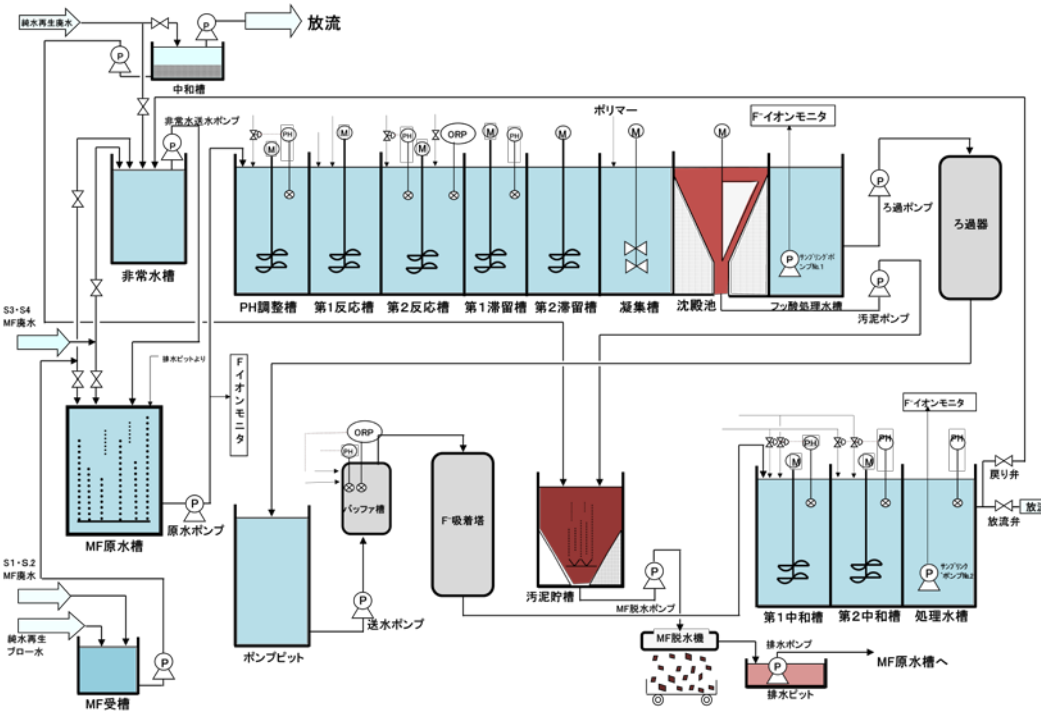


図 4.1 廃水処理プラントの構造

る。これらの情報は月次でまとめて設備保有者である研究所に報告される。

上記の保全業務は，複数の担当者で行われる。担当者間，および取り纏め者との間の情報共有は，朝と夕方にそれぞれ朝礼，夕礼として保全作業事務所で行われる。特に薬剤の消費量が増加している(つまり，研究室での排水の量が増加している)ため，薬剤の追加の時期が近い，あるいは薬剤が不足しそうな場合は発注処理を行うといった申し送り，あるいは，ある弁に異音がするので様子を見ておいてほしいといった申し送りがなされる。薬剤切れ，あるいは機構の故障により廃水処理プラントが停止することは，研究の進捗を止めることになるため，プラントの稼働維持のためには，このような気づきとその申し送りによる，予めの対応が重要である。

このような日次の保全業務の流れを図 4.2 に示す。上で述べたように，この保全業務は紙文書と電子文書の両方を用いて行われている。現場での点検，記録は紙の帳票を用いて行われており，事務所での廃水処理量の管理は電子文書(PC)を用いて行われている。そして，PC へのデータ入力は，紙の運転日誌の記載内容の人手によるタイプ入力という形で行われている。また，薬剤切れや異常などの申し送りについては，口頭での伝達と図 4.3 に示すような事務所に設置されたホワイトボード上に保全担当者が「RO 純水供給(P)異音」といったように改めて手書きしなおすことで行われている。すなわち，現場の運転日誌(図 4.2 下部右)の情報の一部は日々の検査記録のデータベース(図 4.2 上部右)に人手によって

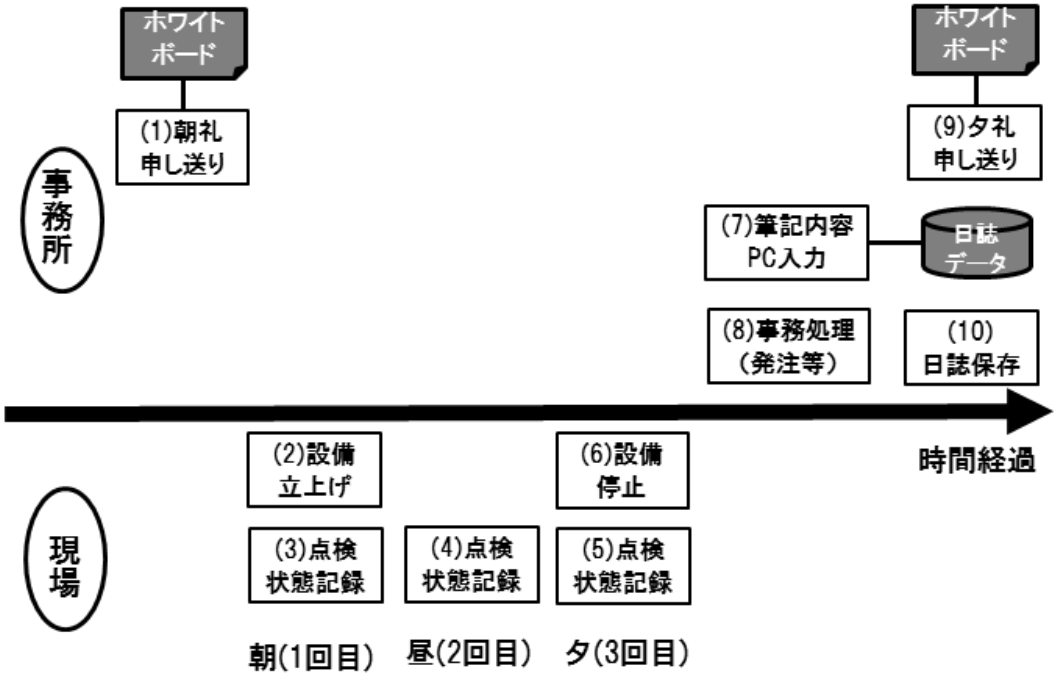


図 4.2　現状の廃水処理設備保全作業の流れ

複写されているが，事務所のホワイトボード(図 4.2 上部)については，これらとは別個の文書として人手によって作成・利用されている。このような状況は，今回例に挙げた保全業務に特別なものではなく，他の現場業務においてよく見られる状態である。

4.2.2.　設備保全支援に関する関連研究

設備保全業務における，設備稼働状況の取得に関して，様々な手段を用いて支援する試みがある。設備の稼働状況に関する情報の取得手段として RFID を用いた研究に，桑名[82]や矢吹[83]らの研究がある。前者は原子力発電プラントの弁ハンドルロック用開閉鍵に，後者は街路樹に RFID を付与することで，多数存在する開閉鍵や街路樹の保全担当者が取り違えることを防ぐ。

渡辺らは音声入力による保全データの入力を提案している[84]。音声を用いることで保全担当者によるハンズフリーの入力が実現される。しかし，一般にプラントでは装置の作動音などが大きく，このような騒音の中でも保全担当者の音声を誤りなく認識できる音声認識技術はまだ実現されていない。

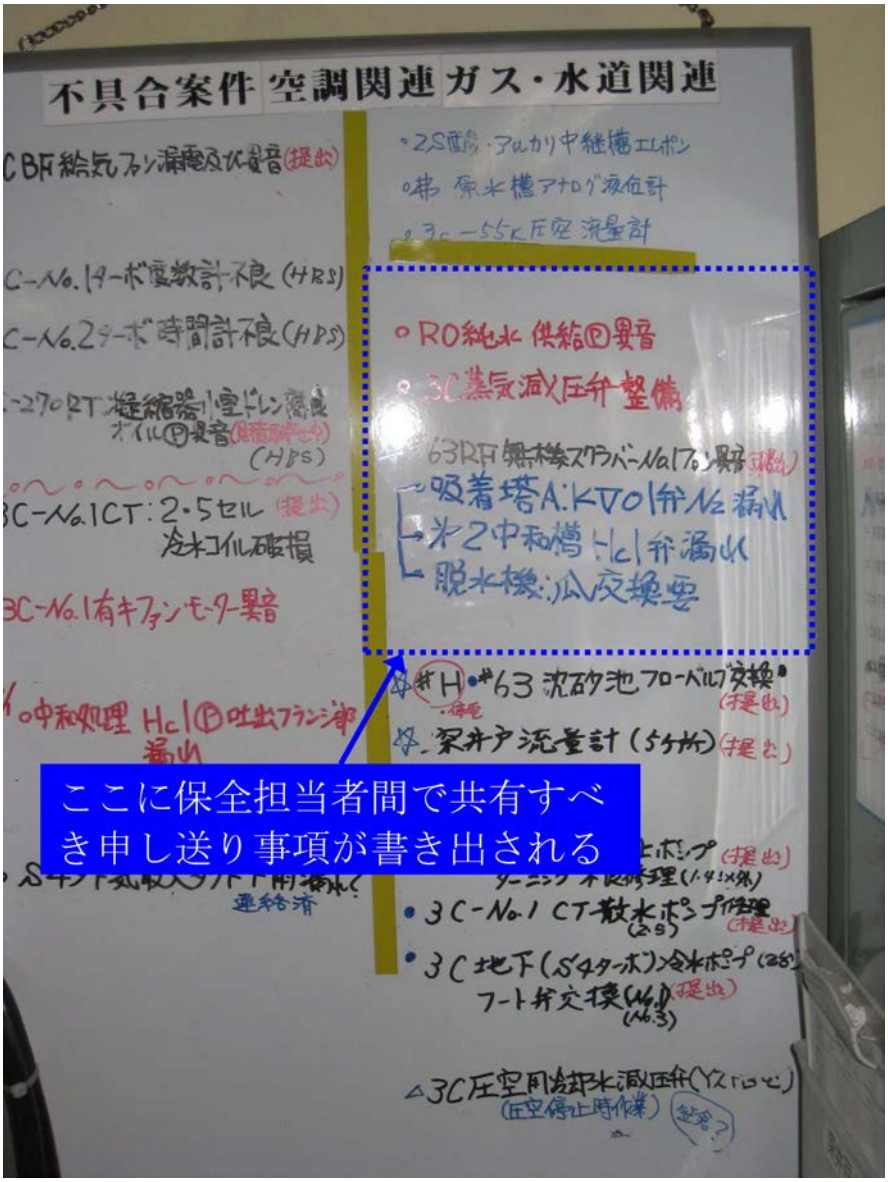


図 4.3　申し送り事項共有のための事務所の白板の例

Ziegler らは, ウェアラブル機器を用いて, ディスプレイへの保全担当者への情報提示と, ダイヤルとボタンを組み合わせた方法による情報入力機能を備えた保全作業支援システムを提案している[85]。携帯性に優れるが，ユーザの入力は提示されたリストからの選択に限られる。

これらの研究では，設備情報の取得に関しては有効な場合があるものの，紙とペンによる直感的な操作に勝る決定的な情報入力手段とはなっていない。さらに，設備の状態に関する情報を保全担当者間で共有することについては考慮されていない。

斎らは，タブレット機器上に表示される地図や CAD 図面に対して，保全現場でメモ書きを行い，これを業務システムに反映させるシステムを提案した[86]。タブレット機器は近年軽量小型化しているものの，取り扱いに注意を要する精密機器のため，作業員の手をふさぐことが多く，また，雨天の屋外や油や粉塵が舞う現場では使用できない。この研究では手書き情報の共有も実現されているが，筆跡情報を業務アプリの画面に紐づけて保持しており，保全担当者が現場で特定の設備に関する情報を参照するには，システムを検索するという操作が都度必要になる。

町保らは，設備保全現場にカメラを設置し，これをネットワークでサポートセンタと接続し，サポートセンタの監督者が音声により遠隔指示を行えるようにすることで，現場との情報共有を行う手法を提案している[87]。しかし，これはサポートセンタに監督者がいる場合にのみ有効である。

4.2.3. 設備保全業務における課題

紙文書を用いて現場の点検作業を行うことに対しては，作業員の慣れ，および追加の機器を持つことにより点検作業への支障がでることを避けるために，運転日誌に手書きで記録するという形態に加え，運転日誌の書式に至るまで，大きく変えるべきではないとのニーズがある。しかしながら，現在の業務形態には，運転日誌に記載された申し送り事項の活用について課題がある。

現在，保全担当者は設備の異常やその気づきを運転日誌に筆記しているが，これを他の担当者との間では, 事務所のホワイトボードに書くか, 口頭で共有されているのみである。しかし運転不可能な故障や運転不可能な状態になる可能性が高いと認識されている事象でなければホワイトボードや口頭で共有されず，潜在的な故障や異常の原因が見過ごされてしまうという問題がある。また，こうした情報は，確実に共有しなければならないが，共有すべき情報をホワイトボードや口頭での共有の際に，保全担当者が現場にて目にすることができないため，担当者が忘却してしまうリスクがある。これらのリスクを回避する手段が必要である。

4.3. 設備保全支援システム

4.3.1. 課題解決に向けたアプローチ

4.2 節での設備保全業務の現状分析を通じて，保全現場で担当者間が共有すべき情報として運転日誌上に筆記された手書き情報が，事務所での朝礼等においてホワイトボード上

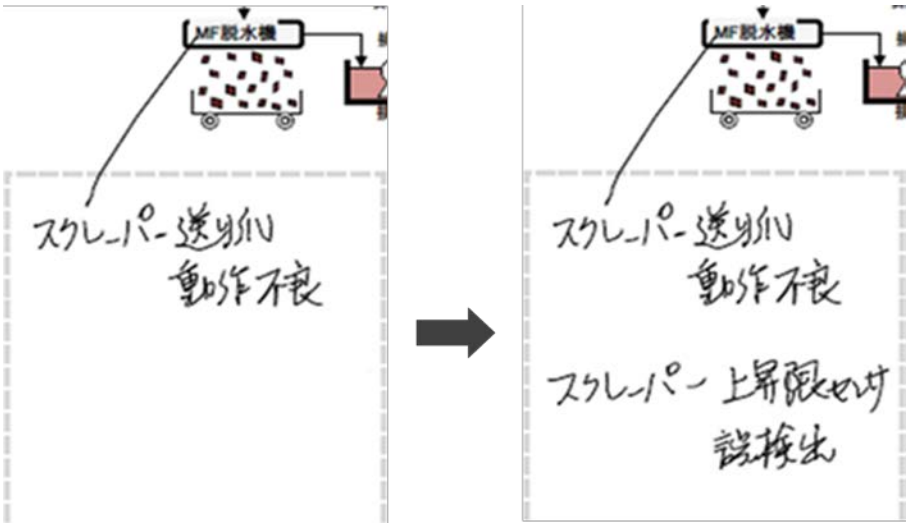
に新たに筆記されていること，さらに，こうして保全担当者間で共有された情報が，別の保全担当者が現場に赴いた際に文字として参照できなくなっていることが分かった。こうした状況に対して，手書き情報を現場と事務所で共有することができれば，設備の故障や異常の見過ごしや，申し送り事項の忘却のリスクの軽減につながると考えられる。そのために，過去に筆記された手書きに保全作業者が注釈(アノテーション)として追記していくことで，手書き情報の共有を実現する手法を検討する。これにより，保全担当者の現場での作業形態を変えることなく，かつ，事務所での人手による書き直しの手間を取ることなく，さらに，保全担当者にとって追記した情報とそれ以前の情報との対応関係の把握が容易になると期待できる。本研究では，紙への手書き入力を起点として，現場での保全担当者によって手書きされた運転日誌の情報，事務所で再度手書きされる情報共有のためのホワイトボード上の情報，そして計算機上で管理される設備稼働状況に関する情報，の 3 つを一体で管理できる，設備保全支援システムを提案する。

筆記データの入力には Anoto 社のデジタルペンを用いる 。市販されている他のデジタルペンと同様，事務所で運転日誌をスキャンすることなく，筆記時刻と紙上の筆跡情報を位置データとして計算機上に取得できる[88]。

保全担当者の現場での気づきの共有のために，既に印刷された情報に対するアノテーションとして筆記された特記事項を電子化し，構造化することで再利用可能にする手書きアノテーション解釈方式を提案する。本方式では，電子文書，およびそれを印刷した紙文書，紙上に追記されたアノテーションといった情報の位置を含む対応関係を保持できるハイブリッド文書[89][90][91][92][93][94][95][96]と呼ぶデータ構造を導入することで，手書きアノテーションと設備保全情報データを紐づける。運転日誌上に特記事項として記入された気づき，申し送り事項を電子化する際に，筆記の位置や引き出し線などから，筆記に対応するプラントの部位を推定し，各筆記を部位ごとにまとめてスレッド化する。さらに，同じ部位に関する申し送り事項を時系列にまとめ，これを後日の担当者のために運転日誌上に再度印刷する。これにより，設備異常の発生後に補修などの対応が取られているかが確認しやすくなり，保全業務の抜けなどが防げる他，保全担当者がその変化を認識し，設備異常やその予兆に気づきやすくすることで課題を解決する。

4.3.2. 手書きアノテーション解釈方式

本論文で対象とするのは，図 4.4(a)に示すような，予め印刷された図面に対して，引き出し線を書いて，「スクレーパー送り爪動作不良」と記載している場合である。これに対し



(a) (b)

図 4.4　手書きアノテーションの例

て，「MF 脱水槽で○月○日に『スクレーパー送り爪動作不良』が発生している」という情報を生成し，保全担当者が参照できるようにする。この記載が行われた以降の保全担当者が携行する帳票には，この記載(アノテーション)が印刷される。そして，図 4.4(b)に示すように，(a)のようなアノテーションに対して，別の担当者によって「スクレーパー上昇限センサ誤検出」というアノテーションが追記された場合，「○月○日に『スクレーパー送り爪の動作不良』が生じている MF 脱水槽で，□月□日に『スクレーパーの上昇限センサ誤検出』が発生している」という情報を生成する。なお，「スクレーパー送り爪動作不良」等の文字列については，今回は文字認識を行わず，プラント設備のどの部位を対象としているかを特定した後，手書きのイメージを保全担当者に見せるのみとする。今回のように自由に書かれたアノテーションの文字認識については，精度面からまだ人手による認識結果の確認が必要となり，業務フローが変わってしまうため，導入効果の評価が難しくなることが理由である。

手書きアノテーションの解釈方法を図 4.5 を用いて述べる。またそれぞれのステップについて説明する。

<S1> 入力された手書きストロークに対して，引き出し線(非文字)と文字列を区別。近傍のストロークの大きさの分布における分散に基づき，手書きストロークの相対的な大きさから閾値処理を行うとともに，この閾値近辺の文字か非文字かを判別しにくいストロークに対しては，文字認識処理を適用することにより文字，非文字を分類する。

このステップは，デジタルペンを用いてある連続した時間内で筆記された手書きストロークが入力となる。図 4.4 のように，引き出し線のストロークは文字のストロークに対して単純であるため，単位時間あたりに筆記されるストロークのサイズ(ストロークを包含する矩形の面積)は大きくなる。あるストロークが文字か非文字かを判別する際，そのストロークの近傍ストロークのサイズと比較する。近傍ストロークのサイズの分布に対して注目するストロークのサイズの比が閾値以上の場合は非文字，それ以外の場合は文字と分類する。

<S2> 引き出し線(非文字)が検出されたかどうかを判定。

<S3> 非文字が検出された場合，印刷された各文字列に対して，引き出し線との距離を計算。

ハイブリッド文書を参照し，元の電子文書から文字列を抽出し，その印刷位置を計算する。これと当該非文字ストロークの端点の紙上の位置を比較し，その距離に反比例するスコアを付与。

図 4.4 の例では，手書きの引き出し線と印刷された文字列「MF 脱水機」や，(図中には含まれないが紙上には存在する)他の印刷された文字列との距離をもとにスコアが計算される。

<S4> 引き出し線の近傍に印刷された文字列が存在するかを判定。

近傍に存在しなければノイズストロークとする。

<S5> 近傍にあると判断されたた場合，引き出し線と同時に筆記された文字列を用いて新しいアノテーションリストを生成。

図 4.4(a)においては，「スクレーパー送り爪動作不良」という手書きアノテーションをもとに新しいアノテーションリストが生成される。

<S6> <S2>で非文字が存在しない場合，既抽出のアノテーションリストが近傍に存在するかを判定。

存在しない場合は，当該アノテーション(文字列)を用いて新しいアノテーションリストを生成。

<S7> 存在する場合，そのアノテーションリストに当該アノテーションを追加。

図 4.4(b)において，「スクレーパー上昇限センサ誤検出」が追記された際に，ここに非文字成分を含まないと判断されるため，近傍の「スクレーパー送り爪動作不良」を含むアノテーションリストに追加される。

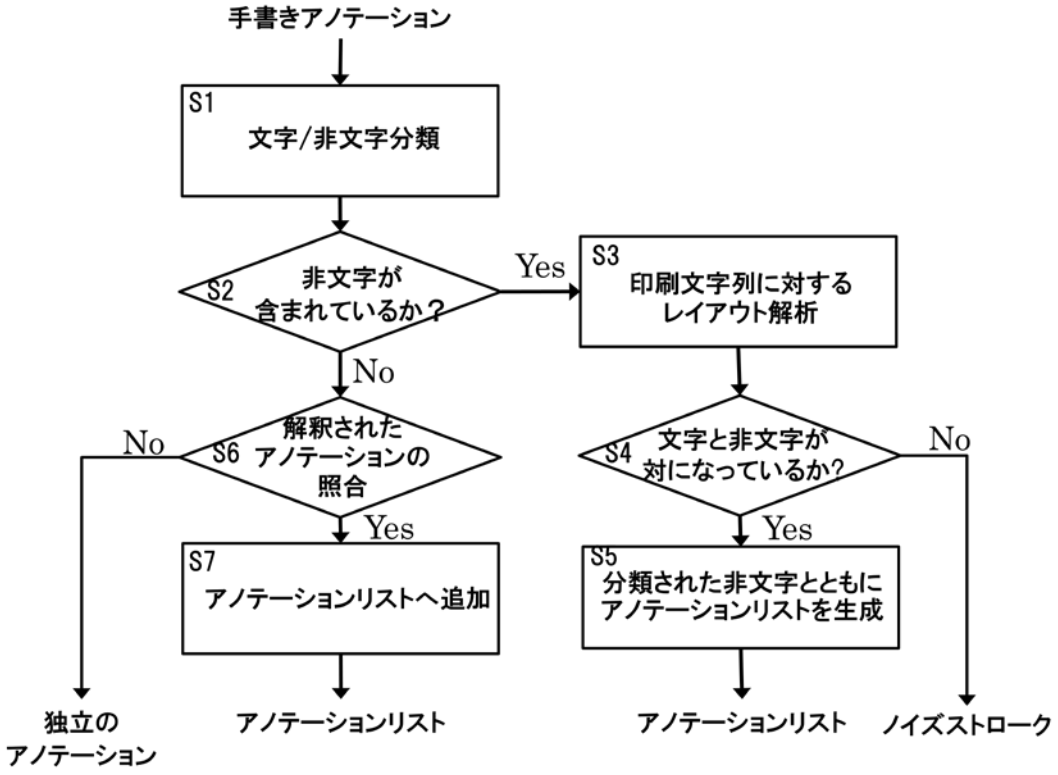


図 4.5　アノテーション解釈アルゴリズム

ここでアノテーションリストとは，プラント設備の部位，手書きアノテーションのイメージ情報，その記入日などから構成される，図 4.6 のようなリスト構造の情報である。リストの先頭は，申し送り事項の対象となるプラント設備の部位，初めてアノテーションが加えられた日，現状の状態(解決済みかどうか)からなり，以降，手書き文字列のイメージ情報とその筆記された日時が追加されていく。アノテーションリストに基づき，朝礼，夕礼時に情報の共有を確実にするため，PC 上に対策未完了の案件の一覧を表示させ，保全担当者で一緒に見ながら申し送りを行う。また，対策未完了の案件を運転日誌のプラント構造図の欄に重畳印刷することで，現場での保全担当者への伝達の確実性を増す。

図 4.6　アノテーションリストの構造

4.3.3.　手書きと紙と電子文書との対応付け

次にハイブリッド文書について述べる。これは，アノテーション解釈を行うために，手書きがどの紙に書かれているか，その紙の情報がどの電子文書に対応するか，さらに新たな手書きは過去に筆記されたどの情報と同じ部位に関するものとして纏めればよいのか，そして，時系列にまとめた申し送り事項(案件)を再印刷した際にどの紙に印刷したかを把握するために，手書きと紙と電子文書との対応付けをとるデータ構造である。ハイブリッド文書の概要を図 4.7 に示す。ハイブリッド文書は PC 上の電子文書，およびそれを印刷した紙文書, さらに紙上に追記されたアノテーション といった情報を全て保持する「原本」である。これは電子文書と紙文書の対応関係として表現され，PC 上で電子文書が初めて印刷される際に生成される。印刷される紙文書に関する情報として，印刷時点での電子文書の写しが生成され, ハイブリッド文書から元の電子文書へのリンクが保持される。また，印刷された紙文書の識別情報として紙一葉を区別する紙文書 ID を持つ。Anoto 社のデジタルペンを適用する場合は，ドットパタンの ID が紙文書 ID となる。さらに，紙文書上に追記されたアノテーションに関する情報が手書き情報として管理される[97]。紙上に追記したアノテーション情報と印刷された電子文書のコンテンツとの対応関係を保持するために，印刷レイアウト情報や印刷者等の情報を印刷属性として持つ。これらに加えて，手書きアノテーションの文字認識結果等を文書メタデータとして管理する。

このように紙への筆記データを計算機に蓄積し，これを加工し，再度紙上の情報として印刷するといったように，紙と計算機の間での情報のサイクルを実現できる。同時に，保全担当者への共有の徹底のために，運転日誌帳票にも対策未完了の気づき，異常を毎日印刷して，現場に出向いた保全担当者にフィードバックを与えるようにする。

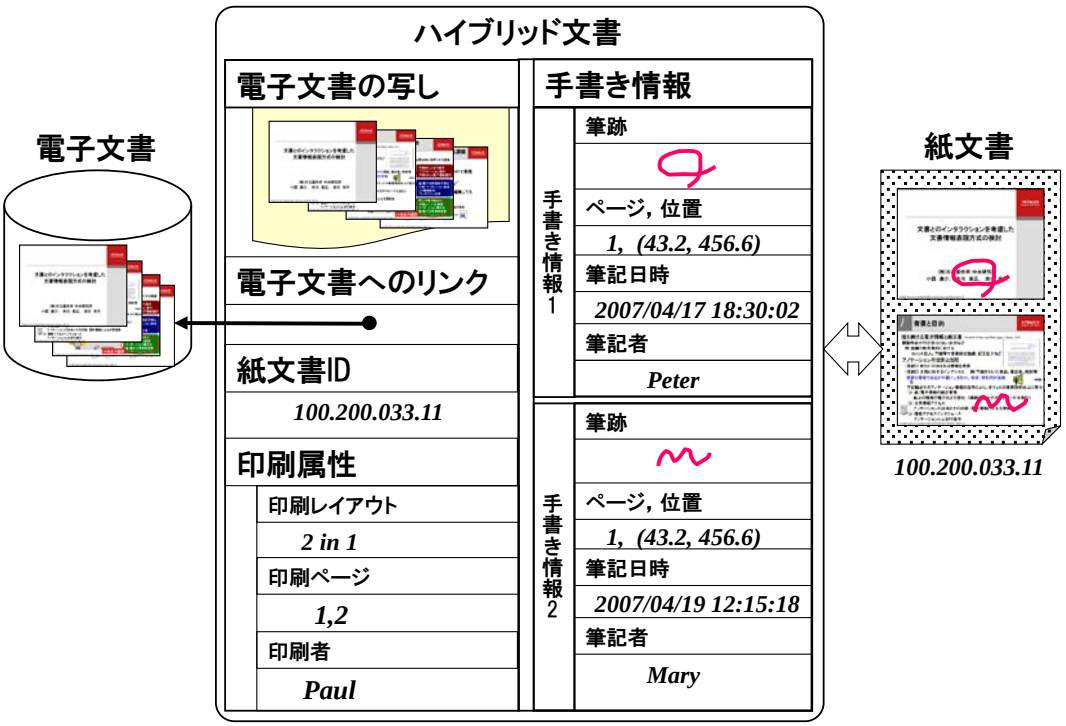


図 4.7　ハイブリッド文書の構成

4.3.4.　提案する設備保全業務の流れ

提案する設備保全業務のフローを図 4.8 に示す。図中(1a)にて参照する申し送り事項は，ある期間にわたって保全担当者が運転日誌に記載した特定装置に関する気づきをシステムが装置毎にまとめてディスプレイに表示する。従来は担当者によってホワイトボードに発生順に書き出されていた申し送り事項が，システムによって自動的に生成されることとなる。図中(1b)にあるように，当日の始業時に運転日誌を都度印刷する。これにより前日に入力した情報を翌日の運転日誌に印刷することができ，これまでのように転記の必要がなくなる。

運転日誌への記載にはデジタルペンを用いる。保全担当者はこの運転日誌が印刷された一枚の紙とデジタルペンを持って現場に行き，従来と同様に装置の計器に示された値を転記し，清掃作業を行ったり，異常を認識すればそれを特記事項として記入する。担当者が事務所に戻ってデジタルペン中に保存された筆記データを PC にアップロードし(図中

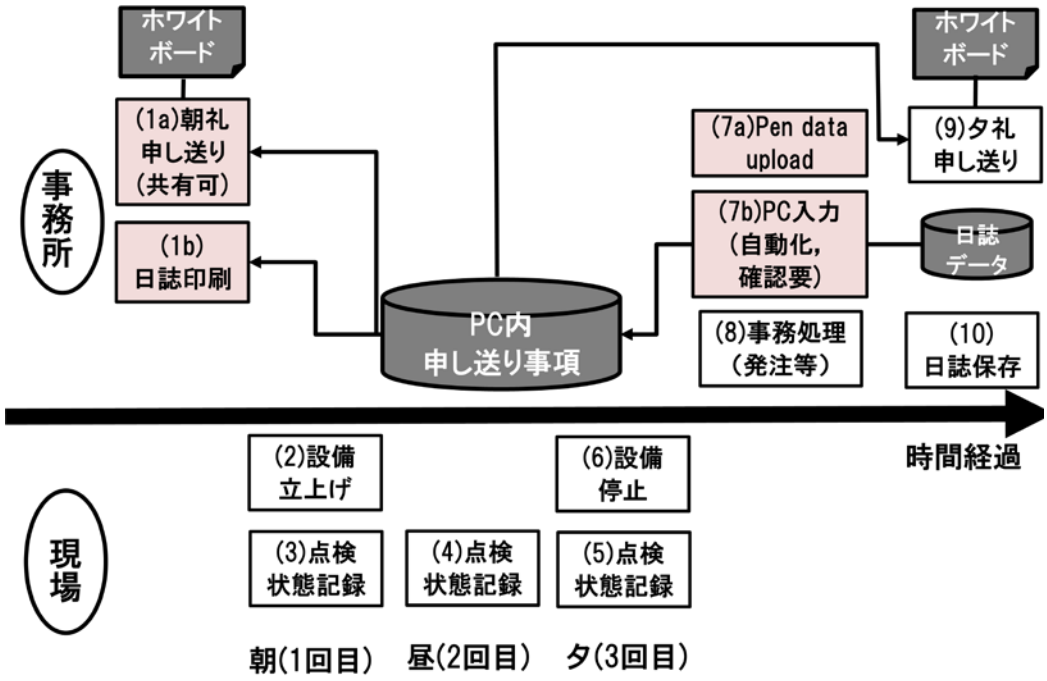


図 4.8　提案する設備保全作業の流れ

(7a))，これに文字認識を適用することで従来タイプ入力していた各種数値情報が電子化できる。担当者は文字認識結果を目視確認し，必要に応じて誤りを修正することで，点検結果の PC への入力が完了する(図中(7b))。

装置の異常など担当者間での共有が必要な情報については，これを運転日誌から抽出し，装置毎に時系列に分類して管理するとともに，事務所の PC 画面上に表示し，朝礼や夕礼で共有できるようにする。さらに申し送り事項に未解決，解決の属性を付与することで，未解決という属性を持つ申し送り事項のみ翌日の運転日誌に再印刷する。これにより，現地で保全担当者が，未解決の申し送り事項を確認しながら，点検作業を行うことができる。

4.3.5.　設備保全支援システムの構成

前節で述べた業務フローを実現する設備保全業務支援システムについて述べる(図 4.9 参照)。システムは，

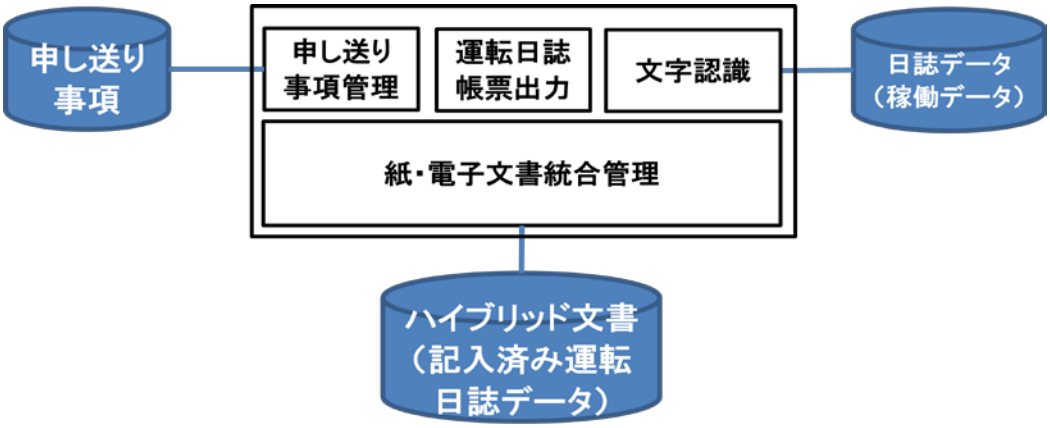


図 4.9　保全作業支援システムの構成

(1) 運転日誌(紙)の ID とそこに手書きされた情報と印刷された電子文書情報を紐づける紙・電子文書統合管理部,

(2) 運転日誌上に筆記された位置から同じ装置に関する手書き情報と判定して，その筆記時刻をもとに申し送り事項として時系列に並べる申し送り事項管理部,

(3) 保全業務を行う日毎に申し送り事項を含めて運転日誌を印刷する運転日誌帳票出力部,

(4) 運転日誌の書式情報をもとに切り出された手書き情報を業務システム上で再利用可能な形式(数値など)に変換する文字認識部,

の 4 つから成る。申し送り事項管理部にて時系列にまとめられた手書き情報は申し送り事項 DB に保存される。文字認識された手書き情報は，日誌データとして保存されるとともに，業務システムとも共有される。

4.4.　設備保全支援システムのプラント保全業務への適用と評価

4.4.1.　プラント保全業務への適用

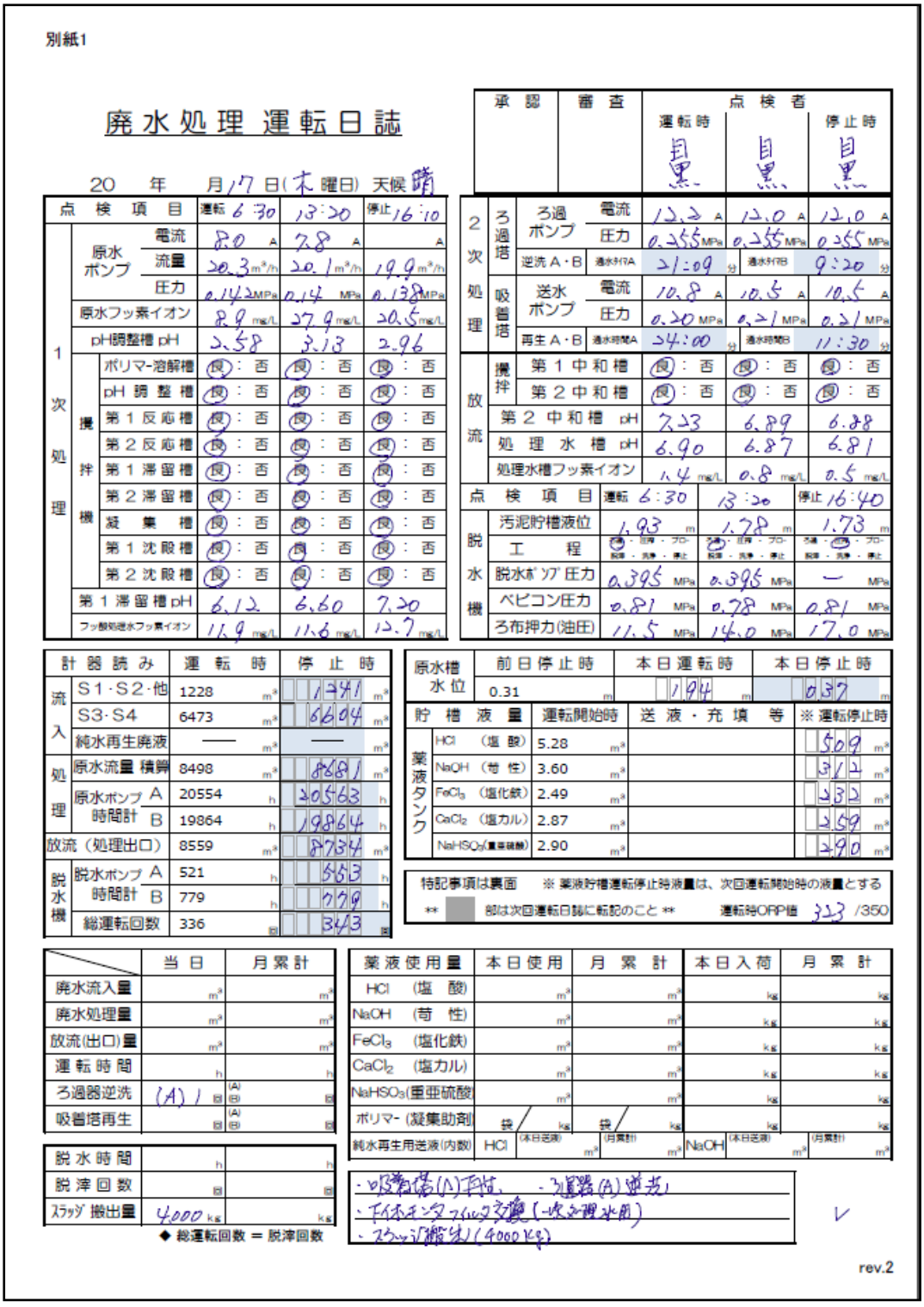
開発した設備保全業務支援システムは，筆者が所属する研究所における廃水処理保全業務に実際に適用した。

今回用いた運転日誌を図 4.10(表面)と図 4.11(裏面)に示す。図 4.10 が従来用いていた各装置の計器の値を記入する帳票である。各装置について，当日朝の値(すなわち前日プラント停止時の値)，当日夕方(プラント停止時)，部位によってはその中間の値，の 2 つない

し 3 つの欄が用意されている。このうち，前日停止時の値については，前日既に数値化して日誌データとして保存しているので，これを DB より取得して，帳票の所定の場所に挿しこみ印刷する。これにより改めて筆記する必要がなくなるとともに，誤記を防ぐこともできる。

図 4.10 の最下部の欄には特記事項を記載する。従来は，「イオンモニタの計器の調整実施」など定常作業の実施記録に加えて，装置の異常など担当者間での共有が必要なものの両方について記載していた。今回こうした申し送りが必要な情報については，図 4.11 に示すようなプラントの全体図が印刷されたページを用意し，プラントの全体図に注釈を付ける要領で記載してもらうこととした。運転日誌の書式と手書き情報はハイブリッド文書として管理されているため，例えば図 4.11 の一番右の手書き「HCl 注入弁漏洩」は同時に書かれた引き出し線により「第 2 中和槽」に対する記述であると認識できる。これらの申し送り事項は，装置毎に時系列に分類されて管理される。未解決という属性を持つ申し送り事項のみ翌日の運転日誌の裏面(図 4.11)の同じ場所に差し込み印刷される。これにより，現地で保全担当者が，未解決の申し送り事項を確認しながら，点検作業を行うことができる。日次で行う運転日誌の印刷のために担当者が行うべきは，PC 上での印刷コマンドの実行のみである。このような現場での保全作業の様子を図 4.12 に示す。

図 4.2 の(7)に示したように，これまで夕方にその日の稼働データを PC に入力していたのが，図 4.8 の(7a)(7b)にあるように，デジタルペン内の筆記データを PC にアップロードし，文字認識により数値データとして PC に格納するだけに簡略化される。筆記データのアップロードはペンを専用のクレードルに挿入するだけであり，データのアップロードの後に文字認識が実行される。運転日誌の紙 ID，および運転日誌の電子データ，その書式情報，運転日誌に記載された手書き情報もハイブリッド文書として管理する。従って，運転日誌紙面の所定の位置(例えば水槽の水位の記入枠内)に筆記された手書きデータは，水槽水位のデータとして切り出すことができる。但し文字認識を誤りなく行うことは困難なため，PC 画面上で人が確認し，誤っている場合は訂正するようにした。事務所でのデジタルペンデータの計算機へのアップロードと手書きデータの確認の様子を図 4.13 に示す。



別紙1

廃水処理運転日誌

承認	審査	点検者 運転時		停止時
		目黒	目黒	目黒

20 年 月17日(木曜日) 天候 晴

点検項目			運転 6:30	13:20	停止 16:10
1次処理	原水ポンプ	電流	8.0 A	7.8 A	A
		流量	20.3 m³/h	20.1 m³/h	19.9 m³/h
		圧力	0.142 MPa	0.14 MPa	0.138 MPa
	原水フッ素イオン		8.9 mg/L	27.9 mg/L	20.5 mg/L
	pH調整槽 pH		2.58	3.13	2.96
	攪拌機 ポリマー溶解槽		良 : 否	良 : 否	良 : 否
	pH調整槽		良 : 否	良 : 否	良 : 否
	第1反応槽		良 : 否	良 : 否	良 : 否
	第2反応槽		良 : 否	良 : 否	良 : 否
	第1滞留槽		良 : 否	良 : 否	良 : 否
	第2滞留槽		良 : 否	良 : 否	良 : 否
	凝集槽		良 : 否	良 : 否	良 : 否
	第1沈殿槽		良 : 否	良 : 否	良 : 否
	第2沈殿槽		良 : 否	良 : 否	良 : 否
	第1滞留槽 pH		6.12	6.60	7.20
	フッ酸処理水フッ素イオン		16.9 mg/L	11.6 mg/L	12.7 mg/L

点検項目			運転 6:30	13:20	停止 16:40
2次処理	ろ過塔 ろ過ポンプ	電流	12.2 A	12.0 A	12.0 A
		圧力	0.355 MPa	0.355 MPa	0.355 MPa
	逆流 A・B	通水時間A	21:09 分	通水時間B	9:20 分
	吸着塔 送水ポンプ	電流	10.8 A	10.5 A	10.5 A
		圧力	0.20 MPa	0.21 MPa	0.21 MPa
	再生 A・B	通水時間A	24:00 分	通水時間B	11:30 分
放流	攪拌 第1中和槽		良 : 否	良 : 否	良 : 否
	第2中和槽		良 : 否	良 : 否	良 : 否
	第2中和槽 pH		2.23	6.89	6.88
	処理水槽 pH		6.90	6.87	6.81
	処理水槽フッ素イオン		1.4 mg/L	0.8 mg/L	0.5 mg/L
脱水機	汚泥貯槽液位		1.93 m	1.78 m	1.73 m
	工程		濃縮・脱水・プロー・脱泥・洗浄・停止	濃縮・脱水・プロー・脱泥・洗浄・停止	濃縮・脱水・プロー・脱泥・洗浄・停止
	脱水ポンプ圧力		0.395 MPa	0.395 MPa	— MPa
	ベビコン圧力		0.81 MPa	0.78 MPa	0.81 MPa
	ろ布押力(油圧)		11.5 MPa	14.0 MPa	17.0 MPa

計器読み		運転時	停止時
流入	S1・S2・他	1228 m³	1341 m³
	S3・S4	6473 m³	6604 m³
	純水再生廃液	—— m³	—— m³
処理	原水流量積算	8498 m³	8881 m³
	原水ポンプ時間計 A	20554 h	20563 h
	原水ポンプ時間計 B	19864 h	19864 h
放流(処理出口)		8559 m³	8734 m³
脱水機	脱水ポンプ時間計 A	521 h	563 h
	脱水ポンプ時間計 B	779 h	779 h
	総運転回数	336 回	343 回

原水槽水位	前日停止時	本日運転時	本日停止時
	0.31 m	1.94 m	0.37 m

貯槽液量		運転開始時	送液・充填等	※運転停止時
薬液タンク	HCl (塩酸)	5.28 m³		5.09 m³
	NaOH (苛性)	3.60 m³		3.12 m³
	$FeCl_3$ (塩化鉄)	2.49 m³		2.32 m³
	$CaCl_2$ (塩カル)	2.87 m³		2.59 m³
	$NaHSO_3$(重亜硫酸)	2.90 m³		2.90 m³

特記事項は裏面 ※薬液貯槽運転停止時液量は、次回運転開始時の液量とする

** ■ 部は次回運転日誌に転記のこと ** 運転時ORP値 323 /350

	当日	月累計
廃水流入量	m³	m³
廃水処理量	m³	m³
放流(出口)量	m³	m³
運転時間	h	h
ろ過器逆洗	(A) 1 回	(A) (B) 回
吸着塔再生	回	(A) (B) 回

脱水時間	h	h
脱滓回数	回	回
スラッジ搬出量	4000 kg	kg

◆ 総運転回数 = 脱滓回数

薬液使用量	本日使用	月累計	本日入荷	月累計
HCl (塩酸)	m³	m³	kg	kg
NaOH (苛性)	m³	m³	kg	kg
$FeCl_3$ (塩化鉄)	m³	m³	kg	kg
$CaCl_2$ (塩カル)	m³	m³	kg	kg
$NaHSO_3$(重亜硫酸)	m³	m³	kg	kg
ポリマー (凝集助剤)	袋 / kg	袋 / kg	kg	kg
純水再生用送液(内数)	HCl (本日送液) m³	(月累計) m³	NaOH (本日送液) m³	(月累計) m³

・吸着塔(A)再生、・ろ過器(A)逆洗
・下水ポンプフィルター交換(一次処理水用)
・スラッジ搬出(4000kg)

rev.2

図 4.10 運転日誌の例(表面)

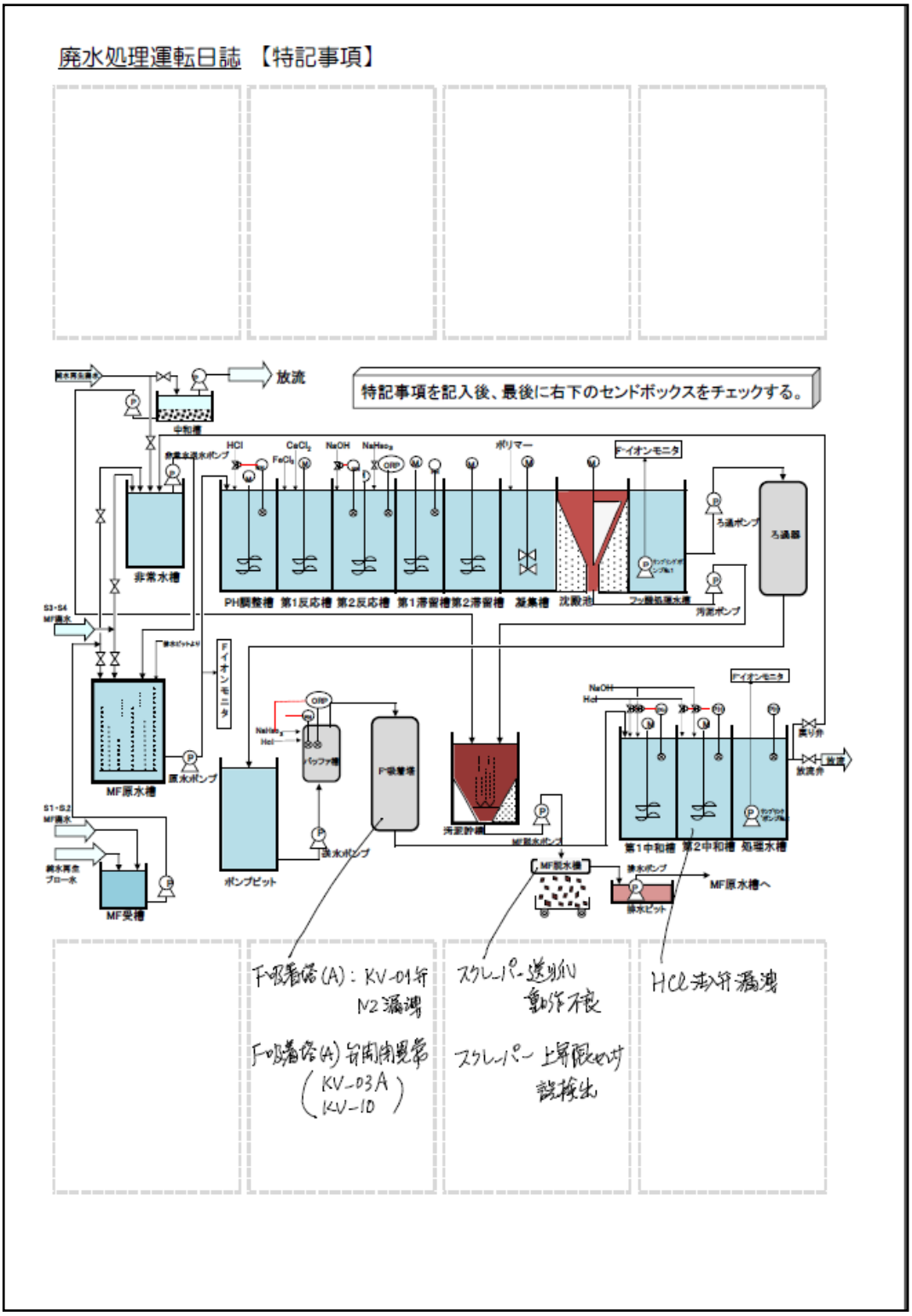


図 4.11 運転日誌の例(裏面)

図 4.12　現場での保全作業風景

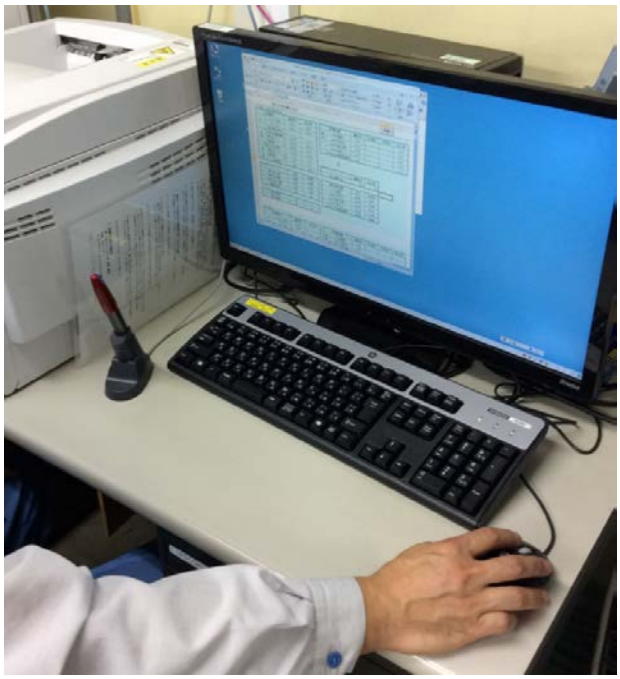


図 4.13　デジタルペンによる筆記データの計算機へのアップロードの様子

図 4.14　保全作業支援システムのディスプレイ

朝礼，夕礼時に情報の共有を確実にするための，設備保全業務支援システムの画面例を図 4.14 に示す。保全担当者は，ディスプレイ上に表示された対策未完了の案件の一覧を一緒に見ながら申し送り事項の共有を行うことが可能になる。プラント内の装置毎の特記事項(申し送り事項)のシーケンスが画面上の表の各行に表示されている。筆記データと共に，それがいつ筆記されたのかが画面中央の列に表示される。例えば部品の交換が済み，装置が正常な状態に戻った時に，画面右端の列にある「済」ボタンをクリックすることで，当該装置に関する対策は完了したとみなされ，以降運転日誌を印刷しても，その筆記事項シーケンスは運転日誌の紙面に印刷されることはない。

4.4.2.　評価

4.3 節で挙げた紙文書に記載された申し送り事項の活用について評価した。申し送り事

項の保全現場での共有について評価するために，実際の運転日誌をもとに，以下の 2 種類の情報共有形態を実装し，これをもとに保全担当者 3 名にインタビューを行い，提案手法の定性的評価を行った。

従来形態: 紙の運転日誌を用い, 口頭とホワイトボードにより申し送り事項を共有する。現場での共有事項へのアクセスはなし。

提案形態: 運転日誌上に特記事項として記載された申し送り事項を設備の部位毎に時系列にまとめて PC 上で管理する。あわせて未解決の申し送り事項を運転日誌に再印刷して現場で情報共有ができるようにした。

2013 年 4 月から 2014 年 6 月までの 15 ヶ月間の運転日誌の特記事項欄に記載された申し送り事項，全 18 件に着目した。ここで同一箇所に関する申し送り事項は，その発生から問題解決までをまとめて 1 件と数えている。特記事項欄への記載は 33 回あった。

これら 33 件について，従来形態の運転日誌と提案形態の運転日誌(申し送り事項が解決するまで印刷され続ける)を時間の経過順に保全担当者に提示した。提案形態の運転日誌の印刷内容が変化(一連の申し送り事項が解決，あるいは新たに申し送り事項が発生) するたびに，保全担当者に従来形態と提案形態との差分，つまり，一方に対する他方の効果を述べてもらった。その結果，提案形態の従来形態に対しての悪影響は述べられなかった一方で，提案形態の従来形態に対する効果がいくつか挙げられた。これを表 4.1 にまとめる。

種別における「終了明記」とは，一連の申し送り事項が修理完了，部品交換完了といった形で終わっているものであり,「様子見」とは保全担当者は異常を見つけたものの，代替手段によって稼働維持が可能であるもの，あるいは再び通常通り稼働したものを表す。運転日誌の特記事項欄に記載された申し送り事項が該当する種別に対して，表中に「○」を記載している。「終了明記」と「様子見」の両者に「○」が記載されているのは，当初経過を観察していたものに対して，後日になって解決が明記されていることを示す。

従来手法でのリスク(提案手法の効果)において，(a)引継ぎミス誘発とは，障害に対して保全担当者がバックアップの稼働や，手動による測定といった代替手段を見つけて，設備を稼働させているため，これが共有できていないと次の担当者が設備を止めてしまう可能性があるというものである。様子見をしている障害 11 件中 6 件がこれに当てはまる。(b)重複報告とは，申し送りができないために代替手段をとれず，後日設備に障害を再発生させてしまい，監督者への報告が繰り返されるというもので，(a)に起因するリスクである。(c)毎回確認要とは，担当者間で経緯が申し送りできていないために，保全作業で異常を認識する度に，継続して注意していれば設備稼働に問題はないという結論に至るための分析作業が毎回必要になるというものである。(d)再発時参照不能とは，様子見をしている部位

に障害が発生した際に，以前の障害の情報が共有できていないので，保全担当者が判断するための情報が不足してしまうというものである。結局前回の障害発生時と同じ手順の確認作業が必要になる。評価対象期間中にFイオンモニタと脱水機の2か所で繰り返し障害が発生しているが，これについていずれも(d)のリスクを保全担当者は挙げた。(e)状況不明確とは，障害が発生し修理までの間代替策での設備稼働を継続しているが，修理完了の情報が共有できていないため，いつまで代替策を継続すればよいのかが分からないという状態である。今回評価に用いた運転日誌では，R「不具合対策完了」として他の障害とまとめて記載されているために，個々の障害の対策状況が分からなかった。(f)定期点検漏れとは，緊急の対応は必要なく，定期点検時に部品を交換すればよいとして様子見をしている障害に対して，申し送りができないことで定期点検の対象から漏れてしまうものである。

AからRの障害場所に関する特記事項欄の記載に対して，表 4.1 の該当するリスクに対応する欄に「○」を記載した。

4.4.3. 考察

提案するシステムでは，運転日誌に記載した測定値等の情報を計算機に入力する際に，デジタルペンと文字認識を適用することで，文字認識の誤りとその訂正に要する時間を加味しても，これにかかる時間を68～78%削減することができた。

文字認識精度を高めるために，各測定値の記載欄には一文字ごとに枠を設けたが，これは筆記のしやすさを妨げるものではない。ただし現在，書き間違いをして訂正線などで修正した場合は文字認識ができない。本システムでは筆跡データは紙面上の位置の時系列データであるため，訂正線を認識することで，書きなおした場合にも認識可能とすることができる。これにより運転日誌の電子化に要する時間はさらに短縮できる。一方，朝礼で共有された情報が印刷されていることを確認した上で，現場に運転日誌を携行したいとの保全担当者の要望がある。そのため，運転日誌は朝礼後にその都度印刷する必要があるが，これに対して面倒であるというコメントがあった。しかし，印刷時間を考慮しても，筆記データの電子化の時間の短縮量の方が大きいこと，印刷時は別の作業をしながら待つことができることから，印刷処理の実行を簡略化することで提案システムの導入に妨げになることは回避できる。

表 4.1　運転日誌上での情報共有の効果

	障害場所	種別		従来手法でのリスク(提案手法の効果)					
		終了明記	様子見	(a)引継ミス誘発	(b)重複報告	(c)毎回確認要	(d)再発時参照不能	(e)状況不明確	(f)定期点検対応漏れ
A	脱水機駆動系					○		○	
B	F イオンモニタ	○				○			
C	脱水機		○		○		○		
D	F イオンモニタかく拌桟		○				○		○
E	吸着塔送水ポンプ		○			○			
F	F イオンモニタ:測定セル		○				○		
G	処理水槽 pH 記録計		○				○		
H	吸着塔(B) B バルブ							○	
I	吸着塔配管・バルブ	○			○				
J	脱水桟パラメータ変更			○					
K	吸着塔(B)弁・バルブ		○	○	○				
L	脱水桟ろ布		○	○	○				
M	一次処理水ポンプ		○	○	○				
N	pH 調整槽薬液自動注入弁	○	○	○	○				
O	吸着塔(A)弁		○	○	○				
P	脱水機スクレーパ		○	○	○		○		
Q	第 2 中和槽薬液注入弁							○	
R	不具合対策完了	○							

表 4.1 の申し送り事項全 17 件のうち 11 件を様子見が占めるように，部品交換や修理をしなければ設備が停止してしまう障害よりも，代替策で設備稼働が可能，あるいは次に障害が起きるまで様子を見る，といった場合の方が多い。後者のような場合，代替策をとること，あるいはある症状は確認できるものの稼働には問題ないことを確認済みであることを保全担当者間で共有することが重要であることが，引継ぎミス誘発(a)や毎回確認要(c)といったリスクが多く挙がったことから確認できた。また，障害が発生した際，過去に同じ個所でなんらかの障害が起きていたことがわかっていると，原因の推定や代替策の選択

に有効であるということが再発時参照不能(d)を挙げた際のコメントからも確認できた。全17件の障害のうち4件が脱水機，3件がFイオンモニタにて発生している。一般のプラントにおいても障害の発生頻度は部位によって異なるが，このような特定部位で頻発する軽度の障害を，より重大な障害の予兆としてとらえ，その情報を共有することには意味がある。こうした観点に加え，保全担当者が突発的に交代するなど事務所での申し送りができない場合があるため，現場で必ず共有できる本提案手法は有効であるとのコメントもあった。一方で完了というフラグだけではなく，様子見というフラグを明示的に付けることができれば，より当該障害の状況が分かりやすいとのコメントもあった。これは今後の改善項目である。

また，実証を継続することで，筆記の量と保全チームとしての障害防止力との相関を実験的に確認すること，さらに，保全担当者間の情報の共有漏れに起因する設備停止リスクをどの程度回避することができるかを定量的に検証することが今後の課題である。

4.5. 結言

プラント等の設備保全業務，特に現場での情報入力の効率向上と保全担当者間の情報共有の確実化を目指して，設備保全作業支援システムを開発した。研究所内の廃水処理プラント保全業務に適用したところ，代替策を講じて稼働を維持したり，次の障害発生まで様子を見ると言った場合に対して，運転日誌上に印刷して現場保全担当者の目に触れるようにしておくことで効果があることが保全担当者へのインタビューにより確認できた。加えて，運転日誌の記載事項の入力時間で68～78%の短縮が確認できた。

保全担当者が現場での気づきをより積極的に運転日誌に書きこむことで，保全担当者間での情報共有が深まり，重大な障害を未然に防ぐことを期待している。今後，実証を継続し,筆記の量と保全チームとしての障害防止力との相関を実験的に確認することを目指す。さらに，提案する保全支援システムにより，保全担当者間の情報の共有漏れに起因する設備停止リスクをどの程度回避することができるかを定量的に検証することが今後の課題である。

第5章

結言

5.1. 本研究のまとめ

本論文では，人間が主に紙文書上に筆記した手書きの情報を計算機上で扱えるような情報に変換，加工し，これを計算機上，あるいは再印刷された紙文書上で参照できるようにするとともに，この情報を用いて業務を効率的に行えるようにすることを手書きの利活用と定義し，その実現のために以下の三つの課題に着目した。

(1) 文字どうしが接触した手書き文字列からの文字切り出し方式

(2) 知識照合による文字列認識精度向上とそのための知識獲得方式

(3) アノテーション解釈方式と手書きによる情報共有の保全業務の現場への適用

最初の二つが筆記された文字列をテキストコード化し，計算機上で利用しやすいようにするための技術であり，三番目が紙文書に引いた線や，追記した手書きを印刷されているコンテンツと対応づけて意味づけ，業務の現場で共有する手法に関する技術である。

第 1 章では，日常の業務における手書きの重要性を示し，手書きの利活用を容易にすることの意義を述べた。次に手書きの利活用のための上記の三つの課題に関連する既存研究の俯瞰を行った。そして，手書きの利活用システムの全体像を示し，本論文で取り上げる三つの課題を位置付けた。

第 2 章では，接触した文字を含む文字列を認識するための接触箇所を含む連結成分の強制切断による文字候補パタンの生成と文字列認識方法について述べた。日本語における文字の接触の仕方を分析し，その結果に基づき，文字領域のストローク成分の形状により接触箇所の推定を行い，この部分の画素を強制的に削除することで文字候補パタンを生成する手法について提案した。生成した文字候補パタンは文字切り出し仮説ネットワーク上の新たな仮説として追加される。一方，日本語の文字は「偏」や「旁」などの複数の要素から構成されるため，接触箇所の切断だけでは正しい文字候補パタンを得ることができない場合がある。これに対して，文字切り出し仮説ネットワーク上で隣接する文字候補パタンをマージすることで，正しい文字候補パタンを生成する手法を提案した。これを地名表記文字列が手書きされた帳票サンプル画像を用いて評価し，提案手法により接触文字列の認識精度が向上するとともに，処理時間の増加も実用上問題のないレベルにとどまることが確認できた。

第 3 章では，認識対象の文字列の集合を知識(言語情報)として用意しておき，文字識別の誤りなどを訂正，補完することで文字列の認識精度を向上させる知識処理において，言語情報として統語的言語モデルを導入し，これを用いた探索的文字列認識方法と，統語的言語モデル構築法について提案した。地名表記や組織名などの固有名詞では，正式な表記に加えて，書き手の違いや，筆記領域の面積の制限から，省略など様々な異表記が存在する。これらをトライ構造のような従来の言語モデルにて表現しようとすると，部分木の重複からオンメモリで実行するには大きくなりすぎるとともに，これを用いた照合を実用的な時間で行うことが困難になる場合がある。この問題に対して，再帰的遷移ネットワークを用いた統合的言語モデルを提案し，メモリ容量，処理時間の両方についてトライ型の言語モデルを用いた照合処理よりも高速に実行できる可能性を示した。また，文字認識結果の候補列によって再帰的遷移ネットワークを探索的に照合する文字列認識手法を提案した。この時，文字切り出しや文字識別の誤りによって，文字切り出し仮説ネットワーク上の文字パタン候補数が増減するために，グラフ探索の途中で入力の文字候補パタンと言語モデルとの対応が取れなくなった場合に対しても，グラフの探索処理を進めるアドバンテージ遷移を導入することにより，文字列認識精度の向上を実現した。さらに，再帰的遷移ネットワークと文脈自由文法の表現力の等価性に着目し，認識対象の文字列の集合を文脈自由言語で表現し，これを再帰的遷移ネットワークに変換することで，人間にとって可視性の高い言語モデルの構築，運用手段を実現した。ここでは，認識対象の文字列の標準的な表記のリストから，文脈自由文法により記述された言語モデルに変換するとともに，異表記の生成規則を用いて生じうる異表記を加えることで，言語モデルの半自動作成を可能とした。次に一般的な異表記を持つ地域と，異表記のバリエーションがきわめて大きい地域を対象に，手書きの地名表記を含む帳票サンプル画像を用いて，提案手法を評価した。異表記のカバー率は，従来のトライ構造の言語モデルに対して，提案する文脈自由文法による言語モデルでは特に異表記のバリエーションが大きい地域程大きく向上させることができた。文字列認識精度も最大で 40pt 以上向上させることができ，提案手法の有効性を確認できた。

第 4 章では，業務における手書き情報の共有について，手書きのアノテーションを再利用可能なように解釈し,電子文書と共に保持する紙･電子文書統合管理方式について述べるとともに，実業務への適用を行った。実業務の例として設備保全業務を取り上げ，紙と手書きが多用されている業務における情報共有の課題について分析したところ，現場や事務所などでの情報共有のたびに書き直しや計算機への入力が行われており，障害の予兆となるようなその時点では緊急度は高くはないが，重要な情報が共有されにくくなっているこ

とが分かった。これに対して，現場で筆記した手書きの意味をその紙に印刷されていた電子文書の情報を用いて解釈し，計算機上に保持し再利用可能にする紙・電子文書統合管理方式について提案した。さらに，デジタルペンを用いて提案した紙・電子文書統合管理方式を実装し，プラント等の設備保全業務，特に現場での情報入力の効率向上と保全担当者間の情報共有の確実化を目指して，設備保全作業支援システムを開発した。このシステムを筆者が所属する研究所内の廃水処理プラント保全業務に適用したところ，代替策を講じて稼働を維持したり，次の障害発生まで様子を見ると言った場合に対して，運転日誌上に印刷して現場保全担当者の目に触れるようにしておくことで効果があることが保全担当者へのインタビューにより確認できた。

以上により，図 1.3 に示すような手書き利活用システムを構成する機能のうち，文字列認識における文字切り出しと知識照合，アノテーション解釈とアノテーション統合紙文書化を実現できることを示した。

5.2. 今後の課題

手書きのさらなる利活用を支援するために，本研究を以下の二つの観点で進めるべきである。一つは，手書きの情報を用いて業務をより効率的に行えるようにするために，手書き情報をより高精度に計算機上で扱えるような情報に変換，加工し，計算機上，あるいは紙文書上での再利用性を高めることである。もう一つは，手書きに関する筆跡以外の情報を用いることで，さらなる業務の効率化を進めることである。

(1) 手書き情報の計算機可読情報への変換の高精度化と再利用性の向上

手書き文字列認識の高精度化に向けて，2 章で述べた接触文字切り出し精度の向上が必要である。ここで，単により多くの連結成分を切断し，生成される文字候補パタンの数を増やすだけでは，文字列読取り処理全体の処理時間が増大してしまう。そこで，接触個所ではない連結成分の切断の増加を極力抑えながら，接触個所を含む連結成分の切断精度を高めていくことが求められる。接触箇所の同定の際に，単一の連結成分を考慮するミクロな視点と，複数の文字候補パタンや文字列全体のバランスを考慮するマクロな視点の併用が必要だと考えられる。

手書き文字列認識の高精度化には，言語モデルの充実が必要である。3 章で提案した統語的言語モデルを用いても認識対象の地名表記のカバー率は 98%あまりにとどまっていた。人間が筆記する際に誤記した場合も含まれていると考えられ，必ずしも統語的

に正しい表記だけを対象にするのではなく，実サンプルの学習により頻繁に生じる誤記を含めて言語モデルを構築する必要がある。その場合，文字列認識において言語モデルの照合時間が増える可能性がある。探索を平等に行うのではなく，言語モデル中に頻度情報を用いるなどして，探索の実行に優先度を付けるなどの工夫が必要である。

手書きアノテーションの利活用として，4 章で提案した設備保全業務の支援システムでは，手書き情報を筆跡画像として共有するにとどまっていた。筆跡内容による計算機上での検索や分類，整理を行えるようにするためには，手書きアノテーションの文字列部分をテキストコードに変換する必要がある。現場で筆記するようなやや乱れた筆跡に対しても高精度に認識できる文字列認識が必要である。文書のレイアウト構造を利用して認識対象の文字列を限定する手法[98]や，文字の幾何的情報や筆記時間に関する情報を用いて文字切出し精度を向上させる手法[99][100]などを提案している。これらを組み合わせて 1 章で述べたような手書き利用システムの実現を目指す。

(2) 手書きに関する筆跡以外の情報の利活用

本論文では，デジタルペンを用いて手書きの利活用を行うシステムを実装した。デジタルペンは静止画像としての筆跡情報だけではなく，筆記具の動きの過程を時系列情報として取得できる。筆記過程の情報が価値を持つ分野として，精神疾患の診断がある。精神疾患患者の回復度合いを測る手段として，一桁の足し算を一定時間内に多数行わせて，作業量の継時的な変化をみる内田クレペリン検査が知られている[101]。川口らはここにデジタルペンを用いて，被験者のペンの動きの時間的変化から集中度の変化を推定することで患者の回復度合いを測ることを試みている[102]。

手書きが有効な別の分野としてアイディア創生, 発想支援がある。思考や判断において，手書きが有効であることは多くの文献において述べられている。例えば，発想法として知られている KJ 法[103][104]のように，アイディアの紙に書き出し，これらをグルーピングして俯瞰することで発想を支援する手法が適用される機会も多い。紙と手書きを用いることで，思いついたアイディアの表現とそのグルーピングを迅速に可能にするため，発想活動を妨げない。また，思いついたアイディアを書くことによって，さらに発想が励起されることも我々は日常の中で体感している。しかし，手書きのどのような情報がアイディア発想に役立っているかについてはよくわかっていない。これに対して，グループによる問題解決型の議論において，情報の整理，主張の構成の際に筆記された手書きデータの分析を行ったところ，アイディアの萌芽が生まれるところ，いわゆる洞察過程においては，筆記の際の筆圧が増したり，ペンの動き(筆速)が相対的に増すことがあることが観察でき

た[105][106]。こうした分析を進めることで，発想過程と筆記行動との関係性が見いだせることが期待される。

このように筆跡以外の様々な情報を分析することで，手書きに潜む筆記者の内部状態や意図を推定し，より多くの利活用可能な情報を手書きから獲得する。こうした情報を分析，解析することで，保全業務だけではなく，医療の診断，新事業企画におけるアイディア創生など様々な業務のさらなる効率化を進める[107][108][109]。

謝辞

本研究の全過程を通じ，懇切なるご指導とご鞭撻を賜りました大阪大学大学院情報科学研究科マルチメディア工学専攻 松下康之教授, ならびに薦田憲久名誉教授に心から感謝申し上げます。

本研究をまとめるにあたって貴重なお時間を割いて頂き，丁寧なるご教示を賜りました大阪大学大学院情報科学研究科 マルチメディア工学専攻 藤原融教授，鬼塚真教授，ならびに前川卓也准教授に謹んで深謝致します。

情報科学研究科博士後期課程において，情報工学全般に関して親切なるご指導とご助言を賜りました，大阪大学 西尾章治郎総長，大阪大学大学院情報科学研究科マルチメディア工学専攻 下條真司教授，原隆浩教授に深く感謝申し上げます。

本研究の機会を与えていただくとともに，温かいご指導とご鞭撻を賜った，(株)日立製作所中央研究所の歴代の所長である，小島啓二博士(現常務執行役研究開発グループ長)，長我部信行博士(現ヘルスケア社 CSO&CTO)，鈴木教洋博士に謹んで深謝致します。直属の上司として筆者が博士号取得を志すに際して, 会社の業務との両立を支援下さいました，(株)日立製作所中央研究所情報システム研究センタ長の歴代センタ長である，山足公也博士(現横浜研究所長)，西村信治博士(現エレクトロニクスイノベーションセンタ長)，同知能システム研究部の歴代部長である，中屋雄一郎博士(現情報通信イノベーションセンタ長)，久光徹博士(現東京社会イノベーション協創センタプロジェクトマネージャ)に，心から御礼申し上げます。会社での研究を進めるに当たり，直属の上司としてまた，共同研究者として御指導頂きました，(株)日立製作所研究開発グループ技師長藤澤浩道博士，(株)日立製作所中央研究所知能システム研究部主管研究員酒匂裕博士(元法政大学大学院情報科学研究科教授)には，真摯に技術を追求する姿から，企業研究者のあるべき姿を学ばせて頂きました。深く感謝いたします。

第 2 章と第 3 章の研究に関して, 共同研究者として様々なご討論ご助言を頂きました(株)日立製作所中央研究所知能システム研究部の先輩，同僚である，故丸川勝美氏，古賀昌史博士，緒方日佐男氏(現(株)日立オムロンターミナルソリューションズ)，新庄広氏，影広達彦博士，嶺竜治氏，永崎健氏(現(株)日立オートモーティブシステムズ)，藤尾正和博士，平山淳一氏に心から御礼申し上げます。

第 4 章の研究に関して，共同研究者としてデジタルペンシステムの開発に取り組んだ，

(株)日立製作所中央研究所知能システム研究部古川直広氏，小西康介氏(現(株)ドワンゴ)に心から御礼申し上げます。また，紙・電子文書統合管理環境，および設備保全業務支援システムの実装に協力頂いた(株)日立ソリューションズ荻原健治氏に感謝致します。さらに，設備保全業務に関する課題のヒアリング，およびシステムの日常業務への導入とその評価にご協力頂いた(株)日立アーバンインベストメント平岡正巳氏，目黒泰行氏，島崎勇氏，佐藤光雄氏，中務晋氏をはじめとする保全業務担当の皆様に感謝します。

筆者が博士号取得を志して以来, 常に励まして頂いた, ドイツ人工知能研究センタ(DFKI GmbH)　Knowledge Management Department Prof. Dr. Andreas Dengel，九州大学大学院システム情報科学研究院内田誠一教授，(株)日立製作所中央研究所主管研究員北原義典博士(現東京農工大学大学院工学府教授)には心より感謝申し上げます。

筆者が京都大学工学部，および大学院修士課程において知能情報処理研究の分野に入った際にご指導を頂いた, 京都大学堂下修司名誉教授, 同大学院情報学研究科西田豊明教授, 公益財団法人京都高度技術研究所主席研究員山田篤博士に感謝申し上げます。

東京大学大学院にて御指導を頂きました，東京大学大学院工学系研究科システム創成学専攻大澤幸生教授に心より感謝申し上げます。

また，本論文をまとめるに当たり，会社業務との調整を行って頂いた，(株)日立製作所中央研究所総務部広知美代子氏(現横浜総務部)に御礼申し上げます。

最後に，いつも体調を気遣ってくれた母と亡き父に感謝致します。また，本論文の執筆に当たり，温かく励まし支えてくれた妻葉子と息子奏太に心から感謝します。

参考文献

[1] J. Bolter: “Writing Space: the Computer, Hypertext, and the History of Writing”. Lawrence Erlbaum Associates, 1991.

[2] 井上幸: “古代木簡研究におけるデジタルデータの整理と集積”, 情報処理学会研究報告 人文科学とコンピュータ研究会報告, No. 11, pp. 1-4, 2014.

[3] A. J. Sellen and R. H. R. Harper: “The Myth of the Paperless Office”, MIT Press, 2001.

[4] 鈴木慶子: “文字を手書きさせる教育―「書写」に何ができるのか”, 東信堂, 2015.

[5] S. Impedovo, L. Ottaviano, and S. Occhinegro: "Optical Character Recognition—a Survey", in P. Wang (Eds.) *Character & Handwriting Recognition*, World Scientific, pp. 1-24, 1991.

[6] H. Fujisawa: “Forty Years of Research in Character and Document Recognition—an Industrial Perspective”, *Pattern Recognition*, Vol. 41, No. 8, pp. 2435-2446, 2008.

[7] Naohiro Furukawa, Junko Tokuno, and Hisashi Ikeda, “Online Character Segmentation Method for Unconstrained Handwriting String using Off-stroke Features”, in *Proceedings of 10th International Workshop on Frontiers in Handwriting Recognition (IWFHR-10)*, pp. 361-366, 2006.

[8] H.S. Baird, S. Kahan, and T. Pavlidis: “Components of an Omnifont Page Reader”, in *Proceedings of 8th International Conference on Pattern Recognition (ICPR 86)*, pp. 344-348, 1986.

[9] Y. Lu: “On the Segmentation of Touching Characters”, in *Proceedings of International Conference on Document Analysis and Recognition (ICDAR 93)*, pp. 440-443, 1993.

[10] K. Ohta, I. Kaneko, Y. Itamoto, and Y. Nishijima: “Character Segmentation of Address Reading/Letter Sorting Machine for the Ministry of Posts and Telecommunications of Japan”, *NEC Research and Development*, Vol. 34, No. 2, pp. 248-256, 1993.

[11] S.K.Paul, B.B.Chaudhuri, and D.D.Majumder: “A Procedure for Recognition of

Connected Handwritten Numerals", *International Journal on Systems Science*, Vol. 13, No. 9, pp. 1019-1029, 1982.

[12] M. Shridhar, and A. Badreldin: "Recognition of Isolated and Simply Connected Handwritten Numerals", *Pattern Recognition*, Vol. 19, No. 1, pp.1-19, 1986.

[13] 仲林清，北村正，河岡司: "あいまい用語検索を用いた高速枠なし手書き文字列読取り方式", 電子情報通信学会論文誌, Vol. J74-D-II, No. 11, pp. 1528-1537, 1991.

[14] M. Gilloux, J.M. Bertile, and M. Leroux: "Recognition of Handwritten Words in a Limited Dynamic Vocabulary", in *Proceedings of International Workshop on Frontiers in Handwriting Recognition (IWFHR III)*, pp. 417-422, 1993.

[15] R.G. Casey, and G. Nagy: "Recursive Segmentation and Classification of Composite Patterns", in *Proceedings of 6th International Conference on Pattern Recognition (ICPR 82)*, Vol. 2, pp. 1023-1026, 1982.

[16] H. Fujisawa, Y. Nakano, and K. Kurino: "Segmentation Methods for Character Recognition: from Segmentation to Document Structure Analysis", in *Proceedings of the IEEE*, Vol. 80, No. 7, pp. 1079-1092, 1992.

[17] F. Kimura, S. Tsuruoka, Y. Miyake, and M. Shridhar: "A Lexicon Directed Algorithm for Recognition of Unconstrained Handwritten Words", *IEICE Transaction on Information & Systems*, Vol. E77-D, No. 7, pp.785-793, 1994.

[18] S. N. Srihari and Y.-C. Shin, V. Ramanaprasad, and Dar-Shyang Lee: "Name and Address Block Reader System for Tax Form Processing", in *Proceedings of 3rd International Conference on Document Analysis and Recognition (ICDAR 95)*, pp. 5–10, 1995.

[19] U. Mahadevan and S. N. Srihari: "Parsing and Recognition of City, State, and ZIP Codes in Handwritten Addresses", in *Proceedings of 5th International Conference on Document Analysis and Recognition (ICDAR 99)*, pp. 325–328, 1999.

[20] D. Chen, J. Mao, and M. Mohiuddin: "An Efficient Algorithm for Matching a Lexicon with a Segmentation Graph", in *Proceedings of 5th International Conference on Document Analysis and Recognition (ICDAR 99)*, pp. 543–546, 1999.

[21] J. Weinman, E. Learned-Miller, and A. Hanson: "Scene Text Recognition Using Similarity and a Lexicon with Sparse Belief Propagation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, No. 10, pp. 1733-1746, 2009.

[22] E. Fredkin: “Trie Memory”, *Communications of the ACM*, Vol. 3, No. 9, pp. 490-499, 1960.

[23] B. Zhu and M. Nakagawa: “Trie-Lexicon-Driven Recognition for On-line Handwritten Japanese Disease Names Using a Time-Synchronous Method”, in *Proceedings of International Conference on Document Analysis and Recognition (ICDAR 2011)*, pp. 1130-1134, 2011.

[24] M. Malburg, and A. Dengel: “Address Verification in Structured Documents for Automatic Mail Delivery”, in *Proceedings of 1st European Conference on Postal Technologies*, pp. 447–454. 1993.

[25] A. Dengel, R. Hoch, F. Hönes, M. Malburg, and A. Weigel: “Techniques for Improving OCR Results”, in H. Bunke and P. Wang (Eds.) *Handbook of Character Recognition and Document Image Analysis*, World Scientific, pp. 227–258, 1997.

[26] 伊東伸泰, 丸山宏: “OCR 入力された日本話文の誤り検出と自動訂正”, 情報処理学会論文誌, Vol. 33, No. 5, pp. 664-670, 1992.

[27] C. Shannon: “Prediction and Entropy of Printed English”, Bell System Technical Journal, Vol. 30, No. 1, pp. 50–64, 1951.

[28] U. Marti and H. Bunke: “Using a Statistical Language Model to Improve the Performance of an HMM-based Cursive Handwriting Recognition System”, *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 15, No. 1, pp. 65-90, 2001.

[29] A. Vinciarelli, S. Bengio, and H. Bunke: “Offline Recognition of Unconstrained Handwritten Texts Using HMMs and Statistical Language Models”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, No. 6, pp. 709-720, 2004.

[30] M. Wuthrich, M. Liwicki, A. Fischer, E. Indermuhle, H. Bunke, G. Viehhauser, and M. Stolz: “Language Model Integration for the Recognition of Handwritten Medieval Documents”, in *Proceedings of 10th International Conference on Document Analysis and Recognition (ICDAR 09)*, pp. 211-215, 2009.

[31] T. Breuel; “The OCRopus Open Source OCR System”, in *Proceedings of Electronic Imaging 2008*, pp. 68150F-68150F-15, 2008.

[32] 永田昌明: “文字類似度と統計的言語モデルを用いた日本語文字認識誤り訂正法”, 電子情報通信学会論文誌 D, Vol. 81, No. 11, pp. 2624-2634, 1998.

[33] B. Schilit, G. Golovchinksy, and M. Price: “Beyond Paper; Supporting Active Reading with Free Form Digital Ink Annotations”, in *Proceedings of SIGCHI Conference on Human Factors in Computing Systems (CHI 1998)*, pp. 249-256, 1998.

[34] M. Price, B. Schilit, and G. Golovchinksy: “XLibris: the Active Reading Machine”, in *Proceedings of SIGCHI Conference Summary on Human Factors in Computing Systems (CHI 1998)*, pp.22-23, 1998.

[35] G. Golovchinsky, and L. Denoue: “Moving Markup: Repositioning Freeform Annotations”, in *Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology (UIST 2002)*, pp. 21-30, 2002.

[36] 伊藤 禎宣，角 康之，間瀬 健二，國藤 進: “SmartCourier: アノテーションを介した適応的情報共有環境”，人工知能学会論文誌，Vol. 17, No. 3, pp. 301-312, 2002.

[37] D. Bargeron, and T. Moscovich: “Reflowing Digital Ink Annotations”, in *Proceedings of SIGCHI Conference on Human Factors in Computing Systems (CHI 2003)*, pp. 385-392, 2003.

[38] A. Brush, and D. Bargeron: “Robustly Anchoring Annotations Using Keywords”, *Microsoft Research Technical Report*, MSR-TR-2001-107, 2001.

[39] M. Schilman, and Z. Wei: “Recognizing Freeform Digital Ink Annotations”, in *Proceedings of International Workshop on Document Analysis Systems (DAS 2004)*, pp. 332 – 336, 2004.

[40] 大賀暁，仙田修司，旭敏之，山田敬嗣: “ペン入力環境における手書き文字アノテーションの活用”，第 2 回情報科学技術フォーラム（FIT2003），pp. 425-426, 2003.

[41] 大賀暁，仙田修司，旭敏之: “ペン入力アノテーション情報の入力/活用システム「アノテーションメモパッド」の開発”，情報処理学会研究報告，2004-DD-43, pp. 9-14, 2004.

[42] 石井大輔，鈴木優，石谷康人: “ペン操作型情報収集とイベント型情報再利用に基づく情報活用システム”， 第 4 回情報科学技術フォーラム(FIT2005), pp. 219-222, 2005.

[43] 鈴木優，布目光生，石谷康人: “ユーザの試行を妨げないペン操作によるインタラクティブな情報検索～意味解析と意図推定に基づく連鎖情報検索～”，インタラクション 2005, pp. 193-194, 2005.

[44] 鈴木優，布目光生，石谷康人: “インタラクティブなペン操作を可能とする検索意図に基づく連鎖情報検索”，インタラクション 2006, pp. 101-108, 2006.

[45] S. Klemmer: “Integrating Physical and Digital Interactions”, *IEEE Computer*, Vol. 38, No. 10, pp.111-113, 2005.

[46] F. Guimbretière: “Paper Augmented Digital Documents”, in *Proceedings of the Annual 16th Symposium on User Interface Software and Technology (UIST 2003)*, pp. 51-60, 2003.

[47] C. Liao, F. Guimbretière, and K. Hinckley: “PapierCraft: A Command System for Interactive Paper”, in *Proceedings of the 18th Annual Symposium on User Interface Software and Technology (UIST 2005)*, pp. 241-244, 2005.

[48] R. Yeh, C. Liao, S. Klemmer, F. Guimbretière, B. Lee, B. Kakaradov, J. Stamberger, and A. Paepcke: “ButterflyNet: a mobile capture and access system for field biology research”, in *Proceedings of ACM Conference on Human Factors in Computing Systems (CHI 2006)*, pp.571-580, 2006.

[49] N. Weibel, A. Fouse, C. Emmenegger, W. Friedman, E. Hutchins, and J. Hollan: “Digital Pen and Paper Practices in Observational Research”, in *Proceedings of ACM Conference on Human Factors in Computing Systems (CHI 2012)*, pp. 1331-1340, 2012.

[50] Hisashi Ikeda, Naohiro Furukawa, Katsumi Furukawa, and Hiromichi Fujisawa, “Toward a Personalized Digital Library for Providing “Information JIT”,” in *Proceedings of IAPR Workshop on Document Analysis Systems (DAS 2004)*, pp. 47-50, 2004.

[51] Hisashi Ikeda, “Human Memory Expansion by Personal Handwriting for Realizing Information JIT”, in *Proceedings of WM2005 Professional Knowledge Management, Experiences and Visions*, pp. 650 – 651, 2005.

[52] 池田尚司，小川祐紀雄，古賀昌史，西村広光，酒匂裕，藤澤浩道，“手書き接触漢字切出しに関する検討”，1998 年度電子情報通信学会ソサエティ大会, D12-14, p. 236, 1998.

[53] Hiromitsu Nishimura, Hisashi Ikeda, and Yasuaki Nakano, “A Segmentation Method for Touching Handwritten Japanese Characters”, in *Proceedings of IAPR Workshop on Document Analysis Systems (DAS 98)*, pp. 130-139, 1998.

[54] Hisashi Ikeda, Yukio Ogawa, Masashi Koga, Hiromitsu Nishimura, Hiroshi Sako, and Hiromichi Fujisawa, "A Recognition Method for Touching Japanese Handwritten Characters", in *Proceedings of 5th International Conference on*

Document Analysis and Recognition (ICDAR 99), pp.641-644, 1999.

[55] H. Ikeda, N. Furukawa, M. Koga, H. Sako, and H. Fujisawa, "A Context-Free Grammar-Based Language Model for Document Understanding", in *Proceedings of IAPR Workshop on Document Analysis Systems (DAS 2000)*, pp. 135-146, 2000.

[56] H. Ikeda, N. Furukawa, M. Koga, H. Sako, and H. Fujisawa, "Context-Free Grammar-Based Language Model for String Recognition", *International Journal of Computer Processing of Oriental Languages*, Vol. 15, No. 2, pp. 149-163, 2002.

[57] 古川直広，加藤陽介，池田尚司，酒匂裕，"再帰遷移ネットワーク型文字列照合方式の高速化"，情報処理学会第 66 回全国大会, pp. 115-116, 2004.

[58] 池田尚司，古川直広，薦田憲久："手書きによる情報共有に基づく設備保全作業支援システムの開発"，電気学会情報システム研究会, IS-14-18, pp. 29-33, 2014.

[59] 池田尚司，古川直広，薦田憲久,"手書きによる情報共有に基づく設備保全作業支援システムの開発とプラント保全業務への適用,"電気学会電子・情報・システム部門論文誌, Vol. 135, No. 6, pp. 580-588, 2015.

[60] M. Koga, T. Kagehiro, H. Sako, and H. Fujisawa: "Segmentation of Japanese Handwritten Characters Using Peripheral Feature Analysis", in *Proceedings of. 14th International Conference on Pattern Recognition (ICPR 98)*, vol.2, pp.1137-1141, 1998.

[61] Y. Lu: "On the Segmentation of Touching Characters", in *Proceedings of 2nd International Conference on Document Analysis and Recognition (ICDAR 93)*, pp. 440-443, 1993.

[62] T. Bayer and U. G. Kuressel: "Cut Classification for Segmentation", in *Proceedings of 2nd International Conference on Document Analysis and Recognition (ICDAR 93)*, pp. 565-568, 1993.

[63] Su Liang, M. Shridhar, and M. Ahmadi: "Segmentation of Touching Characters in Printed Document Recognition", in *Proceedings on 2nd International Conference on Document Analysis and Recognition (ICDAR 93)*, pp.569-572, 1993.

[64] Z. Shi, S. N. Srihari, Y. C. Shiu, and Y. C. Ramanaprasad: "A System for Segmentation and Recognition of Totally Unconstrained Handwritten Numeral Strings", in *Proceedings on 4th International Conference on Document Analysis and Recognition (ICDAR 97)*, Vol. 2, pp. 455-458, 1997.

[65] N. W. Strathy, C. Y. Suen, A. Krzyazak: “Segmentation of Handwritten Digits Using Contour Features”, in *Proceedings of 2nd International Conference on Document Analysis and Recognition (ICDAR 93)*, pp. 577-580, 1993.

[66] D. Nishiwaki and K. Yamada: “A New Numeral String Recognition Method Using Character Touching Type Verification”, *Series in Machine Perception and Artificial Intelligence*, Vol. 34, pp.416-425, 2000.

[67] J. T. Favata: “General Word Recognition Using Approximate Segment-String Matching”, in *Proceedings of 4th International Conference on Document Analysis and Recognition (ICDAR 97)*, pp. 92–96, 1997.

[68] 丸川勝美，古賀昌史，嶋義博，藤澤浩道: “手書き漢字住所認識のためのエラー修正アルゴリズム”，情報処理学会論文誌，Vol. 35, No. 6, pp. 1101-1110, 1994.

[69] K. Marukawa, K. Nakashima, M. Koga, Y. Shima, and H. Fujisawa: “A Paper Form Processing System with an Error Correcting Function for Reading Handwritten String”, in *Proceedings of 3rd Annual Symposium on Document Analysis and Information Retrieval*, pp. 469–452, 1994.

[70] M. Koga, R. Mine, H. Sako, and H. Fujisawa: “Lexical Search Approach for Character-String Recognition”, *Document Analysis Systems: Theory and Practice*, Springer-Verlag, pp. 115–129, 1999.

[71] I. Guyon and F. Pereira: “Design of a Linguistic Postprocessor Using Variable Memory Length Markov Models”, in *Proceedings of International Conference on Document Analysis and Recognition (ICDAR 95)*, pp. 454-457, 1995.

[72] X. Zhou, D. Wang, F. Tian, C. Liu, and M. Nakagawa: "Handwritten Chinese/Japanese Text Recognition Using Semi-Markov Conditional Random Fields", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 35, No. 10, pp. 2413-2426, 2013.

[73] S. Srihari, X. Yang, and G. Ball: “Offline Chinese Handwriting Recognition: an Assessment of Current Technology”, *Frontiers of Computer Science in China*, Vol. 1, No. 2, pp 137-155, 2007.

[74] Y. Li and C. Tan: “An Empirical Study of Statistical Language Models for Contextual Post-Processing of Chinese Script Recognition”, in *Proceedings of 9th International Workshop on Frontiers in Handwriting Recognition (IWFHR 2004)*, pp. 257-262, 2004.

[75] 源城政好, 下坂守: “京都の地名由来辞典”, 東京堂出版, 2005.

[76] 一森哲男: “グラフ理論”, 共立出版, 2002.

[77] W. A. Woods: “Transition Network Grammars for Natural Language Processing”, *Communications of the ACM*, Vol. 13, No. 10, pp. 591-606, 1970.

[78] J. Hopcroft and D. Ullman: “Introduction to Automata Theory, Languages and Computation”, (野崎昭宏, 高橋正子, 町田元, 山崎秀記訳: “オートマトン　言語理論 +論 I”), サイエンス社, 1984.

[79] J. Levine, D. Brown, and T. Mason: “lex & yacc”, (村上烈訳: “lex & yacc”), アスキー出版局, 1994.

[80] 嶺竜治, 関峰伸, 池田尚司, 渡辺成, 酒匂裕: “距離に基づく確信度を利用した文字識別手法”, 電子情報通信学会技術研究報告 パターン認識・メディア理解, Vol. 104, No. 125, pp. 37-42, 2004.

[81] 経済産業省: “産業・社会資本構造物におけるメンテナンス情報の活用に関する調査研究”, 2005.

[82] 桑名諒, 伊藤久幸, 辻本静, 新間大輔, 有田節男, 原田英雄 : “原子力プラント保守支援システムの開発”, 2012 年日本原子力学会大会, p. 218, 2012.

[83] 矢吹信喜, 菊重有輝, 福田知弘, 江端夢歌: “RFID 技術とオントロジーを用いた街路樹管理支援システムの開発”, 土木学会論文集 F3（土木情報学), Vol. 68, No. 1, pp. 13-27, 2012.

[84] 渡辺義大, 野末道子, 佐藤紀生: “音声入力を活用した設備保守データ入力システムの開発”, 電気学会研究会資料. TER, 交通・電気鉄道研究会, Vol. 75, pp. 29-34, 2006.

[85] J. Ziegler, J. Pfeffer, and L. Urbas: “A mobile system for industrial maintenance support based on embodied interaction”, in *Proceedings of the 5th International Conference on Tangible, Embedded, and Embodied Interaction (TEI '11)*, pp. 181-188, 2011.

[86] 斎礼, 広瀬正, 矢島敬士, 薦田憲久: “保守/障害業務支援向け携帯端末用手書きメモ入力システムの開発と配電設備巡視業務への適用”, 電気学会論文誌 C, Vol. 119, No. 12, pp. 1542-1547, 1999.

[87] 町野保, 柳原義正, 南條義人, 河田博昭, 武藤伸洋, 茂木学, 下倉健一朗: “遠隔協調作業支援システム Field-AID の開発と評価”, 日本機械学会論文集(C 編), Vol. 72, No. 716, pp. 1223-1229 (2006).

[88] Naohiro Furukawa, Hisashi Ikeda, Yosuke Kato, and Hiroshi Sako, “D-Pen: A

Digital Pen System for Public and Businesses Enterprises", in *Proceedings of 9th International Workshop on Frontiers in Handwriting Recognition (IWFHR-9)*, pp. 269-274, 2004.

[89] 池田尚司, 古川直広, 藤澤浩道: "紙・電子文書統合管理による手書き情報の活用方式", 2009 年度 HCG シンポジウム, CD-ROM, 2009.

[90] K. Konishi, N. Furukawa, and H. Ikeda: "Data model and architecture of a paper-digital document management system," in *Proceedings of ACM Symposium on Document Engineering (DocEng 2007)*, pp.29-31, 2007.

[91] H. Ikeda, K. Konishi and N. Furukawa: "iJITinOffice: Desktop Environment Enabling Integration of Paper and Electronic Documents", in *Adjunct Proceedings of 19th Annual ACM Symposium on User Interface Software and Technology (UIST 2006)*, pp. 93-94, 2006.

[92] H. Ikeda, N. Furukawa, and K. Konishi: "iJITinLab: Information Handling Environment Enabling Integration of Paper and Electronic Documents", in *Proceedings of CSCW 2006 Workshop Collaborating over Paper and Digital Documents (CoPADD 2006)*, pp. 25-28, 2006.

[93] Hiromichi Fujisawa, Hisashi Ikeda, Naohiro Furukawa, Kosuke Konishi, and Shoichi Nakagami, "Information Just-in-Time: Going Beyond the Myth of Paperlessness", in (Bidyut Baran Chaudhuri and Swapan Kumar Parui Eds.) *Advances in Digital Document Processing and Retrieval*, pp. 35-49, World Scientific, 2008.

[94] 小西康介, 古川直広, 池田尚司, "文書とのインタラクションを考慮した文書情報表現方式の検討", 情報科学技術フォーラム一般講演論文集, Vol. 5, No. 4, pp. 421-422, 2006

[95] 古川直広, 池田尚司, 小西康介, "紙とペンによる情報アクセス方式の開発", 情報科学技術フォーラム一般講演論文集, Vol. 5, No. 4, pp. 423-426, 2006.

[96] 池田尚司, 小西康介, 古川直広, "文書へのアノテーションを活用する文書管理システムの開発", 情報科学技術フォーラム一般講演論文集, Vol. 5, No. 4, pp. 427-428, 2006.

[97] 古川直広, 池田尚司, 小西康介, "デジタルペンを用いた研究ノートの開発", インタラクション 2007, pp. 59-60, 2007.

[98] 古川直広，加藤陽介，池田尚司，酒匂裕，“レイアウト駆動型文字列認識方式”， 電子情報通信学会技術研究報告 パターン認識・メディア理解, Vol. 103, No. 295, pp. 67-72, 2003.

[99] 古川直広，徳野淳子，池田尚司，“自由文読取のための文字切出し方式の開発”, 情報科学技術フォーラム一般講演論文集, Vol. 4, No. 3, pp. 39-40, 2005.

[100] Sandip Rakshit , Subhadip Basu , and Hisashi Ikeda, “Recognition of Handwritten Textual Annotations using Tesseract Open Source OCR Engine for information Just In Time (iJIT)”, in *Proceedings of International Conference on Information Technology and Business Intelligence*, pp. 117-125, 2009.

[101] 外岡豊彦: “内田クレペリン精神検査・基礎テキスト”, 日本・精神技術研究所, 1973.

[102] 川口英夫，川上憲人，有馬秀晃，池田尚司，坂入実，“書字の時間構造を用いたメンタルヘルスの可視化”, 第 38 回可視化情報シンポジウム， Vol. 30, No. 1, pp. 159-160, 2010.

[103] 川喜田二郎: “発想法　創造性開発のために”, 中央公論社, 1967.

[104] 川喜田二郎: “続・発想法　KJ 法の展開と応用”, 中央公論社, 1970.

[105] H. Ikeda and Y. Ohsawa, “Toward Visualization of Insight Process in Concept Creation Focusing Handwriting Features”, in *Proceedings of 7th International Conference on Communication and Information Technology (CIT13)*, pp.136-143, 2013.

[106] H. Ikeda and Y. Ohsawa, “Visualization of Insight Process in Concept Creation Focusing Handwriting Features”, *International Journal of Knowledge and Systems Science*, Vol. 4. No. 1, pp.18-31, 2013.

[107] 池田尚司，額賀信尾，“非構造化データを扱う情報処理基盤の実現を支えるメディア処理技術”, 2012 年人工知能学会全国大会，CD-ROM, 2012.

[108] 池田尚司，額賀信尾，三好利昇，柳瀬利彦, “ビッグデータ解析と社会イノベーションに向けた取り組み”， 2014 年人工知能学会全国大会，https://kaigi.org/jsai/webprogram/2014/pdf/568.pdf，2014.

[109] 池田尚司，額賀信尾，小林義行，神田直之，渡邉裕樹，平山淳一，“非構造化データ利活用のためのメディア処理技術”，人工知能学会誌, Vol. 28, No. 1, pp. 114-121, 2012.