

## 紙とペンによる情報アクセス方式の開発

## Development of Paper-Based Information Access Method using Pen and Printed Objects

古川 直広†  
Naohiro Furukawa池田 尚司†  
Hisashi Ikeda小西 康介†  
Kosuke Konishi

## 1. はじめに

本報告では、紙と電子の両メディアの利点を活用できる直感的な計算機操作機能を提案する。これは人間の慣れ親しんだ紙とペンを用いて、検索や印刷など計算機への操作を紙面上で指示すると、ユーザの意図を理解して指示された処理を実行する機能である。本機能の基本設計およびデジタルペンシステム[1]を用いたプロトタイプ開発について述べる。

## 2. 研究の背景と目的

紙は人間にとって最も身近でかつ便利なメディアの1つである。簡便かつ安価であり、また携帯性に優れ、ペン等により容易に書込みもできる。一方で、計算機などで利用される電子メディアは検索・複製・転送・再利用の容易さや、大容量情報への高速アクセスなど、紙メディアでは実現困難な優位点がある。計算機やネットワーク技術の発展に伴い、我々はこれら電子メディアのメリットを享受できるようになってきている。

1980年代にオフィスオートメーションという言葉が流行した頃、紙の業務すべてが計算機上のワークフローへと置き換わり、ペーパーレスオフィスの時代が到来するという予測があった。しかし、計算機が発達した現代社会においても紙はいまだに利用され続けている。その理由として、紙は(1)書込みが容易、(2)同時に何枚も参照可能、(3)「めくる」という行為によるナビゲーションなどの特長があるためだと考えられている[2]。特に知識ワーカー(knowledge worker)にとってこれらメリットは「考える」行為の支援に重要である。一方で、前述の電子メディアのメリットも作業効率化に有益である。しかし、現状は紙面と計算機内の情報とを関係付けられないため、たとえば紙面上の文字列を計算機内で利用しようとするると人手で再入力が必要となるなどの不都合がある。したがって、紙などの現実世界と計算機などの電子世界との情報がシームレスに連携する文書管理システムが実現できれば、ユーザは紙の高利便性と、計算機の高速、高機能な情報処理の両方を活用できる。たとえば、知識ワーカーが紙をベースに知的活動しているときに、注目紙面への書込みなどの知識ワーカーの思考の妨げとならない簡単な操作によって、必要とする情報を電子世界から簡単に抽出できれば、知識ワーカーの生産性を高めることができる。

そこで、紙と計算機の連携による知識ワーカーの生産性向上を本研究の目的とする。本稿では、人間の慣れ親しんだ紙とペンを用いて、検索や印刷など計算機への操作を紙面上で指示すると、その意図を解釈し、指示操作を実行するという指示解釈機能(CommandInk)の検討結果を述べる。指示解釈機能の基本設計およびプロトタイプ開発による原理の検証と予備評価結果について報告する。

## 3. 指示解釈機能の基本設計

## 3.1. 指示解釈の例

まず、我々の考える指示解釈機能について、紙を用いた業務への適用を例に具体的に説明する。図1は医師が発行した処方箋を元に薬剤師が調剤する場面での本機能の適用例である。薬剤師が薬剤に関する情報を確認したいときに、処方箋上の該当薬剤名をペンで下線を引き、「関連情報検索」を丸囲みで選択する。すると文書管理システムは意図を理解し、その薬剤の関連情報を検索してモニタに表示する。薬剤名は長いスペルのものや類似の名称が多いために、キーボード入力やマウス選択よりも、ペンで紙面上の薬剤名を直接指定できた方が効率的である。このような、紙とペンによる計算機操作が実現できれば、利便性の高い文書管理システムを提供できる。

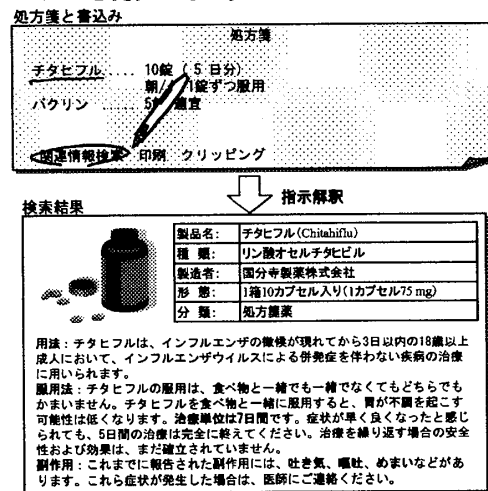


図1 指示解釈機能の適用例

## 3.2. 計算機への命令の形式

次に、紙とペンを用いてユーザが計算機へ指示する命令の形式について整理する。

指示解釈機能はその目的から、知識ワーカーが紙をベースに知的活動する際に計算機上で実行するネット検索や英和辞書引きなど種々の命令に対応しなければならない。これら命令は、図2に示したように、基本的に下記2種類の要素の組合せから構成される: (1)「ネット検索」や「印刷」など命令の種類を表す要素、(2) 検索キーワードや印刷対象範囲などの命令の引数となる要素である。これら要素を以降、命令要素(command element)とよぶ。

ユーザは実行したい命令の各命令要素を、紙面上でペンを用いて該当文字列や領域を囲ったり、文字列を直接書込んだりして指定する(図3)。たとえば、図3の1-Aのように印刷された文字列を指定したり、1-Bのように命令要素の文字列を直接手書きして指定したりする。またクリッピングや類似画像検索など命令種類によっては、文字列ではなく領域そのものをペンで丸囲みして指定する場合もある

†(株)日立製作所 中央研究所, CRL, Hitachi, Ltd.

(図3の2)。指示解釈機能はこれら多様な命令要素の形態に対応する必要がある。

また、命令を表現する命令要素の組合せ方(命令の規則)にも多様性が存在する。たとえば、丸囲み領域を印刷するといった単純な命令でも、最初に「印刷」という命令種類を表す文字列を選択し次に印刷領域を丸囲みで指定するのか、最初に印刷領域を丸囲みし次に「印刷」を選択するのか、といった命令要素の順序の差異がある。また印刷以外にネット検索やクリッピングなど複数種類への対応が必要な場合がある。さらに一度に複数個の命令が記入される場合もありうる。

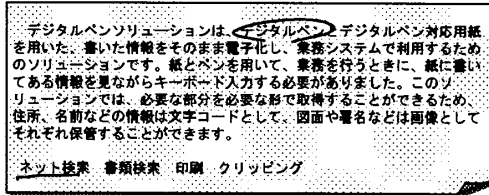


図2 命令要素の例

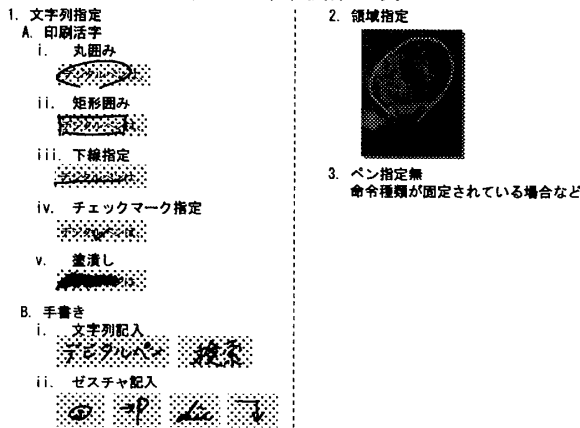


図3 命令要素の指定例

以上から、本研究では、(1)命令要素、(2)命令規則の2種類の多様性への対応のために、指示解釈の処理を下記二段処理で構成することとした。

- (1) 命令要素抽出(command element extraction)
- (2) 命令解釈(command interpretation)

まずペンストロークと紙面上に印刷された文書情報からユーザが意図した命令要素を抽出する(命令要素抽出)。この命令要素抽出は下記3ステップからなる。

- (1-1) 注釈記号切出し  
入力ストローク集合を注釈ごとに分割
- (1-2) 注釈記号識別  
丸囲みや下線など注釈記号の種類を判定
- (1-3) 指示文字列抽出  
注釈記号種類にそってユーザ指定のテキストを抽出

次に得られた命令要素の集合と事前に用意した命令規則辞書とを照し合せ、相応しい命令を決定し、それらを確信度で順位付けしてXML形式で出力する(命令解釈)。命令規則辞書を文脈自由文法で規定することにより、種々の命令規則に対応できるようにする。命令要素抽出で可能性のある命令要素候補を全部出力させ、それらの中から命令要素の列全体として相応しいものを選択していくため、命令要素抽出の不完全性(常に正しい命令要素のみを抽出する

とは限らないこと)を許容でき、全体として高精度な命令解釈が期待できる。

### 3.3. 指示解釈の機能構成と処理フロー

指示解釈実現のための機能構成と処理フローを考える。まずPDF形式の電子書類を入力し、その中から文字列など操作の対象物を抽出する必要がある。この抽出処理を高精度に実行するためには、PDF書類中の文字行の配置などの文書のレイアウトを解析する必要がある(文書レイアウト解析:layout analysis)。また前節で述べたように、命令要素抽出と命令解釈の各部分が必要であり、さらにそれらを実行するための事前知識として、命令要素辞書(command element dictionary)と命令規則辞書(command rule dictionary)が各々必要となる。命令要素辞書は、図3に示した命令の要素の指定記号となる丸囲みや下線、ゼスチャ記号のストローク形状情報を保存する。命令規則辞書は、各命令の規則を文脈自由文法で規定する。これら辞書を容易にメンテナンスできる命令定義ツール(command entry tool)が必要である。また構成要素を直接文字列で記入された場合のために、手書き文字認識部(keyword recognition)も必要である。これら各部分を制御し、指示解釈結果(CI-XML)を出力するのが指示解釈制御部(Commandink manager)である。指示解釈結果を命令実行部(command executor)に渡し、実際に命令が実行されることとなる。

以上から、指示解釈を実現するための機能構成と処理フローを、各々図4、図5のように決定した。

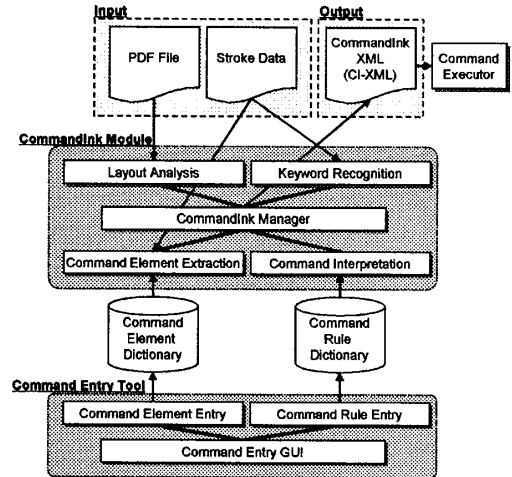


図4 指示解釈の機能構成

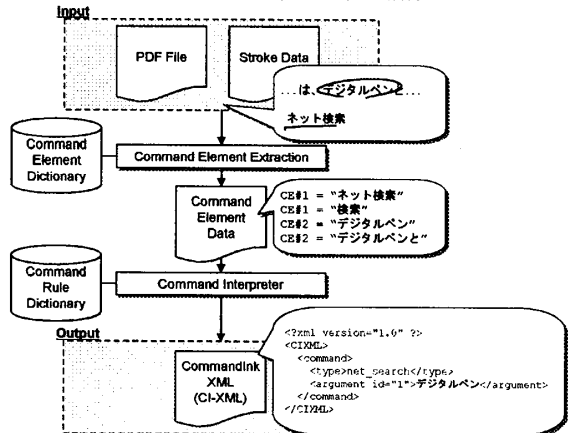


図5 指示解釈の処理フロー

#### 4. プロトタイプ開発

前章に述べた基本設計の予備評価のためにプロトタイプを開発した。プロトタイプの利用シーンとして、ユーザが計算機の前で紙書類を読んでいるときに、紙面上の文字列をペンで指定すると、ネット検索や辞書引きなどを実行するシーンを想定した。プロトタイプで対応する注釈記号は、(1)丸囲み、(2)下線の2種類とした。

今回の想定シーンでは、命令規則は複雑ではないため、プロトタイプ開発に当たっては命令要素抽出部を特に注力して開発した。プロトタイプの全体構成と命令要素抽出処理について以下に説明する。

##### 4.1. プロトタイプ構成

指示解釈機能では、印刷物の内容や書込み情報が入力として必要である。したがって、電子書類とその印刷物、さらに印刷物への書込み情報を管理するシステム上に指示解釈機能を実装することが望ましい。そのシステムとして、我々が開発した iJIT(information Just-in-Time)システム[3][4]を選択した。

iJIT システムは、サーバクライアント構成となっている。各クライアントから印刷要求があると、サーバで印刷処理を実行し、クライアントが指定したプリンタから紙書類が出力される。その紙書類にユーザがデジタルペンで書込むと、ペンが紙面上に印刷されたドットパターンを読取ることにより書込み内容を電子化し、サーバに転送される仕組みとなっている。プロトタイプでは、指示解釈機能を iJITサーバの1プラグインとして実装した。

##### 4.2. プロトタイプにおける命令要素抽出

命令要素抽出を構成する、注釈記号切出し、注釈記号識別、指示文字列抽出の各ステップを順に説明する。

注釈記号切出し方式として、たとえば文字認識向けの文字行抽出や文字切出し処理をそのまま適用すると、注釈記号同士が近い場合、過結合されてしまう問題があった。たとえば図 6A のように上下の注釈記号同士が近接して記入されたとき、文字認識向けの処理ではこのようなサイズにばらつきがあるパターン同士の近接を想定していないため、上下の注釈記号を誤って1つの注釈記号として切出してしまったことがあった。

そのためストロークの分離・結合の判断基準として、パターンサイズと連動した変動閾値を採用した。判断対象の2つのストローク集合に対し、2集合の幅と高さの最大値  $l_{max}$  で正規化することにより、ストローク集合間距離の閾値  $\theta_g$  を設定することとした。ストローク集合間距離  $g$  (図 6B 参照)が閾値  $\theta_g$  よりも小さければ1つの集合に統合する。なお  $\alpha$  は定数であり、大きいと過統合、小さいと過分割の傾向となる。

$$\theta_g = \frac{\alpha}{l_{max}} \dots\dots\dots (数式 1)$$

注釈記号の識別には、文字認識で用いられている統計的  
文字識別手法を適用した。

指示文字列の抽出に関し、ストローク被覆矩形と各文字  
矩形の重なるの度合いのみで単に判断すると、目的の文字  
列に隣接した文字を巻込んで抽出されることがあった。こ  
れは図 7 の例のように、ユーザが注釈をいつも正確に記入  
するとは限らず、注釈記号が指示文字列外にはみ出して記  
入されることがあるため、ユーザの意図よりも長い文字列

を抽出してしまうことがあった。

そこで、文字同士の連続性に着目した。注釈記号に被覆  
された各文字に対し、その前後の文字との連続性を下記 2  
特徴から算出した。

- (1) 言語的特徴：文字種(英数字、ひらがな、カタカナ、漢字等)の揃いなど
- (2) 幾何的特徴：文字行内位置、文字間など

この連続性に基づき、被覆率の閾値を変動させ、指示対  
象文字列を決定する方式とした。

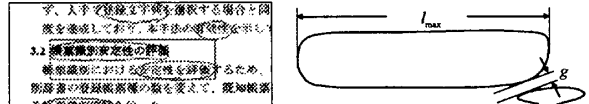


図 6 A:注釈記号切出し誤り例と B:切出しパラメータ

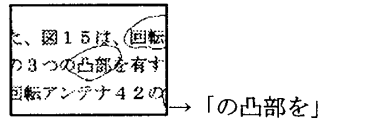


図 7 指示文字列における隣接文字の巻込みの例

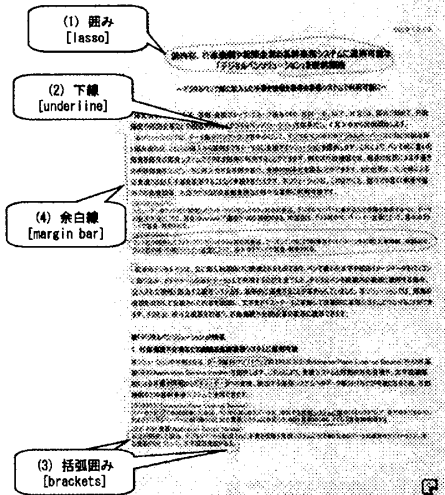


図 8 サンプル書類の例

表 1 評価データセット内訳

#	注釈記号	個数
1	囲み	517 個
2	下線	583 本
3	括弧囲み	119 組
4	余白線	106 本
	合計	1,444 記号

#### 5. 評価実験

プロトタイプを用いて、指示解釈機能の精度評価の予備  
実験を行った。プロトタイプにおいて、命令要素の文字列  
の抽出精度がそのまま全体の指示解釈精度となるため、命  
令要素抽出を構成する 3 ステップ各々の精度を評価した。  
評価に用いたデータセットおよび目標精度と各々の実験結  
果について以下に述べる。

##### 5.1. 評価データセット

指示すべき文字列と注釈記号の型が記入された見本書類  
を被験者に提示し、デジタルペンを用いて収集用書類に注  
釈記号を記入してもらった(例:図 8)。被験者には見本書類  
と同じ文字列を同じ注釈記号で指示してもらうように依頼  
した。被験者は社内から 10 名を選んだ。

書類は想定シーンに従い、ニュースリリースと技術論文、特許、英文 Web ページの4種類を用意した。注釈による文字列指示の従来研究[5]-[9]を考慮して、今回収集する注釈記号を表1に示す4種類と定めた。プロトタイプで対象となる丸囲みや下線以外の注釈記号も集めたのは、プロトタイプ使用時に実際にユーザが紙面に対象以外の注釈記号を記入する可能性があり、それら記号を正しく切出し、除外できるかを確認するために収集した。

5.2. 目標精度

ネット検索や辞書引きの利用シーンから、目標精度を表2のように設定した。想定シーンにおいて、注釈文字列の抽出に誤りがあった場合、ユーザが計算機上で修正することになる。このとき注釈記号の切出しそのものを誤ると結局全ての文字列を再入力しなくてはならないため、注釈記号切出しは100%に近い精度が必要である。そこで注釈記号切出しの目標精度を99%以上と定めた。抽出された文字列の一部をユーザが修正しなければならない場合が1割程度ならば許容の範囲内と仮定すると、命令要素抽出全体の精度は、注釈記号切出しと注釈記号識別、指示文字列抽出の各々の精度の積となるため、注釈記号識別と指示文字列抽出の目標精度を各々95%以上とした。

表2 命令要素抽出の目標精度

	目標値
注釈切出し精度	99%以上
注釈記号識別精度	95%以上
文字列抽出精度(丸囲みと下線)	95%以上
命令要素抽出精度(上記精度の積)	90%以上

5.3. 注釈記号切出し精度

変動閾値採用以前は、過統合(丸囲み同士の二重統合 27箇所、三重統合 2箇所)と過分割(括弧の縦画、横画の誤分割 1箇所)が存在していたが、変動閾値の採用によりそれらを含めて、全件正しく切出せた。

5.4. 注釈記号識別精度

注釈記号の識別に関し、全 1,444 記号を正しく識別できた。リジェクトもなかった。

5.5. 指示文字列抽出精度

指示文字列の抽出において、連続性算出式のパラメータ決定のため、事前にパラメータ学習が必要である。したがって、データセットを5人分ずつA,Bに分け、クロスバリデーションで評価した(表3)。本評価では、助詞や句読点、括弧等を含めて完全一致のみを正解とした。成功例と失敗例を図9に示す。

丸囲みで 96.3%、下線で 94.2%であった。丸囲みと下線の合計の指示文字列抽出精度は 95.2%となり、目標の 95%以上を達成できた。また命令要素抽出全体の精度も 95.2%となり、目標の 90%以上を達成できた。

6. 関連研究との比較

本研究と関連した従来研究として、鈴木らの情報活用システム[8][9]がある。タブレット PC 上で、書類の着目箇所をマウス等で指示し、対話型プロセスにより関連情報を検索する手法を提案している。文書 OCR により、電子書類だけでなく紙書類の入力にも対応する。入力書類内容を解析し、応用規格に基づいた XML 文書に変換する。XML 文書のメタ情報を活用し、ペン操作したユーザの意図を推測し、的確な情報を提示する仕組みとなっている。

本研究と従来研究はビジネスユース、特に知識ワーカーの活動支援という目的は同じであるが、アプローチが異なる。本研究では紙媒体の利便性に着目した実世界指向のインターフェースを提供する。一方、従来研究では電子媒体による対話型インターフェースを提供する。したがって、閲覧性などの利便性や紙との親和性などの点で本研究が優れ、応答性などの点で従来研究が優れる。適用アプリケーションの属性に合わせて、これら技術を使い分ける、もしくは組み合わせることが重要であると考えられる。

7. おわりに

本報告では、紙と電子の両メディアの利点を活用できる直感的な計算機操作機能を提案した。本機能の基本設計およびプロトタイプ開発による評価から、命令要素抽出精度 95.2%を得た。

表3 指示文字列抽出精度

		プロトタイプ精度
学習 A	丸囲み指定	95.4% (248/260)
評価 B	下線指定	92.1% (267/290)
学習 B	丸囲み指定	97.3% (250/257)
評価 A	下線指定	96.2% (282/298)
合計		95.2% (1,047/1,100)
	丸囲み指定	96.3% (498/517)
	下線指定	94.2% (549/583)

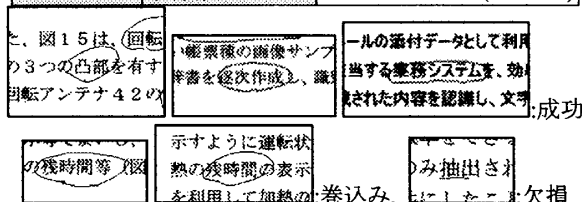


図9 指示文字列抽出の成功例(上段)と失敗例(下段)

参考文献

- [1] <http://www.hitachi.co.jp/tegaki>
- [2] Abigail J. Selten, and Richard H.R. Harper, "The Myth of the Paperless Office," MIT Press, 2001.
- [3] Hisashi Ikeda, "Human Memory Expansion by Personal Handwriting for Realizing Information JIT," Proc. of Wissensmanagement 2005, pp. 650-651, 2005.
- [4] 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," FIT2006, 2006.
- [5] Dan R. Olsen, Jr., Trent Taufer, and Jerry Alan Fails, "ScreenCrayons: annotating anything," Proc. of UIST '04, pp. 165-174, 2004.
- [6] Bill N. Schilit, Gene Golovchinsky, and Morgan N. Price, "Beyond paper: supporting active reading with free form digital ink annotations," Proc. of CHI '98, pp. 249-256, 1998.
- [7] Chunyuan Liao, François Guimbretière, and Ken Hinckley, "PapierCraft: a command system for interactive paper," Proc. of UIST '05, pp. 241-244, 2005.
- [8] 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~, "インタラクシオン 2005 論文集, pp. 193-194, 2005.
- [9] 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," インタラクシオン 2006 論文集, pp. 101-108, 2006.