

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

Development of Document Management System Featuring Utilization of Annotated Information

池田 尚司†
Hisashi Ikeda

小西 康介†
Kosuke Konishi

古川 直広†
Naohiro Furukawa

1. はじめに

PCを用いた書類の作成や検索エンジンによるインターネット上の情報の取得が一般的となり、業務で扱う電子文書は増加の一途をたどっている。一方でオフィスにおける紙の使用量は増加しているという調査結果がある[1]。主に電子文書を閲覧のために印刷する場合が多いこと、あるいは協調作業において情報の記録が紙の場合簡単であることが紙使用量増加の理由となっている。文献中でペーパーレス化の成功例として紹介されている企業においても、設計、開発業務における協調作業の過程で使用される文書は紙である。議論で出てきたアイデアや決定事項を記録するのに紙が便利だからである。このような議論の過程そのものを記述した手書きの情報は、組織知(Organizational Memory)であり、企業における重要な経営資源であるはずだが、電子化されずに紙とともに埋もれているのが現状である。

本稿では、紙文書上の手書き(本稿ではこれをアノテーションと呼ぶ)を電子化し活用するための文書管理方法について検討し、特にオフィスでの利用を例に、PC中の電子文書を印刷した紙文書や、ノートに書かれたアノテーションを活用するプロトタイプシステムについて述べる。

2. アノテーションの役割と活用

紙文書へのアノテーションとは余白への書き込みや、印刷された文字や図を強調するための下線や、訂正のための二重線などがある。手書きアノテーションの役割は2つある。すなわち、(1)新たに付加される情報を表現することと、(2)文書に対するインデックスとなることである。後者の例として、文書中のある単語に下線を引いた場合、その単語が文書の特徴付けるキーワードとなる場合や、そのアノテーションがいつ誰に書かれたかといったエピソード的な情報がその文書の特徴付ける場合もある。

このようなアノテーション情報については、以下の4つの観点での活用が考えられる。

- (1) 筆跡情報の電子化と可視化
- (2) アノテーション情報の共有
- (3) アノテーションの内容、属性情報による文書検索

アノテーションの内容による検索とは、筆記文字列をクエリとした検索の他、下線等によって指示された印刷文字列や図形をクエリとした検索を想定する[2]。属性情報とは、筆記時刻や筆者といった情報の他、アノテーションの履歴情報なども考えられる。
- (4) 手書きした文字列・図形の認識によるコード化

3. 紙文書へのアノテーションを活用可能な文書管理システム“iJITinLab”[3]

3.1. システム概要

前節で述べたアノテーション情報の活用を可能とする文書管理システムについて検討した。オフィスにおける利用を想定して、(1)PC上で作成した電子文書のプリントアウトへのアノテーションと、(2)研究ノートや実験ノート等に書き込んだアノテーションの利用を対象とした。

PC上の文書のプリントアウトへのアノテーションを電子化するためには、図1に示すように、PC上の電子文書を印刷して、そこに書いたアノテーションを電子化するとともに、そのアノテーションを元の電子文書と対応付けて保存しておく必要がある。ノートに関しては印刷の必要はないが、書かれたアノテーションをページごとに保存しておく必要がある。また、研究ノート・実験ノートの場合、手書きだけではなくPCで作成した図表などを切り貼りすることも多いため、貼られた図表データをノートデータに対応付けることも行う必要がある。

なお、今回の実装では手書きアノテーションのコード化は行っていない。

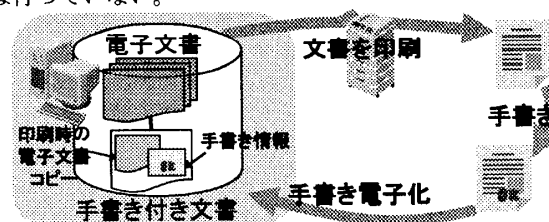


図1 印刷した紙文書へのアノテーションの電子化の流れ

3.2. アノテーションの手段としてのデジタルペン

手書きアノテーションの電子化手段として、Anoto社の方式[4]を用いたデジタルペンを採用した[5]。紙の上の場合毎にユニークなドットパターンが印刷されており、このドットパターンをペン先のカメラで読取ることにより、ペンの軌跡がドットパターン空間における座標データとして表現される。また、電子文書を印刷時に紙毎に異なるドットパターンを動的に割当て印刷することにより、印刷された電子文書と紙上に書いたアノテーションを対応させて電子化するようにした。

3.3. システムが提供する機能

プロトタイプでは、以下の機能を実現する。

- (1) 筆跡情報の電子化と可視化
 - (a) 電子文書のドットパターンを付加した印刷

PC上で印刷命令を実行するだけで、特別な操作をすることなく、ドットパターン付きで紙文書に印刷される。
 - (b) アノテーションの電子化と電子文書との対応付け

† (株)日立製作所中央研究所

アノテーションは筆跡だけではなく、筆記時刻や筆者(ペン ID と対応)などとともに電子化される。電子化されたアノテーションデータは印刷元の電子ファイルに対応付けて保存される。複数ページを1枚に印刷するなど、紙上のレイアウトが変わっても、アノテーションの論理的な位置は保存される[6]。

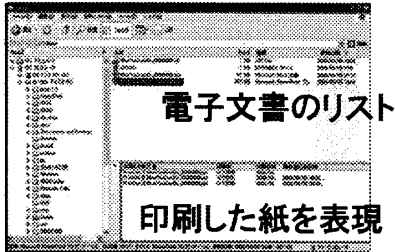


図2 アノテーション情報と電子文書との対応付け

(c) アノテーション付きの紙文書イメージの表示

PC上のファイルブラウザから紙文書のエンタリを選択することにより、手書きアノテーション付きのイメージを表示することができる。



図3 アノテーション付き紙文書イメージの表示

(2) アノテーション情報の共有

手書きアノテーションの情報は電子文書とリンクされて保存されており、元の電子文書をコピーすることによりアノテーション情報も共有することができる。また、アノテーションを含む文書ファイルを生成することにより、電子メール等を介した共有も可能である。

(3) アノテーションの内容・属性情報による文書検索

(e) 書込み日時/筆者/印刷日時等によるアノテーション、およびそれを含む電子文書の検索
電子文書自体の属性ではなく、これを印刷した紙文書上での行為の情報による検索を可能とした。

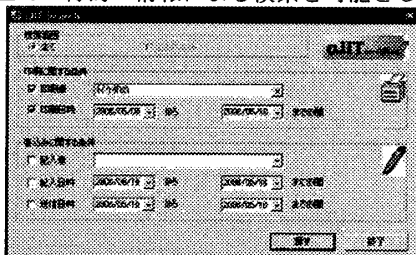


図4 アノテーションの属性情報による文書検索

4.3. プロトタイプシステムの構成

本プロトタイプシステムはサーバ・クライアント構成で実装した。デジタルペン用のドットパターンと各文書の対応関係の管理、および印刷機能と、アノテーション付き文書のDBはサーバ側に実装した。ユーザは自身のPCをクライアントとして、ペンデータの送信、アノテーション付き文書の表示、検索などの機能を実行する。

4. 考察

2006年7月現在、日立中央研究所内でペン200本を用いた試行を実施中である。現在は会議メモの共有、出張報告書コメントの共有などが主な使われ方である。また、予めドットパターンを印刷したノートを所員に配布した。これについては実験結果や研究のアイデアを記録する研究ノートとして使用される。

アノテーションによる文書検索の効果、手書き情報の共有方法については今後試行結果に基づき検討する。

5. 関連研究

Guimbretière[7]らは、デジタルペンで紙に書き込んだ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した。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た。紙文書を用いて電子文書の編集作業を行うことにフォーカスしている。Yeh[8]らは、生物学者がフィールドワーク中にノートに書いた手書きを電子化するとともに、筆記と同時に撮影した写真などを時間と場所でリンクして管理し、専用のソフト上でノートや関連する写真をブラウズ可能にした。印刷した紙文書上のアノテーションについては考慮されていない。

6. おわりに

オフィスにおける紙文書へのアノテーションの利用のために、知識集約型業務の支援を目的として、印刷された紙文書への手書きアノテーションの電子化と、アノテーションを含む紙文書と電子文書を統合して管理する方法を検討し、プロトタイプシステムを実装し、研究所内での試行を立ち上げた。利用シーン等の分析・検討は今後の課題である。

参考文献

- [1] A.J. Sellen and R.H.R. Harper, "The Myth of the Paperless Office," MIT Press, 2001.
- [2] 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," FIT2006.
- [3] Hisashi Ikeda, "Human Memory Expansion by Personal Handwriting for Realizing Information JIT," WM2005: Professional Knowledge Management Experiences and Visions, pp.650-651, 2005.
- [4] <http://www.anoto.com>.
- [5] Naohiro Furukawa, Hisashi Ikeda, Yosuke Kato, Hiroshi Sako, "D-Pen: A Digital Pen System for Public and Business Enterprises," Proc. of IWFHR 04, pp. 269-274, 2004.
- [6] 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," FIT2006.
- [7] François Guimbretière, "Paper Augmented Digital Documents," Proc. of UIST2003, 2003.
- [8] Ron Yeh, Chunyuan Liao, Scott Klemmer, François Guimbretière, Brian Lee, Boyko Kakaradov, Jeannie Stamberger, Andreas Paepcke, "ButterflyNet: a mobile capture and access system for field biology research," Proc. of CHI 2006, pp. 571-580, 2006.