

# OLEによる文書処理系に依存しない高画質ビジュアル 文書管理方式と実用について

2H-2

平山健一† 鈴木友峰†

(株)日立製作所 ソフトウェア開発本部†

鈴木哲也† 田中哲雄‡

(株)日立製作所 システム開発研究所‡

## 1. はじめに

近年のPC市場の拡大により、ワープロの利用頻度が高くなった。これに伴い、文書の量は膨大になり文書名だけの管理・運用は極めて困難な状態になっている。膨大な文書に対し、文書内容をビジュアルに表現し直感的に認識できれば、文書の検索が高速、かつ正確になり、管理・運用が容易になる。

本稿では、各文書処理系から文書内容を取得し、ビジュアル文書管理を実現するための方法を提案する。また、文書内容を一覧表示した際に文書の内容を判別可能とする高画質縮小方式と表示可能範囲の拡大方式について述べる。

## 2. 文書内容取得方式

従来は、文書ファイルのファイル形式を解析・展開して、ワープロなどの文書内容を取得していた。しかし、この方式では文書処理系毎にファイル形式を調査し解析モジュールを開発する必要がある。また、機能追加等によりファイル形式が変更された場合は、解析モジュールを修正する必要があり作業量は膨大になる。この問題を解決するために、文書処理系に依存しないOLE<sup>\*</sup>による文書内容取得方式を採用した。<sup>1)</sup> 更に、将来的にも統一的なインタフェースで印刷イメージの文書内容を取得可能となる。

## 2. 1 OLEによる文書内容取得方法

OLEによる文書内容取得方法について説明する。まず、文書名を引数としてOLEオブジェクト作成関数を呼び出す。OLEは、登録データベースから、文書に対するアプリケーション（以下、APと略す）を特定し、OLEオブジェクトの作成を要求するメッセージを送信する。メッセージを受け取ったAPは、引数で指定された文書のOLEオブジェクトを作成し、そのポインタをOLEに渡す。要求元は、ポインタを元にOLEオブジェクトの中に含まれているプレゼンテーションデータと呼ばれる印刷イメージの表示用データを取得する。

## 2. 2 文書内容取得時間の短縮

あるAPを例にとった場合、文書内容取得時間は、従来のファイル形式を解析する方式の2~3秒に対して、OLEでは約6倍の13秒かかる。この13秒の内、62%の8秒はOLEによるAPの起動(7秒)と終了(1秒)の時間である。

膨大な数の文書内容を取得する場合は、起動・終了時間を短縮することにより文書内容の取得時間を大幅に短縮できる。これを実現するために、最初の文書内容を取得するために起動したAPを終了しなようにメモリ上にロックしておき、2回目以降はこのメモリ上にロックしたAPとOLEを介して文書内容を取得する。これにより、例えば100個の文書内容を取得するのに起動(7秒)×1回+文書内容展開(5秒)×100回+終了(1秒)×1回=508秒となり13秒×

A Method and a Practice of High Quality Visual Document Management System Using OLE without Dependence on Applications

Kenichi Hirayama<sup>†</sup>, Tomomi Suzuki<sup>†</sup>, Tetsuya Suzuki<sup>†</sup>, Tetsuo Tanaka<sup>‡</sup>, Hitachi, Ltd. Software Development Center<sup>†</sup>, Hitachi, Ltd. Systems Development Laboratory<sup>‡</sup>

<sup>\*</sup>OLEは、米国 Microsoft Corp.が開発したソフトウェア名称です。OLEは、Object Linking and Embeddingの略です。

100回=1,300秒の約39%に処理時間を短縮できる。

### 3. 高画質縮小表示方式

文書内容を縮小して表示することにより、一度に複数の文書内容が表示可能となる。文書の場合は、画像と違い、文字や表枠のつぶれ・かすれが起こらない高画質な縮小表示アルゴリズムが要求される。代表的な間引き処理による縮小と独自に開発した細線保存縮小を50%縮小について比較し、図-1に示す結果を得た。この結果、細線保存縮小方式を採用することにより、文字のつぶれ・かすれ及び表枠の消去を極力抑えることが可能となる。2)

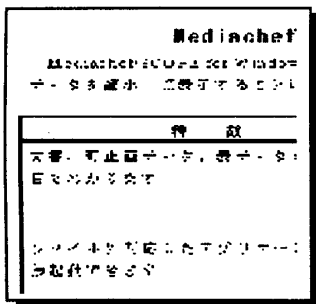


図-1 表(けい線)有文書の縮小表示アルゴリズム比較

### 4. 空白除去による表示範囲拡大方式

通常、文書は、文書の左、上側に空白が存在する。この空白を含めて表示すると表示できる文字数が少なくなる。図-2のように、空白を除去することにより、文字の表示範囲を拡大でき、より文書の内容判別が容易になる。

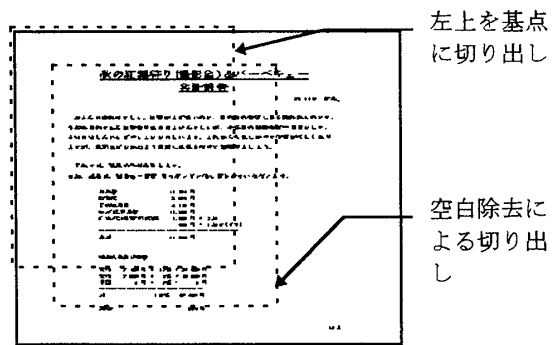


図-2 空白除去による表示範囲拡大

### 5. おわりに

上記方式を具現化し、Mediachef/SCOPEとして高画質ビジュアル文書管理を製品化した(図-3にMediachef/SCOPEの外観を示す)。OLE

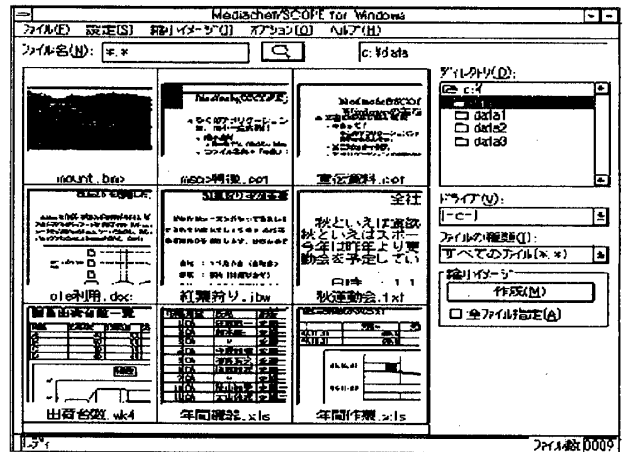


図-3 Mediachef/SCOPEの外観

による文書内容取得方式を採用したことにより1文書あたりの文書内容取得の処理性能は劣るが、まとめて文書内容を取得し、取得した文書内容を保存することにより、実用上は問題なく文書管理ができる。更に、本方式により多種の文書処理系への対応が容易になった。また、細線保存縮小による文書の縮小、および空白除去による表示範囲拡大により、人間が内容を判別可能な縮小表示を実現し、膨大な文書のビジュアルな高速検索を可能とした。

### 参考文献

- 1) INSIDE OLE2:Kraig Brockschmidt 著,マイクロソフト(株)監訳,エー・ピー・ホク/長尾 高弘訳,アスキー刊,1995
- 2) 樋野,宗政,吉村,金澤,“適応サブリングによる線幅保存拡大縮小方式の開発”,情報処理学会第45回全国大会