

電子メールの整理検索における情報視覚化技術の利用

工藤 正人、田中 みどり、古関 義幸
{m-kudo, midori, koseki}@mmp.cl.nec.co.jp

NEC
C&C 研究所

近年の電子メールの急速な普及に伴い、個人の扱う電子メールの数が増大し、電子メールを使いこなすためには整理や検索の効率化が不可欠になってきている。本稿では、情報視覚化技術を適用することで、大量の電子メールを簡単に整理し効率的に検索できるようにした電子メールシステム VisualMail について述べる。VisualMail では、二次元平面上での電子メールの分類状況の視覚化を利用して電子メールを一覧表示することにより、電子メールの分類状況を直感的にユーザに把握させることができ、フォルダでは不可能だった複数の分類先の中の電子メールを一括表示させることができる。また、例示に基づくルール作成を特徴とする電子メール自動分類機能を備えることで、大量の電子メールを簡単に整理でき、更にその自動分類のためのルールを容易に入力保守することができる。

Information Visualization for Electronic Mail Management

Masato Kudo, Midori Tanaka, Yoshiyuki Koseki
{m-kudo, midori, koseki}@mmp.cl.nec.co.jp

NEC
C&C Research Laboratories

With growth of computer networks, electronic mail system has come into wide use. users receive and send a lot of electronic mail about various topics, and an electronic mail system which can manage large volumes of electronic mail efficiently is in great demand. In this paper, two techniques are proposed to classify and retrieve electronic mail efficiently. One is to visualize classification categories. Whole mail is displayed on a two-dimensional plane where two axes represent time and mail contents. The other is automatic classification where classification rules are created by utilizing a user's classification demonstration. The electronic mail system using these techniques, named VisualMail, has been developed.

1. はじめに

近年、電子メールが急速に普及し、業務・教育・プライベートなど様々な目的・場面でのコミュニケーション手段として利用されるようになってきている。これに伴い各個人が送受信する電子メールの量も増大しており、「1ヶ月前に部長からもらった電子メールをどこにしまったかわからなくなってしまった」、「大切な電子メールを読み落としてしまった」といった問題が頻繁に起こるようになってきている。このため電子メールの本来の利用目的であるコミュニケーションや電子メールを用いた業務・作業に支障をきたすようになってきた。このような問題を解決するためには、大量の電子メールを簡単に整理でき、かつ効率的に検索できる必要がある。

電子メールに限らず大量の情報を簡単に整理したり、効率的に検索するための技術として、近年情報視覚化技術が注目されている。情報視覚化技術は、様々な情報を視覚的に表すことによって複雑な情報をわかりやすく表現したり情報を簡単に生成・加工できるようにする技術である。

現在、この情報視覚化技術を適用することにより、大量の電子メールを簡単に整理し、効率的に検索できるようにした電子メールシステムである VisualMail を開発中である。本稿では、この VisualMail について述べる。まず2章で、電子メールシステムを考えるうえで注目すべき電子メールの特徴を述べ、次に3章で、従来の電子メールシステムがどのようにして電子メールを管理しているのかについて述べる。その後4章で VisualMail がとったアプローチについて述べ、更に VisualMail の概要を述べていく。

2. 電子メールの特徴

まず第一にあげられる特徴は、電子メールは手紙やファクシミリなどと同じような非同期分散型のコミュニケーション手段であるということである。電子メールは、コンピュータネットワークを通じて文書を交換し合うことでコミュニケーションを図るものであるから、お互いに離れた所にいる人間同士がコミュニケーションをとることが可能である（分散型コミュニケーション）。また、電話のようにリアルタイムにコミュニケーション

をとることを強制されない（非同期型コミュニケーション）。このため、電子メールによるコミュニケーションには時間的・空間的な制約が少ないという性質がある。しかし同期型や対面型のコミュニケーションと比較すると、一つの話題のスパンが大きくなり、長時間にわたってその話題についての文書がやり取りされる傾向があるため、同時に複数の話題が平行して進行し、それぞれコミュニケーションする相手が異なるということが多くなる。従って、送受信した電子メールを話題や相手ごとに整理しておくことが重要になる。

第二にあげられる特徴は、電子メールは半構造化された文書であるということである。電子メールには規定されたフォーマットに従って書き込まなければならない部分（ヘッダ）とまったく自由に内容を書き込める部分とが用意されている。ヘッダには発信者、発信日時、宛先、タイトルといった情報が記述されるため、その内容を見るだけで電子メールの概要をある程度予測することができる。また、受信した電子メールをヘッダに書かれた発信者名やタイトル別で分類することにより、上で述べたような話題や相手ごとの整理ができる。更に、ヘッダはフォーマットが規定されているので情報を自動的に抽出することが容易であり、抽出された情報に従って受信メールを自動的に分類することも可能になる。

3. 従来の電子メールシステム

大量の電子メールを効率的に検索するためには、前章でも述べたように「送受信した電子メールを話題や相手毎に整理できるようにする」ことが重要である。そしてそのような整理の仕組みを備えた場合、扱う電子メールの量が増えれば増えるほどその整理に手間がかかるようになり、今度はそれが大きな問題になってくる。従って大量の電子メールを簡単に整理するためには、「送受信した電子メールを自動的に整理してくれる」ことも重要になる。

従来の電子メールシステムでは、上で述べた点を実現するために、

- フォルダを利用して話題や相手毎に電子メールを整理する
- 分類ルールやプログラムを作成し自動的にフォルダへ分類する

というアプローチをとっているものがほとんどである。以下ではそれぞれのアプローチについて簡単に説明していく。

3.1 フォルダを利用した電子メールの整理方式

フォルダは電子メールを分類する入れ物であり、ちょうどファイルシステムのディレクトリに相当するものである。フォルダの中に入っている電子メールはそのフォルダが開かれない限り見えない様な構造になっている。このような仕組みにより、話題別や発信者別に関連する電子メールを分類しておくことができ、かつまたそのフォルダを開くことで関連する電子メールだけを一覧することができる。一方個々の電子メールに注目してみると、一覧の中に表示されていることすなわち今開いているフォルダに分類されていることであるから、その電子メールの分類先が直感的にユーザに示されていると言える。

フォルダには以上のような利点があるが、一方で

- 電子メールを業務に利用するようになり送受信する電子メールの量が増えて来た場合、それらの電子メールを管理しやすくするために複数のフォルダを利用して整理することが多いが、その様な場合には電子メールを読んだり、読んだ電子メールに返事を書いたりといった作業は複数のフォルダの間を渡り歩きながら行うようになる。この時頻繁にフォルダの切り替えを行わなければならないため、非常に手間がかかる。
- 送受信する電子メールの量が増えフォルダが増えてくると分類に手間がかかり面倒。

という問題を持っている。また上述の分類の手間の問題の軽減のため、送受信した電子メールのフォルダへの自動分類機能を備えた電子メールシステムにおいても、

- 自動分類がユーザの意図したとおりに行われなかった場合、間違って分類された重要な電子メールを見落とししたり、どこに分類されたのかが解らなくなるためにその電子メールを見失ってしまい、探すのに余計な手間がかかる。
- 新たな電子メールを受信する際に自動分類を行うシステムの場合、受信した電子メールを一括して一

覧できる手段を持たないため、受信した電子メールをまとめて読みたい時には、フォルダの切り替えを頻繁に行わなければならない非常に手間がかかる。

- 上記の新たな電子メールを受信する際に自動分類を行うシステムの場合に、受信した電子メールをまとめて読むのを助けるために、まだ読んでいない電子メールが入っているフォルダをユーザに知らせる仕組みがあるが、送受信する電子メールの量が多くなってくると受信してもすぐには読まないものも出てくるため、それらの読んでいないものと新たに受信したものが混同してしまい、どこに新たに受信したものが分類されたのかが解らなくなってしまう。

といったような問題がある。

以上フォルダの抱える問題を幾つか述べてきたが、それらを眺めてみた場合、フォルダを切り替える際の手間という共通の問題があることがわかる。ではこの手間はなぜ生じるのかというと、それはフォルダ自体の構造によるところが大きい。中に入っている電子メールしか表示しないというフォルダの構造により、複数の分類先の中に格納された電子メールを時間順に一括して一覧することができないのである。これにより、上述の新たに受信した電子メールを自動受信後にまとめて読みたい時などに、電子メールに対する読む、返事を出すといった作業を効率的に行えなくなってしまっているのである。このようなことから考えると、大量の電子メールを効率的に処理、あるいは検索できるようにする上では、「複数の分類先の中に格納された電子メールを時間順に一括して一覧できる」電子メール一覧表示方式が必要不可欠であると言える。

3.2 電子メールの自動分類方式

前節でも少し触れたように、送受信する電子メールの量が増えフォルダが増えてくると分類に手間がかかり面倒になってくるので、その手間を軽減し大量の電子メールをより簡単に整理できるようにするために、決められた書式に従って分類ルールやプログラムを作成することにより、電子メールを受信時や読んだ後に自動的にフォルダへ分類できる電子メールシステムが増えてきている。しかしこの自動分類方式にも、

- 分類ルールやプログラムを書くために特別な知識が必要なことがある。

- 分類ルールやプログラムを書くのに手間がかかる。
- 分類ルールやプログラムの変更が容易にできない。

等の問題があり、従来の電子メールシステムでは自動分類機能がしきいの高いものになってしまっている。

このため大量の電子メールをより簡単に整理できるようにするためには、「ユーザが特別な知識を必要とせず、手間をかけずに簡単に分類ルールを入力し保守できる」自動分類方式が必要不可欠になる。

4. VisualMail のアプローチ

前章では、大量の電子メールを簡単に整理し、効率的に検索できるようにするために、従来の電子メールシステムがどのようなアプローチをとっているのか、またそのアプローチにはどのような問題点があるのかを述べた。更にそれらの問題点を解決するためにどのようなものが必要なのかを検討した。

その結果、

- 複数の分類先の中に格納された電子メールを時間順に一括して一覧できる電子メール一覧表示方式
- ユーザが特別な知識を必要とせず、手間をかけずに簡単に分類ルールを入力し保守できる自動分類方式

が必要であることが分かった。

VisualMail の開発の際には上記の二つの方式の実現を目標とした。そしてそれらの方式の実現のため、

- 電子メール一覧表示に情報視覚化技術を利用
- ユーザの例示に基づいた分類ルールを入力保守

という二つのアプローチをとることとした。以下ではこれら二つのアプローチについて述べていく。

4.1 情報視覚化技術を利用した一覧表示方式

電子メールの分類先がフォルダと同様、直感的にわかり、かつ複数の分類先の中の電子メールを一括して表示することのできる一覧表示分類方式を検討する上で、我々は情報視覚化技術に注目した。情報視覚化技術は、

大量の情報や複雑な情報を直感的に誰にでもわかりやすく表現するための技術である [1],[2]。我々はこの情報視覚化技術を、電子メールシステムに取り込まれている電子メールの分類状況をユーザに示す部分で利用し、それらの電子メールがどんなカテゴリに分類整理されているのかを視覚的に表示するというアプローチをとった。図1に我々のとったアプローチでの電子メールの表示例を示す。

部長			■	
VM Prj.	■	■	■	
A Prj.		■		■
お知らせ	■			
テニス			■	■
		12/10	12/11	12/12

図1 電子メールの分類状況の視覚化

縦軸に分類先、横軸にリアルスケールで受信時間をとる2次元平面上に、送受信した電子メールをアイコンの形でマッピングし、電子メールの分類状況を視覚化するようにした。これにより、複数の分類先の中の電子メールを一括して一覧表示できるようになった。またそれぞれの分類先について表示非表示を切り替えられるようにし、非表示の分類先にのみ分類されている電子メールを画面上に表示しないようにすることで、時と場合に応じて画面上に表示する電子メールを柔軟に切り替えられるようにした。

その後、まとめて送受信する時もあるが全然送受信しない時もある、という電子メールの性質を考慮し、横軸を受信時間から受信順序に変更した。これにより電子メールを受け取った時間の前後関係だけを残り等間隔に電子メールを表示できるようになったため、電子メールがこなかった時間分の無駄なスペースを省くことができた。また電子メールの概要をつかみやすくするために、分類状況と共にヘッダ情報を表示するように改良を加えた。

このようにして電子メールシステム VisualMail の最終的な電子メール一覧表示方式を決定した。その画面例を図2に示す。

図2においては画面の一行が一つの電子メールを表し、画面下から上方向に向かって受信順に表示される。画面の左側には各電子メールの発信者やタイトルなどのヘッダ情報が表示され、右側には分類カテゴリを示す領域

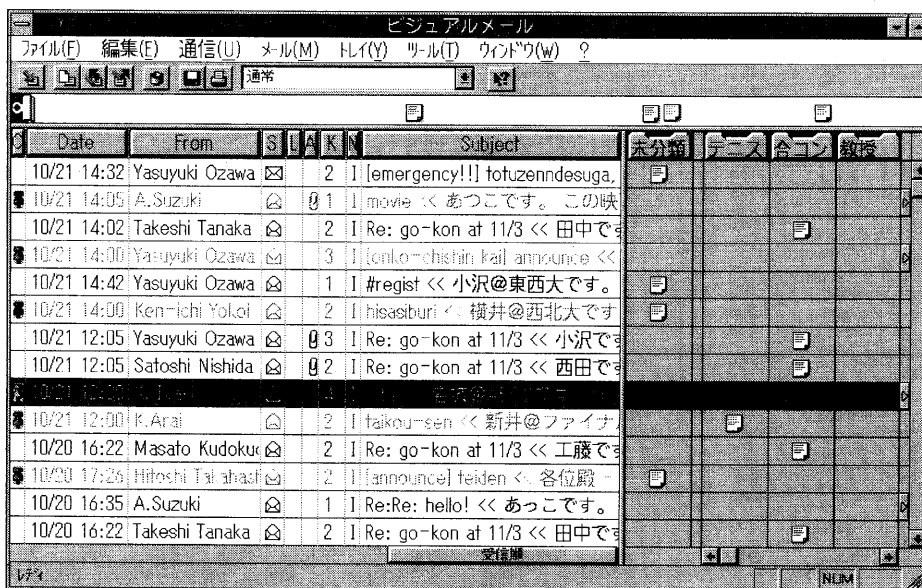


図 2 VisualMail の電子メール一覧画面

(以下本稿ではこの領域のことをトレイと呼ぶ) が複数並んで表示されるようになっている。またトレイ上に置かれた書類の形をしたアイコン (以下本稿では文書コピーと呼ぶ) の配置により、該当する電子メールがどのトレイに分類されているかが示される。一つの電子メールが複数のカテゴリに分類される場合には、複数のトレイに文書コピーアイコンが表示される。文書コピーアイコンは、次節で説明する電子メール自動分類機能によってメール受信時に自動的に配置されるようになっている。また、ユーザ自身がマウスを操作して文書コピーアイコンを移動することにより任意の電子メールの分類を変更することも可能である。

ユーザは、トレイの表示非表示を自由に設定でき、電子メール一覧画面には表示中のトレイに分類されている電子メールのみが表示される。従って、トレイの表示状態を切り替えることによって、特定の話題の電子メールだけを表示する、すべての電子メールを表示する、複数のカテゴリのいずれかに分類されている電子メールをまとめて表示する、など表示内容を簡単に切り替えることができるようになっている。

ここまで述べてきたような方式により各電子メールの内容だけでなくそれらがどのように分類整理されているのかを表示できるようになったが、この結果次のような効果が明らかになった。

- テキスト情報を読まなくても文書コピーアイコンの

配置によって電子メールの分類状況や話題の流れが一目で認識できる。特に、大量の電子メールを新たに受信した時に受信した電子メールの概況を簡単に把握できる。

- 複数の分類先の中の電子メールを一括して一覧表示できるようにし、またそれぞれの分類先について表示非表示の機能を持たせることで、一覧表示させる電子メールを状況に応じて柔軟に切り替えることができる。これにより、フォルダにより整理を行う電子メールシステムが抱えていた大きな問題点の一つである、表示切り替えの手間を大幅に減らすことができ、電子メールを読んだり、返事を出したりといった作業の効率をあげることができる。
- すべての電子メールを時間順に並べるので、どのトレイに分類されているかに関係なく新しく受信した電子メールをまとめて読むことができる
- 電子メール全体の分類状況だけでなく各電子メールのヘッダ情報も表示されているので個々の電子メールの内容も予測することができる

なおこれまでに述べてきた電子メール一覧上での分類情報の視覚表示の他に、VisualMail では更に文書コピーアイコンを自由に置くことのできる領域 (以下本稿で

はボードと呼ぶ)を一覧表示領域の上部に設けてある。この領域は重要な電子メールを常に目に触れるところに置いておくための領域である。受信順の表示領域では、電子メール数が増えて画面上に表示しきれなくなった場合には重要な電子メールもスクロールされて画面上から消えていってしまうが、ボード上のアイコンは常に表示されているので、重要な電子メールを忘れてしまうようなことを防ぐことができる。

4.2 例示によるルール作成を特徴とする自動分類方式

一口に自動分類方式といっても、分類ルールを自動的に学習するもの [3] と、ユーザが定義した分類ルールを利用するもの [4] がある。そこでまずどちらの方法で自動分類を行うかということについて検討した。

電子メールは日常の業務や作業を遂行するためのコミュニケーション手段である。送受信した大量の電子メールの自動分類機能を電子メールシステムが持つことにより、ユーザの負担が大幅に軽減することが期待されるが、その分類が正確でなければかえって業務遂行の弊害となってしまう。また、電子メールでやりとりされる話題や業務内容は時間を追って変化しており、それに伴って、受信した電子メールをどのように整理したいかというユーザの分類基準も変化する。このような変化に素早く適応できなければユーザの満足する分類を行うことはできない。以上のことから、電子メールシステムにおける自動分類機能には「正確性」と「ユーザ適応性」が必要不可欠であると言える。

この視点からこれまで研究されているテキストデータの自動分類方式を見てみると、学習型の分類では、正確性も低く、ユーザ適応のための学習時間が長くなってしまい、電子メールのように素早い適応性を要求される応用分野には不適合である。一方、既存の電子メールシステムでも利用されているユーザ定義型の分類は、正確性は優れているがルールを作成保守のためにユーザの手間がかかってしまうという問題点がある。しかし、「ユーザがいかにか簡単に分類ルールを入力保守できるか」という問題を解決できれば電子メールの自動分類に適した自動分類方式となる。

そこで VisualMail では、基本的にはユーザ定義型の自動分類方式を採用し、上記の手間の問題を解決するために『ユーザの意図する分類がどのようなものなのか、

その一例をユーザ自身に操作させ、そこから分類カテゴリごとに共通するメールボタンを作成する』という手法を取り入れることにした。本手法によれば、ユーザは特別なプログラミング言語やルール書式を覚える必要はなく、分類例を示すだけで自分の意図する分類ルールを作成することができる。また、分類基準を変更しなければならない場合にも一回事例を示すだけですぐに分類ルールを修正することができる。以下では本手法について述べていく。

4.2.1 メールボタン表現

各分類カテゴリには、そのカテゴリに分類されるべき電子メールの特長を表すメールボタンを設定することができる。メールボタンには肯定メールボタンと否定メールボタンがあり、ある分類カテゴリにいくつかの肯定メールボタンと否定メールボタンが設定されている場合、肯定メールボタンのいずれかに一致し、かつ、否定メールボタンのいずれにも一致しない電子メールがそのカテゴリに分類される。図3は分類ボタンの例を示し、

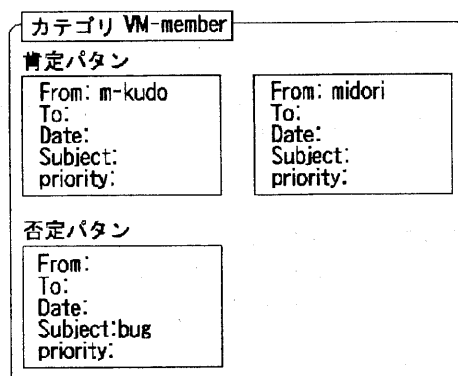


図3 メールボタン

『差出人が m-kudo か midori で、かつ、サブジェクトが bug ではないメールはカテゴリ VM-member に分類する』ことを表している。

新たな電子メールを受信すると、システムはすべてのカテゴリのメールボタンとの照合を行いそのメールの分類カテゴリを決定する。複数のカテゴリのメールボタンとマッチした場合には、そのメールは複数のカテゴリに分類される。逆に、マッチするボタンが一つもなかった場合、そのメールは「未分類」となる。

画面上では分類カテゴリはトレイとして表示され、図

2のようにトレイ上の文書コピーアイコンにより各電子メールがどのカテゴリに分類されているかが表示される。また、各トレイに設定されたメールボタンは図4のようなメールボタン設定画面から設定や修正を行うことができるようになっている。

4.2.2 ユーザ操作に基づくメールボタン作成

図5にメールボタン作成のプロセスを示す。

1) ユーザによる分類操作

画面上のトレイにはそれぞれメールボタンが設定されている。新しいトレイの追加、既存トレイのメールボタン変更など、現在の分類基準を変えなければならない場合、ユーザは画面上の文書コピーアイコンを直接操作して自分の意図する分類を例示する。例えば、「これまでメーリングリスト prj1-ML 宛ての電子メールはトレイ“prj1”にまとめていたが、その中でバグ関連のものは、バグ情報を集めたトレイ“バグ”に分類するように変更したい」場合、トレイ“prj1”に分類されている電子メールの中でバグに関するものの文書コピーアイコンを一つ選びトレイ“バグ”に移動する。複数のアイコンを同時に移動することもできる。

2) ボタン候補抽出

次にユーザの操作した電子メールからメールボタンを作成する。ユーザにより操作されたアイコンが一つの場合には、ヘッダ情報をすべて抽出してメールボタンを一つ作成する。操作されたアイコンが複数であった場合には、それらに共通する項目があるかどうかを調べ、共通項目があるならばそれをボタンとする。共通項目がないならば、操作された各メールに対応するボタンを一つずつ作成する。

3) ユーザによるボタン選択

抽出したメールボタンをユーザに提示し、その中で分類条件として必要な情報だけを選択してもらいメールボタンを決定する。

4) ボタン追加

ユーザが操作した文書コピーアイコンの移動前のトレイに、ユーザの選択したメールボタンを否定メールボタンとして追加する。否定メールボタンを追加することによって、分類例として操作した電子メールと同じボタンを持つ電子メールは移動前のトレイには分類されな

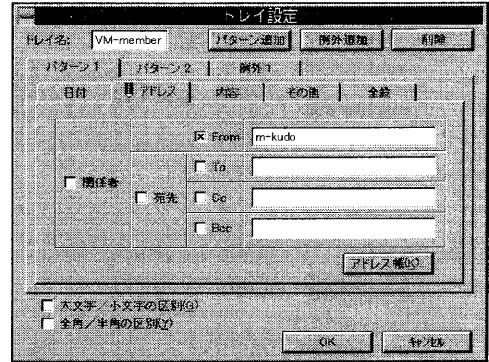


図4 メールボタン設定画面

			未分類	prj1	prj2	バグ
					
10/31	prj1-ML@...	[bug]...				
					
					

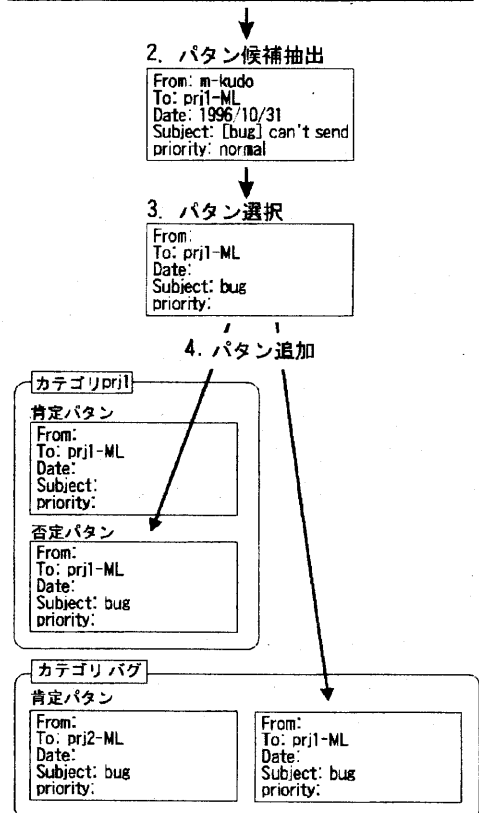


図5 メールボタン設定プロセス

いようになる。

移動先トレイには、ユーザの選択したメールボタンを肯定メールボタンとして追加する。これにより、分類例として操作した電子メールと同じボタンを持つ電子メールは移動先のトレイにすべて分類されるようになる。

これまでに述べてきたような方式によりユーザの分類操作例に基づいて分類基準を作成保守できるようにしたことで、次のような効果が明らかになった

- アイコンを移動するだけで必要な情報を自動的に抽出するので、分類条件をキーボードから入力する必要がない。そのためタイプミスなどのケアレスミスがなくなる
- トレイに分類されるメールボタンを増やす、一つのトレイに分類されていた電子メールをいくつかのトレイに細分化する、など分類条件が複雑になる場合でも、アイコン操作だけで簡単に条件を修正できる

5. VisualMail のシステム構成

VisualMail は、インターネットメールのクライアントアプリケーションとして実装されており、Microsoft 社の Windows NT3.51 以降あるいは Windows95 上で動作する。メールサーバとの通信には、SMTP (Simple Mail Transfer Protocol) および POP3 (Post Office Protocol) プロトコルを利用する。通信処理部分は DLL (Dynamic Link Library) の形でそれ以外の部分とは別に実装されているため、パソコン通信等の他のメールシステム用の DLL を追加していくことにより、VisualMail 上で様々なメールシステムを統合して扱えるようになっている。

6. おわりに

本稿では、現在我々が開発している電子メールシステム VisualMail について述べた。VisualMail では、二次元平面上での電子メールの分類状況の視覚化方式と例示に基づくルール作成を特徴とする電子メール自動分類機能を取り入れたことにより、日常の業務や作業で電子メールを利用している一般ユーザが多種大量の電子メールを簡単に整理でき、かつ効率的に検索できる。

今後についてであるが、VisualMail のこれまでの試用を通して幾つか検討が必要と思われる点が見つかっているため、これらについて検討を進めて行く予定である。大きな問題点としては、

- 画面上に表示できるトレイの数に限りがあるため、トレイを数多く作ってしまうとトレイ自体の管理が難しくなる
- 非表示状態のトレイに対し、例示を使った自動分類ルールの作成ができない
- 分類パターンの記述能力が低いため複雑なルールが書きづらい

等があがっている。またユーザビリティテスト等による VisualMail の評価も今後行っていきたい。

なお、現在 VisualMail の β 版を公開中である。<http://softplaza.biglobe.or.jp/library/doc/tool/visualmail.html> からダウンロードして試用することができるようになっている。

謝辞

本研究の機会を与えてくださった C&C 研究所 後藤所長、阪田所長代理に感謝致します。また、熱心に議論していただき、適切な助言をいただいた C&C 研究所 ターミナルシステム研究部、ソフトウェア研究部の方々に深謝致します。

参考文献

- [1] Ahlberg et.al.; "Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays", *CHI 94 Proceedings*, pp.313-317, 1994.
- [2] Rao et.al.; "Rich Interaction in the Digital Library", *CACM, Vol.38, No.4*, pp.29-39, 1995.
- [3] Lutz et.al.; "MAFIA - An Active Mail-Filter-Agent for an Intelligent Document Processing Support", *SIGOIS Bulletin, Vol.11, No.4*, pp.16-32, 1990.
- [4] 溝口ほか; 「帰納学習による分類機能をもった電子メールクライアントの開発」, *情報処理学会第 51 回全国大会論文集第三分冊*, pp.213-214, 1995.