

解説



ビジュアルインタフェース

—画像による論理情報の新しい伝達法†—

酒 井 善 則††

1. はじめに

—コミュニケーションメディアとしての画像—
人間の扱う機器の高機能化にともない、機械と人間のインタフェース、すなわちヒューマンインタフェース、が重要な研究課題となりつつある。ヒューマンインタフェースは機械と人間の間のコミュニケーションととらえることもでき、コミュニケーションに使用するメディアに応じて種々の形態が考えられている。ビジュアルインタフェースはコミュニケーションメディアとして画像を用いたインタフェースであり、画像のもつ次の特徴を利用している。

(a) 情緒的情報を伝達可能なこと

(b) 国際性、すなわち言語に依存しない特徴があること

(c) 並列性、すなわち複数の画像を同時に表示可能なこと

ビジュアルインタフェースの種類は多く、たとえば、コンピュータ端末の操作性を増すためのマルチウィンドシステム、情報の有効な表現形態であるハイパメディア、人間に感性を伝えることを目標とする人工現実感や臨場感通信などは全てビジュアルインタフェースの代表例と考えられている^{1),2)}。本文では論点の発散を避けるため、これらビジュアルインタフェースのうち情緒的情報の伝達を主目的とするものを避け、人間—機械間で論理情報を伝達するために用いられ、かつ画像情報のもつ特性を生かしたものに対象を絞りその考え方について論じる。すなわち本文で論じるビジュアルインタフェースは

(a) 従来はキーボードやテキスト表示などで行っていた機械—人間間の情報伝達方式の代替あ

るいは補助として用いられるものであり、かつ画像のもつ

(b) 国際性、並列性

を利用しているものに限定する。論理情報を伝達する方向としては機械→人間、人間→機械、双方が考えられる。機械→人間の情報伝達は言語情報に画像を組み合わせる形が多く使われている。これに対して人間→機械の論理情報伝達としては、図形記号や人間の動作・形状を使って各種指示を伝える方法や、アイコンなどのように表示された画像を選択手段として用いる方法が検討されている。本文では以下、人間→機械方向の情報伝達に中心を置きつつ研究の考え方を紹介する。

2. 情報選択手段としての画像

画像により論理情報を表示してこれを選択することにより、等価的に論理情報を入力する方式が数多く用いられている。これは画像情報の汎用性、国際性および並列表示性を利用したものであり、コンピュータの入出力系で適用されている例は多い。

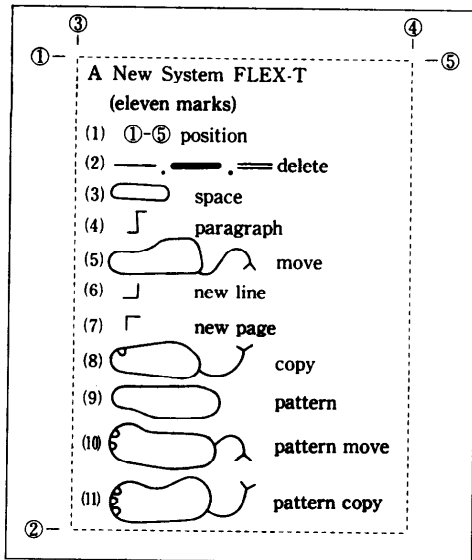
2.1 ピクトグラムとアイコン

情報選択手段として最も多く用いられているアイコンはピクトグラムから発達したものと考えられる。絵文字と訳されているピクトグラムは、意味するものの形状を利用してその意味概念を伝えることを目標としており、学習なしで理解できる、国際的に通用する、即時に理解できる、といった優れた性質をもっている^{3),4)}。

ピクトグラムをシステムとして活用する計画もマンマシンインタフェースを中心に進んできている。図-1は絵文字システムの例であり、単純な幾何形態で単語を作っている。アイコンはコンピュータのインタフェースとして利用されているグラフィックシンボルであり、ピクトグラムを表示して選択することにより、等価的にコンピュー

† Visual Interface (A New Communication Method by Visual Information) by Yoshinori SAKAI (Tokyo Institute of Technology, Faculty of Engineering).

†† 東京工業大学

図-3 FLEX で使用される図形記号^{a)}

(1) ファクシミリを用いた文書編集システム (FLEX)

本システムはファクシミリを用いた文書編集システムとして開発されたものであり^{a)}、図形記号による編集の最初の試みといえよう。編集対象となる原稿と編集用図形記号はファクシミリで別々に入力され、2値画像として蓄積される。原稿についてはセグメンテーションを行い、文字単位あるいは語単位に処理可能とする。図形記号は図-3に示す種類に限定され、コンピュータは入力した図形記号を認識する。次に指定された場所の原稿に対して、図形記号の指示に従い、セグメント化された文字あるいは語単位に移動、コピーなどの編集処理を実施する。

(2) 色マークによる編集

手書きマークによる編集を行う場合、編集対象の原稿と編集用の図形記号をどのように分離するかが大きな課題となる。編集用の図形記号を原稿上に直接記入してかつ区別可能な構成とするため、色マークによる編集システムが開発されている^{a)}。このシステムでは入力した色マーク付き原稿から色マーク部を分離して、その形状および原稿との位置関係より指示内容を認識している。

(3) 自己拡張機能をもった図形記号による編集

筆者らは図形記号に階層構想をもたせて言語として扱い、ユーザが未定義記号を使用した場合は

意味	図形言語		
範囲指定			
複写			
削除			
拡大			
縮小			
回転			

図-4 図形記号の構成例

Mark			
	(Line)	(Hook)	(Circle)
Alphabet			
	(ARROW)	(CROSSING)	(CLOSED)
Word			
	(Copy)	(Delete)	(Rotate)

図-5 図形記号の階層化構成

ユーザに質問してその編集内容を知り、次に使用された場合は正しく理解できるようにする、言語体系の拡張機能をもった編集システムを研究している¹⁾。図形記号は表示された文書の上に透明タブレットを介して電子ペンで入力される。編集者の入力する図形記号の種類は図-4 のようであり、編集範囲とそれに対する動作指示で編集内容が指定される。理解段階では各記号を図-5 のようにマークからワードまで階層化して考える。ここでマークはペンダウンからペンアップまでの一筆で描かれた情報であり、その形状でマークの種類が識別される。アルファベットは複数マークの組合せであり、組合せ方法によりアルファベットの種類が識別される。ワードはアルファベットの空間的位置関係により定まるものであり、たとえばワードが範囲と矢印からなる場合、矢印が範囲の重心にかけるベクトル力の性質に応じて、拡大、縮小、複写などが識別される。図形記号が入力された場合は、マーク→アルファベット→ワードの順に認識して編集指示内容を理解する。ユーザが未定義記号を使用した場合はユーザに質問することにより、その指示内容を知る。次にアルファベットの表現方法を増殖することにより対処する

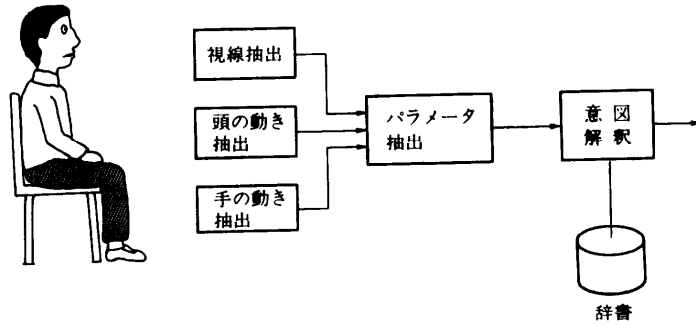


図-6 動作入力システム

ことを試み、対処不可能な場合はマークの種類の増殖を行う。アルファベットあるいはマークの増殖で未定義図形記号の理解が可能な場合は、認識・理解のための辞書内容を更新して、次に使用された場合理解できるようにする。このようにすることにより、次の特徴をもつシステムが構築できている。

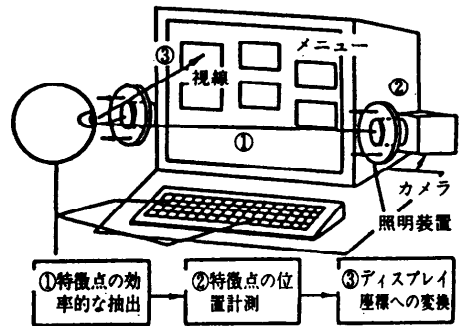
(a) 図形記号理解部はユーザ固有の辞書をもつことができ、かつ同一指示内容を複数の図形で表現することも可能となる。

(b) ユーザが新たに使った図形記号も学習して理解可能とすることができる。

本方式は Pen Windows, Pen Point におけるジェスチャに対して、学習機能をもたせたことに相当して、文化や習慣の違いによるジェスチャの差を吸収する効果がある。

3.2 人間の動作情報による入力

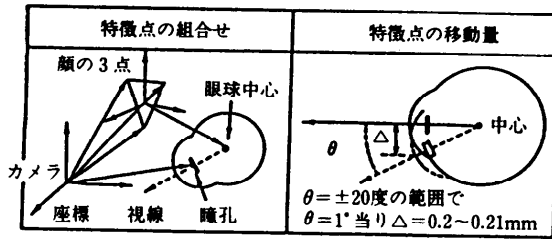
機械に情報を伝える場合まず考えられるのは、音声と手によるキーボード操作あるいは文字、図形の入力である。しかし通常の人間同士の会話の場合、人間自体の形状あるいは動きが情報を表すことが多い。人間の体のもつ情報表現能力はきわめて大きく、たとえば主要な関節の数は10以上あり、1関節の角度で1ビットの情報を表すと考えても10ビット以上の情報表現能力がある。このような人間の動作・形状情報を機械への入力手段として用いる研究が開始されており、その代表例として、●視線による入力、●頭の動きによる入力、●身ぶり手振りによる入力、が検討されている。動作入力システムの概要は図-6のようであり、センサ出力信号より動作パラメータを求め、これを辞書と比較することにより意味を理解する。

図-7 非接触視線検出方式¹²⁾

(1) 視線による入力

人間の視線を検出することができれば、マウスなどのポインティングデバイスを用いずにディスプレイ上のアイコン選択を行うことができる。視線の動きを認識する最も容易な方法はアイカメラを用いる方法であるが、アイカメラは人間に装着する必要があるため、使用する人間が不快感を感じる恐れが大きい。したがってカメラ入力画像の処理による非接触視線認識法が必要となってくる。図-7に非接触視線検出方式の例を示す¹²⁾。カメラにより視線検出に必要な特徴点を抽出する。次に特徴点の3次元位置を高精度で計測して、これより求めた視線方向をディスプレイの座標に変換して視点位置を求める。

視線検出のための特徴点としては、目の反射によってできる第1、第4 Purkinje 像、黒目、瞳孔、顔面がある。これらを組み合わせることにより視線を求めることができる。図-8に顔面と瞳孔による視線の検出法を示す。顔面の3点に分けると、これより眼球の中心を求めることができる。瞳孔の位置と眼球の中心が求まると、これを結んだ線が視線となる。

図-8 顔面と瞳孔による視線の検出¹³⁾

瞳孔の奥には網膜があり入射光を反射するため、入射光と同じ方向からカメラで撮影することにより反射光が観測できる。したがってカメラ方向と離れた位置の光源により撮影した画像とカメラ方向の光源により撮影した画像の差分をとることにより、瞳孔の位置が計測できる。また、顔面上の鼻、唇などの3点を定めこの位置を計測することにより、前述のように眼球中心を求めることができる。

(2) 頭の動きによる入力

人間の頭の動きは YES, NO を表す記号として利用することができる。頭の動きをカメラ入力画像から検出するため次のような手法が考えられている¹³⁾。

- (a) 背景画像と初期正面画像を入力して、正面画像と背景画像の差分をとることによりシルエットを作成する。
 - (b) 首の位置を認識することによりシルエット像から頭部を切り出す。
 - (c) 頭部抽出後、頭部から髪の毛の部分と顔の部分とを分離してそれぞれの部分の重心を求める。
 - (d) 各フレームごとに髪の毛の部分と顔の部分の重心を求め、これより重心の移動量を求める。
 - (e) 重心の移動量より頭部の x 軸、 y 軸上の回転角度を求め、これより意志の解釈を行う。
- これらの検出法によりリアルタイムに頭の動きが認識できれば、機械側の質問に対する回答情報として利用することが可能となる。

(3) 手振りによる入力

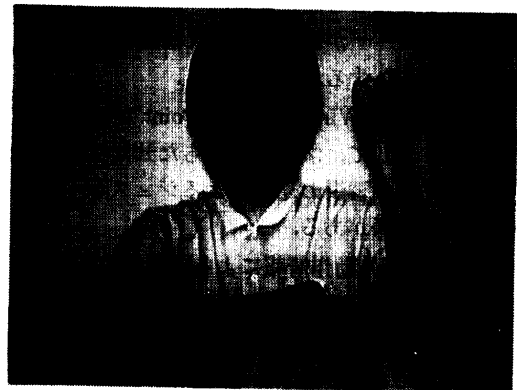
人間の手振りによる情報入力の利用法としては次の二つが考えられている。

- (a) 手の指し示す位置を検出してポイントとして用いるもの。
- (b) 手振り自体になんらかの意味をもたせるもの。

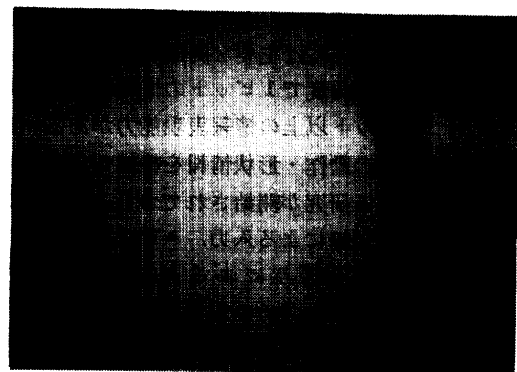
また手振り自体に意味をもたせる利用法として

は、YES, NO の伝達のように機械との会話を目的とする場合と、手話翻訳のように最終的には人間同士の会話を目的とする場合に分かれる。

手振り認識の方法としてはデータグローブを利用する方法が最も容易であるが¹⁴⁾、最近カメラ入力画像の処理による非接触形の手振り認識法が研究されている。筆者らは、色情報により手の位置・方向を検出する方法を研究している^{15), 16)}。色情報による検出法は手の領域抽出に手の色情報を利用する方法であり、輝度情報のエッジ部と手の色領域を組み合わせることにより領域抽出を行い、これをもとに手の位置、方向を検出する。本方式では、あらかじめ計算機に与えられた個人別の手の色情報をもとに識別を行うとともに、背景色の変化を観測することにより、照明光の変化による色情報の変化にも対応する。図-9 に入力画像と抽出された左手の領域画像例を示す。この領域情報をもとに手の軸となる線を求め、翻訳に必要なパラメータを抽出する。



(a) 入力画像



(b) 左手画像

図-9 色情報による手の検出例

認識後の手の情報の翻訳のためには、両手の位置や角度をファジイ変数として辞書に登録しておく、この辞書を検索することにより手振りの意味を知る方法などが提案されている。この方法では、認識した手の位置、動きなどを情報として辞書に入れると、最も可能性の高い意味が辞書より示される。しかし手振りの表す意味をあらかじめ厳密に指定するのは困難であるため、個人対応に別々の辞書をもつこととして、辞書内容をユーザの手振りに合わせて更新する方法も考えられている。

人間の動作によるインタフェースは最終的には意図理解の研究につながるが、福祉用も含めて、リアルタイムで実現できる部分から順に世にできることとなろう。

4. む す び

ーマルチメディアインタフェースへー

人間—機械間のコミュニケーション手段として画像情報を用いたインタフェースであるビジュアルインタフェースの研究動向を、人間—機械の論理情報伝達を目的とする方式を中心に紹介した。画像メディアはコミュニケーション手段として完全なものではなく、今後は音声などの他メディアと組み合わせ、真に画像の長所のみ利用するマルチメディアインタフェースへ発展するものと思われる。

参 考 文 献

- 1) 森島繁生：顔画像によるヒューマンインタフェースの試み，第3回情報伝送と信号処理ワークショップ資料，pp. 43-50.
- 2) 秋山他：21世紀の電気通信—臨場感通信システム—，電子情報通信学会誌，Vol. 73, No. 12, pp. 1355-1359.
- 3) 太田幸夫：3. 情報表現手段としての画像—ピクトグラムによる人間への情報伝達，画像電子学会誌，Vol. 18, No. 4, pp. 178-186 (1989).
- 4) 太田幸夫：ピクトグラム（絵文字）デザイン，柏書房（1987）.

- 5) 市川，平川：4. プログラミングとアイコン，画像電子学会誌，Vol. 18, No. 4, pp. 187-190 (1989).
- 6) 岸本，笠原：データベース検索への応用，画像電子学会誌，Vol. 18, No. 4, pp. 191-195.
- 7) 鈴木：手書き入力 OS 誕生，パソコンの新市場を切り開く，日経エレクトロニクス，1991年4月15日号，pp. 115-132.
- 8) 末永康仁：文書・画像の編集と作成，信学論，Vol. J 68-D, No. 4, pp. 456-465 (1985).
- 9) 林他：色マーク手書き記入による図面の自動編集・校正，情報処理学会論文誌，Vol. 26, No. 4, pp. 740-747 (1985).
- 10) 葛貫他：JIS 校正記号準拠のオンライン手書編集方式，情報処理学会論文誌，Vol. 27, No. 10, pp. 1027-1034 (1986).
- 11) 酒井他：自己拡張機能を持つ文書編集図形言語，画像電子学会論文誌，Vol. 19, No. 3, pp. 138-145 (1990).
- 12) 伴野他：非接触視線検出のための特徴点抽出法，信学技報，PRU 88-73 (1988).
- 13) 境野他：映像による頭の動作の検出に関する検討，信学技報 IE 88-120 (1989).
- 14) 高橋 等：手振り認識方法とその応用，電子情報通信学会論文誌，Vol. J 73-D-II, No. 12, pp. 1985-1992 (1990).
- 15) クンラボン他：学習機能を持った視覚言語処理システムに関する基礎検討，信学技報 HC 90-21 (1990).
- 16) クンラボン他：照明適応能力を持つ色情報の検出法についての一検討，平3電子情報通信学会秋季全大予稿 A-132.

(平成3年7月1日受付)



酒井 善則（正会員）

昭和44年東京大学工学部電気科卒業。昭和46年同大学院修士課程修了。昭和49年博士課程修了。同年日本電信電話公社入社。以来PCM-FDM伝送方式，ファクシミリ通信方式，画像通信会議方式の研究に従事。昭和62年東京工業大学工学部助教授。平成2年同教授。画像通信，画像を用いたインタフェース，通信網の研究に従事。工学博士，IEEE，電子情報通信学会，テレビジョン学会，画像電子学会各会員。