

マルチメディアの本質

—コンピュータ屋の視点から—

舟久保 登

Key words : デジタル表現, コード化, 処理速度, 記憶容量, 論理操作

1 はしがき

筆者の大学における所属学科はメディア・ネットワーク学科である。そしてこの中のメディアなる名称が、多分に最近のマルチメディア傾向に負っていることは、充分推察される所であろう。(ここで「傾向」と述べ、あえて「普及」と言わなかったのに注意されたい。われわれが日常生活に欠かせないテレビは映像と音を同時に視聴させるので、マルチメディア製品と考えられるが、数年後に現在のものと置き代るデジタルテレビの登場をもって私は普及とする。この間の違いの明確化が、本論文の目指す大要の一つである。)そこでこの新しいマルチメディアの本質(中核で基盤をなす事項)を考察し、披露してみることにした*。この種の課題に関心のある人々の議論のきっかけにでもなれば、大変幸いに存ずる。

2 コンピュータと脳の機能

コンピュータ(以下の議論では、われわれにもっとも身近なパソコンを想定して頂

ければ充分である。)は人間の発明した素晴らしい道具の一つであるが、一般に道具とは人間の持っている身体機能の補充、拡大、代替から生れたという素朴な捉え方がある。確かにわれわれは一生懸命足を使っても100mを10秒で走るのは難しく、またこんなことはそう続けて行えるものではないけれども、そこで自動車を用いれば時速40km(100m当り9秒)以上の速度で何時間でも移動可能なわけである。そしてこの文脈で話を進めると、コンピュータは人間の脳、中でもその大脳機能の道具化を実現したものであることが、自然に了解されるであろう。そこで人間の脳はどのような機能を持っているのかについて整理する。

人の脳が右と左の半球から成り立っている事柄は、ほとんどの方がご存知と思う。そしてこのことは生物の体の多くの部分が左右相称にできている事実(顔、手や足、内部の肺など。しかし胃などの消化器官、心臓他は違う。)を想起するとき、そう不思議なものではない。ところで人間を他の動物と区別する大きな特徴とは何だろうか。昔から「道具を使う動物である」と良く言われてきたが、そういう行為ならチン

* 筆者は最近2年間3箇所(豊橋、浜松、岡崎)において、ここでのそれと同じ主題で公開講座を実施した。

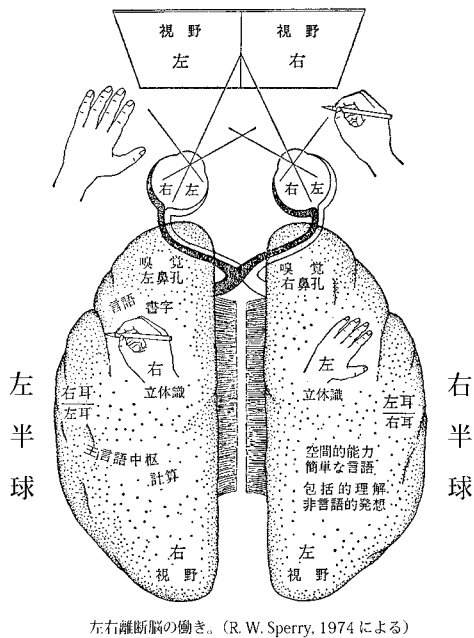


図1 人間の脳の機能

パンジーやゴリラもある程度できるので、その後は「道具を作る動物」と言い変えられたようだ。また「火を制御して使用可能な動物」という主張もある。確かに人間のように自由に火を扱える動物は他になく、この事柄はここで議論しているマルチメディアにも深く関連を有している。それは人が持っている視覚機能の優秀さに依拠する次第で、視覚による外界からの情報取得法におけるその大きな利点のリモート性（直接触れるのではなく、遠隔から観察できる性質）を活用している状況を抜きにしては考えられない。つまり大方の動物のように色覚がなくまた分解能の低い目しかない動物では他の触覚などのメディアに依らざるを得ないが、この場合には情報を捉えると同時にその結果は高熱の火傷のために死に到るといふわけである。

もう一つこれからの議論にもっともつながる重要な特徴は、「人間は言語を使う唯一の動物である」という指摘である。確かにこの事柄は正に人間独自で、深く考えれば考えるほどその意義は非常に大きい。最近のエコロジカル（生態学的）な風潮からは余りさえない言い方であるけれども、他の人との意志の伝達、過去に得た知識の蓄積、加えて言語表現に基づいた明確で正しい推論操作など、「人間は万物の霊長」たらしめている所似はこれにある。ところで大層興味あることは、この脳の一番素晴らしい機能が脳の左半球のある部位に局在している事実である（数年前女性は右半球にもあるというニュースが流れたが、その詳細はさらなる研究を待ちたい）。これは人間という生物にとって言語使用の獲得が進化途上まだ近年の事柄であり、それゆえ両半球を使うほど進歩していない段階という証左であろうか。

さてコンピュータ (computer) は今さら述べるまでもなくその名前が示すように、現在から約50年以上前に数値計算をするための機械として開発された。勿論数値計算は数学における具体的な操作分野に属するが、筆者は日頃数学が嫌いという学生に、「数学は英語などと同様な一種の言語だよ。」と話している。「もし君がアメリカに行ってリンゴが欲しいとき、appleという言葉を知らなかったらどんなに不便か想像せよ。」というわけである。すなわち数学なる言語を使えないことは、人間の最大特権である言語活動のもっとも進んだ部分を非常にもつたいないことに放棄している事態を招いているんだと。以上のような次第で、人間の脳機能を道具化するコンピュータが、まず数値計算の機械化を実現した事

実にはある必然さが存在していたと理解できる。そしてそれからさほど年を経ずしてコンピュータの機能が人間の脳だけが有する言語操作活動（記号処理）に拡張されたことにも、十分に納得がいくのである。

しかしながらこの優れた人間の言語活動もそれだけで閉じているならば、中途半端で不完全な情報処理となる。なぜならばわれわれ人間がそこで生存している外部世界は、決して言語（記号）化なんぞされていない実体だからである。したがってこのようなものからわれわれが上手く生きていくための情報を取得したいとすれば、当然それらがもたらす光や音波、はたまた表面の凹凸、化学的物質の刺激などの物理ないし化学的属性に依らなければならない。そこでこの状況に対して人を含む全ての生物はそれなりの外部情報観測器官を備えており、こうして入力した情報を処理するために、程度の差こそあれ脳全体を用いて対処しているわけである。しかしこの情報の量は大量なもので、特に人間のようにその生活範囲が多様で広範囲な者にとっては膨大過ぎて、たとえ 10^{12} 個の神経素子をもってしても到底対応できるものではない。その結果この問題を克服・解決する手立てとして言語を獲得した次第で（このようになった過程は依然現在における最大なミッシングリングの一つと言われている）、これが実現しなかったら今の人類は存在し得なかったであろう。

こう考えてくると、コンピュータで扱う情報対象が外界の直接呈示するマルチメディアデータに拡大するのは、至極当り前の趨勢である。それではこの状況がどうして最近まで待たされたのか、そしてその実現が果された暁にはどのような事態が起ることに

なるかに、議論を進めることにしよう。

3 情報のデジタル化とその量

コンピュータを詳しく呼ぶ際には、デジタル (digital) なる形容詞を付ける。これは現在のコンピュータが開発された当時もう一方にアナログ (analog) 型があったゆえだが、ほとんどの人が常識的に承知しているようにデジタルコンピュータは取扱う情報を全て2値（通常1, 0と書く）で表現することを前提とする。そこでこの内容について、復習と拡張を引続く議論のために行う。

●数 値

コンピュータは数値計算用というのが最初の目的であったので、数値から話を始めたい。ただし数値は元来離散的なデジタル表現量であり、したがって異なる点は日常ほとんど10進数表示（記号に0, 1, 2, 3, 4, 5, 6, 7, 8, 9の10種を用い、これで足りなくなったら桁上げをする）を使っていることである。そこで必要な事項は単に基数を何に採るかの差だけで、コンピュータの場合はこれを0, 1としたいいわゆる2進数表示を用いている。そしてこの2進数の1桁をビット (bit) と称し、以前は数値を表すのに2進数の36桁、つまり36ビットを割当てた時代もあったけれども、現在は32ビットとするのが次に述べる文字・記号のデジタル表現との関係上、標準となっている。（パソコンのカタログなどで良く見る32bit CPUはこれの直接的な反映である。）念のため10進数との対応を具体的に下書いてみると、普通最上位ビットは正負の符号（0：正、1：負）を表すと約束して、

$$\begin{aligned} & \text{—32個—} \\ (000\cdots 0)_2 &= 0 \\ & \text{—32個—} \\ & \sim (011\cdots 1)_2 = 2,147,483,647 \\ & \text{—32個—} \\ (111\cdots 1)_2 &= -2,147,483,647 \end{aligned}$$

こうして表現し得る整数値は決して小さな値ではない。しかし日本の国家予算に対しては桁数不足だし、われわれが日常でも用いている小数点の付いた表現もできていない。そこでこの対処にいくつかの工夫がなされているが、ここではそこまで触れる必要はない。重要な事項はコンピュータにおいてもっとも基本的な数値情報は、32ビット (=4バイト、この定義はすぐ後で述べる) を使っていることである。

●文字と記号

数値計算のためのコンピュータが実用化されるとほとんど直ぐに文字・記号情報を対象とできるようになったが、この事情は現在におけるコンピュータの利用分野そしてこれに伴う圧倒的な普及に思いを致すとき、非常に大切な歩みを踏み出したと評価される。これはアルファベットのように文字を言葉を記述する記号と素直に捉えられる西洋人だからそれ程抵抗なく発想され、われわれ日本人などのように象形文字文化の人種にはやや飛躍が必要だったのではないかとも思われるけれども、いずれにしろアルファベット、数字それに若干の記号を含めて100種類程度を区別し符号とすれば充分というわけで、8ビットの1, 0が当てはめられた。そしてこの用途は前の数値に比べ比較にならない広範囲なものになったので、8ビットをまとめて新たにバイト (byte) と名付けられ、記憶場所を指定するアドレス (番地) もこの単位で付けられるようになったのである。この影響は大き

く数値のデジタル表現にも及び、上にも記したようにそれまで36ビットを用いていた数値表現が、このバイトの4倍である4バイト=32ビットに短縮されてしまったほどである。ところでわれわれ日本人は言葉を記すのに平仮名、片仮名、漢字を用いており、とても1バイト=8ビット=256種類などでは足りない。そこでパソコンのワープロソフトでおなじみのごとく2バイト=16ビット=65,536種類と2倍に拡張して、文字表現を可能としている。(この2バイト表現を全角と称し、元来の1バイトを半角と呼んで、何となく2バイトの方が規準とみなそうとしているが、この微妙なセンスは外国人にどれだけ感じられているものであろうか。) いずれにしてもわれわれは言語をコンピュータにより扱うために1ないし2バイト/文字を費やしている次第で、私が今この日本語文章を書いている400字詰め原稿用紙一枚当たり約1000バイト弱というところである。

●マルチメディア情報の例としての映像 (動画像)

それではここで議論の対象にしているマルチメディアについては、これらの量はどのようになるだろうか。ここでは具体的にその量を算出するために、その典型であるテレビ映像を取上げる (デジタル化された状態を仮定して)。

映像は外界が光を媒体 (メディア) としてわれわれの目やテレビカメラにもたらす情報であるが、その光の空間上の広がりまた強さは基本的に連続量 (アナログ値) と考えられる。一方コンピュータは0, 1からなるデジタル量しか対象にできないゆえ、この間の違いの関係を若干検討しておく必要がある。ただし読者よ、ご安心あ

れ！ この問題は現在から50年以上前に、既に解決済みなのである。下でこの解決法を直接適用することはしないので、非常に簡単に述べておくと；空間上の広がり連続性については、数年前に死去したシャノンがベル研に所属していた若いときに「通信理論（後に情報理論と呼ばれている）」で解答を与えた。彼の提出した結果によれば、

「もしある信号（情報と同じ）が周波数制限されているならば（現実に存在する信号はその範囲の違いこそあれ、必ずこの条件を満たしているものである）、空間上で飛び飛び（デジタル）に観測した信号値だけから、元の空間上連続な信号を完全に決定できる」ということである（標本化定理）。もう一方の強さについてのデジタル

●数値（1, 0の32個の32ビット）

0 : 0
 8 : 0 1 0 0 0
 2 0 0 5 : 0 1 1 1 1 0 1 0 1 0 1

●文字・記号（1, 0の8個の8ビット=1バイト=半角）

8単位JISコード（X0201）

列 行	0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1								0 0 0 0 1 1 1 1 0 0 0 0 1 1 1 1								0 0 1 1 0 0 1 1 0 0 1 1 0 0 1 1								0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1							
	b ₈	b ₇	b ₆	b ₅	b ₄	b ₃	b ₂	b ₁	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F								
0	0	0	0	0	0	0	0	0	0	NUL		SP	0	@	P	`	p	未定義	-	ク	ミ											
1	0	0	0	1	1	1	1	0	0	!	1	A	Q	a	q			°	ア	チ	ム											
2	0	0	1	0	0	0	0	1	1	"	2	B	R	b	r			「	イ	ツ	メ											
3	0	0	1	1	0	0	0	1	1	#	3	C	S	c	s			」	ウ	テ	モ											
4	0	1	0	0	0	0	0	1	0	\$	4	D	T	d	t			,	エ	ト	ヤ											
5	0	1	0	1	0	0	0	1	0	%	5	E	U	e	u			・	オ	ナ	ユ											
6	0	1	1	0	0	0	0	1	0	&	6	F	V	f	v			ヲ	カ	ニ	ヨ											
7	0	1	1	1	0	0	0	1	0	'	7	G	W	g	w			ア	キ	ヌ	ラ											
8	1	0	0	0	0	0	0	1	0	(8	H	X	h	x			イ	ク	ネ	リ											
9	1	0	0	1	0	0	0	1	0)	9	I	Y	i	y			ウ	ケ	ノ	ル											
A	1	0	1	0	0	0	0	1	0	*	:	J	Z	k	z			エ	コ	ハ	レ											
B	1	0	1	1	0	0	0	1	0	+	;	K	[k	{			オ	サ	ヒ	ロ											
C	1	1	0	0	0	0	0	1	0	,	<	L	¥	l				ヤ	シ	フ	ワ											
D	1	1	0	1	0	0	0	1	0	-	=	M]	m	}			エ	ス	ヘ	ソ											
E	1	1	1	0	0	0	0	1	0	.	>	N	^	n	~			ヨ	セ	ホ	ト											
F	1	1	1	1	0	0	0	1	0	/	?	O	_	o	DEL			ヲ	ソ	マ	ト											

●日本語文字（1, 0の16個の16ビット=2バイト=全角）

創 : 0 1 0 0 0 0 0 1 0 1 0 0 1 1 1 1
 造 : 0 1 0 0 0 0 1 0 0 0 1 0 0 1 0 0

図2 数値と文字・記号の1, 0デジタル化

化は、精度の観点から解決される。いま目の前にある机の寸法を測ろうとしたとき、普通の巻尺ではmmまでしか目盛っていないのでそこまで読取ること満足し、実際その机を配置するなどの目的にはこれで充分であろう。これと同じ趣旨で、目やテレビカメラの強さに対する分解精度には当然限界があるゆえ、それより微細な変化は無視し、有限桁の数値にして（デジタル化した）良いのである。

以上のような暗黙の根拠も背景にあって、現在のアナログ形式であるNTSC規格の普通のテレビは、画面の構成画素数は等価的に700（横）×525（縦）、また明るさについての人間の目の識別レベル数は1視野中では100程度とされているので、コンピュータに都合の良い1バイト＝256レベルを採用して良い。したがってこれに基づき計算すると、テレビ1画面の情報の量は

$$700 \times 525 \times 1 = 367,500 \text{ バイト}$$

である。ただしこの画面は白黒テレビ用であり、現在なら当然カラーしか考えられない。そこで人間は全ての色彩を各々独立の明るさを有する赤（R）、緑（G）、青（B）の組合せで知覚している事実から、上の値を3倍して、

$$367500 \times 3 = 1,102,500 \text{ バイト}$$

さらにいまテレビを想定しているから、画像は勿論動いている動画像（テレビの分野では映像という言い方が多用されている）。このためにNTSC規格では毎秒30フレーム（映像の1画面）を表示している。よって、

$1102500 \times 30 = 33,075,000$ バイト/秒という結果が最終的に得られた。この3千3百万の量がべらぼうに大きなことは一目で分ろうが、念のため先に行った400字詰

め原稿用紙に換算すれば、1頁当たり大雑把に1000バイトとしたゆえ、33,075枚となる。またいま手元にある岩波新書が1頁に日本語で16行×42字、200頁位あるので、

$33075000 \div (16 \times 42 \times 2) \div 200 \cong 123$ 冊に対応する。驚かれたであろうか。われわれがマルチメディア機器であるテレビを見ているとき、音を別にしても毎秒岩波新書を約120冊（積み上げると1m強になる）もの情報を与えられているのだ。人間の眼球内の網膜には全部で 10^8 個の視細胞があるとのことで、確かにこの量の情報は目で観察可能である。一方これを受け取り処理・記憶する脳の神経細胞は 10^{10} 個で、もし先のテレビ画面をそのまま直接覚えておくとすると、

$$10^{10} \div 33075000 \cong 300 \text{ 秒} = 5 \text{ 分}$$

しかもたないわけである。

この理屈から考えると多くの生物が脳を持っているにもかかわらず、条件反射的な生存機能しか果せていないことには道理がある。そして1mにも及ぶ書籍が有する知識の内容について思いを到すと、人間だけが駆使できている言語の重要さに、あらためて感動を覚えざるをえないであろう。

●論理値

この名称にはややなじみがないかもしれないが、現在のコンピュータはこれを表現する素子の膨大な組合せでその機能を実現しているとされる単位である。この値は正式には「真（Truth）」と「偽（False）」、しかしコンピュータとの関連では「1」、「0」と表現した方が分り易いと思う。いずれにしても2つの状態のどちらかを指示できれば充分なので、必要な情報の量は1ビット（＝1/8バイト）である。しかしもう一度確認すると、コンピュータの機能は全てこ

れを元にその組合せで仕上っており、またついでに言及しておけば生物の場合も目であれ耳であれどのマルチメディア感覚器官についても、それ特有の受容器の直ぐ後から情報は0, 1の信号からなるパルス系列により送られ始めるという事実がある（後の図3を参照）。この点では共通しているところがあり、ある自然な秘密が隠されているとの感慨を強くする。

4 コンピュータの性能とその意味するところ

4.1 現在のパソコンの能力

コンピュータといってもわれわれが身近に使用しているパソコンで議論には充分なので、これを調べてみよう。

パソコンは情報処理の機械であるから、まず関心は処理を実行するCPU（Central Processing Unit, 中央処理装置）部分の演算速度に向かう。そこで手近なカタログを眺めてみると、一時ほどの急激な進歩は止ったようであるけれども、クロックが3GHzとある。GはKilo, Mega, Gigaの意であるから1000倍の1000倍の1000倍つまり10億であり、もっとも基礎的な操作を毎秒30億回できる性能である。先の情報の量の計算でテレビの場合1秒間に3千3百万バイトと求められたので、これらを比較すると処理速度の方が100倍、すなわち数十回の処理が可能で、充分適用可能な域に達しつつあると評価できる。

またコンピュータの処理方式の原理は最初から変わらず、プログラム記憶（stored program）方式という用語がそのまま示しているように、処理と記憶は一体のものである。そこでこの記憶装置の記憶容量を見

てみると、処理と同時に動作する主記憶の代表的な大きさは256Mバイト、またディスクなどこの主記憶にデータを供給する役目を果たす補助記憶装置の容量は、およそ100Gバイトに達するほどになっている。この状況について再び先のテレビ映像に対する情報の量との対応を試みると、カラーの1画面が1,102,500バイトであったゆえ、主記憶に対しては100画面以上分、補助記憶に関しては1Gバイト=1000Mバイトであるから

$$100 \times \frac{100 \times 1000}{256} \cong 40,000 \text{画面}$$

にもなる。（ちなみに40,000画面はテレビの30画面/秒を考えると、20分強である。最近売れ筋のハードディスク内蔵DVDの録画時間を思い出すとこの値は数十分の1であるが、勿論DVDの場合は画像データ圧縮がなされているお蔭の長時間さである。しかし他方でこんな仕掛けが容易なものも、DVD装置では画像がデジタル化されているためという理由によることは、言うまでもあるまい。）

以上まとめるとここ最近のパソコンを典型とするコンピュータの性能は、われわれに身近なテレビ映像を代表させて具体的に見積ってみると、処理そしてこれと一体である記憶の機能において、正に取扱いが可能な状況になったという事実である。そこでこの事態を踏まえ、続いてその意味を考察し述べたいと思う。

4.2 処理可能さが意味するもの

ここでもう一度、人間が扱っている数値や文字・記号と、伝統的な写真やテレビによるアナログ型の画像の情動的な違いに注目し、その内容について再確認をしよう。

われわれ人間を含む生物は、同じ仲間は無縁のことは、いろいろな外部世界環境と無縁では生存できない。そこで例えば目を通して自分の周りの状況を観測しているわけだが、そこに見られるものは決して数値や文字・記号になっているものではない。つまりアナログ型ということだが、このアナログ (analog) はまさしく言葉通り類似ないし模擬であって、処理の対象にできないか、少なくとも非常に難い形である。なぜならば対象とこの対象を処理する操作は、アナログの観点からすると、必ずしも関連し統一されるところがないからである。

しかし前述したように脳を持つ動物では、そこでお互いに関連を付け統合を果すためであろうか、個別対応な感覚器官を出た直ぐ後の神経細胞において情報は共通なパルス列の形に置き換えられている。した

がってこれにならってわれわれも、外界ではアナログ的に呈示されている画像を、0, 1 だけから構成されるデジタルデータに変換した次第である。そしてこの結果は膨大な量の情報となるが、最近のコンピュータの性能は充分この量を処理し得る能力を持つという状態になった。

ここでデジタル情報の最小単位である1ビットの1, 0は、論理値の表現なことを思い出す必要がある。そして以前からコンピュータが処理するのを得意としてきた数値や文字などの取り扱いは、これらがやはりこの1, 0による表現を持つゆえに可能となっていることに思い至るべきである。そうすれば下のように考えを発展させるのは、そんなに無理のないところであろう。

それはテレビ映像を1, 0のデジタル表現し得たので、これを数値や文字・記号と同様に扱えるようになったという事態であ

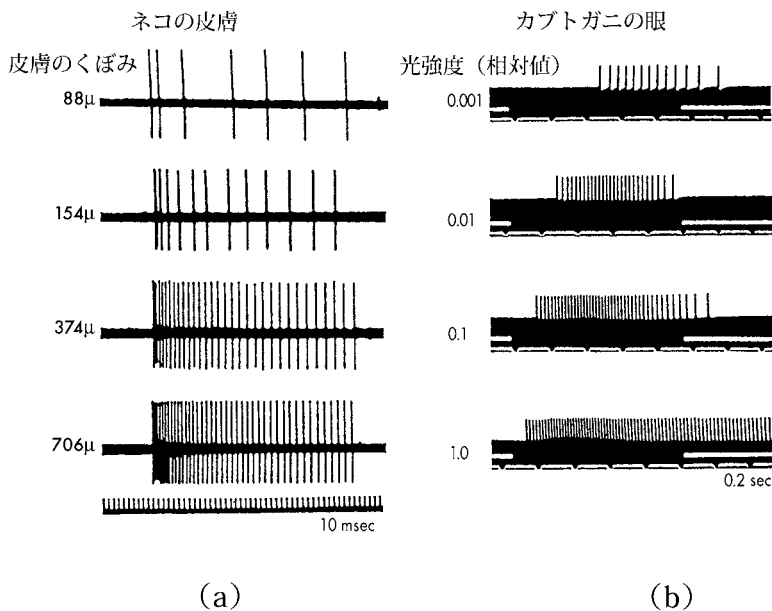


図3 神経細胞におけるパルス列¹⁾

る。つまり映像を1, 0により示したという状態は、それを絵として見せると同時に、潜在的にあらゆる個所で論理操作（この操作は単純なものであるが、前に指摘したように、また現実にコンピュータが行っている如く、全ての処理をなし得る基である）を適用できるわけである。

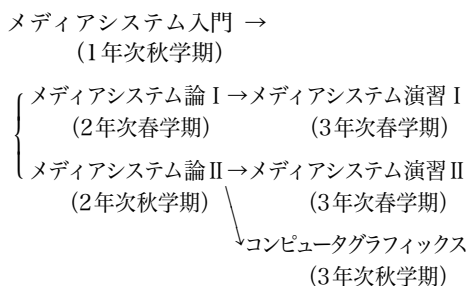
以上ではずっとテレビ映像を例として話を進め、しかも現在ではほとんどこの程度しか筆者には為せないのだけれども、いろいろな種類のマルチメディアを効果的に活用するためには、絶対に欠かせない重大な特性と真に考えている。

5 パソコンによるマルチメディア処理

これまで議論してきたような現状を踏まえて、実際にどのような処理が行われ始めているのかを、豊橋創造大学のメディア・ネットワーク学科の科目構成と内容を例として、具体的に紹介してみよう。

5.1 メディア教育のためのカリキュラム

本大学における直接の科目構成は、下の流れのようになっている。



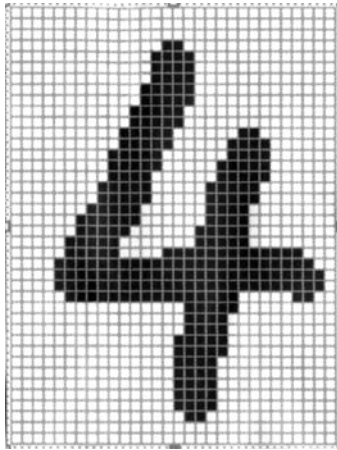
科目名は大きく I と II の系列に分かれているが、マルチメディア分野の現状を踏まえ、I の方は音（音声を含む）関係を取り上げ、II では画像・視覚パターンを扱って

いる。また論は基本的に教室における講義形式により背景となる理論上の知識を与え、演習において理論に基づく実際の操作のためのプログラム作成、この命令によるコンピュータ処理を用いた各種マルチメディア操作の実験を行っている。また現今の学生の関心が深く、興味を抱くコンピュータグラフィックスに対しては独立に授業時間を設け、その分野の性格上この時間は理論と演習の一体化を目指した勉強を行っている。なお現状では、これらの後の方の3科目を筆者が担当している。

5.2 パターン認識

ここから4節を用いて、筆者が行っている授業内容を紹介する。これを通して読者は、マルチメディア処理分野の現況を垣間見ることが可能であろうと意図した積りである。

さて外界が呈示するのは膨大な情報の量になる画像や映像、一方大脳は確かに非常に大量な神経細胞を有するが、それでもこれをそのまま記憶して活用を図るのには無理があり、したがって最高に進化を遂げた人間の場合は言語の使用が生じたとする立場からは、このパターン認識がマルチメディア処理の中心を占めることは当然とも言えよう。ここではもっとも単純ではあるが全ての手法における基礎を与える最短距離法を使った識別²⁾を、手書き数字パターンや笑い顔、怒り顔、悲しみ顔などの顔表情認識に適用している。その例は図4に示すようであるが、折角の機会ゆえにプログラ言語に最新のVisual C++ .netを採用しており、その具体形は近々筆者のホームページでも披露する予定である。



入力数字パターン

4	0	0	0	0	0	0
	0	0	1	1	0	0
	0	0	1	0	1	0
	0	1	1	1	1	0
	0	1	1	1	1	1
	0	1	1	1	1	1
	0	0	0	1	0	0
	0	0	0	1	0	0

入力数字パターンは'4'です。

図4 手書き数字パターン認識の例

5.3 画像処理

ここでは割合パターン認識を重視する立場をとっているゆえ、画像処理においても認識のための特徴抽出や領域分割・抽出に重きをおいている。しかし世間一般では膨

大な情報の量になり勝ちな画像の圧縮などに関心が向いており、この傾向はネットワークの普及に伴って一層その必要性が増加している。そんなわけで、ここでは縁・線の特徴抽出²⁾結果を例示した、われわれ



図5 画像の縁・線特徴抽出

の日常における経験を振り返っても容易に気が付くように、対象についての印象や記憶は色やテクスチャ（表面が見せる一種の様相）よりも形状に基づくことが圧倒的に多い。そしてこの事情を科学的に支持するように、数十年前に行われた猫の視覚神経系路の研究では、系路の最初の方の段階で特定な線分だけに応答する神経細胞が発見されている。またわれわれは新聞などで有名な人の線画似顔絵を見て誰であるかたやすく分るが、この場合写真に比べて印刷などの情報の量が相当少なく済んでいる（圧縮の効果）のにも納得を得られるであろう。言語のように革命的にはではないが、全ての生物においてその進化の程度に応じてではあるけれども、外界の状況を知りそれに適応するための適切なマルチメディア処理が存在している証拠である。

5.4 コンピュータグラフィックス

以上で挙げてきた処理操作が、外界が見せる視覚情報→言語へのある意味で脳の

行っている順当な動作方向のものであったのに対し、コンピュータグラフィックスは逆方向のそれである。そこでこのような事柄について、特に若い人にどうして人気があるのかを若干考えてみたい。その最大な理由の一つは、言語は確かに非常に優れた情報表現ではあるが、やはり元の画像などが持っている内容を完全には表現し得ていないことがあるのだろう。そしてこの事柄は、一方において現代の若者の考えるところが昔と違ってきているため適切な言葉が見出せないためもあるろうし、他方この影響もあってか若者が文章を適切な言葉により作り出す訓練に欠けている事態も無視できない。また芸術としての絵画や彫刻が歴史的に存在してきたように、目は外界に実際に実在するものが見せる呈示を捕えるだけでなく、仮空で想像的な対象を観察して驚き楽しむという役割も持っているようだ。

この分野の教育内容については昨年の紀要に報告した³⁾ので、関心のある向きはぜひ参照されたい。そこでこの経験も踏まえ

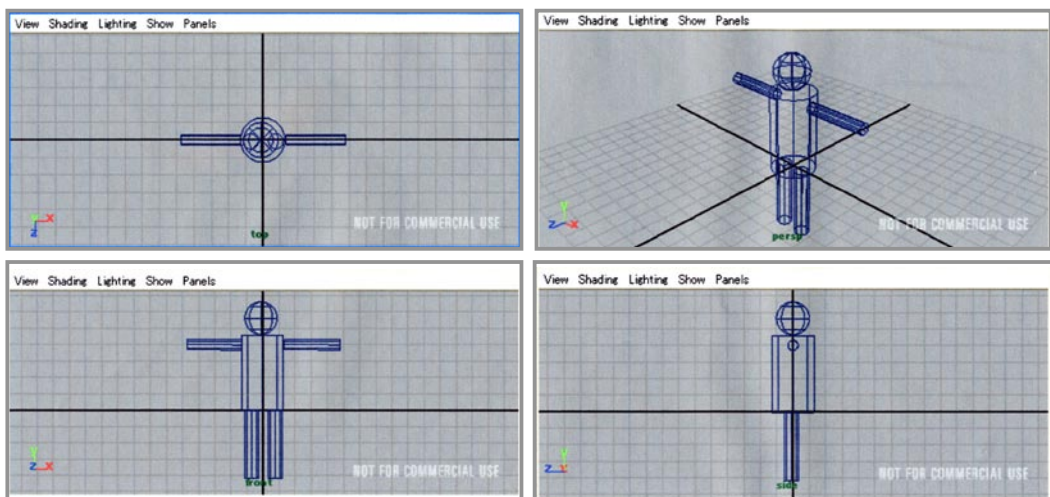


図6 Mayaによる人形状の作成

て、本年度からは本格的な3次元CGソフトであるMayaを導入した⁴⁾。高度なものゆえなかなか私自身使いこなすのに現時点で困難があるのは残念であるが、ここでは球と円柱から構成した人形状をお目にかけることとしよう。

5.5 画像と音声のあるメディア (TVML)

最後にコンピュータによりテレビ番組を制作するよう開発されたTVML⁵⁾を紹介する。マルチメディアの例として前にテレビ映像を対象としたが、現実のものには当然音声を含む音が付いている。したがって今の番組内容からそれほど飛躍したものではないけれども、このTVMLは確かに多様なマルチメディアへの歩みの一歩とは評価できる。それにこれはプログラムによりコンピュータが全てを作成するデジタルなコンテンツであり、もう数年後には現在のものと完全に置き代るデジタルテレビの時代には、一層の活用が十分に予期される範疇である。

6 むすび

今後ますます発展するであろうマルチメディアについて、一番基本的で重要と思われる事項を披露してみたが、理解して頂けたであろうか。もう一度かいつまんで論点をまとめると；この間までコンピュータは、人間にとって独自でもっとも重要な言語活動の機械化であった。しかしその性能の発達により、外界が呈示するアナログ的な情報のデジタル化を通して、人間だけがこれまでなし得ていた言語化をも可能になってきた。これは真の人工知能の実現に対する必然的な帰結であるし、大胆に期待をふくらませれば、人工知能を越える機械知能の出現さえ予想させるものである。

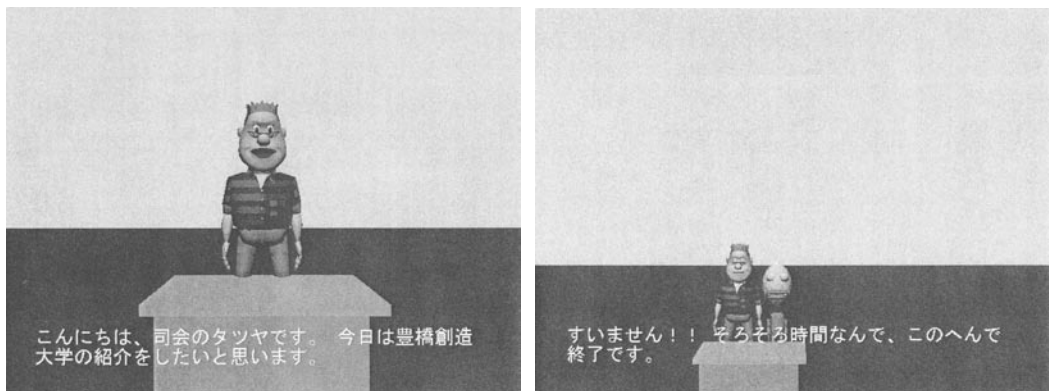


図7 TVMLを用いた大学紹介画面

参考文献

1. 視覚系の情報処理, 永野俊・梶真寿・森晃徳, 啓学出版, 1993, 10.
2. パターン認識, 舟久保登, 共立出版, 1991, 1.
3. 豊橋創造大学におけるコンピュータグラフィックス授業に対する考え方と結果の検討, 舟久保登, 豊橋創造大学紀要, 8, p.85, 2004, 2.
4. Maya 3Dスーパーテクニック, 杉谷泰宏, ソーテック社, 2003, 4.
5. <http://www.nhk.or.jp/str1/tvml/japanese/player2/>, がホームページ.