

II

音楽と映像

聴覚情報と視覚情報についての覚書

はじめに

人はなぜ歌うのか、人はなぜ描くのか（写真を撮るのか）……

人はなぜ音楽を聞くのか、人はなぜ映画やテレビを見るのか……

我々は今日、音楽や映像が氾濫する世界で暮らしているが、この素朴かつ根源的な問いに答えるのは簡単なことではない。おそらく、音楽や映像の創作に携わっているか、あるいはそれに強い関心をもつ一部の人を除けば、「考えたこともない」というのが普通であろう。実のところ音楽や映像の専門家であつてもその答えは人それぞれで、なおかつそれが語られる文脈によつても様々に変化するものである。音楽や映像というものが、非常に広範な関連領域をもつこと、科学的に捉えにくい具体的な存在であること、そしてそれが感情という人間存在の機軸に深く関わっていること、そうしたことが、多種多様な流動的な言説を生むのであろう。この問いに対する答えは当然一つではなく、音楽や映像に関心をもつ人それぞれが、自分なりの答えを模索すべきものと言える。

ここ数年のデジタル機器の急速な普及によつて、いわゆるマルチメディア環境が日常的なものとなり、いつでも・どこでも・だれにでも・あらゆる手段で、情報を取り出す、又創り出すことが可能となつてきた。情報の流れが時間や空間に関してボーダレスになつた

というだけでなく、「オリジナルと複製」「プロとアマチュア」「集団（企業）と個人」など、質の面でも量の面でもその境界があいまいになり、結果、誰にでも音楽や映像を自由に生成し、流通させることが可能になつたのである（特に、インターネットが可能にした「個人から世界への情報発信」は、私たちの世界観を大きく変えた）。

かつて、専門家の間では「今ある技術で何ができるか」ということが重要な問題であつた。しかし誰にでも簡単に高度な技術が使えてしまう今となつては、「何ができるか」ではなく、「何がしたいか」すなわち「何を視たい（聴きたい）」のか、何を創りたいのか、何を伝えたいのか」ということが、より重要な問題としてクローズアップされてくる。その意味でも、音楽と映像に関わっていく場合には、表現上のテクニカルな問題のみならず、「人はなぜ歌うのか、人はなぜ描くのか……」という根源的な問題に立ち返る姿勢が必要となるのである。

このエッセイは、このような問題意識を前提に、音楽と映像に関心を抱きはじめた学生一般の方々を対象に、関連のある様々な分野の知見を、私なりの視点で紹介したいという気持ちで執筆するものである。

ここでの議論の視点は大きく二つある。一つは聴覚と視覚を比較しながら考察するという視点、もう一つはメディアと人間を比較するという視点である。聴覚と視覚の比較というのは、音楽を聴覚情報に、映像を視覚情報に位置づけて、その特徴を比較抽出してみよ

うというもので、目と耳についての「あたりまえの違い」に注目しながら、音づくり・絵づくりのヒントを模索するものである。一方、メディアと人間の比較というのは、音楽と映像の入力・処理・出力に関わる「メディアの情報処理過程」と「人間の情報処理過程」とを比較することで、聴覚情報と視覚情報の入力（認知）や記録（記憶）についてのそれぞれの特性を見出し、モノづくりに関する情報科学的（認知科学的）な理解を深めようというものである。

さて、以上二つが本文の構成に関わる主な視点なのだが、もう一点、「はじめに脳ありき」という考えが考察の基底にあることも銘記しておきたい。今日、バラエティー番組などでも、脳科学が人間の「なぜ」をあつさり説明してしまう場面をよく見かけるようになった。これだけ様々な事実が（そこまで言ったら実も蓋もないということまで）一般に明かされている以上、「脳の話」も交えないわけにはいかない。また個人的になじみ深いということもあって、本文（特に最終章）では頻繁に「脳の話」が出てくることになる。

ちなみに、筆者自身は脳の研究を専門にしているわけではない。もともと「映像美学」が出発点で、現在の専門も認知科学を基礎にした「映像情報のデザイン」である。かつて映像の研究といえば記号論や精神分析が話題の中心で、筆者は学生当時からその難解さと違和感に苦しんでいたのだが、「情報」への関心から認知科学の資料を読み漁っているうちに、90年代になって脳研究の成果をあちこちで目にするようになり、以来、筆者の視点は徐々に「はじめに脳ありき」というものに変わっていったのである。様々な分野ごとに、

バラバラに記憶されていた事柄は、脳科学の知見によって再編成され、おかげで現在「私の脳は非常に快適になった感がある（もちろん、わかった気になっているだけなのだが…）。「脳」を基点に考えることで、「人」と「人がつくりだしたもの」についての考察は、とても見通しの良いものになった。この筆を進める原動力もそれに大きく依存している。

ということ、このエッセイには、芸術の話はもちろん、音響工学・画像工学・情報科学・認知科学・脳科学まで、様々な領域の話が登場することになる。もちろん、これだけ広範な領域の議論を学術的にきちんと書くことは筆者の力量にあまるもので、厳密さを欠く表現や、説明不足な点が多くなることは否めない。しかし、これから新しい世界に足を踏み入れようという段階であれば、多少あいまいであっても先に全体像をつかんでしまう方が頭の整理がつきやすく、また好奇心も湧きやすいのではないだろうか。考えているのが「私の脳」であるから、どうがんばっても厳密さには限界がある。ならば、ということ、であえて領域を広くする方針をとった。

「人生、楽しむ程度に賢ければいい」という愉快な言葉を聞いたことがある。大学に籍を置く身で、そういうのは無責任だとお叱りを受けそうだが、まずは楽しめる程度に全体をつかんでしまおう、というのがこのエッセイのスタンスである。拙い文章ではあるが、ここでの様々な知見の紹介が、音楽と映像に関心を抱きはじめた読者にとつての創造的な思索のヒントとなり、また今後の創作活動の足がかりとなれば幸いである。

II 音楽と映像

目次

はじめに

II・2

● 音楽と映像の歴史

II・8

音楽と映像の領域／芸術のはじまり／音楽の起源／映像の起源／先史時代
エジプト文明／ギリシャ文明／キリスト教文化／複製と大衆芸術の幕開け
アナログメディア時代／デジタルメディア時代／音楽・映像と文化

● 発想法

II・32

混沌から秩序へ／発想の前提・表現の限界／「鍵」を意識する
加算型の発想と減算型の発想／リフレイン／選択的制約
対称性の破れ／手続き記述と美的選択
聴覚情報と視覚情報の比較／「音響・音楽」と「画像・映像」

● 音・音楽の生成

II・52

音声／楽器／電気・電子的手段による音の生成

● 画像・映像の生成

II・62

身振り／画材／コンピュータによるイメージの生成

● 人間とメディア

II・76

感覚・知覚・認知／ノイズ・かたち・ことば／人間の記憶／メディアの記憶
記憶の過程／記憶と想像力／「情報への構え」と「視聴覚の相互作用」

● メディアの聴覚

II・94

音響入力装置の基本構造／音響入力装置の機能／デジタル化と音声認識

音響のアナログ記録／音響のデジタル記録

● メディアの視覚

II・104

画像入力装置の基本構造／画像入力装置の機能／デジタル化とパターン認識

画像のアナログ記録／画像のデジタル記録

● 人間の聴覚

II・122

受容器／可聴周波数と最小可聴限／音程・音色の心理／聴覚野／音声の識別
言葉・メロディーの認知／聴覚のフレームオブリアレンス／聴覚の人工現実

● 人間の視覚

II・136

受容器／可視光線と色彩／色彩の心理／視野／形態の知覚／幾何学的錯視
奥行き知覚／運動知覚／視覚のフレームオブリアレンス／視覚の人工現実

● 結び — 「過剰な脳」の独り言 —

II・158

● 参考文献

II・168

● おわりに

II・170

音楽と映像の歴史

1 音楽と映像の領域

本題に入る前に、まず我々が普通に用いている「音楽」と「映像」という言葉について、また、それをメディアの問題として技術的に扱う場合の「音響」と「画像」という用語について、辞書的な理解をふまえつつ、その輪郭を明確にしておきたい。

音楽 music / 音響 sound

「音楽」とは、文字通りに解釈すれば「音を楽しむ」ということであるが、辞書には「音による芸術。拍子・節・音色・和声などに基つき種々の形式に組み立てられた曲を、人声や楽器で奏するもの、楽」（広辞苑）と記されている。「芸術」を「情報のありかたのひとつ」として位置づけると、「音による芸術」としての音楽は、聴覚に訴える「情報」すなわち「聴覚情報」の一形態であり、生成・処理・伝達・記録の対象として考えることができる。情報の発生源には、声や楽器の生音の他にも、FM音源・サンプリング音源・バーチャル音源など様々なものがあり、また情報の伝達に関しても、直接的な伝達（ライブ）以外に、有線放送・ラジオ・テレビ・Web・各種プレーヤなど多くのメディアがある。今日、

我々の周囲には「音楽」という言葉に関わるモノ・コトがあふれている。

一方、「音響」という言葉は「音」と同義で、「聴覚をおこす空気中の粗密波、あるいは空気中の粗密波によって引き起こされる聴覚」というように二面的に定義されている。それは音というものが、物理的な現象としても、また、心理的・内的現象としても考察できることを意味している。「音響」を我々の身近にあるメディアに関連づけて考える場合には、さらにその形式をアナログとデジタルとに大別する視点も必要になる。アナログ情報としての音は、空気の振動である縦波（粗密波）を機械的・電磁気的な手段で横波に置き換えたもので、文字どおりアナログカルな（類似的な）連続波形として表現される。一方、デジタル情報としての音は、アナログの波形を時間軸上で細かく区切って、各時点での波の高さを数値化して並べるという形で表現される。我々が普段聴いているCDに記録された情報も、片チャンネルの1秒分だけで、4万4千個もの数値データの集まりとなっており、およそ直感的には理解できない、気の遠くなるような存在に変わり果てている。

さて、ここで補足的に確認しておきたいことがある。それは、音楽にせよ音響にせよ、この音というものが流れる存在であって、「時間」の概念を抜きにしては捉えようがないということである。絵を描くのに時間制限はないが、音楽の演奏はリアルタイムでなければならない。絵は見えるモノとして持ち歩くことができるが、音楽は聴こえるものとして持ち歩けない。それを聴くには動くプレーヤが必要なのである。音は常に時間とともにある。これが音に関わるすべてのものに共通の宿命であることを銘記しておきたい。

* FM音源

FM放送のFMと同じ周波数変調（Frequency Modulation）を複雑な倍音生成のアイデアとしてに利用したシンセサイザーユニットで、PCの音源や携帯電話の音源として、現在も利用されている。

* バーチャル音源

物体の素材や形状のデータから、それが作り出す音を物理的にシミュレートするというもの。架空の存在から具体的な出力が得られるという点では、コンピュータ・グラフィックスのサウンド版とも言うべきものである。

* 機械的記録

波の形をそのまま目に見える形で記録したもので、アナログレコードの溝にその典型を見ることができ、「割り箸に画鋲」という簡単な道具でも音を聴くことが可能。

※傷がつくので貴重なレコードでは試さないように：

* 時間軸上の情報

時間軸を必要とする情報は、その記録や再現に「安定して動作（主に回転）する機構」を必要とするため、その意味での制約が多く、扱いが難しい。歴史を遡っても、音の記録はごく最近のものに限られるし、個人のレベルでも「自分の幼少期の声」が記録されているのはごく最近の人に限られる。

「映像」とは比較的新しい言葉で、辞書には「光線の屈折または反射によってうつしだされた物体の像」(広辞苑)と記されているが、「イメージ」にまで幅を広げれば、写真・映画・テレビジョンにおける物体の像のみならず、不可視情報を視覚的に表現したものや「夢」その他、頭の中に思い浮かぶ「想像」にまで意味が広がってくる。ところで筆者は先ほど、音楽を「聴覚情報」の一形態と考えた。議論の構成が対称性をもつよう、映像も同様に「視覚情報」の一形態と考えてみたい。「情報」という概念は、本文のように人間の認知の問題とメディアの技術的問題とをリンクさせながら考えるのに非常に親和性の良い概念である。定義の難しい「映像」という言葉に輪郭を与える鍵概念として、本文ではこの「(視覚)情報」という概念を主軸に据えたい。

さて、「画像」という言葉であるが、これは「像」を工学的な対象として扱う場合の用語で、「2次元平面上に濃淡あるいは色が分布したもの」と説明することができる。メディアに関連して画像を考える場合には、音響の場合と同様、アナログとデジタルの2つの形式の区別が必要となる。アナログ情報としての画像は、フィルムや印画紙で言えば銀粒子の分布として、またCRTや液晶で言えば画面をジグザグに走る電圧の強弱分布として表される。一方、デジタル情報としての画像は、画面を微小領域に区切った画素(電光掲示板で言えば電球)という単位を用いて、その画素ごとの輝度または色(RGB3値の組み合わせ)

のデータを縦横に配列するかたちで表現される。つまりデジタル画像は、縦の画素数×横の画素数×3個の数値データの集合として表され、例えば「30万画素(640×480)」と呼ばれるサイズの画像の場合、 $640 \times 480 \times 3 = 921,600$ 個の数値データの集合となる(約1MB、非圧縮ではフロッピーに1枚が限度)。さらに動画像となれば、一般的には1秒あたり30枚の静止画が時間軸に沿って1コマ1コマ連続して並ぶかたちをとるため、そのデータ量は(同一時間・非圧縮で)音楽の160倍と、膨大なデータとなる。

今日我々は、音楽や映像をデジタルデータとして簡単に扱うことができるが、そのデータ量と処理スピードは直感的な理解の域を越えている。デジタルカメラで撮った写真は、たった1枚ですら、(圧縮しなければ)文字データ換算で本数冊分に匹敵するのである。膨大な量のデータの生成・処理・記録が簡単にできるということが、高度に発達した今日のメディアに支えられているという事実は、日頃から強く意識すべきことであろう。電力の供給が途絶えた途端に「ただの箱」になってしまうコンピュータ、手作業では再現のしようがないデジタルデータ、それが今日の音楽と映像を支える基盤となっている。

さて、音楽と映像、聴覚と視覚、音響と画像、そしてアナログとデジタル、重要なキーワードが揃ったところで、本題に入ろう。広く浅く、舌足らずの説明ばかりになりそうであるが、無理を承知で大きなテーマから斬り込んでみたい。

2 芸術のはじまり

音楽と映像はその誕生以来、どのように変遷してきたのだろうか。

そもそも音や画像はいつの時代から我々に身近な存在となったのだろうか。

人類がチンパンジーと枝分かれした時代であれば約500万年前、現代人と同じDNAを持つホモ・サピエンスのはじまりという意味では約20万年前である。しかしその年代の遺跡にはまだ絵画のようなものの痕跡はない。現在までの調査では、石に刻まれた模様の痕跡が南アフリカのブロンボス洞窟で発見されており、これが約8万年前。壁画というレベルの痕跡であればフランスのショーベ洞窟で発見された約3万5千年前のものが最も古い。ちなみに、3万年前まで共存していたと言われるネアンデルタール人の遺跡からは、石器などはあっても、壁画のような「ビジュアルメモリ」は発見されていない。

高度な言語と抽象的な思考で「予見と計画」を行い、外の世界に様々な情報の痕跡を残す「脳」、すなわち現生ホモ・サピエンスの「脳」の誕生をもって、音楽と映像の起源と考えるのが最も妥当なところだと考えられるが、いずれにせよ「石」か「DNA」に刻まれた情報からしか推察のしようがない遠い過去の話である。

したがって、ここからの考察の多くは推察の域を出ないものとなるが、それでも現代のような便利なメディアが存在しない時代の精神活動を推察することは、創造の原点を考え

*ホモ・サピエンス

近年の発掘調査とDNA鑑定によって、我々と同じ人類は約20種ほど存在したこと、そして、その中で生き残った唯一の種が、大きな脳約1300ccを持ち、高度な言語を操る（言語遺伝子と言われるFOXP2が大きく関与する）現生のホモ・サピエンスであることがわかつている。

*ネアンデルタール人

約30万年前に枝分かれした我々に最も近い存在と言われる。氷期を生き抜いた優秀な狩猟民で、脳も我々と同レベルだが、複雑な言語が操れなかったらしく、絶滅してしまった。ネアンデルタール人の遺跡からは壁画の類が発見されており、彼らが絵を描いたかどうかは定かではない。ちなみに、進化の隣人であるチンパンジーは、ペンを渡せば、落書き程度の行為はする。もともとこの種の問題は、どこかで一線を画したいという我々の発想のほうに無理があるのかもしれない。

るという意味でも重要なことである。そこに表現された内容、伝達の仕組み、さらに社会における位置づけについて、その歴史をおおまかに確認しておきたい。

3 音楽の起源

太古の昔、人が出会ったであろう「音楽」あるいは「音」を、ここでは以下の四つに分けて考察してみたい。すなわち、人や動物の音声、雨や風などの自然の音、人の道具がつくりだす音、そして幻聴という当人にしか聴こえない内的な音の四つである。

音声（動物・人間）

動物の音声を、単純に外敵の存在を知らせる音あるいは食料の存在を知らせる音と見なせば、それはヒトという生物の個体維持のためのシグナルに止まるもので、そのレベルでは音楽の起源としては考えにくい。しかし、人間がこれを真似るとなると話は変わってくる。なぜならば「真似る」という行為は「複製する」ということであり、それは「反復」という音楽的なものと発展する前段階と考えられるからである。

反復される音声は一定のリズムや音程を伴うことで十分音楽となり得る。「ワンワン・」「ケロケロ・」など、擬声語の反復がそのまま歌詞となる例は多く、それを子供たちが好んで口ずさむ様子をみれば、動物の声態模写が音楽の起源と深く関わっていたであろう

ことが想像できる。

人をとりまく環境に存在する様々な音声は、反復によってそのゲシュタルトが意識化され、言語表現の素材となる。人の脳がリズムを意識する最初の体験は、そうした音声の反復であったのではないだろうか。

自然の音

風・水・火・木。人や動物の音声以外の自然の音は、通常は単なるノイズとして無視されるものであるが、特に夜間など視覚刺激の少ない時間になると、様々な感情を呼び覚ますものとして意識されてくる。

風は空気の動きであり、樹間・建物のすき間・あらゆるものの空洞で直接的に突発的な音を発生させる。

水は雨音・川の流れの音・波の音として持続的・周期的な音をつくる。

火は空気を対流させてゴーという持続的な音をだすとともに、ときにパチパチという薪の音、ガサツという炭のはぜる音を出す*。

木は風や水や火の力で間接的に音をつくる。雨に打たれる葉は打楽器、幹や枝の空洞は管楽器、長い蔓は弦楽器とその表現は多彩である。

風の神、水の神、火の神、あるいはまた神の宿る樹木が登場する神話は世界中いたるところに存在する。おそらく原始の人々にとっては、太陽や月と同様にそれらは重要な意味

をもつ存在であり、その奏でる音は神秘的な感情をともなって人々の耳に届いていたものと思われる。

道具がつくりだす音

人が使いはじめた道具は、本来の使用目的以外に、音を発する遊びや祭祀の道具とすることができた。それらは本来の機能を除外することで、そのまま現代の楽器の原型となるすなわち：

棒・器 ↓ 打楽器 (Percussion Instruments)
管・瓶 ↓ 管楽器 (Wind Instruments)
弓 ↓ 弦楽器 (String Instruments) である。

これらの発明は、原始の音楽の発展にとつては非常に大きな出来事であったと思われる。なぜなら、一定の音程（周波数）で音が鳴る楽器、すなわち「楽音」を発生させることのできる楽器を扱うには、まず「チューニング」という周波数を同定する能力が必要であり、さらに音程の差から組み立てられる「音階」というものを成立させるには、聴覚がそれなりの分解能をもつ必要があるからである。太鼓と音声によるリズムだけの音楽に加え、より深化した能力や知識・技術がここが必要になるのである。幼児の音楽能力の発達を観察すればわかるように、まずリズム感、そして音程感の順にその能力は明瞭になってくる。管楽器や弦楽器の誕生の背景には、飛躍的な聴覚系の変化があったと考えられる。

*「火」の存在

落雷や火山の噴火、あるいはガスの自然発火でもないかぎり、火は自然現象としては身近なものではない。むしろそれは、「人がつけて、人が見守っている」ものであり、風や水とは違った意味で「人を安心させる音」であったと考えられる。焚き火の経験があればわかるであろう。その音と光と匂いと暖かさはとても心地よいものである。

*能力：メディアと人間

メディアの性能と人間の能力については、どちらがどちらを規定したかについて述べることは難しい。すなわち、聴覚の分解能が楽器の音程ステップ幅を決めたのか、それとも楽器の音程間隔が人間の聴覚の音程分解能を規定したのか、そのいずれの向きも考えられるのである。

この問題は「人間は寒さゆえに衣服を身につけたのか、それとも衣服を身にまとったために寒さを感じるようになったのか」という問いとも似ている。生物としてのヒトの欠損が文化という松葉杖を必要としたのか（岸田秀）、それとも過剰としての文化が人間の内なる自然を破壊したのか（丸山圭三郎）。とりあえずここでは「物事の因果関係には常に逆の発想も可能であり、そうした逆転の発想が視界を明るくすることもある」ということを強調しておきたい。メディアをつくったのは人間だが、そのメディアによって人間が現在のかたちにつくられたと考えられると見通しがよくなることも多い。

幻聴（怪音）・内なる音

幻聴（幻覚）とは、物理的に空間に音響が存在しないにも関わらず、本人の耳には「ものの音」が聴こえるという現象である（音は「聴こえてくる」ものでもあるため、本人が幻聴と気付いていない場合も多い）。

科学の時代に生きる現代人は「幻」というものを、非合理的なもの、精神的な病のなせるものとして意識下に封じこめているようだが、「神々との交流」が生活の大きな部分を占めていた太古の人々の世界観では、この「幻聴」も音楽的行為の契機となり得る重要な音であったと思われる。現代においても、アルコール、薬物、あるいは激しい舞踏などによるトリップによつて、意識下から自動的に湧き上がってくる言葉やメロディーを拾いあげる作業が時と場所によつては当然のごとく行われている。人の創造活動を語る際にはこの「意識下からわき上がる音」も無視できないものである。

また、幻聴とは異なるが、聴覚からではなく他の感覚器官（たとえば視覚）から感じとる「内なる音」・「内なるリズム」というものも存在する。例えば柱が林立する寺院の中を、一点を見つめたまま歩いてみる。空間に並んだ複数の柱が見る者の移動に伴って一斉にその相対的位置関係を変える。その時視覚は、まさに音楽的なリズムを感じるのである。

ゲーテが「建築とは凍れる音楽である」と言ったように、空間的・視覚的な秩序は、それを見る者の動きによつて、時間的なリズムすなわち音楽となるのである。

4 映像の起源

映像というより、画像の起源と言うべきかもしれないが、とりあえずここではイメージという原語で幅広くそれを捉えてみたい。音楽の起源と同様、四つに分類してみた。

影（人・動物・植物）

影は、オリジナルの複製であると同時にその代理をするもの（アナログン）として最も原始的な存在である。それは光によつてネガティブに生成される平板な形であるが、絵画や写真のように対象から分離することがなく、そのことが影とその主との心理的距離を縮める。影踏みという遊びも、影がその主に最も近い代理物であるという心理的な前提によつて成立しており、「イメージを所有することはその対象そのものを所有することに等しい」という映像に対する人の根源的な心象を物語っている。

自然の造形（そこに想像されるイメージ）

動かない岩石や樹の幹、徐々に形を変える雲、流れる水、ゆらぐ炎・煙など、自然が偶然につくりだす造形は、あるときは動物の形であったり、人の顔であったり、あるいは神や仏のイメージであったりと、様々である。

* 炎：囲炉裏とテレビ

人が管理するものとしての炎は、また別の意味でも映像の起源として興味深いものである。例えば、囲炉裏。その炎は夜間の居住空間の照明として、そこに集う人々のまなざしを集める機能をもつ。それは言ってみれば2000年代以降のテレビであり、「空間の力関係を決定づける」という映像メディアの一つの側面の起源と言えよう。

特に、岩石の表面や樹木の表面などの場合は、外輪郭だけでなく、全体の形がそれと似ているという点で、また実体としてそこに存在するという点で、影よりも独立性の高いイメージとなる。人はそうしたものに對し「そこに靈的な力が宿っている」あるいは「その対象と神秘的な力でつながっている」と考えがちであり、そこに呪術や信仰の対象としての「像」の起源を想像することができる。

道具がつくりだすイメージ

日常的な道具をイメージとの関係で考える場合、大きく2つのタイプに分類することができる。一つは筆記具となるもの、すなわち人の頭と手を使って像を生成するための道具で、もう一つは鏡や針穴など自動的に像を形成する道具である。

前者には大地や岩を引っ搔いて描くための棒や鋭利な鉋物、ふりかけて形にする砂（砂絵）や顔料、植物からとった染料とそれを塗る筆・刷毛の類の組み合わせ、といったものがあり、これは現在我々が用いる各種の画材の原型である。

後者は水を張った器（水鏡）、磨き上げた金属（銅鏡）、そして室内に外界の倒立像をつくる針穴と暗い部屋（カメラオブスキュラ）などで、これは後に写真機に発展するものとなる。ちなみに音の場合は、音を反射させるものとしての壁面や、音を封じ込める（というイメージのある）壺や瓶などが存在するが、それは像の場合の鏡や針穴空間とちがって、録音機に発展するようなものではない（単にそう解釈されていたに過ぎない）。

*鏡

鏡の存在はまた別の意味でも重要である。J・ラカン（1903）流に言えば、人としての出発点はこの鏡像に同一化することであり、この自分自信の鏡像こそが人間にとって最も関心のあるイメージであることはまちがいない。

*靈的な力

例えば、ボディーパーペインティングというものも、それを画像と考えれば、その起源は装飾品と同様、壁画より古いことが想像される。それは、例えばアマゾン源流域に暮らす人々の暮らしからも想像できるのだが、彼らの暮らしの中では、特に祭事において「精霊の姿を真似る」あるいは「精霊が宿る」といった意味づけでボディーパーペインティングが行われている。それは遊びであると同時に呪術的な行為なのだ。また、アメリカ先住民が描く砂絵「蟻男」なども、雨乞い神事に關わるなど、極めて呪術的な画像である。「描く」という行為にはそうした靈的な存在が關わるという心象がつきまとっているようにある。

描かれたイメージ、とらえられたイメージ、原始の心象ではイメージはその主（対象）と神秘的に結びついており、像が対象から靈的な力を写取ると考えられたり、代理物としての像に対して行う行為が現実起こると考えられたりする場合が多い。日本に写真術が伝わった幕末、「写真を撮ると魂が抜かれる」という迷信が流布したが、このような思考は、現代人の感覚にも無いわけではない（平気で写真を踏んだり破いたりできる人は少ない）。

幻覚・夢

通常、視覚には視野という限界があつて、映像はその外から目に飛び込んできたりはしない（音とは違う）が、幻覚と夢は唯一「勝手に見えてくる」ものである。それは意識から解放された脳活動のなせるわざであり、「内なるイメージ」である。

人はよく「夢は経験があるが幻覚は経験がない」と言うが、入眠幻覚という言葉があるように、眠りに落ちかけた時、あるいは睡魔と戦っている時に人はそれと気付かぬままに幻覚を見ていることも多いはずである。超自然の存在として話題になる妖怪や幽霊も、その9割以上が幻覚であると言われ、その意味では、夢や幻覚というものも、異界からのメッセージとして大切に受け止められていたと考えられる。

起源の問題としては余談になるが、幻覚や夢は、今日我々が普通に映像を見るという場合にもつきまとっている。映像には「知覚」と「想像」の両方が作用しており、この「想像」の部分が（無自覚的に付加されるという意味で）幻覚や夢と同類なのだ。映像を見る

*幻覚

ペンフィールド（1962）による側頭葉の電極刺激実験もそうだが、外界からの電磁的な刺激によって、「光」のみならず「鮮明なイメージ」が浮かぶことがあることが近年の脳研究でも確認されている。ちなみに、心霊スポットと言われる場所が地質的に磁場の変動が大きい場所であるということも最近よく知られるようになった。

脳科学による種明かしが進んでしまうと、世の中若干味気ない。

というとき、ほとんどの場合我々はカメラの枠の外を勝手に想像して「確かにそんな風景を見た」と思い込んでいるが、例えばテレビに映しだされる風景も、枠の外はこちらの勝手な想像なのである（テレビで見た場所に実際に行って、思い描いていた風景との違いに愕然としたという経験が誰にでもあるのではないだろうか）。

映像の面白さは「覚醒したまま夢が見られる」という点にもある。無自覚的にわき上るイメージというものは、現代の我々にも身近な存在なのである。

5 先史時代

ラスコー・アルタミラ・ショーベの洞窟に代表されるように、先史時代の生活は氷河期のピークということもあつて洞窟の中を拠点としていた。

音に関して言えば、ここでは石琴らしき石片のセットが発見されており、音楽が何らかのかたちで奏でられていたことは、ほぼ確かである。「楽器」の存在から「音」は直接再現可能だが、音律や音階のシステムについては推測の域を出ず、また「楽譜」に相当するものがないため、「音楽」は再現のしようがない。

画像に関しては、その時代の洞窟遺跡から数々の壁画が見つかつており、約3万年前には、明らかに人間がイメージというものを操作していたことが想像できる。壁画に描かれているのは、自らの存在証明であろう「手形」（現代の子供たちの遊びにもそれとよく似

た行為が観察される）、スパゲッティ状の線や丸・三角などの抽象的な図形による幾何学模様、狩猟・食の対象であつた動物（生命維持）、そして、通常「ヴィーナス」と呼ばれる裸婦像（種族保存）である。言ってみれば、自意識や抽象的思考の痕跡を外の世界に残すという人間の最大の特徴が、生命維持と種族保存という生物としての2大条件と並んで遺されていることになる。

しかしこれらが、現代の画像が担っているような「情報の伝達」に関わったかという点に関しては、その基礎となる言語の形態も含めてわからないというのが現状である。洞窟の壁画でも、そこには「枠」や「順序」がなく、画像として独立した世界を構成するためのパースペクティブ（ものの見方）も存在しない。つまり「システム」が存在しないのである。通常、コミュニケーションを成立させるためには、要素間を関係づける「システム」が必要であり、その痕跡が希薄である以上それらはコミュニケーションのための画像というより、きわめて遊びに近い行為として描かれたものと推測せざるを得ない。

むしろ重要なのは、彼らの残した画像には、今日我々があまり意識しない意味が含まれている（らしい）ということである。一般に先史時代は、遊びも生産も呪術も未分化^{*}の時代であると考えられており、絵を描くことも音を出すことも、遊びの行為であると同時に、呪術的な行為でもあつた。彼らは、それを操ることで、外部に対する恐怖に抵抗し、外敵や超越的な存在をコントロールしようとしていたのではないだろうか。その意味では、先史時代の画像や音とは、人と人との間のコミュニケーションの問題というより、もっぱら

* 氷河期

現在壁画が見つかる先史時代の遺跡は氷河期がピークを迎えていた約2万年前（その当時は氷期、現在は間氷期と言われる）頃のもので多く、その代表的なものは、東アフリカ（我々ホモ・サピエンスが発祥した場所）からそう遠くないフランス・スペイン・南アフリカのものである。

アメリカ大陸ではその当時の人類痕跡がないことから、我々の祖先がベーリングア（おそらく地続きであったベーリング海峡）を越えてアメリカ大陸へと拡散したのは、氷期の終わり、「北の回廊」と呼ばれる氷河の裂け目が北アメリカに生じて以降のことと考えられている。

我々は優れた言語能力に基づく予見と計画によって、極寒の時期を乗り越えた。肌の色は様々であるが、人類拡散のルートを通れば、同じDNAをもった生物に行き当たる。

* 未分化

今日我々は、様々な物事を科学的に区分けして考えているが、現代人の意識下にもそうした未分化な原始的心象が存在することは否定できない事実である。現に子どもたちの会話にも「言魂の発想」（声に出された言葉が超現実的な力をもって他に影響を及ぼすといった発想）が見えとれる。

人と動物、あるいはアニミズム的な意味での神々とのコミュニケーションのために存在したと考えられる。

6 エジプト文明

いわゆる4大文明の時代、紀元前4000年以降の話となると、現存する様々な遺産から当時の状況はある程度推察できる。なかでもエジプトに残された多くの遺産は、人間の関心が「生と死」、「天と地」、「神・人間・動物」に向けられていることを物語っており、あらゆるものが混沌としていた先史時代の精神からは大きな跳躍があったことがわかる。

「生と死」をテーマとする墓やミイラ、「天と地」をテーマとする巨大な建造物、神々の関係を表わす壁画（神の表現は半身獣）。これらはいずれも根源的な「象徴（シンボル）」であり、人々の心に「時間」と「空間」の概念が芽生えたこと、そして、それらを核とするシンボル操作で、人々が共有できる「世界観」が構築されていたことがうかがえる。

また、そのようなシンボルの誕生と同時に、それらによる情報の伝達という側面でも歴史上大きな飛躍が見られる。形に関する共通の理解と視覚コミュニケーションの成立を意味する象形文字と記録・伝達媒体パピルスが存在、そして、リズムや音階のシステムなど音楽コミュニケーションの成立を意味する数々の楽器（ナイルの象徴である葦でつくった葦笛、器を2枚重ねたシンバル状の打楽器や太鼓、ハープやギターのような形状の弦楽器

など）の存在。エジプト文明期には明らかに情報の伝達に関わるハードとソフトが誕生していたのである。

特に象形文字は、先史時代のリアルな壁画と異なり、情報の記録・伝達を目的とした最初のものとして飛躍的な発明であった。また、その約6000の記号のうち、約800程度が音声を伴うと言われ、この形と音声とが同時に結び付く言語こそが、視聴覚コミュニケーションの原点であり、おおげさに言えば音楽と映像の関係の原点でもあるのだ。我々は「言語」というものをあたりまえのものとしてなおざりにしがちであるが、形と音声の両方を同時にともつメディアというのは考えてみれば不思議な存在なのである。養老孟司流に言えば、それは聴覚と視覚を結ぶ脳の連合野の状態を反映している。

*時間の概念・空間の概念・コミュニケーションシステムの成立そして聴覚と視覚の連合、音楽と映像が世界に大きく関わっていく準備が整った。

7 ギリシャ文明

紀元前1000年ごろからとされるギリシャの時代には、建築・彫刻・絵画といった空間芸術はもちろん、神話・演劇といった物語りのなもの、すなわち時間芸術も登場してくる。

この時期の特徴は、哲学に代表される「体系的思考」と神話に代表される「物語り」の登場であるが、これは、「人間」あるいは「主体」という概念の発生を意味するものであ

*神・メディア・人

音楽・映像、いずれも人類の起源にまで遡れば、非日常的に神々の世界と交感するための媒体（メディア）としての意味合いが強い。

日本の芸能で言う「道」の概念も、それを究めることによって、「神」の域に到達することを目指している。

*時間と空間

時間と空間の概念については、哲学・現象学といった領域に限らず、自然科学・社会科学の領域でも様々な議論があるが、一つ簡潔に言えることは、それが人間の聴覚と視覚に関わる概念だということである。人の脳の情報処理系ではこの2大派閥が大きくその思考に関与しており、いわゆる垂直思考・通時的思考・機能主義などの時間を基軸とした思考は「聴覚主義」的であり、水平思考・共時的思考・構造主義など物事を空間的に捉える思考は「視覚主義」的であると言える。

る。エジプト文明が、神の視点からあらゆるものを形にしたのだとすれば、ギリシャ文明は、人間の目から見た主体的な世界観でそれらを表現しており、それが今日の芸術観に結びつく作品を生んだのだとも言えよう。古代オリンピックの競技には芸術のオリンピックもあり（図書分類で「780スポーツ」が「700芸術」に含まれるように、本来は、芸術の方が上位概念）、今日我々が一般に「芸術」と呼ぶ、様々な領域の土台ができあがる。

音楽に関して補足すれば、この時代にピタゴラス音律（5度の周波数比が2…3という規則で楽器を調律する）が存在していたことがわかっており、現在我々が一般に西洋音階として用いているドレミの音階でメロディーが奏でられていたと推察される。

8 キリスト教文化

ギリシャ以後、今日に受け継がれる音楽・絵画の最も大きな流れは、ヨーロッパの大部分を支配したキリスト教文化の下で成熟する。時代が近代に至るまで、音楽も絵画もそのテーマ・内容は宗教とともにあり、今日で言う「芸術家」も、事実上は宗教のための、あるいは特権階級のための「お抱え職人」であつたと言うのが実情である。

音楽の伝達と記録に関して言えば、その発展の過程でいくつかの系統立った音階（教会調という現在の我々に一番馴染み深い音階もその一つ）が定められたこと、複数の旋律を重ねる和声の技法が完成したこと、そして記譜法の原型が完成したことが、歴史的に大き

な出来事である。音律に関しても、和声の響きをよくするために、ピタゴラス音律から純正律（3度の周波数比を4…5とする）そして中全音律へ、さらに転調における和声の濁りを解消するために、最終的な妥協点である平均律（すべての半音程の周波数比を一定）へと変遷する。ちなみにこの時代、音楽の発展に関わった主たる楽器は、キリストの生誕時にすでに存在していたと推察されているオルガンであつた（我々に馴染みの深いピアノがほぼ現在の形に完成するのはベートーヴェンの時代すなわち18世紀後半である）。オルガンという楽器が多声部音楽と和声の技法を生み、それが音律の再構成を進め、それがまた楽器の構成を変えた。ハードウェアとソフトウェアは相互に影響し合いつつ、お互いを再構築するものであることを銘記しよう。

絵画に関しても同様で、画家は宗教画を描くための職人（正式な画家には資格が必要とされた）であり、画像はその「文法・辞書」とも言える様式の決まり事に忠実に描かれることで、文盲の人々に宗教的な物語りを伝えるための重要な手段となっていた。図像学（Iconology）という言葉もあるほどで、この時代の絵画によるコミュニケーションは、今日の我々の絵画観とはかなり異なるシステムティックなものであつたと言える。ダビンチの頃には透視遠近法も確立し、写実的な空間の描写がシステムティックに行われるようになる。視覚コミュニケーションは一つのパラダイムで安定期をむかえていた。

さて補足になるが、歴史を振り返るという場合、記録として残る資料の豊富さと今日への影響力の大きさから、我々は一般にキリスト教文化圏の歴史を中心に考えることが多い。

*「芸術」について

「芸術」という用語は、定義が難しい語である。ここではとりえず「美を創造・表現しようとする人間活動、およびその作品」という一般的・辞書的な意味で用いているが、深く掘り下げたい読者は、模倣論・均斉論・表出論・現象論・異化論といったキーワードで文献を探してみるとよいだろう。

*5度の音程

例えばドとソの音程間隔である。この2つという比率は、ドの3倍音とソの2倍音が同じ周波数の音になるということの意味する。音程が調和する（俗に言うハモる）のは周波数がそういう単純な整数比になる場合に限られる。

現代の楽器が主に用いている平均律は、1オクターブ（周波数比で2倍）の間を等比的に12等分して得られるもので、ギターの1フレット分（ピアノで言えば白鍵と黒鍵の間）は約1:20という周波数比になっている。この音律でもドとソの間隔はほぼ2:3になる。

しかし、それとは異なるスタイルの芸術が「野蛮なもの」「劣ったもの」というわけではない。例えば音楽の場合でも、西洋音楽以外にも様々な民族の様々な旋法による音楽があり、それぞれに優れた側面が存在するのである。

「文化」に優劣はない、「慣れ」が異文化に対する違和感を作り出しているだけである。

9 複製と大衆芸術の幕開け

時代は前後するが、15世紀、グーテンベルクの活版印刷術の発明によって文字や楽譜、図版の大量複製が可能になると、社会には様々な変化がおこりはじめた。複製が大量に作られるようになると、文字が身近になり、それを読みたいという欲望が芽生え、文盲率が減り、社会のコミュニケーション体制が変わる。そして文字の代役も兼ねていた絵画は、それ自身の目的へと解放された。多種大量な情報の流通は、人々の知識を豊富にするだけでなく、表現形式と思想の一般化・標準化を促進し、さらに印刷媒体特有の表現スタイルと思想を生む。印刷物という情報の伝達媒体が、人間の思考回路と社会の構造を変え、という、おそらくこれが最初のメディア革命だったと言えよう。メディアは単なる伝達手段ではなく、それに関わる人間の思考回路や社会の構造に影響し、また表現の内容そのものにも影響するものである。万年筆をワープロに変えることで、できあがる文章の表情が変わるように、メディアの違いはその時代を生きる人々の「脳の状態」と密に関係する。

*様々な音階

一般に音楽のシステムとしての音階には5音階のことが多い。身近な例では、ロックなどが用いるペンタトニックスケール(文字通り5つ)や、四七ノキ音階と言われる演歌の音階(ドレミソラ)、また二六ノキの沖縄音階(ドミファソシ)など、西洋音階の7音階とは異なるシステムは身近なところによくある。

10 アナログメディア時代

第2のメディア革命は19世紀中頃の写真術の発明からである。写実を売り物にした多くの職業画家は職を失い、あるいは写真家という職を生み、あるいは抽象芸術家を生む。ここでもメディアが人と社会を変えていくということが、当然のごとく起こるのである。

19世紀末には映画の技術、さらに20世紀のはじめに蓄音技術、そして1930年代に映画がトーキー化されて、原始以来別々にしか記録することができなかった音楽と映像は、ここではじめて共同作業を始めるようになる。同じころラジオ放送も開始されており(NHK,1925)、1950年代にはテレビも登場した(NHK,1953)。

もちろんこの時代になると、規制のかかる一時期を除いては、表現の内容は自由で多種多様なものとなる。アナログ電波によるマスメディアの時代が到来し、出来事はリアルタイムに伝えられ、地球上のあちこちから異文化が紹介される。時間と空間のボーダーがいまになる世界で、音楽と映像も、ごく日常的なものとして存在するようになった。

さて、アナログメディアの時代(全盛期という意味で)は、つい最近までの約150年であるが(といっても、デジタルメディアが完全にそれに取って代わると思えないが)、その150年を一つの時代として片付けるには、あまりにも出来事が多すぎる。その様子を一言でまとめれば、「新たなメディアが登場しては、人と社会をかき回して、消えていく」

そんな時代であったと言える。ラジオ・テレビの放送に加え、様々な規格のフィルム・レコード・オーディオテープ・ビデオテープ・レーザーディスク（機器はデジタルだが扱っているデータはアナログ）など、実に多くのメディアが登場し、そして、やがて消え去ろうとしている。地上波放送もデジタル化されることが決まった今日、結局、アナログメディアの原点である「写真」の形式、すなわち（途中のプロセスはデジタルになったが）最終的にプリントという形で紙の上に定着する形式のメディアだけが、相変わらず大きな勢力を誇っている。

さて、ここで付け加えておきたいのは、このアナログメディアの象徴である写真（プリント）が、生き残る理由である。それは「身につけることができる」という点にあるのではないだろうか。絵画や映画は室内で鑑賞するもので、普通、外には持ち出さない。テレビは携帯用のそれを持ち出すにしても、「イメージを身につける」と言うより「機械を持ち歩く」と言ったほうが近いだろう。

しかし写真だけは、「イメージを身につける」という感覚で、行動を共にすることが可能なのである。アイドルの写真、恋人の写真、子供の写真、日常的に画像を身につけている人は多い。これはおそらくイメージのありかたとして非常に原初的なものであり、イメージと人間の距離に関する最終目標なのである。1979年の発売当時、あつて当然の録音機能を排して、音楽を「身につける」ことに徹したウォークマンも、登場すべくして登場したメディアであり、アナログメディア時代の一つの結論であったとも思われる。

11 デジタルメディア時代

1990年代後半からは、パーソナルコンピュータの高性能化ともあいまって、音も映像もデジタルメディアの管理下に入る。

情報の内容に関して、アナログのメディアと比較すれば、まずアナログの限界であった情報劣化がなく、また編集加工が100%自由にできるようになった（プロとアマのボーダーも消滅）というのがその特徴で、現代は「音楽や映像の表現において不可能がなくなった時代」だと言える。

さらに重大な変化として、音楽や映像の時間展開にインタラクティブに関わることができるようになったという事も挙げられる。アナログの時代は音楽にせよ映像にせよ制作者の意図にそって一方的に展開するものであったが、例えば家庭用のゲームに代表されるように、我々は映像の進行に関わり、その結末を自分で決めることができるのである。もちろんそれは「作品」の可動範囲内のことではあるが、音と映像のヴァーチャルな体験ができるというのは、今までにないことであり、時間展開に関して決定権を失った「作者」の立場からすれば、大幅な発想の転換がせまられている時代だと言える。

伝達媒体に関しては、媒体の物理的種類と扱われる情報のタイプとが無関係である（言葉も音楽も映像も同質のデータとなる）というのがデジタルの特徴で、記録形式さえ

それぞれの標準的なフォーマットに従えば、情報の記録・再生・転送・複写が様々な媒体間で簡単にできてしまう。

表現に関するあらゆる制約からの解放、そして情報の内容に関与しない非個人的な記録・転送媒体、「あとは容量とスピードの問題だけである」ということを考えれば、これがメディアの最終形態だと考えることもできよう。

12 音楽・映像と文化

音楽と映像の社会における立場は、問題とされる場面によって様々である。例えば「高度に実験的な音楽表現」と評価されたものが「はみだし者のシンボル」として大人から敵視されたり、多くの視聴者を寄せ付ける（スポンサーから見れば）有益な番組が、俗悪番組として害虫扱いされたりと、音楽や映像に関わる評価は、その存在が大きいほど、良いとも悪いとも言われる非常に両義的なものとなる。

一般に、音楽にも映像にもそうしたトリックスター的なところがあつて、それは社会の常識や日常性といった、いわゆる制度的な枠組みには納まりきれないという性質をもつ。それらは常に不安定な境界領域から情性化しがちな文化の枠組みをゆさぶりにくるのだ。

楽曲の歌詞における直接的な社会批判、言葉をたわむれにつなぎあわせて使った言葉文化の攪乱、映像表現における俗悪・醜悪な（と言われる）部分の提示、そしてテレビに登場

*トリックスター

「トリックスター」は文化人類学等でよく用いられる語で、「あちらの世界」と「こちらの世界」の境界を行ったり来たりして事件をおこす「いたずら者」という意味。

トランプゲームにおけるジョーカーのように「あるときは有益な存在であるが、あるきつかけで危険な存在にもなる」という両義的な存在のことを言う。

するタレント（少なくともありきたりの常識人はテレビ出演には向かない）、それらはいつも我々の常識に対するはみだしものである（能や歌舞伎に代表されるような「芸能」も、もともと社会の中心からはみだした人たちが「河原」という境界領域を舞台にはじめたものであり、創造的なものには共通してそういう性質があると言えそうだ）。

さて、はみだしもの（常識に縛られない自由なもの）が意識をゆさぶるということに関して、異なる観点から補足すると、忘れてならないのが、カメラという「機械の目」がもつ、我々の視覚に対する根本的な異化作用（慣習化した物事の関係を異なる関係へと変化させる作用）である。「機械の目」は、我々の目とは違って「見たいところ」を意識しないつまり意識に縛られずファインダーに入るものすべてを無差別に写し撮るのである。だから日常の風景を写真に撮ることで、普段見過ごしていたものの存在に気付くことがあるし、また例えば、普段気付かないホクロが、写真を通して発見されることもあるのだ。カメラの映像を介すると、我々の日常的（情性的）な視覚世界は、一旦解体されて改めて組織化されるのである。これがカメラ映像に特有の「ゆさぶりの力」である。何にも縛られず何の関係づけもしない自由奔放な「機械的メディアの視覚」、我々の意識は、日々その力によって様々にゆさぶられている。

「情報の力」そして「メディアの力」は、良きにせよ悪きにせよ、我々の文化の周縁からやってきて、それを活性化する。

発想法

人がメロディーやイメージを創造するとき、そこにはどのような創作の心理がはたらいているのだろうか。音楽や映像作品が生まれる際の作者の心の動きを、単に「センス」とか「才能」とかいった言葉で片付けてしまったのでは仕方がない。創造的な意識活動とは何か。我々の頭の中でほとんど無自覚的に起こっていることを、でき得る限り意識化して整理してみよう。

1 混沌から秩序へ

人は世界を「秩序立てて」捉えようとする生き物であり、また自ら「秩序」をつくる生き物でもある。多くの神話が物語るように、世界のはじめは混沌としていて、境界・区分・目印といったものが存在していない。それが神々の力によって、天と地・昼と夜・善と悪といった2項対立的な区分を与えられ、秩序立ったものになっていく。しかしそれはユクスキュル（1973）の言う環境世界、すなわち生物が本能のプログラムによって対応・生成する秩序ある世界とは異なり、人の意識の中で恣意的なものとして秩序立っているに過ぎない、一種の幻想的な世界である（現代人が科学的に把握している物理世界とて、それ

が言語という恣意的記号による一つの世界観であることにはかわりはない）。

人間の創造行為を考える場合、まずこの点の理解から始める必要がある。なぜなら、音楽も映像も含め、いわゆる過剰な秩序の創造行為は他の生物に無縁のものであり、逆にそのことが人間の創造行為の根底にあるものを知る手がかりとなるからである。

誰にでも経験があるであろう「コレクション」という行為を例にとってみよう。今ここに1950年の年号が刻まれた一つのコインがある。そこへ1951年製のコインが偶然ころがり込むと、その規則性の拡大のために1952年製のものが欲しくなる。そしていつのまにか引き出しの中には年代順という「基準」で秩序づけられたコインの幾何学的配列ができていく。

しかしこの秩序は、蜜蜂がつくる巣の幾何学的秩序とは違う。蜜蜂の巣をつくるためのプログラムは、すべての蜜蜂のDNAに記述されているが、コインのコレクションのための分類プログラムは、コレクターの脳の中にしか保存されていない。年代順という「基準」を知らない者には、ただ「コインがいっぱいある」ようにしか見えないのである。人が何らかのコレクションにのめりこむ場合、まずこのような基準となる「視点・視軸」なり「知識ベース」なりが存在するもので、それを共有し得る人が集えば、いわゆるマニアが胸躍するような秩序ある（意味のある）世界ができるのだが、その道の専門的知識を持たない門外漢には、ただ物が並んでいるようにしか見えないのである。

切り分け方が理解できなければ、意味も見えてこない。コレクションを秩序ある一つの

*コレクター

話が脇にそれるが、筆者の周囲には世代的に「ガンダムコレクター」が多い。幸か不幸か筆者は放送当時、そのすばらしさに出会う機会を逃したために、今も彼らの会話の節々に現れる笑顔の意味がわからない。「シヤア専用」って何？ 知識の共有部分が一部欠けただけでもコミュニケーションには不自由を感じるものなのである。

情報媒体の多様化と個に分散するライフスタイルは、まちがいになく人と人との知識の共有部分を少なくする。二緒に映画を見る（一つの物語を共有する）こと、「同じ教室で同じ時を過ごす」こと、人が人間関係に秩序を見出すには、まず時間と場所を共有することが必要なのだということ、ここであらためて強調しておきたい。

「学校」というものが存在する意義もこの点にある。

世界として他者に理解させる（コミュニケーションを成立させる）には、その世界を見る

ための「視点・視軸」や「知識ベース」を共有しなければならないのである。他者とのコミュニケーションが成立し得る世界をつくるには、その世界が何らかの基準で恣意的に秩序立てられていることと、その基準に関わる知識を集団の成員全員が共有する必要があるのだ。

どのようなタイプのものであれ、人間のつくる秩序というものには、この前提がつきまとい、後付けで被せたプログラムが構造化した秩序であり、したがってその「仕様」を集団の成員が何らかの方法で共有しなければコミュニケーションは成立しないのである。

人間の芸術的な活動の大半も実用性・機能性を離れて、純粹に「秩序」そのものを指向する。ということは、「作品」を発想するという場合も、それが何らかの仕様に基づいて秩序立てられていて、同時にそれが他者（鑑賞者）にも共有される必要がある。コミュニケーションを前提とした創造行為というものは、この事抜きには考えにくい。

すべては、生命維持や種族保存といった本能的必然性に止まらない精神の過剰な働きによるものである。今日、人間のこのような過剰な行為が地球の環境にとってはマイナスに作用する（無ければ無いで済むものを「便利」を理由に開発し、結果的に多くの問題を抱え込む）ものであることは誰もが認めることであるが、この過剰な精神活動こそが人を人たらしめるものであり、この点を抜きにしては「創造」は語れない。過剰な脳が「後付けで被せた」秩序は、自然・必然といったことから自由なのであり、何度でも「組み換え

* 成員が共有する知識

自然的・必然的根拠のない恣意的な知の体系であっても、集団の成員にとってそれが「常識」となり「真理」となることがある。このような集団が無自覚に前提としている秩序は、その意味で「共同幻想」とも呼ばれる。

* 過剰

書店などで「シリーズものの単行本を自分の仕事でもないのに番号順に並べなおしている人」を見かけることがあるが、人間は本来「秩序」をつくることに關して神経症的な生き物であり、「必要のないことにこだわって、時間を浪費してしまう」という不思議な生き物である。

「PCの管理」がその最たるものである。意味もなくソフトのインストールを何度もやり直したり、不具合を調整するのに夢中になったり、結局生産的なことをしている時間は僅かである。メディアが増えたことで、現代人には「ただ管理するため」とそれと関わっている」時間が非常に多くなった。：何とかしなくては。

「ことができる可能性をもつ。秩序という枠組みが変われば、その中にある要素の役割も変わる。同じ素材でも枠組みを変えれば、違った生かし方ができるのだ。そう考えれば、人間の過剰が成せる「秩序」というものも、創造の基本原理として、非常に重要性なものであると感じられるのではないだろうか。「無駄」なものこそが大切なのだ。

さて、楽器を前にして、あるいは白いキャンバスを前にして何かを創ろうとする場合、具体的にはどのように発想すればよいのだろうか。それは「混沌から秩序へ」と意識を働かせることであり、H・リード（1966）『芸術の意味』の言葉を借りれば「心楽しい形式をつくる試み」なのであるが、それには様々な具体化の方法が考えられる。以下にその発想法について整理してみよう。

2 発想の前提・表現の限界（制約）

どのようなものに対してであれ、そこに何らかの「秩序」が見い出せれば、人はそれを「何者かの創造物である」と理解する。そして確かに、蜂の巣のような自然界の幾何形態も含めて、人間がそれを「秩序」あるものとして見聞きすることのできるものは、ある意味で美しいと言いうことができる。つまり、何かを発想するときにはまず考えるべきことは、とりあえず、「何らかの基準に沿って秩序ある状態をつくる」ということである。そのアイ

デアに賛同が得られるかどうかは別として、それが人特有の創造行為であり、できあがるものに何らかの意味なり価値なり芸術性なりが与えられることは確かだろう。

しかし音楽や映像の制作を考えると、何の制約もなく自由に「秩序」をつくることができるかと言うと、事はそれほど単純ではない。作品づくりの自由度にも様々な制限があるのである。具体的な「秩序」構成の話に入るまえにその点を述べておこう。

スケール

第一に、人間の視聴覚に関わる空間的なスケールと時間的なスケールの限界がある。速すぎて聞き取れない、遅すぎて旋律がつかめない、音階のステップが細かすぎて（12音階音楽など）つかめない、小さすぎてよく見えない、大きすぎて全体の構図がつかめない、あるいはまた長時間すぎて疲れてしまう、など、音楽も映像も時間や空間に関して的外れなスケールでは作品になりにくい。

バランス

第二に、人間の体の運動バランスと空間的なバランスとに関わる制約が挙げられる。音楽では「拍子」が大きな問題で、原則2・3・4拍子を基本としてその倍数であればリズムもとりにやすいのだが、5拍子（『Take Five』は有名）や7拍子といった変拍子となると一般的には「ついていけない」ということになってしまう。本来、時間軸上の秩序という

意味では4拍子も7拍子も同等であるが、相手は機械でなく人間であり、体がついてこないものは成功しにくいというのが現実である。空間芸術では、あらゆるものの基軸としての垂直・水平線が作品のバランスに大きく作用している。我々の体は重力とバランスをとっていて、視覚は重力方向に重さを感じるかたちで世界を捉えている。身のまわりのものはすべて建築空間の床面を基準とする垂直・水平の格子におさまっていて、絵画もポスターもテレビの画面もそこに収まらざるをえないのだ。したがって大半の画像・映像は垂直・水平で枠取りされた長方形の空間内にレイアウトされなければならない、またその中では本来同等の意味をもつはずの1本の線も、斜めになっていることで「不安」・「緊張」といった役割を必然的にもたされてしまうのである。見ているのは正立した人間である。この方向は絶対的なものとしてあらゆる空間的な造形に制約を与えている。

環境

第三に、気候・風土の違いも見逃せない。例えば欧米と日本を比べた場合、気温はほぼ同じでも湿度の違いは大きく、これが「楽器の鳴り」や「絵の具の乾き具合」といった物理的な違いとなって、表現の可能性に制限を加えてしまう（気候・風土が異なれば、絵筆の使い方にも必然的な差が生じるものである）。また気候・風土の違いは、当然「快適さ」や「美しさ」に対する意識の違い、すなわち文化の違いも生むもので、それは当然無視できない問題として作家の創作意識を左右している。気温・湿度・気圧といった自然界のパ

ラメータが創作活動に無関係でないことは厳然たる事実であろう。

文化

第四に、年齢差・性差、そして様々なレベルで人間の意識に食い込んでいる「文化」の差異の問題がある。年齢や性の違いで、作品のもつ秩序が見えない、誤解が生じる、あるいは呼び覚まされる感覚が異なるということは当然あることで、さらに、社会の中での立場や役割、生活環境、その民族に特有の「音階」や「色彩嗜好」、そして世界認識の根底にある「言語体系」、そういった諸々の「文化」的基盤というものも、情報の捉え方に決定的な差異を生じさせるのである（むしろこの「文化」の問題が一番大きい）。

その意味では作品を享受できる、言い替えれば、作者と同様の世界観に基づいて作品を共有できる人というものはある程度限られてしまうと見えよう。老若男女問わず楽しめるものというものはそう多くはない。作家は創作の過程で、出来上がった「情報」の受け手が誰であるかを無視するわけにはいかないのである。

制作の現実

最後に、これは議論の本筋からは逸れるのだが、創作物に関する物理的・経済的・法的制約が挙げられる。「芸術家」には皆無に等しい制約だが、企業活動に位置付けられて、経済的に成功することが前提となる音楽プロデューサーやグラフィックデザイナーなどに

とって、この問題は避けようのない制約である。

情報の表現媒体・記録媒体また伝送媒体に関する物理的制約、制作費という経済的制約（素材の妥協・技術の妥協が強いられる）、印刷コストの問題から生じるサイズや色数や解像度の制限、さらに（著作権法は当然としても）広告表現などでは、効能書きの文字サイズに関する制約など一般には知られていない法的制約もかかる。そして「大衆受け」・「スポンサー受け」せねばならないという最大の検討項目があつて、プロと呼ばれる作り手の大半はこれら数々の制約をトータルに計算したうえで、作品を発想せねばならないのである。我々をとりまく大部分の音楽や映像がこうした制約と計算の上に成り立っていることを忘れてはならないだろう。

逆転の発想

「芸術的発想」というものは様々で、「A」という作品があれば、あらゆる意味での「非A」もまた作品になる可能性を持つことも事実である。「作品をつくらない」という「芸術」も、そうすることで「作品とは何か」を考えさせる、あるいは逆説的に「作品」を定義するものであるし、「反秩序」・「破壊」・「脱構築」・「実験」など一見否定的なキーワードを掲げる「芸術」も、「新しい価値観」・「新しい形式」を模索して「非・旧芸術」↓「新・芸術」として、新たな時代の秩序へと組み換えを促す運動だと言える。

人間がすることは、表向きは無為であつたり破壊であつたりしても、結果的には何らか

の「秩序」を志向しているように思われる。それは言ってみれば「無秩序なネットワークが枝を落しながら秩序化していく」という人の脳の宿命なのかもしれない。

3 「鍵」を意識する

何をつくるにせよ、「秩序」を貫く鍵が必要である。それは「視点・視軸」「準拠枠」「方向性」「順番」「階層」「規則」「法則」「テーマ」「コンセプト」など、様々なレベルのものが考えられるのだが、要は「何がしたいのか」「何を基準にまとめるか」を明確に意識化することが重要だと言えよう。

単に文字を並べるという行為にも、「画数で並べる」「音（あいいうえお順）で並べる」「形で分類する」「文章になるようにならべる」など様々な基準が考えられる。また「現在の並びに対して、ある一定の演算を行って新しい並びをつくる」ということをすれば、新旧の間に歴史的な法則ができるし、「とりあえずコンピュータでランダムに並べることを繰り返して、面白い並びのものを採用しよう」というオートマティス的な発想もいい。とにかく「秩序のつくりかた」にはいろいろな鍵があり、その発想の可能性は無限にあるのである。

もちろん、場合によっては法則もテーマもなしにその場の勢いで作って成功する（この場合、あとから法則やテーマが発見される）ということもあるが、我々の身の回りにある

テキスト・音楽・映像、さらには工業製品・建築物・都市空間まで、勢いで偶然的にできたというものはほとんどない。^{*}ものづくりに携わる大半の人間が、何らかのかたちで自分なりのあるいは企業なりの発想の鍵を定め、それに従って秩序ある構築作業に関わっているのである。

そして当然のことだが、そうしたもののづくりの鍵とは「自分は何が作りたいたのか、もともと何が好きなのか、そして何故それが好きなのか」といったことを自問自答することとしか見えてこないものである。

4 加算型の発想と減算型の発想

さて、ここから具体的な「秩序をつくる」ための手順の話になるのだが、まず源初の「混沌とした状態」（すなわち何もない状態や素材が山積みされた状態）をどのように攻めるかについて、二つのタイプの攻め方が考えられる。一つは何もないところに何かを加えていきながら次第に秩序ある状態に組み立てていく方法で、もうひとつは逆に山積みの状態から何かを削除していくことで次第に秩序ある状態に整理していくという方法である。前者を加算型の発想、後者を減算型の発想と呼ぶが、人間が何か事を起こすときは大体このいずれかの発想をとるものである。

例えば、「話をする」という身近な行為でも、頭の中では、テーマにそった話題を集め

^{*}予見と計画

P・J・ウィルソン（1983）の有名な言葉を紹介しよう。

「最も巧みな蜜蜂と最も無能な建築家の違いは、建築家が設計図にもとづいて仕事をするところである。」
『人間—約束するサル』佐藤俊訳

ながら、テーマにそぐわない部分は削除するということをしているし、「写真を撮る」というときも、ファインダーを見ながら足りないものを視野の中に加えたり、またじゃなものをごかしたりしている。また例えば、都市計画という大きな事業を例にとつても、「現状の景観をそのままに、そこへ当てはめ得る建造物を次々に加えていく」という加算型の性格のもの（一般にアジアに多い）もあれば、「沼地を埋め立て、小高い土地を削り、まがった河川をまっすぐに矯正してすっきりさせていく」という減算型の性格のもの（一般に西欧に多い）もある。

一般的なことで言えば「選択」という日常的な行為も、加えることと引くことの両方で成り立っているわけで、その意味では人間の意識的な行動はすべて加算と減算で成り立っていると言ってもいい。我々が日常あたりまえに行っていることを意識して見直せば、アイデアはいくらでもこがっているのである。

5 リフレイン

リフレインすなわち反復という行為は、あらゆる素材を秩序あるものへと構成する基本的かつ重要なテクニックである。ノックの音は2回であることで、それが人の行為であるとかかるし（秩序あるもの・意味のあるものとして聞こえる）、単なるインクのしみも折り畳んで複写すれば、左右対称な模様になり様々なイメージを喚起する（ロールシャッハ・

テスト）。ひとつのリズムパターンも、音程を変化させながら繰り返せば楽曲になり、簡単な図形も並べればタイルパターンになり、また全体を部分へと再帰的に反復すれば雲や海岸線の美しい輪郭となる。

反復すること・複製をつくること・まねること^{*}、それは未知の生々しいもの・名付けようのないものにひとつの「基準」を付与して意識化する。我々の意識の中では、何らかの観点で同様とみなされるものが二つ以上あつてはじめてイメージが一般化され名付けられるのである。混沌としたものが秩序化されるということは、このようなイメージ・概念の生産をも意味している。

^{*}「まねる」は「まねぶ」すなわち「学ぶ」の語源とも言われ、人間が何かを秩序づけながら意識化するときの基本である。

^{*}まねる

6 選択的制約

制作の過程で「あれもしたい、これもしたい」と、いろいろな可能性が見えてきて、収集がつかなくなることもある。「あれもこれも」を秩序化する発想で攻める場合は別として、素材やそれを扱う技術に無限の可能性があると、かえって首尾一貫しなくなつて「濁った」ものが出来上がるものである。

そこで逆に、自ら素材や技術に制限を加えることで骨格を見えやすくする、という発想も大切である。小学校低学年で習う言葉だけを使つてみる（作詩）。楽器の数や音域を制限し、重複する音は省いてみる（作曲）。色数を数種類に制限する（絵画）。文字のフォ

ト・スタイル・サイズを限定する（ポスター）。モノクロにする（写真）。撮影に使うカメラやレンズを一本にしぼってしまう（映画）。特殊効果のパターンを数種類に限定する（テレビ）。生成アルゴリズムやパラメータに制限を加える（CG）。そうした制限をかけると自分を取り組みたい主たる骨組みに意識を集中することができ、結果的に「要素間の関係がわかりやすい」すなわち「伝わりやすい」ものができることになる。

特にコンピュータによるデジタル加工処理が一般的になった今日、何でもできてしまう道具を前にすると「あれもできる、これもできる」と発想の鍵がしぼり込めなくなることが多い。そこで逆に自ら利用する技術を制限し、作品に一貫性をもたせるという発想は必須のものとなるだろう。

7 対称性の破れ

我々は、創造的思考において確かに「秩序」を指向するものなのだが、いわゆる「シンメトリー（対称）」という完全なる秩序に対しては、それを忌避する傾向があることも事実である（シントロフォビアと呼ばれるシンメトリーに対する忌避症もある）。古代ギリシアの伝統をひく文化圏では非対称なものは美的によくないものとして退けられる傾向にあるが、特に我々日本人の文化について言えば、あえて非対称とするような造形が多い造園や建築空間の構成を見ても完全なるシンメトリーは少なく、むしろ対称を意識しなが

らわざとそれを破るというつくり方をしたものが多い。もともと自然界に存在する樹木も動物の体も、基本的な私たちはシンメトリーで力学的にバランスをたもってはいるが、ディテールは左右異なっていて、それがそれぞれの表情を豊かなものにしている（例えばもし人の顔が中央を軸に完全に左右対称な「ウルトラマン顔」であつたとすると、それはなんとなく気持ち悪いものになるだろう）。

情報量という観点から言っても、非対称なものの方が豊かな情報をもつことが説明できる。完全なる左右対称なものは、例えばそれが画像だとすると、送信側が左半分の情報を転送するだけで受信側では全体を復元できる（単純に1/2に圧縮可能）。すなわち半分だけですがすべてが表現されているわけで、情報量は半分しかないのである。一方非対称なものはある程度の秩序感・バランスを保ちながらも半分に圧縮することは不可能で、その分大きな情報量をもつと言える。

これは冗長度という言葉で言えば、シンメトリックなものは冗長度が高く、アン・シンメトリックなものは冗長度が低いということである。一般的に「冗長度が高い・伝わりやすい・情報量が小さい・つまらない」ということに對して、「冗長度が低い・伝わりにくい・情報量が多い・驚きが大きい」ということは、情報の観点から対置できるものであるが、創造された秩序というものを美的なる情報と考えれば、冗長度の高すぎるものは、相手に通じやすいけれども、驚きがない、それ以上の関心と呼ばない、つまらない、というものになってしまう。秩序はそれがあまりにも単純な言葉で説明できてしまうようである

* 「対称性の破れ」

この言葉は、高エネルギー下で物質が粒子と反粒子に分離する際、本来同等であるはずのそれぞれの振る舞いに違いがある（対称的でない）ことを意味する物理学用語であるがそれはこの世界に反物質が存在しないことを説明するキーワードであると同時に、有機的な生命誕生の契機に関わるキーワードでもあり、ある意味で「創造」に関わる根源的な法則なのかも知れない。

対称ではない、どちらかが優位であるということは、実は人間社会でも大切なことである。もし拮抗する二つのものが、どちらも同等という状況であれば、結論が出せずに話が先に進まなくなってしまう。どちらかに優位な部分があるからこそ、物事スムーズなのである（実際には、理論的に差がない場合でも、人間の判断には「感情」というものがある。最終的にはそれが決着をつける。「利き手」もそうである。優位な方が決まっていることで、我々は必要な瞬間に手が出遅れることなく対応できるのである。

と、意識を素通りしてしまうのである。例えば、先に述べた「反復」によって形成された「秩序」も、それがあまりにも簡単な規則性で成り立っていると意識を素通りしてしまう。したがって、逆に反復の過程に若干のズレが与えられれば、「反復」という秩序構成の「基準」はより生き生きと意識化されるのである。

情報はそれがわかりやすいだけでは退屈する。^{*}かといって驚きの連続では疲れてしまう。対称性を破るという発想には、創造物の秩序とそれを破壊する力との微妙なバランスを考えることが重要である。

8 手続き記述と美的選択

コンピュータ・グラフィックスには、何らかの造形的規則と生成パラメータを与えてあとは機械に描き出してもらおうという発想のものがある。数学者の（美的）実験から発見されたジュリア集合やマンデルブロ集合といったフラクタル図形もその典型で、単純なプログラムで生成されているにも関わらず、無限の奥行きをもったディテールで美しい秩序を表出させている。同様の発想で、剰余や三角関数のもつ周期性を利用した幾何学的な造形も、そのプログラムに何らかの工夫をこらすことで思いがけない変化を生むことがあり、このような数式という生成規則による造形も、その応用は無限の可能性をもっていると言える。これはもちろん音楽の領域でも応用されており、語の生成（作詞）や旋律・和声の

生成（作曲）に規則を与えたうえでランダムに複数のサンプルを自動生成し、その中から気に入ったものを選ぶという発想はコンピュータの登場以来一つの作曲の手法として現実に利用されている。

さて、こうした手続き記述型の発想法には大きく二つのタイプがある。ひとつは「全体を統一的な生成規則で描く」もので、もうひとつは「要素の相互関係にのみ規則を与えて、複数の要素から成る全体を描く」というものである。

前者の典型的な例は数理曲線や再帰図形などであるが、これらは一般に「いかにも機械的で表情に欠ける」という性格をもつため、生成の過程で若干の乱数を加えることで表情を豊かにするという方法をとることが多い。この乱数は先の話で言う対称性を破る存在で、例えば単純な2分木の再帰図形でも、適度な乱数で十分豊かな表情を醸し出す。

後者は、いわゆる複雑系の現象のシミュレーション、例えばフォン・ノイマン（1957）のセル・オートマトン、クレイク・レイノルズ（1989）のボイド（鳥もどきのアニメーション）などにその典型を見ることが出来る。多くの要素が絡みあう複雑な系（Complex System）において、我々の目に面白く見える部分というのは、秩序の領域と混沌の領域がせめぎあう「カオスの縁」と呼ばれる領域である。たばこの煙に例えて言えば、たばこの先から一直線に上昇している部分が（決定論的に記述される）秩序の領域で、天井に拡散した部分が（確率論的に記述される）混沌の領域。そして、その中間にある煙の「乱れそめ」の部分が（カオス理論の対象となる）見ていて面白い領域である。コンピュータで

^{*} 退屈と驚き（ツッコミとボケ）

漫才には「ツッコミ」と「ボケ」という役割分担がある。話の筋道を秩序立てていこうとするのが「ツッコミ」の役割で、話の筋道をはぐらかしていくのが「ボケ」の役割である。「ツッコミ」はあたりまえの話を繰り返して冗長度を上げ、「ボケ」は意外な（出現確率の低い）言葉で話の情報量を上げる。

一般に「笑い」は、人の頭の中で凝り固まった常識や習慣的思考がゆさぶられる時に生ずるものである。

要素間の関係に規則を与えて、そのふるまいを観察した場合も同様で、安定した秩序に落ち着くか、あるいは無秩序な攪乱になるか、あるいはその中間的状况として複雑で有機的な形を次々に展開するかになるのだが、もちろん最後の状態が見えて面白いものであり、造形的にも応用が効きやすいものである。ただしこの場合の条件の与えかたは難しく、その試行錯誤はゲーム感覚ですらある。

いずれにせよ、こうした発想法の最大の特徴は、「何ができるかは、結果を待つしかない」という点と「数多く生産させて、美的に良いとおもわれるものを選択する」という作品に対する姿勢（進化論的な発想）で、これは機械（特にコンピュータ）が登場する以前にはできなかった発想である。もちろん自然界にある偶然の産物のなかから面白い部分を抽出するという発想は古くからあるものかもしれないが、偶然を機械的に制御しながら面白い形が出来るまでひたすら作らせるといのはまったく新しい発想と言えよう。

9 聴覚情報と視覚情報の比較

さてここで、次の節で話題となる、音響・音楽や画像・映像の生成に関する物理・工学的な議論の準備をしておこう。^{*}音楽と映像による情報のデザインを考える際には大前提となる事項でもある。

まず、音響・音楽すなわち聴覚情報についてであるが、基本的かつ重要な事実は、我々

の可聴範囲^{*}の音の波長が約17mmから17mと、ほぼ我々を囲む物の寸法と同じだということである。このことは視覚情報と同様な、空間的な「聴覚像」の生成が根本的に不可能であるということの意味している。もちろん複数の音源を空間的に配列して、音源から独立した情報源からの信号によつて「音像」をつくるという発想はあるが、上記の理由でそれは人のスケールでは不可能で、また、人の聴覚がたかだか二つの受容器をもつにすぎないことから、限界がある。したがって聴覚情報は、刺激の時間変化として表現される他はなく、音源と情報源を同一とするかたちで生成・発信される。言葉・音楽・物音、聴覚情報はいずれも時間軸を基軸とした音の配列であり、時系列の情報であるという事実を再確認しておこう。

次に、画像・映像すなわち視覚情報について考えよう。視覚情報の担い手である「光」を、聴覚情報の場合の「音」と比較すれば、その波長^{*}は380nmから760nmと非常に短い。我々の視覚にはこの波長の差が色相の差として知覚されるが、その色相や輝度の時間的変化で「音」の場合のような時間情報を生成（例えば、光通信）しても、人の能力でそれを直接処理することは難しい。すなわち、我々「人」としては、多くの色点の空間的な分布形状が情報の担い手となるわけで、当然、我々の視覚の受容器はそのために2次元的な配列をなしているのである。

また、このことと関連して言えば、視覚情報の生成には「照明という一次光源と情報が載る二次光源との区別が存在する」という事実が重要で、情報の生成・伝達には光源と情

^{*}音速と光速

音速 331.5m/s (1気圧0℃)
光速 3.0×10^8 m/s

^{*}可聴周波数

正常な聴覚で聞こえる周波数の範囲は通常20Hzから20,000Hzで、
波長＝音速÷周波数
を計算して本文中の波長は求まる。

^{*}可視光線

本文中の数字は、可視光線の波長であり、単位のnmは1mの1/1,000,000,000を意味する。
この範囲の外側は、一方が紫外線、もう一方が紫外線である。光はラジオやテレビ、携帯の電波と同じ電磁波の一種である。

報源の両方を関連づけて制御する必要がある。これは、音源と情報源が同じ場所にある聴覚情報とは決定的な差である（マイクは音源に向けるのが常識であるが、カメラは直接光源に向けても意味がない）。

文字・画像・風景、すべての視覚情報は、情報源における点の空間的な配列に、光源からの照明があたることによって発生することを確認したい。

10 「音響・音楽」と「画像・映像」

「音響・音楽」と「画像・映像」という情報の差異の問題は、すでに述べたように、前者の場合、情報は音源からやってくる時間的なもので、後者の場合は二次光源としての物的な形状を情報源とする空間的なものであるという違いがすべての議論の骨子となる。

この事実を起点として、「音響・音楽」と「画像・映像」について以下のような特質がクロージアップされる。

まず、前者は時間的な変化が情報であり、その処理形態はシーケンシャルなものとなるが、後者は空間的な配列が情報であり、その処理形態は一括処理的になるという点である。これは、前者の場合その情報が時間とともに消えてしまうため、次々にやってくる音の情報を、その都度記憶しながら処理せねばならないが、後者の場合は情報が長時間保存可能なものとなるため、必要な部分ごとに時間をかけて情報を処理できるということを意味する。

るものである。

また、「音響・音楽」では、その波長が我々を取り囲むサイズであり、すべての音が重なりあっている、特定の音源からの音のみを耳やマイクに取り込むということが不可能であるのに対し、「画像・映像」の場合は、短い波長の光の配列が問題であるため、その配列を捉えるすなわち結像するためのピンホールまたはレンズが必要だが、特定の場所（情報源）を選択的に見たり撮影したりすることが可能である。

以上を具体的な話でおきかえると、「耳は開閉しないが、目はまばたきする」、「テープもCDも回っていないれば用をなさないが、フィルムはコマずつでも意味がある。」「音は音の状態、持ち歩くことはできないが、絵は画像の状態、持ち歩ける」、「マイクには感音部があれば十分だが、カメラには結像と感光の二つのシステムが必要である（カメラはボディだけでは用をなさない）」、「目覚まし時計は音を使う（視覚には訴えようがない）」、「耳ざわりなものが騒音公害となることはあるが、目ざわりなだけでは公害とは認定されにくい」、といったことが「音響・音楽」と「画像・映像」の大きな差となるのである。

これらはメディアの抱える必然的な差異の問題として、様々な想像的発想に深く関わるものであることを銘記しておきたい。

音・音楽の生成

人は自らに備わった音源（声帯）を利用して音声を発することができると同時に、機械的あるいは電気・電子的な手段で様々な音・音楽を生成することができる。以下にその具体的な手段と原理の概要を紹介しよう。

1 音声

人は声帯を振動源として（男声 110Hz ～ 女声 220Hz ～）、また喉から唇までの声道を共鳴腔として、複雑な音声を構成して情報を発している。声帯は声の基本周波数を決め、声道（特に口の開きぐあいと舌の位置）が共鳴の性質を決める。例外として、ささやき声の場合は声帯は振動せず、空気の流れを雑音源として共鳴のみで音を作っている。そしてもちろんその情報の大半は、話し言葉としての言語情報である。

我々は通常、話し言葉を構成する音節を単独に生成・識別することができる。つまり音韻的な音色についての絶対音感をもっているわけで、たとえば日本人の場合、約 100 種類の日本語の音節による音韻体系をもつて、言語情報を生成処理している（参考までに述べる）と中国では 400 以上、英語だと 3000 以上と言われる）。

*話し言葉

人間の言語ほど複雑な音声合成は現生ヒト以外の生物では構造的に不可能であると言われる。約 20 種ほどいた人類の中で、絶滅した他の種と現生人（ホモ・サピエンス）との運命を分けた決定的な要因がこの声帯の構造にあったということも最近の研究で明らかになっている。

リチャード・ドーキンス (1976) が言うように、「言語」という情報体は、遺伝子という情報体とは異なる、物質に依存しない新たな進化のシステムであり、人類はそれを操ることで、かつてないスピードでその世界を変えている。

この話し言葉の単位音節の生成・識別には、特に母音のホルマント (Formant、音を特徴づける成分音) の存在が重要で、例えば「イ」の音では 300Hz と 2000Hz、¹「エ」の音では 500Hz と 1700Hz の成分が特に強いというような特徴がある。歌い声 (Singing Voice) の場合も、この特徴成分の発振を保持すれば、音程とは無関係に「イ」・「エ」の発声ができるというしくみである（その意味では歌い声は、一種の楽器として捉えた場合、音の出しかたの自由度が大きい、非常に可能性の大きい楽器であると言える）。

我々は、このような音韻体系を聴覚系の形成と並行して（遺伝的ではない）習得し、生後 18 ヶ月ごろにはほぼその基礎的な生成・識別能力を獲得する。

2 楽器

楽器は、弦振動や気柱管振動を利用して周期的な振動波を放射するものと、衝撃による自由振動波を放射するものとに分類される。前者はいわゆる「楽音」として旋律や和声を形成するメロディー楽器群を意味し、後者は（振動に周期性がないため「音程」をもつことができない）「非楽音」を発生するリズム楽器群を意味する。

音にも色と同様に三属性があり、それぞれ音の大きさ・音程・音色という。

物理的には、「大きさ」は振動の振幅の大きさであり、「音程」は振動の基本周波数のことである（例えば 440Hz は「ラ⁴ a3」の音）。つまり、音が大きいほど波の揺れ幅が大きく、

音程が高いほど、揺れの密度が細かいということである。さて「音色」、つまり「ピアノの音色」とか「バイオリンの音色」とか言う場合の音色であるが、これは時間軸上にグラフ化すれば「波形」として、周波数を軸としてグラフ化すれば、スペクトルパターンとして物理的に捉えることができるものである。つまり、我々の耳が聴き分ける楽器の音色というものは、基本周波数とその倍音列^{*} (Overtone Series) の強度分布によつて特徴づけられている。電子楽器がピアノに似た音やギターに似た音を生成する場合も、この倍音の分布パターンを様々な方法でコントロールすることによつてそれが実現される。

弦振動と打弦楽器・撥弦楽器・擦弦楽器

一般に弦の振動周波数は $f = \sqrt{T/m} / (2l)$ で求まる。Tは張力、mは単位長さあたりの質量、lは長さである。この式の意味するところは、弦長(l)が長く、また単位長さあたりの質量(m)が大きい(つまり弦が太い)ほど周波数(f)は低くなり、張力(T)が大きいほど周波数(f)が高くなるということである。身近な楽器であるギターをイメージすれば理解しやすいであろう。弦が細いほど、弦を強く巻くほど音程は上がる。また、12フレット目は、弦の長さの半分の位置にあるという意味で、周波数が2倍、つまり音程が1オクターブ高い。

弦振動は、基本振動以外にもその2倍(中央に振動の節)・3倍(1/3のところに振動の節)：といった倍音を含むことで様々な音色を作り出す。したがって弦楽器の音色は、

振動の腹や節の位置に関わる「弦をはじく場所」や「弦に触れる場所」を変えることで様々にコントロールすることが可能である。

さて、弦楽器はこのような弦振動により音を生成するが、弦そのものは表面積が小さい(すなわち放射抵抗が小さい)ため、直接大きな音を出すことはできない。そこで、この弦の振動を駒(ブリッジ)を介して共鳴板に伝え、この板を強制的に振動させることによつて音を放射する^{*}。いわゆるアコースティックな弦楽器では、この共鳴板が適当な容積をもつ箱に結合して、箱の中の空気をも共鳴させるかたちで音を放射している。エレクトリックな弦楽器の場合は、弦と共鳴板(ソリッドボディが大半)の振動を、電磁形変換器やピエゾ抵抗変換器などで電気的な振動に変換して利用する。いずれの場合もボディの質量・形状・材質などがその音質に与える影響は大である。

楽器の分類としては、ピアノを代表とする打弦楽器(Struck String Instrument)「ギター・ハープ・チェンバロなど弦をはじく撥弦楽器(Plucked String Instrument)」バイオリンのように弦をこする擦弦楽器(Rubbed String Instrument)に分けられる。

気柱管振動と管楽器

開管の共鳴周波数は、最も低いもの(基本周波数)が、 $f = c / (2l)$ で、この整数倍の周波数の発振が可能である。cは音速、lは管長であり、管長が短いほど共鳴周波数は高くなる(閉管の場合は $f = c / (4l)$ で、この奇数倍の周波数が発振可能)。管楽器の振動は、

*倍音

音程のある「楽音」は、通常基本周波数の振動と、その2倍の周波数、4倍の周波数など、整数倍の周波数の振動が重なっている。つまり、同じラ(a3)の音と云っても、具体的な楽器の音では、440Hz、880Hz、1760Hz・・・と複数の振動の組み合わせになっている。我々の耳は、その各成分の強さの分布の違いを聞き分けることで、それがピアノの音かバイオリンの音かを識別している。

*音再生と振動板

音の最終的な出力に際して、アナログ信号で板を振動させるという方法は百年來変わっていないが、物の振動を空気の振動に変えるというのは、実は非常に変換効率が悪く(のれんに腕押しの状態)、音になるのは1%程度であると言われる。また特に小さな面積の振動板では、それを低い周波数で振動させても空気が回り込んで音にならないため、低い音を再生しなければ大きな面積の振動板が必要になる(コントラバスのサイズが大きいのも同じ理由である)。これは能率的ではないことから、現在振動板を用いずに直接空気を駆動する技術も試されている。

基本的には息を吹き込むという直流エネルギーの供給によって持続する振動で、これを自励振動という（自励振動は管楽器の他、バイオリンのような擦弦楽器にも見られる）。

管楽器はその大半が両端が解放された開管で、気柱の縦振動が音源となり、指孔や管端から音波を外部へ放射する。管の一端には直流のエネルギーを振動エネルギーに変換するきっかけをつくるリード（Reed）が必要で、その種類によってエアリード楽器・ダブルリード楽器・シングルリード楽器・リップリード楽器に分類される（ちなみに、シングルリード楽器は閉管とほぼ同様で奇数倍音列の発振となる）。エア・ダブル・シングルリード形式の三者は木管楽器（Wood Wind Instrument）に分類され、例えばフルート・リコーダー・尺八・日本の各種横笛など、オーボエ・ファゴット・雅楽のひちりきなど、クラリネット・サキソホンなどがそれぞれに該当する。またリップリードの楽器はほぼ金管楽器（Brass Wind Instrument）とイコールであり、トランペット・トロンボーン・ホルンなどがそれにあたる。

剛体・膜の振動と打楽器

すべての「物」は力学的な衝撃を加えると振動し、音を出す。弦や管が発生する「楽音」以外のこの衝撃音はみな「非楽音」であり、打楽器はこれを原理としている。

打楽器の振動体は棒・板・塊・膜であり、それぞれ例えば、トライアングルや木琴、シンバルや鐘、カスタネットやウッドブロック、ドラムや鼓などが、それに該当する。

この種の楽器が発生する「非楽音」は、周期性のない振動を基にしてい、音程は特定できず、スペクトルパターンも広範囲の連続的な分布かあるいは非整数倍の成分を多くもつ離散的な分布をなす。「楽音」のように整数次の倍音が並ぶものではないため、いわゆる和音も濁ったものになる。

3 電気・電子的手段による音の生成

さて次に、電気・電子的な音の生成、つまりシンセサイザーのような電子楽器を用いた音の生成について述べよう。ここではまず、その音色を特徴付ける二つの要素について概説する必要がある。一つはスペクトルパターン、もう一つは時間経過パターンである。

まずはじめに「スペクトルパターン」すなわち「周波数分布」であるが、これは、楽器の音色を特徴づける複数の倍音（基本周波数の整数倍の音）をどのように分布構成させるかという問題である。これをコントロールするには、原信号の変調の仕方で倍音の出方を制御（FM音源など）してもよいし、単純な正弦波発信器を複数並べて加算制御、あるいは豊富な倍音をもつ信号から成分を減算制御してもよい。要は、各倍音成分の発振強度が調整できさえすれば、様々な音色が作れるのである。ただ実際には、非整数倍の成分が音色を特徴付ける場合も多いため、より個性的な音作りをするには、わずかにチューニングを狂わせた倍音などを合成する必要がある。さらに、ドラムのような「非楽音」の場合は、

連続スペクトル、すなわち「楽音」のように倍音が等間隔に並ぶ離散的スペクトルとは異なる音色が必要で、その生成には発振音を変則的に変調するなどの工夫が必要である*。

一方の「時間経過パターン」であるが、これは、音の鳴り始めから鳴り終わりまでの音量（厳密には音程や音色も含む）の時間的な変化の問題である。音の立ち上がり、減衰の早さ、余韻の残り方などもその「音」を特徴付ける重要な要素であり、例えばピアノの音を真似ようとした場合、単に倍音構成を真似るだけでなく、特に音の立ち上がり部分の倍音構成の時間変化パターンを上手く真似ないと「ピアノらしい」音にはならない。

シンセサイザー

では、電気・電子的に音を生成する楽器の代表であるシンセサイザー（音を総合するという意味）について、その基本的な構造と発音の仕組みを見てみよう。

もちろん実際の楽器ごとに音源や編集加工のプロセスは異なるのだが、基本的なところは同じで、そこには発振器（Oscillator）・フィルタ（Filter）・増幅器（Amplifier）の3つのブロックがあつて、発振器から出た様々な倍音を含む信号が、フィルタによって加工処理（一般的には通過制限）され、最後に増幅されて出力されるという流れになっている。

その際、それぞれのブロックにはEG（Envelope Generator）による時間経過パターンの制御がかかり、例えば発振器にかかれば音程（Pitch）の時間変化、フィルタにかかれば波形（Wave Form）の時間変化、増幅器にかかれば音量（Amplitude）の時間変化とい

*打楽器音の再現

もともと周期的な振動をつくる発振回路からでは、楽音以外の音をうまくコントロールするのはむずかしいため、今日では、もっぱらサンプリング音源が利用されている。シンセサイザー等で一般のメロディー楽器とドラム音が別扱いとなる理由の一つもこの点にある。

うように音を制御できる（それぞれをPitchEG FilterEG AmplitudeEGとよぶ）。

このEGは一般にADSRという四つの時間区分を用いるもので、A（Attack）は立ち上がり、D（Decay）は減衰、S（Sustain）は伸び、R（Release）は残響と、それぞれの設定によって、音程・波形・音量それぞれの時間経過パターンが制御される。

さらにLFO（Low Frequency Oscillator）がそれぞれのブロックにかかることで、ヴィブラート、ワウ、トレモロに相当する効果をかけることも可能で、以上のすべての要素の総合的な制御によって、あらゆる自然楽器のシミュレーションはもちろん、自然楽器では生成できない音も加工生成することが可能となる。

かつてアナログ回路しかなかった時代には、発振・フィルタリング・時間経過パターンの制御、いずれも技術的に制限があったため、音のバリエーションはそう多くはなかったが、現在では原波形をサンプリングで得られることと、デジタル回路による演算加工の自由度の高さが相まって、無限の音作りが可能になっている。

MIDI (Musical Instruments Digital Interface)

ここで補足的にMIDIにも触れておこう。MIDIは、シンセサイザーやPCなどのデジタル機器において音楽情報を交換するための規格（1983制定）で、これを用いると、タッチの強弱を含む発音のON/OFF・ベンド・音色切り替えなど、リアルタイムでの演奏制御が可能になる。MIDIのデータは、ステータスバイト（情報の種類）とデータバイト（内容）

の二つをセットにした計2バイトが1単位で、通信速度31250bpsで送受信される。

これによって、様々な電子楽器間での演奏情報のリアルタイム交換や、楽曲のデジタルファイル化などが可能になった。楽器を直接演奏できない人にも、作曲や自動演奏を楽しむ機会が与えられることとなり、絵の苦手な人にとつてのCGと同様、創作活動におけるハンディを解消するものとして、非常に意義のある存在である。その後、各種のメディアにおける演奏情報の統一の必要性が生じたことから、GM規格が制定(1991年)され、共通音源仕様(異なる機種でもほぼ同じ音色で演奏が再現される)が確立されている。

さらに今日では、PCの処理速度の向上と音楽ソフトウェアの機能充実によって、MIDIデータ(楽譜データ)とサウンドデータ(直接的な音声波形データ)を統合して処理・演奏する環境が実現し、安価なパーソナルのDTM(Desk Top Music)機材だけでも、作曲からレコーディング、そして音楽CDの作成までが可能になった。伝える中身と伝え方の問題は別として、ただ単に市販品のかたちにできるかどうかというレベルでは、もはやプロとアマチュアの境界はなくなっている。

スピーカ

電子楽器が出力する電氣的な振動(交流電流)を物体の振動に変換し、最終的に空氣の粗密波としての音を生成するには、スピーカ(電氣音響変成器・繼電器)が必要である。

スピーカには、コイルをつけた振動板を磁場に置いて電流に応じた振動をつくるという

動電型(ダイナミック型)と、電圧によって伸縮する圧電型とがあり、スピーカ・ヘッドホンの大部分は動電型、効率重視の携帯電話や薄さが必要になる壁掛けスピーカなどには圧電型が用いられている。動電型の場合、電氣から音響へ、また音響から電氣への相方向への変換が可能で、例えば動電型のスピーカはマイクロフォンとしても利用できる。^{*}スピーカを形状で分類すると、コーン型・ドーム型・ホーン型・リボン型などがあり、また再生帯域に関して分類すると、人間の可聴域(20Hz~20000Hz)を一つのスピーカユニットでカバーする「フルレンジスピーカ」と、可聴域を帯域別に分けて、各帯域を専用に振動させる「スーパースーファ(超低音用)・「ウーファ(低音用)・「スクーカ(中音用)・「ツイータ(高音用)」などがある。

再生のシステムには、モノラル・2元ステレオ(立体音響)・多元ステレオ(5・1チャンネルが主流)などがあるが、2元ステレオが最も一般的な方法である。2系統の音を2個のスピーカから出すというこの方法は「音場に立体感ができる」・「音源の移動が再現できる」などの理由で1950年代以降普及しはじめ、あらゆる音響機器もそれに伴って2チャンネル(L⇄R)を基本とするようになった。

音は聴覚がとらえ得る範囲の空氣の振動である。この厳然たる事実がある以上、音を生成する楽器や音響機器の基本的な仕組みは変わらない。子どものころに作った糸電話を思い出してみよう。基本に立ち返れば何か面白い発想も浮かんでくるのではないだろうか。

*可逆変換器

トランシーバなどでよく用いられる仕組みだが、動電型のスピーカは、そのまま音声をぶつけてマイクとして使う事もできる。このような、どちらの向きにも変換できるものを可逆変換器と言う。
ちなみに「変成器」と言えば可逆変換器、「繼電器」と言えは不可逆変換器を意味する。

画像・映像の生成

人は音を発することはできるが、「光」を発することはできない（人は音源をもっているが光源はもたない）。もちろん、気功術などを修得することで、遠赤外線の一つを自在にコントロールしつつ放射することが可能になるという例もあるが、これはきわめて稀な例であり、ここでの議論の対象ではない。「光」は数百nmという非常に波長の短い電磁波であり、我々の身体スケールではこれを音声と同様の仕方でもコントロールしながら放射することなど、とうてい不可能なのである。したがって、我々が視覚情報を生成するという場合は、もっぱら頭と体を使うことによって直接あるいは間接的に空間的な形状（情報源）をつくり、それを光源で照らすという手段をとることになる。以下にその具体的な手段と原理の概要を紹介しよう。

1 身振り

ペンをもつこともできない乳児が最初に発信する視覚情報、それが身振りである。

脳科学の分野で、近年ミラーニューロンという概念が話題になっているが、それは「他者の身振りはイコール私の身振りである」ということを説明する概念である。人間が他者

の行為を見る際の脳の状態を調べると、自分自身がその行為を行う場合と同様の部位が興奮している。つまり人の身振りというものは、見る人を同一の感覚に駆り立てる視覚情報であり、ただ単に「体の動く形が見えている」という以上にコミュニケーションを体感的なものにする特殊な視覚情報なのである。ビデオ映像でもあくびがうつる。身振りは、それが映像に映し出された人物のものであっても、非常に重要な情報を担っている。

さて、身振りという視覚情報の中でも最も重要なのは、「手の動き」と「顔（まなざし）の動き」である。このことは、手品師^{*}の動作を見ればよくわかる。観客は手品師の行う指差し動作や、小道具を見つめるまなざしに最も誘導されやすい。アニメ・キャラクターの動き、パントマイム、手話、いろいろな表現を思い出してみよう。身体を情報源とした視覚情報の大半はまなざしと手の動きが担っている。

蛇足だが、確認の意味で付け加えておきたい。視覚情報には照明が必要である。身振りも当然「光の下」でなければ見えないし、光の方向が変われば印象も変化する（能面・あごから懐中電灯）。立体である身体（顔）も、形と照明の両方に依存した視覚情報である。

2 画材

次に、人間が最も古くから一般的に用いている画像生成の手段について確認しよう。いわゆる画材による絵画的な画像の生成である。

^{*}手品師の動作

手品師が、指差し動作やまなざす動作によって観客の意識をタネから遠ざける操作をミスリーディングという。右を指さすと観客はほぼ間違いないくそちらに注目する。その間に左側で細工をするというものである。「あっち向いてホイ」という遊びも同じ現象を利用したものである。

鉛筆・色鉛筆・パステル

鉛筆は、筆記具としては最もなじみ深い存在である。木製の軸のなかに黒鉛を主成分とする芯を埋めこんだもので、芯そのものは、微粉碎された黒鉛と粘土を練り合わせ、脱水、押し出し成型、乾燥、焼成、油脂・樹脂の含浸という手順で形成される。工業製品として世界的に知られるようになるのは18世紀ドイツのバヴァリア鉛筆からで、現在では、描線の反射率（すなわち黒さ）によって、9Hから6Bまでの17種類に分類されている。

色鉛筆は、色彩表現に用いられる画材の最もポピュラーなもので、芯の形成は基本的には鉛筆と同様であるが、着色成分として白亜・アルミナを染料で着色したものや着色合成樹脂などを用いる点と、高温での焼成を行わない点が異なる。一般的には、硬質（ルモグラフ等グラフ作成や製図用）・中硬質（一般的な絵画デザイン用）・軟質（ダーマトグラフ等マーキング用）の3段階のものがある。

パステルは少量のメデイウムに顔料を混ぜ圧力をかけて固めたもので、そのアイデアは17世紀ごろに生まれ、18世紀のはじめにはパステル画というジャンルでパリを中心に普及した。パステルはメデイウムが少ないため、色が鮮やかで、そのものの色と紙に塗ったときの色に差がないといった長所があるが、一方で混色が難しい、定着力がないなどの短所もある。メデイウムの濃度の差で、ソフト・セミハード・ハードに分類され、色名と明度番号（0から8の階級）で特定できる。

インクとペン

インクの原型は、油煙を膠やゴム質などのメデイウムに混合したものである。それは中国・日本など東洋地域では墨汁として古くから利用されてきたが、ヨーロッパでは、古くは主な成分に木材のタールを用いる「ピスタ」や、コウイカの墨をベースにした「セピア」なども用いられていた。12世紀以降は主として「ブルーブラックインク」（空気に触れて酸化すると青黒く変化する化合物溶液）と呼ばれるもの、今日ではカーボンを水に溶かしてつくる「ブラックインク」や水溶性染料による「色インク」が使用されている。

ヨーロッパで発達したペンは（鳥の羽根 *Pen* に由来する）このインクを鋼鉄のペン先につけながら描く筆記具で、その形態には、一定の太さで描くのに適した標準的な「かぶらペン」、漫画家がよく利用する「Gペン」、製図やテクニカルイラストレーションに利用される「丸ペン」などがある。

ちなみに、アルファベットの書体はこの「ペン」の描線と密接な関係にある。アジアにおける毛筆と漢字、ヨーロッパにおけるペンとアルファベット、情報を記録するメディアは、そうした文化の根源的な部分にまで大きく影響している。

絵具と筆

水彩・アクリル・油彩などに大きく分類されるが、基本的には顔料は同じで媒質がそれ

ぞれ異なると理解すればよい。

水彩絵具は、18世紀後半からイギリスを中心に広まったもので、メディウムとしてはアラビアゴムが利用されるが、メディウムに対する顔料の割合や顔料の粒子径をコントロールするなどして透明性を変化させ、透明水彩・ポスターカラー・グワッシュなどと区別する。透明水彩はその名のとおり透明性が高く、重ね塗りをすると下の色が透ける。ポスターカラー・グワッシュはともに不透明であるが、その差は、グワッシュのほうが品質が高く、粘性が高いといった点にある。

アクリル絵具は、ビニル絵具などと同様に合成樹脂をメディウムとするもので、1930年代ごろからの壁画創作運動に関連して発達した。絵具の接着力が強いため、紙以外でも布・木版などの吸収性のある媒体が相手であれば描画が可能で、また調整剤によって様々な表現効果を得ることができるため、その応用範囲は広い。

油絵具は、植物性の乾性油をメディウムとするもので、一般に溶剤の調合によりその乾くスピードをコントロールしながら様々な表現効果を得る。15世紀のヨーロッパにおいて考案されたと言われ、はじめは白の地塗りをした板に描いていたが、17世紀後半になるとキャンバス（画布）に描くのが主流となり、今日のスタイルが確立する。

さて、筆であるが、これは中国では文字の歴史とともに、ヨーロッパでは美術の歴史とともに発達した。今日では洋画用の筆として、フラット（平筆）・フィルバート（筆先が楕円）・ラウンド（丸筆）・ファン（扇筆）などの区別があり、日本画用としては平筆（平塗り）・

削用筆（一般の線描）・面相筆（細線描）・則妙筆（墨描き）・彩色筆（彩色）・隈取筆（ぼかし）といった区別がある。

上記以外にも、俗にマジック・フェルトペンなどと呼ばれるマーカーや、インスタントレタリング・スクリーントーンなどの転写型の画材、カラーペーパー・グラフィックテープといった直接貼るタイプのものなど、様々な画材がある。

キャンバス・紙

最後に、画像を定着させる土台にあたるもの、すなわちキャンバスや紙について、概説しよう。キャンバスは、麻布にニカワを塗って乾燥させ、その上に白く地塗りをほどこしたものである。肖像・静物画（F）、風景画（P）、海景画（M）、正方形（S）など縦横比の違いで区別があり、そのサイズは号数で表わされるが、紙のように倍々で考えられる単位ではないため注意が必要である。ちなみに100号がほぼ等身大である。

我々が用いる紙は、一般に和紙と洋紙とに分けられるが、和紙はその耐久性の良さから、長期保存すべき文書や紙幣、障子・襖・照明器具などの建築・工業デザインの材料として用いられ、一方、洋紙は描画や印刷などに用いられる。^{*}洋紙はその原料として、植物から抽出したセルロース繊維の集合体すなわちパルプを用いる。パルプには機械パルプ（原木を機械的にすりつぶしたもの）と化学パルプ（薬品を混ぜて圧力釜で煮てつくるもの）とがあるが、その製紙のプロセスなどによっても様々な分類される。

*洋紙

欧米の図書館で問題となっているように、従来の洋紙はそれに含まれる酸性物質のために、やがてはポロポロになってしまう。そこで保存に適した耐久性のある紙対策が考えられ、現在は中性紙や微アルカリ紙が多く生産されるようになってきた。

さて、ここで紹介したような画材は、今日どこでも手に入るものなのだが、逆にそのことが、我々の創作活動を惰性的にしていることも否めない。例えば、よく見かけるB1のパネルや、A4のケント紙。これらはルート2矩形（縦と横の比が $1:\sqrt{2}$ ）と言われるが、このサイズの矩形にボスターや絵を描く理由はどこにあるのだろうか。作者自身は、この矩形を美しいと感じているのだろうか。この縦横比は「半分に切っても同じ縦横比の矩形となり、サイズのシリーズ化がしやすい」という理由によるもので、工業的に都合の良い比率であるにすぎない。そう考えると、一枚の絵を描くのにこの比率に縛られる必要はないということがわかる。商業的なデザインワークでは予算や流通の問題から結果的に様々な規格ものに依存せざるをえないが、モノ作り一般においては、既製の版形や各種材料・道具といったものから、もっと自由になつていいのではないだろうか。画面は丸くてもよいし、また紙である必要もない。ペンの代わりになるものなどいくらでもあるし、植物の葉やワインでも色は染まるのだ。画材の歴史をふまえた上で、あえてそれを遡ったり、路線を踏みはずしてみたりすることも、新たな発想への契機として大切なことと言えよう。

3 コンピュータによるイメージの生成

いわゆるCG（コンピュータ・グラフィックス）は、従来の画材によるイメージ生成とは異なり、「数値データから自動的にイメージが生成される」という特徴をもっている。

数式がつくりだす抽象的なパターン、生物・物理・化学におけるミクロ・マクロな世界のふるまい、カメラでは撮影不可能な架空の世界の構築など、かつて視覚が体験した事のないイメージの生成が可能になった。

CGの歴史は、サザランド（1963）によるスケッチパッド（対話型の描画ツール）の開発にはじまり、その後、図形の描画のための各種アルゴリズムの確立と、それに関連するハードウェアの進歩が相まって、現在では、建築・工業・出版・芸術・娯楽映像・ゲーム・Webと、わずか40年程度で「もはやできないことは無い」という域に達している。

手続記述型のCG

描画の手順が機械的に表現できる場合、それをプログラムしてコンピュータに自動生成させることができる。そんな発想で作るCGを手続記述型のCGと言う。この種のCG制作では、出来上がるイメージを事前に予測するのは難しく、何ができるかわからないというゲーム感覚の面白さもある。対話型ツールが普及した今日では少数派の存在であるが、テレビ等で見かける自然現象のシミュレーションや、幾何学的抽象的なアニメーションなどを作る場合には必須の手法である。以下、いくつかの簡単な事例を紹介しよう。

数理曲線：描画空間の座標（ x, y, z ）を、高次の多項式や三角関数を用いてパラメトリックに表現すると、複雑な幾何学模様*をつくることができる。縦方向の振動と横方向の振動が合成してできるリサージュ図形などがその典型である。

*幾何学模様と乱数

数式が単純な場合、できあがる形状はそのままでは機械的で表情に欠けることが多い。そのような場合は、プログラムの一部に乱数を発生させる仕組みを加えると、変化ができて面白くなる。

フラクタル：部分の形が全体の縮小形になっているというような「自己相似性」をもつ図形をフラクタル図形という。樹木・雲模様・海岸線の形状など、いずれの形も、基本となる「図形を描く」プログラムをその図形の一部を描く部分に再帰的に利用することで生成することができる。例えば「一本の幹の先端に二本の枝を描く」というプログラムを、その枝を描く部分に適用すれば、枝の先端から二本の子枝ができる。これを子枝・孫枝：と繰り返せば、枝の生い茂る木ができるというしくみである。フラクタル図形は、きつかけとなる手順やデータが単純であっても、できあがる全体は非常に複雑で、またわずかな初期値の違いが最終的な結果を大きく左右するために、最終的なイメージは予測が付きにくい。このようなシステムを一般に「複雑系 (Complex System)」と言うが、実は、我々生命体を含む多くの自然現象がそのようなふるまいをしている。要素単位で見れば単純な規則でも、複数の要素が影響しあうと非常に複雑になるのである。膨大な計算を必要とするものであり、コンピュータの存在無くしては、このようなイメージには出会えない。

形状記述型のCG

頭で思い描いた形を、マウス等を用いて対話形式でコンピュータに入力していく。こうしたタイプのCGを形状記述型のCGと言う。このタイプのものは、従来の絵画や彫刻の場合と同様の思考方法でモノづくりを行うもので、絵筆をマウスに置き換えたようなものと言いうことができる。最近ではそのツールも非常に身近なものとなっている。

ペイント系ツール：画面を画素（ピクセル）の2次元的な集合と考え、各画素に色情報をもたせるかたちで像の生成を行う。ブラシや消しゴムを使う感覚で利用できるものだが、写真などの静止画を取り込んで様々な画像処理を行える点は、従来の画材にはない特徴である。ペイント系ツールの上では、絵画も写真も区別なく同質のデータとして扱われる。

ドロー系ツール：画面を図形（オブジェクト）の集合と考え、図形単位に頂点座標と結線情報および色彩などの属性情報をもたせるかたちで画面を構成する。データの性質上、ペイント系のツールのような画像処理はできないが、文字も含むすべての図形要素について、拡大・縮小における画質の劣化がないため、印刷物を作成する場合の最終的なツールとしての利用価値が高い。

3DCGツール：3次元すなわち（x,y,z）の軸をもつ架空の世界でデータを定義し、その形状を代数的な計算によって2次元のスクリーンへ投影して表示するものである。3次元空間内でのデータの移動・回転・拡大・縮小、視点をカメラの位置へ移動する視野変換、隠線・隠面の消去、光源・物体・視点の関係に基づく影と明るさの計算、テクスチャ画像のマッピングなど様々な計算プロセスを経由して、最終的には架空の世界の写実的なイメージが生成される。

さらに、その位置・回転・サイズなどのデータを時間というパラメータにしたがって変化させれば3DCGアニメーションができあがる。その制作は従来のセルアニメーションの場合と同様のキーフレーム法、すなわち、動きのキーポイントとなるフレームのみを定

義し、その中間部分は計算で補間していくという方法が採用されている。

3DCGアニメーションには、全体の動きを骨格の動きとして定義するスケルトンアニメーション、関節の回転角に制限を与えて先端から中心へ（指先から肩へ）のなめらかな関節の動きを定義するインバースキネマティクス（逆運動学）、さらに、対話型操作では定義が難しい複雑な動きを、実際の人間に取り付けたセンサーから取り込むモーションキャプチャなど、キャラクターに振り付けを行うための様々な技法が用意されている。

オーサリングツール：画面上に配置されたアイテムを、ユーザーが操作をすることで、物語が進行する。そのようなインタラクティブなムービー作りを可能にするのが、このオーサリングツールで、文字・図形・静止画・動画・音声とあらゆるタイプの素材を用いた総合的な演出が可能である。ここで実現されたインタラクティブティーは、ビデオゲームなどではすでに当然のことなのであるが、視聴者が動画像の進行に関与できることはCGの技術以前にはなかったことであり、まったく新しいタイプの映像体験を生んだと言える。

プリンタ・ディスプレイ

コンピュータで生成された静止画や動画は、最終的にプリンタやディスプレイといった媒体で可視化される。この2つは、前者がCMYK 4色の減法混色、後者が直接「光」を素材としたRGB 3色の加法混色で画像を生成するという違いがあり、それぞれの発色の限界に関しては注意が必要である。

プリンタにはインクジェット式・電子写真方式（トナー画像転写式）・熱転写式（溶融型と昇華型）などがあり、データ形式の面ではポストスクリプトデータを処理するPSプリンタがデジタル印刷の主流と言える。その他プリンタではないが、フィルムレコーダなども出力装置として用いられている。

CGの画面出力にはラスタスキャン型のCRT・液晶・プラズマディスプレイなどが用いられるが、テレビとは仕組みが異なり、ノンインターレース方式（上から下への順次走査）で、画面のリフレッシュレートも高い。アスペクト比は4:3が一般的で、画面の画素数は1024*768、1240*1024、1600*1200など、段階的に選択できる。

プリンタの出力も画像の出力も基本的には平面であるが、音響のシステムと同様、立体視のためのステレオグラフィックスシステムもある。レンティキュラ板やパララックスバリアで表示面を加工する方式、2チャンネル分の複合画像を偏光や時分割シャッターで左右の目に分離する眼鏡方式などがあり、VRシステムの最終出力として欠かせない技術となっている。

さて、コンピュータによる画像・映像の生成について概観したが、便利な環境が整った今日の状況に関して、これまでの経験をふまえた私なりの「覚え書き」を残したい。筆者が、コンピュータというものと出会った時、それはまだ「画面」を持たない存在、つまり入力も紙（パンチカード）で、出力も紙（連続伝票）というものであった。

80年代にはパーソナルコンピュータも台頭し、モノクロ、8色、16色、4096色、90年代に1670万色と、あつという間にフルカラー表示が実現し、現在では動画像のリアルタイム処理も可能になった。かつて、動画はもとより、大きなサイズの画像や色数の多い画像はモニターに表示する事ができず、CGで作品を作るには、生成した画像データ（画素ごとのRGB値を羅列した汎用画像ファイル）を専門のラボで出力してもらうしかなかった。そんな時代と比較すれば、3DCGを含めた映像の制作がPCの画面上でストレスなくできてしまう現状はまさに夢のようである。

しかし、このコンピュータに依存した今日の制作の環境では、余計なことに必要以上の時間がかかったり、目的と手段が逆転してしまったりすることも多い。デザイン教育に携わる中で、日頃感じている「大切なこと」をいくつか挙げておこう。

第一に、ツールは手段であって目的ではないということ。優れた機能をもつソフトウェアを前にして「すべてを使いこなさなければ・・・」という強迫観念が、制作の手を鈍らせてしまうことが多いのである。確かに、競争原理で動くデザインの現場では、知らないことで作業能率を下げるようなことがあってはならない。しかし、それも度を過ぎると結局「使い方を覚えるためにソフトを使っている」という状況に陥ってしまう。今やるべきことは何か、無意味なことに時間を使いすぎているか、常に自分に問いかける必要がある。

* ツール

「競争原理が優先する現場」を離れば、もともとモノ作りのツールは何でもいいのだ。その意味では、有志が集まって開発を続けるオープンソースのツールは、無料であるという以外にも、極端な仕様変更やライセンス管理のわずらわしさがなく、なにより「自由な配布」が永久に保証されているという信頼感がある。純粋なモノ作りには、本来このようなツールが向いているのかもしれない。以下を参照されたい。

<http://www.gnu.org/home.ja.html>

第二に、画面に向かう前にしっかりと設計（スケッチ）が必要だということ。特に、紙に書くことは重要である。例えばメモ。エディターで打ち込んだメモは確かに美しいが、すべての文字が均一に打ち出されるため、思考のメリハリが見えなくなる。一方手書きのメモは、文字の大小や強調・省略、さらに文字と文字の位置関係から、すぐにその主題と思考の図式が想起できるのだ。直接画面上で編集すれば修正と同時に完成であり、それは一見能率的に見えるのだが、切り貼りの堂々巡りで何を書いているのかわからなくなってしまうことも多い（これはプログラムの開発でも同じである）。直接画面に向かう方が能率的に見える時代だが、設計段階では紙と鉛筆の方に軍配がある。

第三に、「何を伝えたいのか」を忘れてはならないということ。作業がゆつくりと進行していた頃はそれを常に意識する余裕もあった。しかし、「ファイル・新規作成」だけで新しい紙が準備されてしまうような環境では、いろんな技術を試してみたくなる。そして思考が追いつかない。技術の誇示ばかりが優先されて肝心の「伝えたいこと」が脇に追いやられてしまうことがないよう心がけたい。たとえば線一本でも十分に存在感はあるのだ。

最後に、CGもひとつの手段であってその領分があるということ。絵の具がもつ物質的な存在感や写真映像がもつ「聖なる一回性」の感覚は、CGが勝負できる領域ではない。大事なのはCGでなければできないことを模索することである。メディアの「らしさ」が十分追求されたとき、それはツールの良し悪しに関わらず強い存在を生む。

人間とメディア

ここでのテーマは、情報という概念を鍵に、入力・処理・出力および記憶を行う存在としての、「人」と「メディア（情報機器・媒体）」を比較考察することである。

人間が、目で見て・脳で処理して・記憶するように、「メディア」もまたカメラでとらえて・CPUで処理して・記録する。もちろん、意識をもつて思考する生物である人間と、意識をもたない機械メディアとを、同じ「ものさし」で比較することはできないのだが、音楽も映像も同じ情報であるという観点で、「人」と「メディア」の情報処理過程を比較すれば、普段意識されない人間の特徴、またメディアの特徴を抽出することができるのではないかと考えた。本文では、このような観点から、今後「メディアの視覚」とか「メディアの記憶」といった多少違和感のある表現も常用しつつ、様々な考察を行いたい。その内容は音楽・映像という言葉からずいぶん遠ざかる感もあるが、筆者の経験から言えば、このような視点に立った考察は、音楽や映像の制作に欠かせないものである。

さて、後の文章につなげる準備として、はじめに「メディアの視聴覚」と「メディアの記憶」について、その特徴を概説しておきたい。

「メディアの視聴覚」の特徴の一つは、「感覚器（入力装置）」から「記憶（記録媒体）」へ至る経路が一本のケーブルで実現できることと、そこを通過する情報が「信号」というかたちで単独で取り出せるということである。当然だが「人の視聴覚」ではこれはできない（視神経を伝わる画像情報を取り出してモニターで見たり、直接他人の脳と接続したりはできない。だから人は、様々な表現手段を用いて頭の中の情報を移動・複写しなければならないのである）。

「メディアの視聴覚」のもう一つの特徴は、アナログ処理系とデジタル処理系の2つが明確に分離されている（中継にはA/D変換器が必要）ということである。現在の大半のシステムでは、音や光の刺激そのものをとらえるという「感覚」レベルの処理はアナログ処理系が行い、情報を識別・解釈するといった「知覚・認知」レベルの処理はデジタル処理系が行っている。この図式は「人」の知覚システムでも同様ののだが、アナログとデジタルの境界が明確であること以外にも、信号の伝達が電気・電子的で化学物質の受け渡しが行われないこと、また一般に「トップダウン処理」がないこと、他のモダリティーとの相互作用（共感覚）がないことなども、「人の視聴覚」とは異なる特徴と言える。

「メディアの記憶（記録）」についても述べておこう。アナログ式とデジタル式の二つがあるが、いずれも、引出し式であること、一回で正確に記録されること、容量に明確な限界があること、自発的動機がなく自己組織化もしないこと、情報がそのままの形で取り出せること、といった特徴をもっている（もちろん人の記憶の裏返しとしての特徴である）。

以下、これらをふまえて「人の視聴覚と記憶」の特徴を考察してみよう。

* トップダウンとボトムアップ

情報の入力に際し、脳の知識ベースを利用して刺激を待ち受ける、つまり、上から下へ降りてくるプロセスをトップダウンプロセス（概念駆動型処理）、逆に刺激情報が目から脳へと上がっていくプロセスをボトムアッププロセス（データ駆動型処理）と言う。

* 共感覚

音を聴くと色が見える（色聴）場合や、「黄色い声」という表現など、ある感覚が他の領域の感覚と関係をもつ場合を言う。

1 感覚・知覚・認知

心理学の教えるところによると、我々の情報の読み取りには、感覚 (Sensation)・知覚 (Perception)・認知 (Cognition) の3つの段階がある。

「感覚」とは、外部（あるいは体内）からの刺激に対する対象性のはっきりしない「感じ」のことである。聴覚で言うとき音の大きさ・高さなど、視覚で言うとき明るさ・色などがそれにあたる。この感覚については一般に、1次元刺激（音の大きさ・明るさなど）について、 7 ± 2 段階程度の弁別能力があると言われ、また感覚の大きさは刺激の物理強度の対数に比例する ($R = \log S$) ということが知られている。さて、「メディアの感覚受容器」であるセンサーや各種測量器はどうだろうか。このような、物事を物理量で捉えるレベルの問題では、「メディアの視聴覚」は非常に優秀である。それは、我々が通常「明るさ」を測ったり、「音の大きさ」を測ったりするときに、測定機器に頼ることを考えれば当然とも言える。1次元刺激の弁別・評価に関して、正確さを必要とする場面では、人間は常にそれを機械に委ねている。

「知覚」とは、受けつけた感覚刺激により構成される対象性のはっきりした経験である。聴覚で言うとき時間的な刺激の配列関係である「旋律」や「リズム」が、視覚で言うとき空間的な刺激の配列関係である「形」がその対象である。知覚は、刺激の関係性によって対象

*感覚遮断

あらゆる刺激が0あるいはコンスタントな空間で、自らも安静にすることを感覚遮断というが、我々は常に何らかの刺激を受けながら生活しているため、そうした極端な実験空間に長時間おかれた場合、耐えられなくなってしまうか、または幻覚などの異常な心理現象を生ずる。

人間や動物は本来刺激のないときに刺激を求めようとする「知覚欲求」をもつ生き物である。快適に安らげる空間とは、とりあえず無刺激・無感覚の空間ではないことがわかる。

化されるので、物理量に変化があつたとしても関係が同じならば同一のものとみなされる。

例えば、キーが変わつても同じ曲に聞こえるとか、明るさが変わつても一つの部屋を見間違ふことはないということなどが、それにあたる。さて、「メディアの視聴覚」の場合はどうであろうか。物事を関係性で捉えるということは、生命体に特有のものであつて、「ボトムアップ処理」を基本とする情報メディアにはそれは難しい。四角い形をカメラから入力して、それを丸や三角とは違う「四角」と判定させるだけでも、そのパターン認知処理は非常に複雑なのである。もし機械にこのようなことをさせたければ、「人」の知覚における情報処理プロセスの模倣、すなわち、知識データベースにもとづいたトップダウン型の処理システムが必要になる。

「認知」というのは、受けつけた知覚対象を記憶に照合して、自分の世界像における位置づけを行う段階である。感覚・知覚までは、他の生物でも見られる比較的低次のプロセスであるが、認知のレベルは、「言語」や「文化」といった知識ベースも関与する人間特有の非常に高度なプロセスである。このレベルの処理は、他の生物や人工知能を持たない通常のメディアには難しいものとなる。

2 ノイズ・かたち・ことば

我々をとりまく刺激あるいは情報は、この感覚・知覚・認知という3段階の概念をもち

* 7 ± 2

人間のセンサーとしての限度を物語る数字であるが、日常の刺激は多次元的であるため、実際にはさらに弁別能力は上がる。

ちなみに、 7 ± 2 というのは短期記憶に関する話にも登場する。一般に人の短期記憶の容量には限界があつて、一度に覚えらるる項目の数は 7 ± 2 程度であるというものである。G・A・ミラー (1956) はこれを「マジカルナンバー7」と呼んだ。小説やマンガの主人公がいずれも7人程度であるというのも、このことと無関係ではない。

いて分類することができる。まずは単なる刺激に止まるものの例として「ノイズ（雑音）」、次にかたちのレベルに止まるものとして「外国語」・「音楽」・「図形」、そして最後に言語的な意味にまで到達する「母国語」・「映像」などが、それぞれ感覚・知覚・認知の各段階に相当する。

まず感覚レベルまでの刺激情報であるが、一般に我々は純粋な感覚というものを得ることは難しい。なぜなら大半の刺激はすぐに秩序化されて「形」として知覚され、またそれはすぐに言語的に了解されてしまうからである。したがって、このレベルの対象となり得るのは「ノイズ」のような対象化されない刺激か、あるいは一部の「芸術的創作」に見られる「人間の自動的な「読み」を遅延させる実験的作品」のようなものである。人にとって秩序化して把握しにくいもの、言語化不可能な生々しさをもつものは、不快であると同時に新鮮であるという両義性をもっている。

さて次に、知覚のレベルまでのものについてであるが、例えば「外国語（その内容が理解できない場合の話である）」は、「パターンをもつ音」として聴覚に、あるいは「規則性のある図形の並び」として視覚に与えられるもので、知覚レベルまでの対象となる。外国語による歌（Song）は、他の楽器と同様に、純粋に声という楽器が奏でるメロディー（形）として聞こえてくるし、また例えば英字新聞は細かな図形の作る模様に見える（包装紙としても違和感がない）。秩序あるかたちとしては捉えられるが、言語的に何かを意味するわけではない（言い替えれば意味を求めようという欲求を生じさせない）純粋な「形」、

そのようなものが知覚レベルまでの対象である。抽象的絵画や建築は、よく音楽に例えられるが、それも、それらを知覚レベルで捉えた場合の話である。

言語的に認知されるレベル（あるいは言語的な理解を要求するレベル）の情報としては、「母国語」による言語情報がその代表としてあげられる。言葉や文字は、聞く・見ると同時にただちに認知されるもので、我々のコミュニケーションは、音声・文字の聴覚・視覚刺激から自動的に感覚・知覚・認知へと進む「読み」にささえられている。これは逆に言えば、母国語の言語情報の場合、認知された意味内容は記憶されるが、そのときの音声や文字の感覚・知覚レベルの情報は記憶に残りにくいということを意味するものであり、言葉や文字を、純粋にその音や形の水準で伝えようとするには「美的」な配慮が必要となるのである。

さて、やっかいな存在が「映像」である。「映像」はそれ自体では言葉や文字を含んでいないにもかかわらず、我々はそれを言語的に了解しようとする。たしかに、そこに映しだされるものは、我々の現実世界に存在するものであって「名付けられるもの」であるから、言語的に認知できるのであるが、しかしスクリーンに映し出される「鳩」は、我々が街で見かける「鳩」以上（以外）の何かを意味する場合があるし、同じ「鳩」のカットでも前後のカットとの関係でその意味する内容が変わってくる。「映像」は単なる現実のコピーとは違うのである。さらに、知覚される「形」が同じでも、その解釈すなわち認知のされかたが、見る人によって様々であることを考えれば、「映像」という情報は、それを受け

取る場面で意味が生成するという性質のものだとも言えるのである。これは、記号の「形」や「配列規則」が同じであれば、一般的にその意味が一義的に定まる（と考えられている）「言語」とは大きな違いである。

3 人間の記憶

人間の記憶は、コンピュータのメモリーのような「引き出し」に知識項目が一つずつ入っているというイメージでは捉えられない。詳細を後にして、先に一般的な事柄を述べると「人」の記憶は、細胞イコール一つの記憶単位と考えるより、神経細胞同士の結合の「関係」が記憶の「構造」をかたちづくっていると考える方が、あらゆる点で説明がつきやすい。人間の記憶は、「複数の神経細胞が複数の事象についての情報を重層的に担う」という意味で、「ホログラム式の記憶である」とも言われる。

人の脳は140億*個ほどの神経細胞 (Neuron) で構成されており、それらのシナプス (Synapse) 結合によって、細胞同士の興奮の伝達が行われている。マツカロとピッツ (1943) が提唱した神経細胞のモデルによると、細胞のそれぞれは、静状態と興奮状態の2状態があつて、興奮状態においては電気パルス列が出力されるのだが、この場合、ひとつの細胞の出力は、それに結びつくシナプス (約1万個) からの興奮信号の重み付きの総和が、あるしきい値を超えるか超えないかで「1 or 0」に決まるという。

したがって結合の強い（重みの大きい）細胞間では興奮状態が一斉に伝わり、結合の弱い細胞は静状態という、脳全体でみれば一つのパターンが生じる。この興奮パターンがある一つの概念なりイメージなりに相当すると考えられるのである。

また、興奮パターンが、自己組織化する、すなわち「人」がある事象を記憶するというプロセスをうまく説明する仮説として、ヘップ (1949) の「シナプス強化法則」がある。その仮説によると、神経細胞が興奮する際、その細胞に刺激を伝えたシナプス結合部については、その結合がより強化され、結果としてその後の刺激は以前に増して伝わりやすくなるというのである。

この考えをふまえると、我々の記憶では、複数の神経細胞の同時興奮パターンという「結合関係」が重要で、脳内でその興奮パターンが繰り返されるたびに（反復学習にあたる）、その「関係」がひとつの記憶単位として組織化*していくと考えられるのである。

ただしこの場合は、記憶単位といっても、その同時興奮する細胞群のひとつひとつは、それ以外の刺激に対しても他の細胞との関係で興奮することがあるわけで、その意味では一つの神経細胞が複数の事象の記憶に関わっていると言える。これは、ある部分の細胞が欠落しても、一つの事象の記憶がすっぽり抜け落ちるのではなく、その部位に関わる記憶全体がぼやけるということをも意味するもので、人の記憶が「ホログラム」*的であると言われるゆえんである。

さて、こうした知見によれば、物事は一つ一つの項目としてではなく、「関係」として

*記憶単位

細胞が一つの単位になっているという例としては「顔細胞」や「手細胞」というものが知られている。これはサルの実験で明らかにされたもので、「顔（目が必須）」や「手」という視覚刺激に対して特異的に反応するニューロンのことである。身振りに関して述べたように、我々にとって「顔」と「手」は特別な存在と言える。

*細胞の数

一般にこのように言われるが、これは大脳皮質の話であつて、実は小脳にはこれを大きく上回る1000億以上のニューロンがある。「頭で覚える記憶」は大脳皮質、「体で覚える記憶」は小脳、というふうに考えられている。

*組織化

脳ほど細胞が活発に変化するところはない。特に樹状突起を伸ばして他との接合関係を更新するという活発な動作には、がん遺伝子のようなものが関わっているとも言われる。脳は遺伝とはあまり関係がないようにも見えるが、実はタンパク質をつくる遺伝子の大半は脳で発現している。脳の中では常時DNAが活動し、大量の化学的プロセスが高速で行われているのである。

*ホログラム

3次元の空間情報を2次元の板に記録するもの。空間領域の情報を周波数領域の情報に変換して記録するため、板そのものには形は見えないが、参照光をあてると像が再生する。板の一部（特定の周波数領域）が欠落してもボケる程度で情報が完全に消滅するわけではない。あらゆるものが重層的に記録されるという意味で脳の記憶に例えられる。

一挙に構造化されて記憶されているということになる（構造主義言語学のF・ソシュール（1916）も同じことを言っていた）。例をあげてみると、我々は新しい言葉を覚える際に、反対の意味の言葉や、対になる言葉とともに「二項対立」的に記憶する方法をよくとる。これは単独の項目よりも二つの対立項目で記憶するほうがその関係の問題として記憶に位置付けやすいことを意味している。さらに言えば、我々の日常的な用語には単独では用をなさない「上」とか「左」とかいう概念があつて、辞書の「左」の項には「右の反対」、「右」の項には「左の反対」と記されており、要するに関係の問題でしかない概念も多い。

4 メディアの記憶

この章のはじめに触れたように、メディアの記憶動作は、直接的・間接的な指示による1回の動作で行われるもので、その記憶の状態は「引き出し式」で正確である。さらに言えば、そこには自発的な「動機」は必要なく、「反復学習」も必要ではない。そして人間との違いで強調されるべきことは、それが自己組織化しないということである。情報機器の典型であるコンピュータのふるまいを見れば、これらの事は一目瞭然であろう。

さてここで、人間の学習や記憶のモデルを応用した人工知能のプログラム、すなわち一般的なメディアの記憶とは異なる、人間に近い仕組みを持った記憶について簡単に紹介しておこう。ここで触れるのは、学習認識装置「パーセプトロン」(F.Rosenblatt, 1960)と、

* 右脳と左脳

我々の体は対側制御、すなわち体の左半分を右脳、右半分を左脳が制御している（左視野は右脳、右視野は左脳に入る。また感覚的・空間的・音楽的な情報処理を右脳、論理的・言語的な情報処理を左脳、というふうに、右脳と左脳ではそれぞれタイプの異なる処理がなされている。右脳が活発に動く人と左脳が活発に動く人では、情報の受け止め方や、発想の仕方が異なることが予想されるのである。もちろん左右の脳は連携しているのだが、言葉・音楽・映像の受取り方には「右からささやかれたか、左からささやかれたか」と同様の問題が無関係ではない。

ちなみに「子音＋母音の音声」[か] [き] を用いる日本人は、子音中心の言語を用いる西洋人と比較して左脳（言語脳）で処理される音声の種類が多く、このことが「虫の音・母音と構造が似ている」をも文学的な素材とする」日本人特有の文化を作り出したとも言われている。

連想記憶装置「アソシアトロン」(Nakano, 1969) というもので、マツカロ・ピッツの神経細胞モデルとヘップの法則を応用したモデルであるが、これらは「人」の学習・記憶をうまく説明するモデルとして興味深い。

パーセプトロンは一般に小脳の記憶モデルとして知られるが、その仕組みは、神経細胞群にあたる複数の素子の興奮パターンを入力として、その刺激が何であるかを識別し、結果を出力細胞に相当する一つの素子の興奮の有無によって得る、というものである。出力素子には複数の素子からの「シナプス」が結合していて、それらからの入力のみ付き総和が出力を決める。学習はこの結合の各々の重みを修正していくことで行われるのだが、具体的には、あるパターンを入力させて、その識別結果が正しければなにもしないが、間違えた場合はそれに関与したシナプス結合の重みと出力素子の判定のしきい値を「間違った判定をおこしにくい方向へ」修正する。これを繰り返す（反復学習）うちに、同じ間違いをしなくなる、すなわち正しい識別ができるようになるというものである。その際、入力のパターンの与えかたと判定の修正が上手であれば、すなわち「良い教師」につけば、短時間で識別能力がつく。この人工知能は「人」の記憶システムのモデルであるから、間違ふこともあるし、教え方の上手下手も関係することになる。

一方、アソシアトロンは、入力が複数の素子の興奮パターンで与えられる点はパーセプトロンと同じだが、すべての素子についての相互の結合強度が、同時に興奮している細胞間で正の方向へ、興奮している細胞としない細胞の間では負の方向へ（排他的に）修

正されるという点がその特徴である。刺激パターンの入力度にこの操作（記録）を行うとすると、入力パターンごとにそれに関与する（そのとき興奮状態にある）素子同士の結合が強化されることになる。

こうして出来上がった「結合強度情報をもった細胞群」に対して、「ある細胞が興奮すると、それと強く結合している細胞が同時に興奮する」ように動作させると、一部の興奮から記憶されている興奮パターンが再現されることになる。これが「なにかをきっかけに全体を思い出す」ことであり、「ある事象の記憶から、それに近い事象の記憶が想起される」ということである（連想記憶モデルと言われるのはこのためである）。

このモデルでは、複数の神経細胞群が全体として複数の事象を重層的に記憶しており、きっかけとなる部分的な興奮の与え方で、うまく全体のパターンが想起されたり、よけいなものまで同時に想起したり、あるいは異なるパターンがあざやかに浮かび上がったりする。そのふるまいはまさに「人」の「記憶の呼びだし」そのものである。

5 記憶の過程

さて、おおまかに「人」の記憶とそのモデルとしての人工知能の発想について述べてきたが、さらに人間の記憶の過程についても触れておきたい。

心理学の知見によれば、人間の記憶には複数の領域・段階があり、それぞれ、感覚

登録器 (sensory resister) ・ 短期記憶 (short term memory) ・ 中期記憶 (middle term memory) ・ 長期記憶 (long term memory) と呼ばれる。

感覚登録器は、視覚で1秒以下、聴覚で数秒の記憶で、聴覚刺激・視覚刺激などの感覚刺激をそのままのパターンですべて記録すると言われる。瞬間的に目を見開いて閉じたときに「目の前の情景が焼き付いている」という感じがするのがそれである。しかしこの情報は次々に捨てられる運命にある。

短期記憶は、感覚登録器の内容から知覚された意味のある情報を数分という短い時間の間記憶する領域で（同時に記憶できる項目数は 7 ± 2 程度）、ここで「興味をもって理解する」といったことが伴えば、次の中期記憶に送られる。

中期記憶は、脳内の「海馬」に1時間から最大1ヶ月程度保持される（大半は9時間ほどで消滅する）記憶で、この間に複数回のアクセスを受けたものが重要なものとして「側頭葉」に送られ、それが最終的に長期記憶になると考えられている。

長期記憶にはさらに、宣言的記憶（言葉で記述できる事実に関する記憶）と手続記憶（クルマの発進のしかたなどの手続きに関する記憶）との区別があり、宣言的記憶はまたさらに意味記憶とエピソード記憶に分けられる。前者は反復学習による体系的な知識ベースで指標（どこで覚えたかという情報）のないもの、後者は特定の時間・空間に関する具体的な体験の記憶である。ちなみに「記憶を失う」という場合は、このエピソード記憶の喪失であることが多い。

6 記憶と想像力

ここで、このような記憶に関わる、いくつかの身近な問題を考察してみよう。

はじめに、「記憶の容量」についてである。「メディアの記憶系」の場合、記憶が「引き出し式」であることから、おのずと限界が生じるのであるが、人の「ホログラム」式記憶の場合は、どこまでが限界というものではなく、知識の構造化が能率的であればあるほど、より多くのことを記憶できる。人の脳細胞の数はほぼ同じで、実際にはその数%しか働いていないという報告とあわせれば、「頭のよい人」というのも「容量」の問題ではなく、知識の構造化がうまくいかどうかの問題であると言えるだろう。例えば、ある社会現象を説明するモデルが、過去に学んだ物理現象を説明する数式モデルと似通っていると気付いた場合、その知識はパラレルに重ねあわせながら記憶（既存の神経細胞の結合関係が流用）されるわけで、この場合の記憶は能率的である（実際、学問に興味をもった場合、このような学習の転移がおこることは多い）。ある分野について学習すると、異なる分野の知識の飲み込みも早くなるのはそのためである。

次に「記憶の正確さ」についてであるが、「メディアの記憶」は当然与えられた精度の範囲で正確に再現されるもので、なんらかの障害によって間違う場合は、もとの情報は見る影もないというのが普通である。一方、人間の記憶は、基本的な言葉の意味や日々の生

活に関わる範囲では正確だが、そうでない部分については、あいまいであるか、欠落しているか、^{*}まちがった記憶になっているかである。これも結局は、記憶の構造が「ホログラム」式であることに由来するもので、新しい情報が記憶を再構造化する過程で、言い替えれば、神経細胞同士の結合関係があちこちで強まったり弱まったりする過程で、古い記憶に関する結合が弱まって薄れたり、あるいは別の記憶に関わる部分の結合関係を変えてしまったたりすることによって説明がつく。

最後に「進歩」という観念について。上述したことのくりかえしになるが、新しいことを覚えるということは、過去の記憶の「関係」を修正することで、これはすなわち「何かを覚えるとき、気付かぬうちに何かを忘れていく」ということを意味する。人は成長する過程で確かに新たな知識を蓄えていくように思えるが、これは「知識を確実なものにしていく」ということで、神経細胞のレベルで言えば「頻繁に駆動する一部の知識・思考回路に関してはその結合が強化されて、その他の結合は断ち切られていく」、簡単に言えば「頭が固くなる」・「思考がワンパターン」になるということなのである。

なにも知らない子供が、ユニークな発想で大人を笑わせたり、すぐれた想像力を発揮したりするのは、このことの裏返しである。

その意味では「進歩」は幻想であり、人の社会化（大人になること）とは、あらゆる可能性の放棄の上に成り立っていると言いうこともできるのではないだろうか。

一面的な見方で、人の脳に優劣がつけられないことは言うまでもない。

* 記憶の違い

連即記憶モデルに関連して述べたように、我々の記憶は、連結の状態があいまいになると、複数のイメージが渾然となって想起されてしまう。「このあたりにこんなものがあったはずだ」という、本人にとっては鮮明な「昔の旅先での記憶」も、それが長い時間のあいだに複合・再編された記憶である場合もあるのだ。

7 「情報への構え」と「視聴覚の相互作用」

一般に、「耳」はマイクに、「眼」はカメラに対応させて考えることができるのだが、ここで、「人」と「メディア」の決定的な違いである「情報への構え」の問題と「視聴覚の相互作用」の問題をクローズアップしてみよう。

情報への構え

「情報への構え」とは、簡単に言うと、情報を見聞きする場合の「予備知識」や「先入観」がつくりだす心的状態である。我々は通常、「今からどんな情報が送られてくるか」についての予備知識や先入観を持っている。それらは送り手に関するものであったり、情報の属性や内容についてであったりと、様々な側面があるが、これが情報の読み取りに大きく関わってくるというのが、「メディア」の場合にはない「人」特有の現象である。

ボトムアップ（データ駆動型処理）とトップダウン（概念駆動型処理）という概念をもちいてくりかえすと、我々「人」が情報の読み取りを行う場合は、ボトムアップがすべてではなく、自分の知識ベースを手がかりに「こちらから予測をつけながら情報を迎えにく」というトップダウンがそれに関わっていると考えられる。これは読み取りの大半をボトムアップに依存する「メディア」とは大きな違いである。

例えば、多義的に解釈可能な図形（ルビンの壺など）でも、人は「情報への構え」のありかたしだいで、無意識的に一つの「読み」を選択し、他の解釈を捨てる（この点については後に詳しく述べる）。また例えば、話を聞くという場合も、風景を見るところの場合も、我々は日常的な経験から、時間的に次にくる「音」や、空間的にその周囲に見えるはずの「形」を事前に予測できるのが普通であり、送られてくる情報を「こちらから迎えにく」というかたちでスムーズに（ある意味では惰性的に）処理することができるのである。いずれの例も（人工知能を除く）機械には無縁の話である。

視聴覚の相互作用

次に「視聴覚の相互作用」についてであるが、マイクやカメラといった「メディアの視聴覚」の場合、音と映像を同時に記録するにせよ、音は音専用の領域に、映像は映像専用の領域に記録される。音と映像は相互に干渉しあうことなく、独立に処理・記録されるのである。ところが「人の視聴覚」では、聴覚中枢と視覚中枢の区別はあるものの、音と映像の独立性は完全ではなく、感覚のレベルでの色聴現象をはじめ、優位なモダリティーへの統合、読み取り支援、干渉による異次元の感覚情報の生成など、様々な相互作用がある。例えば色聴。これは「音を聞くと色が見える」すなわち音程と色相とのあいだに感覚的な結びつきが見られる現象で、カール・ジーツ（1931）の説によると、おもちゃのピアノのように、ドレミ・・・の音階が赤橙黄・・・に結びつくというタイプのものである。

*マガーク効果

マガーク（1976）の実験は、視覚と聴覚の相互作用を説明するものとして有名である。それは、「が」を発音する口の動きを視覚的に見せ、同時に「ば」の音を聴覚的に聞かせると、被験者には「だ」と捉えられるというものである。

「空のマッチ箱を振ると音が聞こえる（袖の中に中身の入ったマッチ箱が隠されている）」という手品も同様で、視覚情報（振る）という動作に聴覚が誘導された（心理的に音源の位置が移動した）ものと説明できる。

また例えばテレビ。音と映像が複合した情報は、時間的には聴覚が優先し、空間的には視覚が優先するかたちで、一方が他方に追従・融合するかたちで捉えられる。例えば、アフレコされた足音と素材映像の足の動きが合っていない場合でも、足音に映像がなじむように違和感なく捉えられるし、また逆に、ニュース解説者の口元（画面上）には実際の音源（スピーカ）がないにも関わらず、音はそこから聴こえているかのように捉えられる。前者は音のリズムの問題で聴覚が優先している例であり、後者は空間的な位置の問題で視覚が優先している例である。いずれにせよバラバラに与えられた音と映像は、自然に一体となって捉えられるのである*。

さらに、遠くで会話する2人を双眼鏡でクローズアップして見ると、声まで聞こえる（正確には聞こえやすくなる）という例がある。視覚からの情報がトップダウン的に聴覚に作用して、話者の音声が増強しやすくなった結果と考えられる。これは、騒がしい状況のなかで、特定の人物の声のみを聞き分けられるという「カクテルパーティー・エフェクト」などと同様に、視覚が聴覚の読み取りをサポートしていると考えられる現象である。

我々は、視覚と聴覚のこのような相互作用を、「あたりまえ」と感じていたり、あるいは気付いていなかったりするが、音楽や映像の制作においては無視できない問題である。

視聴覚の情性化と日常性

我々の世界は、日常的で予測可能な出来事のくりかえしであり、「予期せぬ出来事」と

いうのはその言葉どおり非常に小さな確率でしかおこらない。たとえば、電車を待つホームにセスナ機がすべりこんでくるとか、机の引き出しを開けたら魚が泳いでいるなどということは、まずあり得ないことであって、もし日常がそうした予測のつかない事態の連続であれば、我々のトップダウンは効力を失い、すべてを視聴覚のボトムアップに依存せざるをえなくなる。おそらく人はそのような状況に長くは耐えられないであろう。

我々はこの世界の日常性（予測可能性）ゆえに、情報の読み取りに際して部分的なボトムアップだけで状況を理解し、スムーズに世界に適応することができるのである。

しかし一方で、この日常性は、人の情報の読みを情性化させてしまうものでもある。そのせいで我々は、はじめて何かに出会った時の新鮮な感覚や、物のもつ物質的・具体的な存在感を、意識下に埋没させたまま、その多くを忘れ去ろうとしている。

街のノイズに聞き過ごしたものの、街の路上で見過ごしてしまったもの。埋没した情報をゆさぶりおこすには、すべてをボトムアップ的に吸い上げる「メディアの耳」・「メディアの眼」の登場を待たなくてはならない。

次章から具体的に視聴覚の問題に入ろう。聴覚と視覚、人間とメディア、2つの基軸に沿って、それぞれを比較しながら考察してみたい。

*視聴覚のスレ

音楽番組などで、実際には演奏していない「クチバク」シーンの場合でも、通常は音楽を聴く聴覚が優先して、全体が違和感なく捉えられるのだが、歌っている「顔」や、楽器を演奏する「手元」がアップになると、視覚情報処理の活性が上がり、視聴覚のコンフリクトが生じる。そうすると「合っていない」ということがすぐにバレてしまう。

メディアの聴覚

「機械」は、そこで生じているすべての音波を無差別にとらえる。もちろん周波数に関する上限・下限と音圧の感度の限界はあるが、「人」の聴覚のように、刺激の種類によって態度を変えたいというようなことはない。

1 音響入力装置の基本構造

「メディアの耳」つまり音響情報の入力装置には、マイクロフォンとピックアップがあるが、マイクロフォンは空気中の粗密波としての音をとらえるもの、ピックアップは空気以外の振動をとらえるものである。以下、その構造について概説する。

マイクロフォンはスピーカと同様、基本的には電気音響変成器（可逆変換器）あるいは継電器（不可逆変換器）で、動電形・電磁形・静電形・圧電形・炭素形・半導体形などがある。特に動電形と静電形が一般的で、前者は電磁誘導すなわち、小さなマグネットのつくる磁界中で、電気導体である振動板が音波により振動することで起電力が生じることを利用するもので、後者はコンデンサの充放電電流を利用する、すなわち、導電性の振動板と固定極が向かい合ったコンデンサ様の構造に直流電圧をかけておき、振動に応じた静電

容量の変化から生じる充電・放電電流を得るというものである。それぞれダイナミックマイク・エレクトレットコンデンサマイクなどと呼ばれる。また、かつての電話（黒電話）の送話器や放送用に使用された炭素形のカーボンマイクは、炭素粒の振動接触による抵抗値の変化を利用するもので、変換率が非常によく（ただし不可逆）、イヤホン程度ならば増幅なしで直接音を再生できるという性能をもつ。

ピックアップは、電流の変化・磁界の変化・物体の動きなどを振動として拾う場合に利用されるもので、動電形・電磁形・半導体形などがある。電磁気的な振動は「マグネットの芯のまわりにコイルを巻く」という簡単な仕組で、電磁誘導による起電力を得ることができる、また物体の振動は、圧力をかけると抵抗値が変化するというピエゾ抵抗効果をもつ半導体（不可逆）を利用することで取り出すことができる。

さて、人間の聴覚が左右2つの耳から情報を入力するように、「メディアの聴覚」にも2つ（あるいはそれ以上）のマイクを用いる方法、すなわちステレオ（マルチ）録音の発想がある。セッティングの方法には、レベルや位相の差を考慮して、2個の単一指向性マイクを適当な間隔をあけて配置する「レベル・位相差方式」、2個の単一指向性マイクを一点に配置する「レベル方式」、多数のモノラルマイクの入力をそれぞれ適当に左右のチャンネルにふって、架空のステレオ音場をつくる「分岐方式」の3つの方式がある。2つ並べれば立体的に聴こえるという単純なものでもなく、特にマイクとマイクの距離を開ける（位相差を利用する）場合は、その距離には注意が必要である。

* 位相差

波の1周期を360度として、揺れのタイミングのズレを角度で表したものを位相差と言う。例えば同じ振動数の2つの波が位相差180度で揺れていれば、一方が山の瞬間に他方は谷という関係になる。

位相差180度というのは、ステレオのスピーカの一方をプラス・マイナス逆につないでしまった場合が、それに相当する。この場合、2つのマイクを音源に対して半波長分離してステレオ録音したような、左右が分離した聴こえ方になる。

2 音響入力装置の機能

さて次にその機能の問題であるが、マイクは音響を電気信号に変換する最初の砦であり、その良し悪しは後のすべての処理・記録に影響する。処理・記録系がどんなにすぐれていても、入力段階で十分な素材が得られなければその先に限界を生じてしまう。

マイクの特性を表わす値には、感度・雑音・周波数特性・指向特性・最大無歪み音圧レベル・公称インピーダンス・共振周波数などがある。

感度というのは聴覚で言う最小可聴値とは意味が異なり、 $1\text{kHz} \cdot 1\mu\text{bar}$ の平面波を正面から与えた場合に発生する電圧で、 1V を 0dB とした dB 単位で表示する。通常 80dB から 150dB の間の値をとる。ただし、単に感度がよければよいというものではない。適切な情報の記録には必要外の音を捨てるということも重要だからである。

雑音はそれ自体から発生するものを言い、音圧レベルに換算して表わす。これは当然小さな値であるほうが望ましい。

周波数特性とは、周波数領域における感度分布のことである。一般的には低中音域でバランスのとれた静電形が良く、次に動電形、その他はいわゆる良質な録音は望めないとされる。しかし、全体にバランスが良いということが美的に良い音につながるとは限らない。「機械」であるマイクは、その構成のしかたで様々な帯域に対応するものが実現で

きるわけで、録音する用途によって適切な帯域をもつマイクを選択するということの方がむしろ重要である。例えば、帯域の広い音楽の録音の場合は $40\text{Hz} \sim 15000\text{Hz}$ とほぼ可聴域全域が必要になるが、音声の録音では $100 \sim 8000\text{Hz}$ に重点を置くなど、必要のない帯域をカットすることが S/N 比向上の点からも望ましい。

指向特性とは、音源の方向と感度との関係を表すもので、無指向性・単一指向性・両指向性・ライン・放物面（パラボラ）集音などの分類がある。無指向性は全方向について均等な感度分布をもつもの、単一指向性はマイクを向けた側の約 180 度の範囲について感度が高いもの、といったぐあいで、最も指向性の強いパラボラ集音では前面 20 度以下で、前方の集音距離は無指向性マイクの 3 倍以上となる。

最大無歪み音圧レベルは、出力の高調波歪み率が 1% 以上になるときの入力音圧レベルである。人間の耳では鼓膜を破く危険があつて測定できないが、メディアの場合は「正常に機能する範囲」を示す値として測定できる。ダイナミックレンジという言葉もこれに近いもので「ノイズレベル（下限）から歪みが生じるレベル（上限）までの範囲」を言う。

* インピーダンスとは、いわゆる交流抵抗（電圧／電流）であるが、これは機器間での接続の際に重要な値で、当然整合をとった接続にしなければ信号は回路をうまく流れない。

共振周波数は、物体としてのマイクの固有振動数で、マイク自身を叩いたときに発生する音の周波数と考えるとよい。マイクやピックアップの「個性」を生むものであると同時に、バランスの良さを追及する場合にはやっかいな存在となる。

* インピーダンス

振動は、つなげば伝わるというものではない。例えば、空気と水では抵抗（インピーダンス）の違いがあるため、空気中の振動は、そのままでは水の中には伝わらず、通常は反射されてしまう。音響機器の場合も同様で、各機器の入力と出力に明示された $\text{O} \sim \text{Q}$ の値が一致するようにつなぐ必要がある。

こうして見てくると、マイクは「人」の聴覚とは異なり、ものによって様々な「個性」が存在するため、録音に際しては個々のマイクの特性をいかに適切に生かすかということが重要な問題となる。室内であれば残響の程度、屋外であれば風によるノイズの問題、テレビ・映画のロケではマイクの位置の制限など、状況に応じた録音計画が重要である。

補足になるが、エレクトリックギターのピックアップなどは周波数特性のアンバランスさや歪み具合まで含めて楽器の個性を左右する重要な要素として位置付けられており、音を拾うというより音をつくるための要素と考えたほうがいい。分類上はたしかにピックアップであり、増幅処理を行う機器からみれば、その「聴覚」に相当するのであるが、構造的にもギターの一部として音源に「寄生」している事実から、それは楽器の「声帯」として機能していると考えるべきであろう。「メディアの耳」はその所属をかえることで「声帯」の一部にもなるということである。

3 デジタル化と音声認識

マイクロフォンやピックアップからの入力値は、その処理・記録がデジタルとなる場合はA/D変換が必要となる。俗にサンプラーと呼ばれる音声のデジタル化マシンは音楽の分野から登場したが、現行の大半のデジタル機器はアナログ音声入力端子を持ち、その内部でA/D変換（俗に言うキャプチャー）ができるのが一般的である。

音声のデジタル化では、原音波形をどこまで忠実に再現するか、その必要性に応じて標本化周波数と量子化数と呼ばれるパラメータを決める必要がある。

標本化定理 (sampling theorem) によれば、標本化周波数は、信号に含まれる最高周波数の2倍以上必要である。「人」の聴覚は20,000Hz ≒ 20kHzまでの帯域をもつので、標本化周波数が40kHz以上あれば、耳には十分忠実に再現されているように聴こえる。実際には 44.1kHz・22.05kHz・11.025kHzとったレートのものが一般的に採用されており、順にCD・ラジオ放送・電話のクオリティーと考えるとよい。

量子化数は音の波形の振幅を何段階に分けるかに関わるもので、例えば量子化数16ビットであれば、音の強弱が $2^{16} \equiv 65,536$ 段階で表現される。一般にCDでは16ビット、マルチメディアコンテンツでは16ビットや8ビットが採用されている。

もちろん、さらにモノラルかステレオかという選択が必要で、ステレオであれば当然2倍の情報量となる。

音声がデジタル化されていると、コンピュータを利用することで高度なレベルの処理が可能になる。例えば特定話者方式の音声認識では、まず特定の話者の声で言葉の入力を行い、あらゆる音節についての特徴分布をもとに分類識別の枠組みを形成して、これを知識ベースとして記憶する。識別の段階では、入力された音が、記憶された分類枠組みの中でどの音と強い相関を示すかによって、一つ一つの音節を識別する。もし、「いくつかの命令語を区別するだけ」という問題であれば、言葉を音節単位に区分するという手間が省け

る分、処理は容易になる。いずれにせよ、「メディアの脳」の場合にも「人の脳」と同様に一種の知識のデータベースをもたせてトップダウン的な処理をさせなければ、高度な認知的判断はむずかしい。

4 音響のアナログ記録

今日ある様々なメディアを見れば、音響は画像よりも簡単に扱えるような気がするが、音響の記録は、実は画像の記録としての写真よりも歴史が浅い。蓄音機の発明者エジソン(1847-1931)も、彼が生まれたとき写真はすでに存在していたのである。映画についても始めはサイレント(無声映画)であり、いわゆるトーキーは1930年代から後のことである(トーキー第1作は、1927年アメリカ映画の『ジャズ・シンガー』)。音響は時間軸を基軸とする情報であるため、その記録再生にはリアルタイムで動作できる仕組が必要で、これはアナログにせよデジタルにせよ、高速で動作させるという技術の成熟を待たねばならなかった。余談になるが、歴史上の人物についても、写真はあがるが、声の記録はない。最近の人でも声の記録はめったに存在しない。我々はもともと、音声を記録するということを写真ほどには重視していないようである。

アナログ記録の原理は文字通りアナログカル(類似的)で、要するに空気の振動の様子を、それに似た形で、物理的・化学的・光学的・電磁気的狀態に置き換えて、線的に再現

できればよい。ただ、動きを止めるわけにはいかないので、記録には必然的に回転する機構が必要となり、現に我々が眼にするアナログ記録媒体はすべてそのようにできている。

具体的な媒体としては、(現在では一部の人にしか縁がないが)硬質塩化ビニルに機械的に溝を掘ったアナログレコード、映画のサウンドトラックに見られる光学録音の媒体となるフィルム、そして(これも現在では利用者が少なくなったが)強磁性材を塗ったテープに電気信号を磁気の変化として記録する磁気録音テープなどがある。磁気テープ媒体には、ノーマル・ハイ・メタルなどの一般的なボジション区分で呼ばれる、ガマヘマタイト系・クローム系・メタル系のものがあり、順に高域での特性がよくなる。

5 音響のデジタル記録

標準化・量子化によって数値データとなった音声は、「0・1」に対応する2値状態で媒体に記録される。デジタル音声情報を記録する媒体としては、オーディオCD(Compact Disc)、MD(Mini Disc)・DAT(Digital Audio Tape)とった音楽情報の記録を主目的とした記録媒体と、USBフラッシュメモリーやHD(Hard Disk)など、汎用の補助記憶媒体とがある。特にポータブルオーディオに関しては、PCのデジタルデータの持ち運びにも兼用できる、汎用のリムーバブルメディアの利用が日常的なものとなった。

まず、音楽情報の記録が主目的の前3者について、概要を説明しておこう。

CD:「0・1」に対応する結晶の向きを生成することで、レーザーによる非接触読み取りを実現した媒体で、デジタルのディスクとしては最初のものである(1982年発表)。標準化周波数44.1kHz、量子化16ビットである。

MD:光磁気記録方式のディスクで、MOと同様の媒体である。レーザーによって小磁石を加熱し、磁性変化が可能になったところで「0・1」に対応する向きに磁場をかけるという方法で記録を行い、逆に読み込む場合は、レーザーの反射光の偏光角が小磁石の磁極の向きによって異なることを利用する。

DAT:幅3.81mmの酸化鉄テープやメタルテープを使用する磁気記録方式の媒体で、録音にはいくつかのモードがあるが、標準で標準化周波数48kHz、量子化16ビットで、テープソフトの再生は44.1kHz、16ビットとCDと同様である。

一方、汎用のメディアに記録される音声データについてであるが、非圧縮型のものではAU (UNIX)、WAVE (Microsoft/IBM)、AIFF (Apple)など、また圧縮型のものではMP3 (Fraunhofer/IIS)、WMA (Microsoft)、ATRAC3 (SONY)など、複数の形式がある。クオリティーについては、標準化周波数44.1kHz、量子化16ビットはもちろん用途に応じて数段階の選択肢がある。

さて最後に、アナログとの決定的な違いとなる情報の質の問題に触れよう。デジタル記録媒体における情報のクオリティーは、先に述べた標準化周波数・量子化数の値、および圧縮の方法とその程度によって一義的に決まるもので、記録される媒体の物理特性とは無

関係である。したがって、媒体間での情報の交換や複製(著作権の問題を考慮して転送先や回数に制限がかかるものもある)を何回行っても、情報の質が落ちることはない。この点は、音質が媒体に左右されたり、複製によって確実に音質が劣化したりするアナログ媒体とは大きく異なる。

PCの画面で、音声ファイルを開いてみよう(録音や編集のできるソフトで、wavファイルを開いて、部分を拡大するとわかりやすい)。「メディアの記憶」の中では、サウンドは、ただのギザギザの波として存在していることがわかる。意識や感情といったものとは当然無関係で、また「意味」を担うような分節が与えられることもない。空気の弾性波を電氣的に横波におきかえただけのギザギザである。我々は、言葉や音楽を聴くということに特に難しさを感じていないが、この連続波形から、言葉の要素としての音や、特定の楽器の音、ましてや音楽の「美しさ」などというものを取り出すようなプログラムが記述できるであろうか。

人間の脳は、こんな状態のものを、様々なレベルの意味のあるかたち¹に体制化しながらそれを捉えているのだ。人はなぜ音楽を聴くのか、そのま²えに、そもそもこんなギザギザからなぜ言葉や音楽が聴こえてくるのか。メディアに記録された音の正体を見ると、そんな素朴な疑問も生まれてくる。

メディアの視覚

「メディアの視覚」には、世界を一挙にとらえるもの（例えば35mmカメラ）と、それをさらに水平、走査と、垂直、走査とによって線的な電気信号に変換してとらえるもの（例えばデジタルカメラやスキャナ）とがある。この違いは、要するに像を形成する仕組みが化学的か電気的かで決まるもので、電気信号という時間軸上の情報を利用するのは、おのずと走査変換という仕組みが必要になるのである。

いずれの場合も、像が記録されるまでのプロセスはすべてボトムアップであり、見えるものすべてが無差別に記録されることは言うまでもない。

1 画像入力装置の基本構造

「メディアの目」すなわち「カメラ」の基本構造を、歴史をふまえつつ確認しよう。

カメラの歴史の第1段階は「針穴をあけた暗い部屋」、いわゆるカメラ・オブスキュラである。中世の画家（例えばダヴィンチ）はすでに風景絵画の補助手段としてその原理を利用していただが、暗い部屋の一方の壁に小さな穴を開けると、外の風景が穴の反対側の壁にさかさまに写し出される（倒立像ができる）という現象は、おそらくそのずっと

以前から偶然的に知られていたと思われる。

針穴のかわりに凸レンズをつけると集光面積が大きくなり、出来上がる倒立像も明るくなる。像のできる面をすり硝子にしてトレースすれば風景画ができることから、レンズとすり硝子をもつカメラ・オブスキュラは18世紀ごろから画家のあいだに普及するようになる。これがカメラの歴史の第2段階である。

歴史の第3段階は、1839年、像を自動的に定着させる写真術の発明で幕をあける。この発明によって多くの肖像画家や風景画家が転職を余儀なくされるというほど、それは画期的なできごとであった。そして、後のイーストマン・コダック社のロールフィルム（1888）によって、ほぼ現在のスチールカメラの原型が完成する。

スチールカメラの構造は、一般に外界側から順にレンズ・絞り・シャッター・フィルム面となっており、シャッターの存在を除いては「人」の眼球とほぼ同様である。現在のスチールカメラは、35mm^{*}・APS（約20mm）・中判（6×6・6×9cm）・大判（4×5・8×10インチ）などフィルムのサイズによって分類できる。

光学系の構造をそのままにして、フィルムを一定の速さで断続的に動かせば映画撮影用のカメラとなる。1895年にはリュミエール兄弟によってその上映が行われており、写真から50年後に映画の誕生ということになる。35mm他様々なサイズがあるが、現在ではCFや映画撮影といった業務用のもののみで、一般には馴染みのないものとなってしまった。

スチールカメラの光学系に加えて、CCDなどの光電変換系を用い、光の二次元像を線

的な電気信号に変換して順次転送する仕組みを持たせることで、テレビ・ビデオカメラができあがる。テレビ放送として実用化されるのは1936年であり、ビデオという記録メディアが実用化したのが1956年。つまり記録できるカメラとしては、半世紀の歴史しかもない新しいカメラである。レンズ群による光学像の生成、イメージセンサによる光電変換・蓄積・走査、さらにビデオテープへの記録で電子的信号処理・電磁気的変換・機械的動作。ビデオカメラでは像の記録に至るまでに、多くのプロセスを経由せねばならない。

2 画像入力装置の機能

「カメラ」の機能について、レンズの性能・レンズの焦点距離・絞り・シャッタースピードといった撮影パラメータを中心に考察してみよう。

レンズの性能

カメラという機械にとつて最も重要な光学系を構成するのがレンズである。レンズの材質は光学硝子という良質の硝子であるが、一般的にはクラウンガラス(K)、それに鉛を加えて屈折率を上げたフリントガラス(F)の二つになる。光学硝子は当然無色透明で均質であり、光の透過に関して等方であること、またレンズの設計に必要な光学常数、すなわち精巧な屈折率と分散率をもつことが要求される。実際には、一枚の凸レンズだけでは

光の波長による屈折率の差、いわゆる色収差が避けられないため、複数のレンズを群に構成して単体のレンズに見立てる。

レンズにはFナンバーという数値があり、それらもレンズの性能に関係する。略説すると、Fナンバーは「焦点距離／レンズの有効径」を表わす値（暗さの尺度と言える）で、レンズの口径が大きくなりFナンバーが小さくなるほど「明るいレンズ」ということになる。レンズは口径の大きなものほど分散・収差が大きくなりその分良質のレンズの製造が難しいため、一般的なカメラではFナンバー1:4〜2.8程度のレンズがついている。

「メディアの目」であるカメラの意義が、世界を客観的にボトムアップすることにあるとすれば、世界を細部まであざやかに写取ることのできるレンズはそれだけで価値がある。

しかし、「メディアの耳」であるマイクの場合と同様で、ムラがなくバランスのよい空間周波数特性が得られるということ、美的な映像を写し出すということとは必ずしもイコールではない。技術的な成果が美的要求をすべて満たすものではないということ、ど
*の世界にも言えることなのである。「レンズの味」という言葉もあるように、様々な収差によるボケぐあいというものがレンズの個性であり、我々はそうした「味」を必要に感じ
て選択するという思考法も捨ててはならないのである。

レンズの焦点距離

焦点距離はレンズの主点（後ろ側主点）から焦点面（フィルム面）までの距離のことだが、

*空間周波数特性

一般に、レンズの描写力はMTF (Modulation Transfer Function) と呼ばれるもので評価されるが、これは画像の細かさに対してどの程度までコントラストが正確に再現できるかをグラフにする形で提示される。単に解像力（本数／ミリ）の高さだけでなく、レンズの空間周波数特性の分布が重要だということを意味する。

カメラの機能の問題として焦点距離が重要なのは、この値が画角（具体的には主点から画面の対角線の両端とを結ぶ線のなす角）に直接関わるという点である。例えば35mmフィルムの場合は、サイズ36×24mmで対角線43.2mmであるから、焦点距離50mmで画角46度となる。

焦点距離が短くなれば画角は大きく（ワイドに）なり、長くなれば画角は狭く（望遠に）なる。人間の眼に自然に見える角度がほぼ50度であることから50mmのレンズは標準レンズ、28mmや35mmは広角レンズ、じつと見つめる画角にあたる85mmはポートレートレンズ、135mmや200mmなどは望遠レンズと呼ばれる（ただし、同じ焦点距離何mmと記載されていてもフィルムの大きさやCCDのサイズが変われば画角も変わるという点には注意が必要である。例えば6×6cmのフィルムでは焦点距離80mmが標準画角となるし、35mmよりサイズの小さいCCDを用いるデジタルカメラでは焦点距離が非常に短くても標準画角となる場合がある）。この値が固定的なレンズを単焦点レンズ、この値を一定の範囲で変えられるものをズームレンズと言う。

絞り

絞りはレンズの使用面積つまり明るさを調節する単純な機構であり、ものによる優劣の差がほとんど無い部分である。絞り機構はレンズ群の中間にあつて、複数枚の金属羽根で構成されている。レンズ鏡胴の絞りリングで開閉を調節するが、リング上の目盛はレンズの解放F値から順に公比ルート2の等比数列で並んでいる。すなわち目盛を1段増やすこ

とに有効径が、 $1/\sqrt{2}$ ずつ小さくなる（採光面積が半分になる）ことを意味する。これは主としてフィルムにあたる光の量を適正に調節するためのものであるが、これは人間の目の虹彩と同様、絞れば被写界深度が深くなり前後のピントも合いやすくなるという映像表現上の効果が大きい。

シャッタースピード

シャッターは、世界をとらえる「一瞬」というものにどの程度の時間を与えるかを決める機構であり、その選択可能性が大きなものほどカメラとしての機能は優れていると言える。一般的なスチールカメラでは4秒から1/4000秒までの間を1/2倍間隔で選択できるようになっていて、これは絞りの1段に対応してフィルムにあたる光の量を1/2ずつ調整する目的をもつ。「動くものを止めて写すか、動きを軌跡として写すか」といった人間の目では直接見ることでできない視覚世界の表現に関わるものであり、写真メディアに特有のものであることをあらためて銘記しておきたい。

フレームレート

動く映像をとらえるカメラとなると、1秒あたり何枚の画像を撮影するかということも重要である。ただこの点については、高速度撮影カメラのような特殊なものを除いては、規格として決まっていて、フィルムを使う映画の撮影カメラでは秒間24フレーム（コマ）、

*デジタルカメラの画角

様々なサイズのCCDを使用するデジタルカメラのカタログでは、物理的な焦点距離の記載が画角を説明するものとはならないため、従来のカメラの感覚で理解できるよう「35mmカメラ換算で50mm」などと記載されている。

*使い捨てカメラ

いわゆる「使い捨てカメラ」は、感度100のものを例にとれば、Fナンバー11の25mmレンズで、絞りはF11固定、シャッタースピードも1/100固定というもので、日中晴天の屋外で2〜3mの距離でほぼ適正に写るようになっている。

NTSC信号をベースにしたテレビやビデオなどでは秒間30（正確には29.97）フレームである。人間の目がチラツキを感じずに自然な仮現運動が生じるのがこのあたりだと考えればよい。これより遅いと、動きが飛んだようなギクシャクしたものに見え、逆に早すぎても残像がダブって見えてしまう。

イメージセンサ

フィルムを使うカメラであれば、光学系（レンズ・絞り・シャッター）までが入力システムで、フィルムは記録の問題になるのだが、像を電気信号に変換してから記録する必要があるカメラでは、光電変換系であるイメージセンサまでが入力システムの一部となる。

イメージセンサはテレビ・ビデオカメラ・デジタルスチルカメラ等で光学像を電気信号に変換する部分であり、レンズ同様に最終的な画像の解像度や色調に大きく影響する重要な部分である。撮像管と固体撮像素子とがあるが、現在ではCCDやCMOSに代表される固体撮像素子が大半を占めている。

撮像管は、イメージセンサとしてはテレビの初期から存在した電子管で、画素という区切りが存在しないため、固体撮像素子より解像度の高い高画質のものが実現でき、テレビ放送用のカメラなどは現在でもこれを用いている場合がある。光学系からの情報を電気信号に変換する手法の違いで、ビジコン・サチコン・プランビコン・イメージオルシコン・SEC管などがある。

固体撮像素子は大半のビデオカメラやデジタルスチルカメラが採用している半導体板のセンサで、1次元（ラインセンサ）または2次元（エリアセンサ）に並ぶ光センサの集合体を感光面としてもつ。これは網膜の状態に似ており、各センサが1画素という最小単位に相当するかたちで、光電変換・蓄積・走査の3つの役割を担う。解像度は並んでいる画素の数で決まるので、カメラの性能は「〇〇万画素」というスペックを見れば簡単に区別できる。

撮像管・固体撮像素子ともに関わる問題だが、カラーの映像信号は基本的にRGB3種の信号に分離して取り出すことが必要で、その方法については、1系統のイメージセンサから取り出す方法と、光学的にRGBに分離して3系統のセンサで取り出す方法（一般に3管式・3板式と呼ばれる）とがある。当然3系統もつものが高画質であり、業務用には後者が使われる。

3 デジタル化とパターン認識

光学系あるいはさらに光電変換系によって得られたアナログのフィルムや映像信号をデジタル化して処理・記録する場合は、AD変換系が必要で、フィルムスキャナ・イメージスキャナ・ビデオキャプチャーなどと呼ばれるものがそれを行う。今日ではデジタルスチルカメラやデジタルビデオカメラなど、光学系・光電変換系及びAD変換系を合わせ

持つ機器が一般化しており、アナログからデジタルへというプロセスを意識することが少なくなりましたが、世界を光の像としてとらえる光学系と、それを電氣に変換する光電変換系、そしてさらにそれをデジタル情報へ変換するA/D変換系、この3つのブロック全てが、画像のクオリティーに関わっているということは、常に意識しておくべきであろう。

さて、画像のデジタル化においても音声同様、標本化数と量子化数を決める必要がある。画像の場合、縦横の画素数を決めることが標本化にあたり、この数はA/D変換系の限度内で自由に決めることができるのだが、問題は今述べたように光学系と光電変換系がどの程度のものであるかが重要である。例えば原画像が35mmフィルムによるものであれば、その画質は約1000万画素程度に相当し、また通常のビデオ画像であれば約30万画素程度に相当する。原画像を超える解像度をもたせても特別な目的がないかぎり意味がないため、おのずとそれが標本化数の上限をきめることになる。

次に量子化数についてであるが、これは色数に関わるもので、例えば量子化数8ビットであれば256色、24ビットであれば1677万色となる。RGB各8ビット(256段)で24ビット量子化したものが、一般にフルカラー画像と呼ばれ、我々の視覚の能力はカバーされる(ただし1ビット異なるだけでも、ある程度の面積で並べて見れば差は分かる)。

画像がデジタル化されていると、コンピュータを利用することで高度なレベルの処理が可能になる。古くから実用化されているOCR (Optical Character Reader) 等の数字認識システムに加え、手書きの文字認識も一般化し、さらにテレビカメラの画像から空間

を認識するシステムなどが実用化されつつある。

文字認識では、音声に比べて識別の手がかりが多く、例えばペン入力のものでは、ペンのアップ・ダウンの情報から筆順や画数の情報を得ることができ、特定の筆跡に限らず、ある程度クセのある文字でもほぼ正確に変換されるようになった。もちろん「機械」はあらゆる文字についての分類識別の枠組みを知識ベースとして記憶しており、記憶された分類枠組みのどの文字と相関が強いかをみることで、一つ一つの文字を識別する。また、文字単位の識別では限界があるため、単語や文法レベルの情報(前後の文脈を参考にする)も知識ベースとすることで、文字識別能力を上げる方法も取られている。

カメラによる空間認識に関しても視野内の特徴的な水平・垂直線や、消点へ集約的に向かう斜めの線などについての知識ベースを手がかりにしており、「知覚」「認識」のレベルではトップダウン的な処理が必要不可欠であると言える。

視覚情報の認識マシンが実用化されれば、自動車のオートクルージングや盲導ロボットなどが実現し、また「人」の視覚はあらゆる単純作業から解放されるであろう。

「人」はこれまでに、遠くを見る(望遠鏡:1609)・微小なものを見る(顕微鏡)・「一瞬」を凍結する(写真:1839)・「動き」を記録する(映画:1895)・遠くの出来事を実時間で見る(テレビ:1936)・疑似的な空間に介入する(ゲーム・VR)など、視覚とその周辺領域がかかえていた多くの限界を「メディア」によって拡張してきた。その分だけ視覚はセンサとしての負荷を減らし、その機能を組み換え(例えば、活字世代の線的眼球制御から、

漫画・テレビ世代の面的眼球制御へ）、部分的にはその能力を退化させた（例えば、視力^{*}）のである。そして近い将来、「人」はその得意領域である「知覚・認識」をも「機械」に委ねることで、より新しい視覚を手に入れようとしている。我々はそろそろ「自由になった視覚でなにを見るか」という新しい問題への取り組みを考えなくてはならない。

4 画像のアナログ記録

静止画像の記録に関しては、音の場合のような時間軸の制約がなく、またランダムにアクセスできるものであるため、カメラオプスキュラなどによって「じっくり観察して描き写す」という手作業でのアナログ的記録が可能であった。現に我々は、歴史上の人物の肖像を記録として見ることができるのである。その意味では画像の記録は原始洞窟絵画の時代にまでさかのぼることになるのだが、それはさておき、ここでは1830年代におこる「像の自動的記録」というものを出発点としよう。

銀塩写真

当時、銀の光化学反応に着目してカメラオプスキュラの像を定着させようとしていた化学者や画家の試行錯誤は、1826年ニエプス (J.N.Niepce) のヘリオグラフィ、1835年タルボット (W.H.F.Talbot) のネガポジ法、1839年ダゲール (L.J.M.Daguerre) のダゲ

*視力

視力とは、眼で2つの点を区別しうる能力のことで、「区別できる2点の視角(分)」の逆数で表わされる。例えば、視角で1分に相当する2点間の区別ができるものは、視力「20(標準値)」である。

*ネガポジ法

反転を利用するという、このネガ・ポジ法は「一枚のネガから複数のポジが得られる」という点で、グーテンベルクの活版印刷とならんで、今日の「複製芸術」社会のきっかけとなる重要な発明であった。その意味ではネガ・ポジ法を考案し、世界最初の写真集「自然の鉛筆」(1843)を作ったタルボットの業績は偉大である。

レオタイプ、1841年タルボットのカロタイプ^{*}の発表というかたちで、多くの成果を得た。そして1888年イーストマン (G.Eastman) のロールフィルムの発明以来、現在にいたるまで「銀塩写真」はアナログ記録の主役として重要な位置を占めている。

銀塩写真は、ハロゲン銀 (AgX ：塩化銀・臭化銀など) が光によつて黒変することを利用するものである。ハロゲン銀を塗布した面に光像を結ぶと、光のあたった部分に銀核とよばれる黒点が形成される(潜像)。このままでは像として使える状態ではないが、これをさらにモノメチルパラアミノ・フェノール硫酸塩やハイドロキノンを主薬とする現像液に浸すことで、銀核の周囲にAgが集結(数千万倍に増幅)して目に見える像が得られる。一定時間の現像反応の後、酢酸などによる反応の停止、チオ硫酸ナトリウム(ハイポ)などによる未露光 AgX の除去、すなわち定着処理を終えて、像は永久化される。一般にこの像はネガ像であるため、別の感光面(印画紙)を用意し、そのネガを透過する光を感光させて同様の処理を行うことでポジ像を得る。以上が写真の記録のおおよその原理である。

カラー写真の場合は、カプラーによる発色現像を行う。これは光をBGRの順に分離し、各々の補色にあたるYMCの3層の発色でネガ像を形成するというアイデアで、例えば、被写体の青い光は、フィルム上でB (Blue) に感光する層をY (Yellow) に染めてネガとなり、プリント段階でそれに白色光をあてると、RとGが透過して印画紙上のCとMを染め(すなわち印画紙上では青の色が出て)、ポジとなる。

一般的にはこのような原理で、ネガフィルムにネガ印画紙の組み合わせでプリント写真

とするが、カプラーの発色をポジティブにするというタイプの、リバーサル（ポジ）フィルムもあり、スライド上映用や印刷物を作る場合の写真原稿として利用されている。

フィルムを使った動画の記録（映画）では、一般に35mmロールフィルムを縦に秒間24コマというかたちで撮影するが、この場合の1コマは、スチールカメラが用いる1コマのサイズの半分（ $24 \times 18\text{mm}$ ）でパーフォレーション4穴分にあたる。一般に普及した8mmなどはポジフィルムであったが、オプチカル処理を前提とする映画では、ネガフィルムが用いられる。

さて次に銀塩写真方式の特性の問題に触れよう。ここでは、感度と粒状性・色温度などがその特性に関わる選択肢として問題になる。

フィルムの感度はISO感度というかたちで数値として与えられ、カラーネガでは感度100・400等が普及している。一般にこの感度と粒状性は相反するもので、感度の低いものほど粒子が細かく色の再現性も良くなるという傾向があり、感度をとるかディテールをとるかはユーザの選択によることになる。ただ注意が必要なのは、フィルム感度はカラーフィルムの場合と白黒フィルムの場合では、同じ感度でも多少意味が異なるという点である。カラーの場合はフィルム感度はフィルムそのものの固有の感度で、その値に無関係に現像処理が行われるのに対し、白黒の場合は表示感度は撮影時の推奨感度設定値であり、現像の際は設定した感度値に応じて増感現像・減感現像が行われるのである。例えば白黒で感度400のフィルムを、カメラの感度設定値を1600や3200に設定して（フィルム

に当たる光量は $1/4$ 、 $1/8$ となる）撮影した場合は、それに応じて増感現像（現像時間を長く）することで適正なネガが得られる（カラーの場合、これは基本的には難しい）。

色温度は写真のホワイトバランスに関するもので、どのような光源のもとで撮影するかに応じて選択されるべきものである。「人」の視覚は光源に応じて自動的にホワイトバランスをとっているが、「機械」の視覚には「白」は光源の色温度によって赤みをおびたり青みをおびたりする。太陽光のもとでの撮影を前提としたデイライトタイプが一般的でほぼ6000Kを白色としているが、フィルムメーカーによって±200Kほど前後する。一方白熱電球のもとでの撮影を前提としたタングステンタイプは、ほぼ3000Kを白色としている。色温度値は当然、ホワイトバランス適正にたもつたものであるが、「朝の青白い雰囲気を出すために、太陽光のもとでタングステンタイプを使う」など、表現上の理由で逆の選択がなされる場合も多い。

以上は主としてフィルムの問題であるが、印画紙に関して補足すると、表面の質感について光沢・半光沢などの区別があり、またコントラストの強さの区別として硬調・軟調などがある。

磁気テープ

電気的な映像信号によるテレビジョンの場合は電磁気的な記録方式をとる。音声に比べるところその周波数が4.25MHzと大きいために、より高速に読み書きする機構が必要で、は

じめてのものはビング・クロスビー研究所の固定ヘッド型白黒で1951年、そしてテレビ局で最初に実用化されたのはアンペックス社の回転4ヘッド2インチ機で、テレビ放送よりおかれて1956年のことであった。家庭用のものでは1964年ソニーのオープンリール、1975・1976年のカセット式1/2インチのベータとVHS方式、1985年ソニーの8mm方式、そして1996年デジタルビデオと次々に発展して現在に至っている。

媒体としては、1/2インチのビデオテープでは一般にコバルト添加ガンマ酸化鉄の磁性体をポリエステルフィルムベース上に塗布したもの、また8mmやデジタルではメタル磁性体のものが用いられる。記録構造はそれぞれ的方式で若干ずつ異なるが、基本的にはテープの走行方向に対してヘッドを斜めに回転させ、1フレーム分の映像信号を2つのトラック（2フィールド分という意味で、テープ上では斜めの線分奇跡2本になる）に記録する、いわゆるヘリカル走査方式がとられる（NTSC対応のビデオではヘッドは毎秒30回転している）。

さて、磁気テープの特性の問題であるが、磁気方式ではヘッドとテープの相対速度・周波数特性が関わる水平解像度、ビデオS/N（輝度）、カラーS/N（色相・彩度）などが問題となる。水平解像度は、一画面で何本の縦線を再現できるかを数値で表わすもので、画質競走のもともわかりやすい指標である。垂直方向に関しては走査線の数以上の分解能は物理的に無理であるが、水平方向については信号（特に輝度信号）の高域が忠実に再現できれば解像度は上がる。しかしどちらかというとこれはディスプレイの特性に依

存することが多い。ビデオS/Nは45dB以上あれば十分と言われるが、これが低いと画面がざらつき、輪郭が不鮮明になる。対策としては磁性体の超微粒子化や表面性の向上があげられる。カラーS/Nは、色相がズレするというPMノイズと、彩度すなわち色の飽和度に関するAMノイズの問題があるが、特にビデオテープに関しては後者の問題が大きい。NTSCでは輝度（Y）と色度（IQ）は分離され、色信号IQは3.58MHzの色副搬送波にのっている。色彩が鮮やかに再現されるためには、やはりテープがバランスの良い周波数特性をもち、全体で4.25MHzという広帯域の映像信号に十分追従できる能力が必要となる。

5 画像のデジタル記録

標本化・量子化そして何らかの方式で圧縮された静止画・動画のデジタルデータは、「0・1」に対応する2値状態で媒体に記録される。したがって、記録媒体は文字や音声・音楽をデジタル記録する媒体と同一で、従来の磁気記録方式を用いたテープ、レーザーによって「0・1」に対応する結晶の向きを読み書きするディスク、レーザによる加熱と磁性変化を利用したディスク、そして半導体を利用した固体のメモリーなどが使用できる。

デジタルビデオカメラが採用しているDVテープやDVD、デジタルスチルカメラ等が採用しているメモリーカード、そしてPCを経由すれば、あらゆる汎用の補助記憶媒体が静止画・動画の記録メディアとなる。

アナログメディアの記録の場合は、媒体の品質特性が直接記録される情報の質に関わるが、デジタルの場合は0と1がきちんと区別されて、安定して記録されるかということが重要で、それさえ正しく識別できれば、情報の質には影響しない。そのクオリティーは標本化周波数・量子化数の値および圧縮レートによって一義的に決まるもので、媒体の物理特性は無関係なのである。

さて、記録されるデータの形式についてであるが、現在デジタルで静止画・動画を扱う場合の形式はPC上で読み書きできるものを含めるとかなりの数になる。デジタルビデオが採用しているDV形式、DVDが採用しているMPEG形式、デジタルカメラが採用しているJPEG形式が代表的なものになるが、PC上で各種ソフトウェアを利用して扱うことができる形式となると、一般的なもののだけでも静止画ではBMP・GIF・PNG・EPS・TIFF、動画ではAVI・QuickTimeなどがあり、これに各種のソフトウェア特有の形式が加わると膨大な数になる。したがってPCを経由してデータ交換する際は、双方の環境で正しくファイルの読み書きができるかといった確認が必要になる。特に圧縮の方式を把握しておくことは重要で、例えば同じAVI形式（ファイル名が（avi）でも圧縮方式の違いで相手側では再生できないといったことも生じる。最近では大半のソフトウェアがあらゆる方式をカバーできるようになってはいるが、この種の問題はデジタルデータを扱う場合に避けられない問題であり、ある意味では常識と考えて接した方がよいであろう。

静止画・動画に関連するメディアの充実、特にデジタルカメラと携帯電話が一体化することで、カメラを日常的に携帯できるようになったことは、非常に画期的なことである。撮り直しは何度でもできるし、現像代もいらぬ、PCで管理すれば検索にも時間がかからない。面白いものを見かけたとき、資料が必要なとき、時刻表をメモする必要があるとき、一瞬で記録が済むことは非常に便利なことである。

しかし一方で、我々は、自分の目で見て記憶することや、ペンを持つて（手を動かして）記憶するということをおろそかにしはじめたのではないだろうか。これは、必ずしも喜ばしいこととは言えない。とくに創造的な精神活動にとつては、過去の記憶や様々な感情を呼びさましながらゆっくりと見る事、ペンを走らせながら頭の中を整理すること、そういった時間のかけかたが必要なものも多いはずである。

メディアの視覚は純粋なボトムアップでそこにあるものを正確に記録する、一方、人の視覚には、様々なトップダウンが関与していて、見ているその場で情報の「編集」が行われていく。見ることにしても記憶することにおいても、メディアと人間は、その特性を生かすかたちで共存すべきであろう。

人間の聴覚

聴覚が捉える音は、自然環境のものと社会環境のものとに分類できると言われる。前者は自然界や異種の生物を情報源とする音で、障害物や餌、外敵に関わるものであり、後者は同種の生物同士でのコミュニケーションに関わる音である。もちろんここでの関心事は、コミュニケーションに関わる音の方であり、聴覚に関する心理学的基礎をふまえながら、人間が音声や音楽といった情報をどのように処理しているのかを考察してみたい。

1 受容器

「人」の聴覚の受容器は、外耳・中耳・内耳の3つで構成されるが、外耳と中耳は音響振動の伝達を、内耳は感覚細胞をもち、刺激による興奮を起こす。

まず外耳であるが、これは耳介・耳殻・外耳道からなり、外耳道の終端には鼓膜があった。これが空気中の粗密波を最初に体内へ伝える役割を担う。耳介が3kHz付近、耳殻が5kHz付近、外耳道が25kHz付近にそれぞれ共振点をもち、これと頭全体や胸・肩等の共振もわずかに関与して、鼓膜の付近では2〜6kHzの帯域で約10dBほど音圧が上がっているという。この帯域はすなわち人の音声の、音素識別や話者識別にもっとも重要な領域であり、これは「同種間のコミュニケーションに対する優位性」を物語っている。

中耳は鼓膜の振動を内耳の前庭窓に伝える役割を担う。その中空を鼓室というが、そこには3つの耳小骨があつて、空気の「大きな振幅・小さな圧」を体内の「小さな振幅・大きな圧」へ変換する「てこ」の役割を担っている。すなわちここでは、空気と水のインピーダンス整合が行われている（ちなみに、水中に生活する生物の場合は「水の振動」→「体（水と同じ）の振動」で音が伝わるため、この仕組みは不要である。したがって魚やイルカなどでは、内耳が直接頭部に埋ったようなかたちになっている）。

内耳は三半規管・前庭・蝸牛で構成されるが、聴覚に係るのは蝸牛で、その螺旋状の管を2分する基底膜と呼ばれる膜の上に音の感覚細胞が配列されている。直線的な配列である。ベケシー（1943）の観察によれば、鼓膜から伝達されてくる振動は蝸牛内で進行波をなし、この進行波は周波数の高い音では蝸牛の入り口に近いところ、低い音では奥のほうでその振幅が最大になるという。この基底膜のふるまいによって、我々は入力された音の高低を感じ取っているということになる。

さて、その基底膜上には音の感覚受容器であるコルチ器官がのっているが、それを構成する有毛細胞は、ゆれによって電位を発生し、それが聴神経（第1次ニューロン）に伝達され、あとは2次、3次とシナプスを介して中継されて4次で大脳皮質の聴覚領に至る。1次ニューロンは基底膜上のいずれかの位置の興奮を伝達しており、その位置によって、各ニューロンは特徴周波数（最大感度を示す周波数）をもつことになるが、その上位では

*耳介

耳介は正面向きに優位な形状をしており、そのために、音の聴こえ方（周波数特性）は方向によって変化する。人間はこれを利用して、音が前方から聴こえているか後方から聴こえているかを知ることができる。ちなみに、左右に関しては、両耳が捉える音量の差や位相差がヒントになっている。

多数の入力に対して1つの出力というかたちで神経回路網をなしていく。興奮は単純に聴覚領へ上向するだけでなく、遠心性（上から下ってくる）のニューロンのフィードバックが干渉するなどして、複雑な機能が実現されていることになり、この回路網全体でおこっている興奮の空間的・時間的パターンが、我々の音の聴こえ方を決定づけている。

最終的には、大脳皮質の側頭葉にある聴覚領が情報処理をおこなっており、ここは音の感覚という単純なレベルの問題ではなく、空間や時間の知覚に関わる高度な情報処理機能を実現している。

2 可聴周波数と最小可聴限

「人」の音波に対する可聴周波数範囲は、ほぼ20Hzから20,000Hzで、音楽の話で言うと（1000 Hz¹⁰ Hzであるから）10オクターブに相当する。内耳の特性のみでは100kHz程度まで感じるらしいが、伝達系である中耳の限界で上限がきまつてくる。これは当然、構造・サイズの異なる聴覚器をもつ生物では異なるもので、例えば犬は上限が50kHzとなる（調教用の犬笛はこの帯域を利用しており、人間には聴こえない）。

音の強さの範囲については、音圧では $2 \times 10^{-4} \mu\text{bar}$ から $2 \times 10^2 \mu\text{bar}$ と約百万倍のレンジをもっている。もちろん最大限界の方は、これ以上では鼓膜が破れるという意味の数値である。最小値の方は、正常聴力者での最も感度の高い3kHz付近での音圧である。

一般に音の強さをいふときは SPL (Sound Pressure Level: 音圧レベル) が単位として用いられ、これは $20 \log \left(\frac{\text{実効音圧}}{\text{基準音圧}} \right)$ / (基準音圧: $2 \times 10^{-4} \mu\text{bar}$) dB という式で与えられる。ちなみに、百万倍は120dBにあたる。対数が用いられるのは、人の聴覚のダイナミックレンジが百万倍もあること、また一般に感覚が物理量の対数に比例することが主な理由で、メータ等はこのdB表示を採用することで感覚的に理解しやすくなっている。

さて、ここで音の高さ（周波数）や強さ（音圧）について、その差をどの程度まで細かく区別できるか（弁別能力）という点にも触れておこう。この分解能を示す値を弁別閾 (difference threshold) というが、これには一般に「弁別閾の値は刺激のレベルに比例する（弁別閾／刺激量＝定数）」というウェーバーの法則があてはまる。ここで定数とは、感覚の種類によって異なるものでウェーバー比とも呼ばれる。この法則は刺激のレベルがある程度以上大きくないと成立しないが、周波数、強さ、音の持続時間などあらゆる感覚刺激についてあてはまる。

3 音程・音色の心理

視覚心理の分野では色彩がもたらす様々な心理的効果が知られているが、聴覚の領域でそれに相当する音程や音色についても同様の心理的効果が考えられるのだろうか。空間を基軸とする色彩と時間を基軸とする音とは何を何に対応づけるかということ自体が難し

*10オクターブ

ピアノは88の鍵盤を持つが、これは、 $88 \div 12 \approx 7$ オクターブ強である。「鍵盤を増やせば表現力があがる」ようにも思えるが、それもと2オクターブ程度で、それ以上は意味が無いことになる。

*超音波を利用する生物

コウモリやイルカなどの生物では、自らの発する超音波の反射音の時間遅れパターンを情報として利用しており、かなりの精度で空間把握ができると言う。例えばコウモリは1mmの針金を2mの距離で認知できるといふ。

*マスキング現象

妨害音によって、最小可聴値が上昇する（ようにするに聴こえなくなる）現象をマスキングという。一般に「妨害音のレベルが上がればマスキング量は増える、マスキング量は妨害音の周波数に接近するほど大きくなる。低音は高音をもマスクするが、高音は低音をマスクしない」などのことがわかつている。

い問題なのだが、可能な範囲で考察してみたい。

音嗜好：色彩の場合は、波長の違いである色相に対して各々の文化的背景も絡めた嗜好が見られるが、音の場合は波長の違いとしての音程（大半の人には相対的なもの）には嗜好があるとは考えにくい。つまり「私はレの音よりソの音が好き」とか「高い音が好き」とかいった発言は稀にしか聞かない。それよりも音色（これは倍音の構成比の問題で視覚の領域では「色の濁りぐあい・色調」とでも言うべきものだろうが）のほうが明確で、「ピアノの音が好き」とか「歪んだギターの音が好き」とかいった発言は聞かれやすい。しかし、これは個人的なものであったり、その時の状況に応じて変化したりするもので、一般的な傾向は見られにくい。あえて言えば「爪で黒板を引っ掻く音」のような高周波を多く含む非楽音については、大半の人が「不快である」という点で一致しているぐらいである。

感情効果：音程や音色が感情と結びつくかという点、これも嗜好の問題と同様で、音程よりも音色のほうがその関係は緊密である。「ファットな音／やせた音」「メロウな音／ブライトな音」などいずれも音色に関する言葉で、倍音の含まれかたや時間経過パターンが聴く者の感情と無関係ではないことがわかる。音楽情報の場合は、リズム・メロディー・和声が強く感情と結びついており、したがって楽曲の主旋律を行く楽器（音色）が適切に選択された場合、その伝えは最も効果的になる。

誘耳性：色彩の心理では誘目性という言葉があるが、誘耳性という言葉はない。ただ、明らかに「全体の中で目立つ音」というのは存在する。例えば声に関して言えば、大衆の

中で「よく通る声」や「マイクにのりやすい声」が存在するし、音楽の演奏の場合にも、主旋律を演奏するのに向いている楽器というのがある、バイオリン・ギター・サキソフォンなど、いずれも「高い音程域で豊かな倍音を含む」という特徴をもっている。

音程の調和：これは視覚という色調和の問題で、一般には耳に心地よく響く音程の組み合わせ、すなわち「和声」の問題になる。和声というものは、実は物理的に簡単に説明できるもので、例えば最もよく用いられる3度の和声は、その周波数比が4対5、あるいはトライアド（3和音）の機軸である5度の音程は周波数比2対3というように、周波数が簡単な整数比で表わされる複数の音は馴染みやすい（ただし現在主流の平均律ではこの比は完全ではない）。これは例えばド（1度）の3倍音がソ（5度）の2倍音と等しいというふうに、それぞれの音の整数次倍音が共鳴するためである。このような物理的整合性の理由から、主要な3和音や4和音は、原則として3度間隔の積み重ねで成立し、逆にそれ以外の音程の組み合わせを用いると、不協和で緊張感のある和音となる。

4 聴覚野

視覚には視野というものが存在し、自分の位置を中心とした世界の一部の情報が読み取られているが、聴覚の場合、基本的にはすべての方向の音が耳にはいつており、その点が視覚と大きく異なっている。耳は開閉しないのであって、常に（時間的にも）あらゆる方

向からの（空間的にも）情報に対して「待ち受け」の状態にある。「呼びかけ」や「警告」、「目覚まし」が視覚ではなく聴覚に訴えるのはそのためである。

もちろん、外耳の構造から言っても後方よりは前方に対して感度が高くなることは言うまでもなく、また耳が2つあることから音源の方向をある程度特定する能力があるわけで、我々の聴覚は完全なる無指向性ではない。

5 音声の識別

音の種類について、それらを区別し言葉で言い当てる能力を識別能力というが、「人」の聴覚には絶対音の識別、音声の識別、話者の識別、和音・音程の識別など、様々な能力がある。我々は通常こうした能力をあたりまえのものと感じているが、その大半が生後の学習により修得されるものであることを考えると、開発の可能性のある能力として、非常に興味ある問題を提供してくれる。

まず絶対音感であるが、これはある音を単独で聴いてその音名を言い当てるという能力で、これは周波数分析器などの機械にとっては簡単なものであるが、「人」の聴覚の場合には、臨界期の間（8歳ごろまで）に適切な訓練を受けた者のみが持ちうる能力である。音楽に携わる者にとっては有利な面も多いが、逆に「どんな音を聴いてもいちいち音名が浮かんでしまう、中途半端な音程だと不快に感じてしまう、音楽が純粹に楽しめない」と

いったマイナスの面もあり、その能力を修得させるべきかどうかについては音楽関係者の間でも賛否両論ある。

音声の識別能力は、我々が皆持っているもので、「ア・イ・ウ・」という音節を聴き分ける能力、言語によるコミュニケーションを成立させる基盤となる能力である。これもやはり後天的な（臨界期はもちろんある）もので、日本で生まれ育った者には日本語の音節（約百種）を聴き分ける能力、英語圏で生まれ育てば英語の音節（3千種以上）を聴き分ける能力が身についている（脳の聴覚領に組織化されている）。音声識別の手がかりとなるのは、音節の短い時間内でのスペクトルパターンとその時間経過パターンという2つの物理的パラメータで、例えば母音を特徴づけるホルマントというものも、すでに述べた通り、音を構成する周波数成分の問題である。

さらに言えば、この場合最も重要なのは「ア」の物理パラメータというより、「ア」と「イ」の物理パラメータの「関係」である。我々は「ア」の音だけを聴いて「ア」の識別能力を修得したのではなく、日本語のあらゆる音節を聴くなかで「ア」の立場を覚えたのである。したがって機械に音声識別をさせる場合も、単に入力音に関する物理パラメータのボトムアップのみでなく、音節のデータベースからのトップダウンを用いることで、識別はよりスムーズなものになる。

次に話者の識別であるが、これは声の質に関する識別で、子供の声・大人の声・男声・女声・誰々の声などと聴き分ける能力である。楽器の種類を識別する能力も同様で、やは

* 臨界期

脳神経回路の基本的な組織化が完了するまでの重要な時期を臨界期という。人の脳は、4・5歳ごろまでにニューロンを最大に増やし、あとは必要な部分（脳が刺激を受けて組織化を行った部分）だけを残して捨てていく、という進化的プロセスを踏む。したがって、臨界期（脳機能ごとに異なるが、ほぼ8歳ごろまで）に刺激を受けなかった、あるいは訓練されなかった事柄に関しては、後からその処理能力を身につけようとしても間に合わないが、非常に努力を要することになる。幼少期の教育が大切なのもそのためである。

しかし、英才教育を受けなかったからといってあきらめてはいけない。人は臨界期以前に体験したことを（忘れてはいるが）それなりに身につけている。眠っている才能に気づくためにも、広い視野で自分の可能性を試してみるべきであろう。

筆者は、長く18歳以上の教育に携わっており、「自分の探し方」次第で、学校の成績とは関係なく才能が発揮されることを知っている。そして思うに、「学力」とは「答える力」ではなく「問う力」のことである。

* ホルマント

日本語の5母音のホルマント、すなわち音を特徴付ける周波数成分は以下のとおりである。

イ	300	2000
エ	500	1700
ア	800	1200
オ	500	800
ウ	300	1000（単位Hz）

り学習によって修得される能力である。

和音や音程などの識別は、音楽的な訓練をある程度受ければ比較的簡単に持つことのできる能力である。これは絶対音感のように周波数に相当する物理的絶対値を言い当てるものではなく、あくまで音の高さの相対的な関係を問題とする。その修得訓練は臨界期以降でも遅くはなく、また能力の修得がマイナスに作用することもない。

6 言葉・メロディーの認知

音を識別するというレベルの問題に加えて、その情報内容を読み取って記憶するというレベルの問題にも簡単に触れておこう。

まず、言葉の認知ついてであるが、それも知覚の問題である以上、重要なのは言語要素間の関係・構造である。言葉を介したコミュニケーションでは、同じ内容を「話し手を変えて・大きな・高い声で・ゆっくり話す」というように変化させても、伝わる内容には（美的な側面を除けば）変化はない。我々が「言葉」を思い浮かべるという場合も、確かに音が付随するが、その際それが「常に誰某の声で思い浮かぶ」というような具体性はないこれは視覚の場合で言うと、読んだ内容は覚えているが文字がどんな書体であつたかは覚えていないということと同じである。言葉は美的な表現の場合を除いて、常に音そのものの具体的な側面を欠く音列のパターンとして読み取られているのである。

* 錯語

視覚には錯覚という現象があるが、聴覚の場合は錯語ということが起こる。これは音節の識別に混乱が生じた場合と、聴き手のトップダウンの作用で日頃使い慣れた音節列として聴いてしまう場合があるが、一般的に「耳慣れない単語」というものは何度も聴き返さないと正確に音節を拾いきれないものである。

メロディーの場合も同様に構造をもったものであり、我々のメロディー認知にはその私たち（ゲシュタルト）が重要である。すなわち、音の刺激情報について重要なのは、各音の音程関係とその時系列パターンであり、音の絶対的周波数ではない。したがって言葉の場合同様、メロディーの再生の際に「楽器を変えて・音量を変えて・移調して・テンポを変えて」という操作をしても、我々はそのメロディーを同一のものとして認知することができる。アレンジに大幅な変更を加えても、その中から主旋律を聴き分け、メロディーを認知し、曲名を言い当てることのできるというのは、「人」の聴覚の能力の強みである。「メディアの聴覚」には「曲名当て」は難しい。

7 聴覚のフレームオブリアレンス

フレームオブリアレンス(frame of reference)という言葉がある。日本語では「参照枠」あるいは「関係づけの枠」などと訳されるが、これは我々の視聴覚情報の認知を理解する上で重要な概念である。すでに述べたとおり、人は言葉にせよ音楽にせよ、その音そのものを直接処理するというより、それらにある観点から自分の知識ベースへ位置づけるかたちで認知する。その際に有効になっている意識の構えなり知識ベースの構造なりが、フレームオブリアレンスに相当する。聴覚に関するフレームオブリアレンスは、物音・言葉・音楽とあらゆる情報の認知において機能するもので、逆にこれが十分に機能しないような

状況では、我々の聴覚は情報の読み取りに失敗したり、あるいはストレスを感じたりする。まず物音の場合は、人の頭の中の言葉の辞書がフレームオブリアレンスとなる。これはいわゆる「カテゴリー態度」という人特有の態度すなわち、すべての刺激情報に対してそれをカテゴライズし言語的に了解しようとする態度に由来するもので、耳に入る物音は無意識のうちに「○○の音」と了解されている。逆に名付け得ぬ音に対しては、「何の音?」「誰の声?」と緊張が生じ、通常とは異なるボトムアップに集中する聴覚態度が生じることになる。

次に言葉に関して言えば、頭の中の「単語辞書と文法構造」や「聴き慣れた表現構造」というものがフレームオブリアレンスとなる。「単語辞書と文法構造」は、人にとつてはあらゆる情報処理の基底に位置するものであり、そこに位置付けられない言葉は一般的なコミュニケーションを成立させない(それはいわゆる「詩的な表現」の部類に入る)。また、「聴き慣れた表現構造」ということに関して言えば、例えば「日常会話での言葉使い」や「小説の文体」といったものがそれにあたる。初めて読む作家の小説に、はじめのうちは読み辛さを感じるのもそのためである。ほとんどの場合、それを読み続けていると作家の言い回しや文章運びのテンポのようなものが、読む人の認知的なフレームを形成し、以後それが参照枠となることでスムーズな読み取りが可能になる。慣れ親しんだフレームが有効に機能すれば、我々は情報の読み取りを失敗なくスムーズに行うことができるのである。補足的に、聴き慣れない表現構造が、情報の読み取り能力を下げる例を示そう。

「ジュウ、キュウ、ハチ、シチ、ナナ、ロク、ゴ、ヨン、サン、ニイ、イチ」。指を折って数えると11ある。数を降順に数えることに慣れていない我々の聴覚は、この間違い(7がだぶっている)に気付きにくい。

さて、最後に音楽についてであるが、音楽の場合、曲の鳴りはじめのフレーズから了解されるその曲のキーとスケール(音階のセット)が、その曲を聴いている間のフレームオブリアレンスとなる。ただしこの場合、キーやスケールは言語的にハ長調とかイ短調などと了解されるわけではない(もちろん絶対音感のある人は別であるが)。そして一般にはキーはその音楽が鳴っている間は有効であるが、曲が鳴り止むと消滅する性質のもので、例えば数分後に同じ曲を口ずさもうとしても、メロディー(ゲシュタルト)は思い出せてもキーは正確ではなくなる。またスケールに関して言えば、我々は一般的な西洋音階に慣れているため、西洋音楽に関しては曲を聴いている間は長調や短調といったフレームが有効になって、音楽の聴き取りはスムーズになり、音の間違いに気付いたり、次にくるフレーズがある程度予測できたりするが、聴き慣れない音階を用いた曲の場合は耳が慣れる(頭の中にスケールのフレームオブリアレンスができる)までは聴き取りがうまくいかず、例えば音はなくても気付きにくい。とくに現代音楽などでクロマチックスケール(すべて半音間隔で12種類の音を使用するものでカテゴリー数が7±2を超えている)などが使用されると、そのメロディーはなかなか了解できないものとなる。当然記憶もできない。このように、フレームオブリアレンスとは我々の音の聴き取りに際して重要な役割を

果たしているのであるが、言語のように長い時間をかけて成立したものと、音楽のキー・スケールのように短時間で成立して一時的に機能するものもあり、そのふるまいは音情報の設計にとって忘れてはならないものであることを確認されたい。

8 聴覚の人工現実

人工現実是一般に2つに分類され、一つは仮想現実 (Virtual Reality) 、もうひとつは遠隔臨場感覚あるいは遠隔現実 (Tele Reality) と呼ばれる。

仮想現実とは、アミューズメントや芸術の領域で話題となるもので、今ここに実際には存在しない架空の世界を出現させる技術が必要になる。音の仮想現実ではスピーカが主たる情報発生源になるのだが、視覚情報の場合のディスプレイとは異なり、今聴いている音が本当の音か、スピーカから出ている音かについての物理的な差異がほとんどない点で、それは比較的簡単に実現される。よく観光地などで、スピーカからうぐいすの鳴き声が聴こえてくる場合があるが、本物の声とだまされるケースも多い。これも音源と情報源が同じ位置にあるという聴覚情報の性質によるもので、視覚情報の場合には不可能な話である。

遠隔現実とは、ロボットを媒介として別の空間内へ臨場する技術を用いるものだが、具体的には人間が直接入れないような場所や危険な作業現場などにロボットを派遣し、あたかも自分がロボットになったかのごとく遠隔地からそれを操作するという場合がこれにあ

る。音の遠隔現実ではヘッドホンが主たる音源となるが、この場合はロボットの顔や耳にあたる部分の音響特性を人間のそれに近似させ、両耳に相当する位置にマイクを取り付ければ (バイノーラルヘッド)、十分な臨場感が得られることになる。ただしヘッドホンだけでは現実存在する「体を感じる振動」が再現されないため、その点の違和感が残る。

一般に聴覚世界の人工的構築は、視覚世界のそれと比べるとはるかに簡単で、その臨場感も現実と区別できないまでに近似させることが可能である。サウンドスケープという言葉があるが、それは生の音に限らず、スピーカだけでも十分に音の景観を構築できるのである。

すべては、音源と情報源が同じところに位置し、生成されるものが物理的に同一の弾性波であるという聴覚情報の特性に由来する。視覚情報の場合でも、光源そのものが情報源 (例えば星座や UFO) となっていれば、「スピーカのうぐいす」と同じように人を欺くことも可能であろう。

* 気導音と骨導音

空気の振動として伝わる音を気導音、物体の振動として直接骨を伝わる音を骨導音と言う。インナーイヤ型のヘッドホンなどでは、低域が骨導音として伝わるために、小さな幕面にも関わらず、低域から高域まで、バランス良く伝わる。ちなみに、蛇は骨導音で周囲の音をとらえている。

人間の視覚

我々は、世界の認知の大半を「視覚」に依存している。自然環境という意味での外部世界はもちろん、人間社会における大半の情報も、文字・図形・静止画像・動画像という形式で視覚に訴えてくる。さらに言えば、「人」特有の「シンボル化能力（イメージ喚起能力）」というものも、「不在の現前」すなわち、目の前に無いものを頭の中にイメージ（視覚化）することであり、頭の中の「視覚像」が世界の認知に果たす役割は非常に大きいと言える。人間の「視覚」について、あるいは「イメージ」について考えるということは、「人」そのものについて哲学することでもある。

1 受容器

生物の体には感光細胞というものがある。文字通り、光を感じる細胞のことで、下等な生物の場合では体表にそれが分布しているが、高等になるにしたがって感光細胞の集合は窪みの中に入り込む（例えばオウム貝の眼はこの段階のもので、ピンホールカメラと同様の仕組みで視覚像を形成している）。さらにそれが進化すると、その窪みへ入る光量の調節機能をもったり、透明な膜によって覆われたりしてくる。人間の場合では、窪みの中の

感光細胞の集合が網膜であり、光量の調節するのが虹彩であり、透明な膜が角膜である。また水晶体は角膜の派生として、硝子体は窪み内部の空間を外圧から護るために発生したと考えられている。ちなみに「目」は発生学的には脳の一部分である。以下、虹彩・水晶体・網膜について、その機能を確認しておこう。

虹彩は外部からの光量を調節する機構で、強い光によって感光細胞が破壊されないようにするのが本来の目的であるが、カメラの絞りと同様に被写界深度の調節機能を副産物として与えてくれる。すなわち、虹彩が大きく（瞳孔が大きく）なればレンズの使用面積が大きくなるため、ピントの合う範囲が狭くなり（対象の前後がボケる）、逆に小さく絞られるとピントの合う範囲は広くなる（対象の前後もくっきりと見える）。曇り空の下ではぼんやり見える風景が、明るい日差しの下ではすっきり見えるのというも、日差しによるコントラストの問題だけではないのである*。

水晶体はカメラのレンズに相当するもので、それを支える毛様体筋の弛緩・収縮によって厚みを変えることで、焦点距離を調節し、網膜上にピントを合わせる。通常のレンズの理屈と同様で、遠方を見ている場合は薄く、近くを見る場合は厚くなる。

網膜はカメラというフィルム面に相当し、感光素子の2次元的な配列で像をとらえるようにできている。網膜に倒立像が写っていることを最初に考えたのはケプラー（1604）と言われており、その事実はシャイナー（1625）によって（牛の眼球で）確認されている。網膜上の各感光細胞は、それぞれに入ってくる光の量や波長に応じて化学物質を放出

*不在の現前

生後まもない乳児は、大人の指さし動作に対して「指先そのものを見る」という反応をするが、言葉の獲得とほぼ平行して「指が指し示すものを見る」というふうに変わってくる。また「母親の不在に気づいたとたんに泣き出す」というのもほぼ同時期である。

指が指し示す方向に何かが「あるであろう」こと、さっきまでそこにあったものが「ない」ということ、このいずれもが、不在のものをイメージするという能力を必要とする。「ない」ことがわかるには「あった」ことがイメージできなければならぬのである。これは「目の前にあるものがすべて」である動物にはない能力であり、人間が予見と計画によって世界を切り開いていくきっかけとなった最も基本的な能力だと言える。

*不在の現前 補足

もちろん、「目の前にあるものがすべて」という動物（真猿類以外）でも餌を隠せば探す。しかしそれは「餌があった」ことをイメージしての行動というより「においがするのになにがない」という目の前の疑問に反応している場合が多い。人間との比較において身近な動物の行動を観察する際は、動物が我々の予想以上に嗅覚情報に依存しているということに気をつけるべきである。

*絞りとシャープネス

眼鏡をかけている読者であれば、それをはずして時計のベルトの穴越しに風景を見てみるとよい。絞りこめば（像は相対的に暗くなるが）ピントが合いやすくなるという事実が確認できるのであろう。

ピンホールカメラのようにレンズを使わず結像するものは、原理的にはボケとは無縁であり、すべての距離にピントが合う。

し、それが視神系の細胞へ伝わる。このとき、網膜の中心付近では感光細胞と視神経細胞の連結が1対1、周辺部では多対1となっていて、中心部の像が重要であることを物語る。ちなみに、人の網膜を「画素数」で例えると、(中心部が高解像度になるのは当然として) 視神経の数から、ほぼ1000×1000画素程度であると考えられる。

さて、眼球をカメラに例えて考えると、「倒立像」という言葉からひとつの疑問が湧いてくる。「我々が見ている世界はさかさまなのか?」という疑問である。答えは、網膜上の物理的事実としては「YES」である。これは、カメラにおけるフィルム面と同様で、網膜上には倒立像が写っているのである。この問題について、レンズやプリズムによる網膜像の逆転実験を行ったストラットン(1896)の報告によると、「逆さ眼鏡」をかけはじめて3日目ほどで違和感がうすれ、1週間ほどで以前と同様の視覚が確立して、重力方向や触覚、聴覚との矛盾がなくなるといふ。また生まれてすぐのネコに「逆さ眼鏡」をかけさせるといふ実験でも、ネコの成長過程ではなんら有意な現象は見られなかったという。要するに問題は、脳の中で網膜像が他のあらゆる感覚とどう関連づけられるかにあるのであって、網膜上で正立か倒立かはどちらでもよいのである。

もともと我々には、(自分自身の手足も含めて) 自分の眼に入った世界しか見えていないわけで、網膜像が世界のすべてである。したがって「手の見えているところに手の感覚のある場所が結びつき、足が見えている方向と重力を感じる方向が結びつく」ようになれば、我々は矛盾を感じずにすむのである。脳の中で視覚像と体勢感覚が同一化していれば

問題はない。感覚的には理解しにくい事実だが、我々の世界認知にとって視覚像がいかに優位な立場にあるかということを示す重要な事実だと言える。

ここで、脳の視覚領に関する脳科学の知見を補足紹介しておきたい。

網膜が捉えた像は、ほぼそのままのイメージ配列を保って大脳に向かうのだが(V1野からV3野までは、ほぼ網膜の配列がそのまま)、大脳には視覚に関わる複数の領野があって、領野ごとにかなり明確な役割分担がある。例えば、V1野(第一次視覚野)では初期的な情報の処理と振り分け、V3野では方向・線すなわち「形」の検出、V4野では「色彩」の検出、MT野では「運動」の検出など、それぞれタイプの異なる処理が異なる視覚領野で行われている。また、その処理の流れにも分担があつて、例えば、右視野の像が左脳へ左視野の像が右脳へと分岐していること、「空間視」に関わる情報と「形態視」に関わる情報がそれぞれ大脳の背側と腹側に分岐していることなど、大脳は、かなり複雑に機能分化していると言える。もちろん、それらがどのように連合されるのかといった複雑な問題は未解決だが、このような視覚に関わる領野が少なくとも30以上あつて、大脳皮質の60%以上が視覚情報処理に関わっているという事実は銘記しておくべきであろう(ちなみに、霊長類は皆同様の「視覚動物」である)。

さて、今の段階で確認されている非常に重要な知見は、「視る」ことにも「想像する」ことにも、ともに側頭葉連合野の連想記憶ニューロンの活性化(すなわちイメージ表象の活性化)が関わっているということである。

目からのボトムアップ信号（視覚）、そして前頭葉からのトップダウン信号（想像）、この2つのタイプの信号は、同様のふるまいで我々のイメージ表象を活性化している。すでにサルトルは「想像力の問題」の中でこのことを哲学的に考察していたが、我々の視覚と想像が、脳の中でおこる「イメージ喚起」のプロセスを共有しているという知見は、特に「映像」に関する領域では、あらゆる考察の根幹をなすものとして注目すべきである。

2 可視光線と色彩

人の視覚がとらえることのできる光とは、波長が380nmから760nmの範囲の電磁波である（ちなみにその外の領域は、波長の短い方へ紫外・X・ガンマ；、長い方へは赤外・無線通信の電磁波領域：となる）。我々はその波長の違いを「色」という現象としてとらえており、波長の長い方から「赤橙黄緑青藍紫」というふうに対応づけることができる。もちろんこの「色」は物理的に存在するものではなく、我々一人一人の頭の中で生じている心理現象であって、また「色名」も本来連続的に分布するものを言葉で区切ったものにすぎないため、「赤と橙の境界は？」と問われてもその答えは定まらない。

「人」が見ている「色」をさらに詳しく理解するためには、錐体（Cone）と杆体（Rod）という2種の視細胞の存在を知る必要がある。それぞれの特徴を簡単に言うと、錐体は感度は悪いが色彩を感じる能力があり、杆体は逆に高感度であるが明暗しか感知することがで

きない、というものである。「人」の網膜上には中心部（黄斑部とも言う。凝視点の像ができる部分で、視角で言うと10度以内）に錐体が多く分布し、一方周辺部には杆体が多く分布している。我々が「色」を感じて読み取ることができるのは、まなざしを向けている限られた範囲ということになる。

では、この錐体はどのようにして「色」を見分けるのか。錐体には、R（611nm）・G（529nm）・B（462nm）に感受性のピークをもつ3種類の細胞があり、入射光の波長により生じるそれぞれの反応の割合で、色を感じていると言われる。画像を表示するディスプレイはRGBの3原色によってすべての色をつくっているが、本来3種類の波長の光をまぜて単一の波長の光にするというのは物理的には無理な話で、これはすなわち「我々の錐体がRGBの組み合わせで反応しているために、本来の単一の波長と、3種の波長のまざったものの区別ができない」ということを逆に物語っているのである（この3原色説はすでに、ヤング（1801）・ヘルムホルツ（1860）によって提唱されていたものである）。

3 色彩の心理

さて、この心理現象としての「色」には、色彩嗜好・感情効果・対比効果・誘目性・演色性・色彩調和など、視覚情報のデザインの際に考慮すべき問題が多くある。以下にそれらを簡単に列挙しておこう。

色彩嗜好：好まれる色は、感覚的、情緒的、経験的に好ましい色で、「明るい・健康的・上品・自然・新しい」など様々に形容されるものであるが、その好みも性・年齢・時代・民族などによって異なるものであり、その「人」個人の関わる文化の問題であると言つてよい。ただし総体的に紫系や暗濁色は敬遠されがちである。

感情効果：我々は色彩によって感情を刺激されるが、色彩と感情の対応関係は一般にくつつかの軸で示される。すなわち、「暖色⇄寒色」・「動の色⇄静の色」・「重い色⇄軽い色」・「強い色⇄弱い色」・「硬い色⇄柔らかい色」・「好きな色⇄嫌いな色」・「心地よい色⇄不快な色」などである。しかしこれらの軸の上で、寒暖・静動・重い軽い以外はその色彩の位置付けに個人差が大きく標準的な関係付けは困難である。

対比・同化効果：色同士が影響しあつてその差異が強調されて見えることを色の対比といい、逆に色同士が誘導によつてまとまつて見えることを色の同化という。一般に「図」の色に注目した場合は対比が、「地」の色に注目した場合は同化がおこりやすいことがわかつている。人は「図」に注目するのが普通であるから、対比効果は特に顕著で、例えば、白に囲まれたグレーはより黒く見え、黒に囲まれたグレーは明るく見えるなど、周囲の色の影響で「図」の色は実際の色からずれて見えているということに注意が必要である。

誘目性：より明度の高いもの、より彩度の高いものが人の目を引き付けやすく、色相に関しては青系よりも赤・黄系が有効である。しかし実際的には周囲の色との対比、特に明度の対比効果が問題で、同じ黄色でも黒と組み合わせられた場合が最も誘目性は高くなる。

黄色と青というふうには色相の上での反対色も効果的だが、赤と緑のように明度が近くなるとギラギラして見づらくなる（リーブマン効果）。

演色性：特殊な照明（高速道路のトンネル内ネオンなど）を除けば、照明光は連続スペクトル（可視光の広い範囲の波長を連続的に含む）か、あるいは複数の線スペクトルから成る光で、その分布のかたよりによつて赤みや青みを帯びている。この照明の色みは、一般に色温度という概念で表わされるもので、例えば白熱球やろうそくなどは3000K以下、太陽光は6500K、国内用テレビは9000Kなどとなる。したがつて色温度の異なる照明の下では同じ白でも異なるものとなり、当然物の見え方の印象などは変わつてくることになる。このような光源の性質を演色性と言い、様々な状況下で適切な色温度の照明を計画することが望ましい。学習などの作業に向く照明と、食卓を照らす照明を使いわけけるなど、日常的にも経験のあることであらう。

色彩調和：複数の色面を組みあわせて画面を構成する場合、その配色が美しく調和して見えるかどうかが問題となる。このことについて一般的に次のような解決案が考えられる。①色相について類似するものを配色する（近似色調和）、②同色相で明度・彩度の異なるもので配色する（同系色調和）、③色相に関して反対となる色で配色する（補色調和）、④彩度の高いもの同士は一般に色がぶつかりあうため、多くの色相を配色する場合は低い彩度のもので組み合わせる。⑤高い彩度の色同士が隣接する場合には中間に無彩色を配置する、などである。ただしこれはあくまで一般論であり、構成法や各色の面積比で様々な調

* 色温度

人間の視覚は、通常照明の色温度に対して自動的にホワイトバランスをとりなおしているため、その赤さや青さをあまり感じていないが、フィルムで撮影すると、その差は歴然とする。

照明の色温度の違いによる色かぶりの問題を解消するには、写真を撮る際に、白い紙等を同時に写し込んでおくとうい。後からフォトレタッチツール等で、簡単にホワイトバランスを取り直すことが可能である。

和のさせかたが追及できるはずである。

4 視野

頭部を固定して1点を凝視した状態で見える範囲を視野というが、視野計による調査では、「人」の視野は左右約200度・上下約140度と言われる。この場合左右については大差ないが、上下に関しては上60度・下80度と下の方が広く、経験からもわかるとおり、日常生活ではさらに下の方が優位になる。我々の身の回りの物は大部分眼の高さより下にあつて、とりあえず注意を要するのは足元なのだから、これは当然のことであるが、都市生活では「頭上注意」も常識であり、人の視野と生活環境とは決して無関係ではないことは銘記すべきであろう。

さて、上に述べた視野は「見える範囲」であるが、このうち実際に情報の読みに関わる領域というのは、左右20度・上下10度の範囲で、これは網膜上と言うと錐体が集中的に分布する域にあたる。この範囲の情報は意味あるものとして捉えられているわけで、例えば星座のように我々がかつてに群として見ているものも、ほぼこの視角内におさまっている。視覚情報をデザインするという観点から言えば、この「見える範囲」「読む範囲」という二つの視野のもつ意味は重要で、例えば駅のホームから見える電照看板を計画するという場合でも、まず普通にホームに立った状態で、視認されうる範囲内に設置される必要

があり（気付かれなければ意味がない）、またそこから眺めた場合に、看板のデザイン全体がまとまって見える視角範囲におさまるかたちで判読されることが望ましい（これを超える大きさのものでは、構図や配色といった画面内の設計が的はずれなものになってしまう）。視野と視角の問題は色彩や形態の知覚の問題に比べて忘れられがちであるが、作業環境の計画、鉄道や自動車道路など交通システムにおける案内・標識の計画、公園や都市環境全般における景観の計画、あるいは又聴覚障害者のためのコミュニケーションシステムの設計など、我々の生活に関わる様々な場面で考慮されるべきものである。

補足的に他の動物との比較にも触れておこう。大半の動物は眼が顔面の両側にあつて、各々の視野が独立するかたちでは360度の視野をもつのに対し、「人」場合は両眼とも前方を向いていて、左右の眼の視野の共通領域が広い。つまり両眼での全体の視野は狭いが、両眼視による奥行き知覚（後に詳しく述べる）が有利になるという点が特徴的である。

このことは、「人」以外の動物が障害物や外敵といった自然環境に関する情報を重視するのに対し、「人」はそうした情報よりも相手の表情やしぐさ、あるいは文字や画像情報といった同種のもの同士でのコミュニケーションに関わる情報を重視することを物語る。

5 形態の知覚

形態の知覚について考察する場合、まず「地と図」というキーワードの理解が必要であ

る。「地と図」とは視覚心理学の用語で、「地」は背景、「図」はそこにあるまとまりをもって現出している対象を言う。一般に「図」として見えやすい傾向にあるのは、閉じた領域、面積の小さな領域、垂直・水平方向にそろった領域、上下で言えば下の方、幅が一定な領域、動いているもの、誘目性の高い色彩、輝度の高いもの、とされている。

日常の視覚では、背景と視覚対象との関係は、ほぼ明白であるが、紙面やディスプレイといった2次元の視覚では教示のしかた次第でこの反転が生じやすい。「杯にも見えるし、向かい合う二人の顔にも見える」という有名な「ルビンの杯」なども、この図地の反転現象を応用した図形である。

さて、「図」になるものとは、もともと物理的にひとつの個体として存在するものなのであろうか。実はそればかりではない。我々が通常ものを見る場合、本来無関係のものでそれらを「群化」させて見ている場合が多い（例えば星座）。創作という能動的な行為のみでなく、「ものを見る」という一見受動的な行為の場合にも、人はなかば自動的に物事を「秩序」だてて捉えているのである。

物理的にはバラバラな視覚刺激を、我々の視覚がこちらの都合に合わせてまとめてしまう。このような性質を、心理学では「ゲシュタルトの法則」と言う。ウェルトハイマー(1923)は、「バラバラなものがまとまって見えるための要因」を、近接・類同・閉合・よい連続・よい形・共通運命・客観的構え・過去経験の8つに分類して説明している。例えば、近くにあるもの同士がまとまって見える(近接の要因)、 \wedge \vee や \bigcirc など「閉じた形」はワンセッ

トに見える(閉合の要因)、^{*}群集の中を2人の人物が同じ速さで走っていると、その二人が主役として浮き立って見える(共通運命)、「三角形に注目して下さい」と言われるとバラバラな3点でも三角形に見える(客観的構え)、「朝顔を見ると思ひ出す」という文の解釈(「朝の」顔「か」朝顔「か」)が人によって異なる(過去経験)など、ウェルトハイマーのまとめた群化の要因は、様々な現象を説明できる。

我々の視覚は、その認知において、視覚情報量の経済効率を考えているであろう。バラバラなままで記憶するより、要素の関係を見出して、その関係を記憶する方がはるかに効率的である。例えば「○○○○○○○○○○」は「○が十個」と書けば6文字分圧縮できる。また例えば文章も、一文字一文字見るよりは、前後の文脈をたよりにあり、そのような「単語」に予測をつけて単語単位で読んでしまう方が早い。

情報量が少なくなるようにまとめて見る。要素ではなく要素間の関係を把握する。それが人間の知覚の基本方針と言える。

6 幾何学的錯視

ここまでの話でも明らかのように、ものの見え方というのは、対象となっているものの物理的な特徴と、それを見る我々の脳の働きの両方に原因があることになる。ここで、さらに我々見る側の問題が大きく関わっている幾何学的錯視その他の現象をみてみよう。

*図地の反転

特に視力が低下している場合、意識の構え方によっては、その逆転は簡単に生じるものである。

例えば、心霊スポットなどで「出る出る」という意識で見ていると、背景の影のほうに図になりやすくなって、結果「何か」を見てしまうことがある。数人で同じ景色を見ていても、自分だけが、「何か」の影らしきものを見ることが十分にあるにあり得ることなのである。

*群集(モブ)シーン

この場合、走っている2人をカメラが固定点でとらえるようにフォロー撮影すると、さらに主役が浮き立つ映像になる。この場合2人の役者とカメラとの3者が共通運命にあることになる。

幾何学的錯視 (Geometrical Optical Illusion) とは、大きさ・形・方向などの幾何学的パラメータが実際の値とは異なって見える現象である。最も簡単な例は、「正方形が縦長の長方形に見える」というもので、我々の視覚では水平線より垂直線の方が長く見える（垂直・水平錯視）ことが確認できる。幾何学的錯視には様々なタイプのものがあり、はじめに発見・報告した学者の名で○○錯視などと呼ばれる。一般に、錯視現象は「細く見せる」「高く見せる」などの目的に応じて視覚情報のデザインに応用されている。

その他、物理的には描かれていないのに輪郭線が存在するように見える現象（主観的輪郭）や、物理的にも主観的にも見えていない存在が視覚情報処理に影響するという現象も、見る側の心理に大きく依存した現象である。特に後者は重要で、例えばレイアウトグリッドのような「不可視のガイドライン」は、無意識のうちに我々の視覚に捉えられて、それが全体の構図や秩序感を大きく左右している。例えば、教室という空間の中でも、通常は机が「不可視のガイドライン」に沿ってならんおり、それからずれるものがあると、そのずれた机の輪郭線の延長にさらに新しい「不可視のガイドライン」が生じて、空間は雑然と見えてくる。描かれた線のみならず、図形の線の延長に感じられる「非在の線」も、画面全体の構図・秩序に大きく関与するものであることを銘記しておきたい。

7 奥行き知覚

本来2次元である網膜像をもとに、我々は奥行きを含む3次元の世界を認知している。

第3の軸である、この奥行きを知る手がかりには絵画的要因・生理的要因・運動要因などがある。

絵画的要因とは、大きさ・上下・重なり・きめの勾配・色調・コントラスト・明暗・影のできかたなどで、配置や描きかたによって奥行きを知る手がかりが得られるというものである。

生理的要因には、水晶体の調節や視線の収斂といった筋肉の動きに関わるものと、両眼の網膜像のズレによるものがあるが、特に両眼視差 (Binocular Parallax) はステレオグラムの基本原理でもあり、他の要因とは違う生々しい立体感を得ることができる。

主として視点が移動している場合には運動要因が効いてくる。視点を移動しながら風景を眺めていると、近くの風景と遠くの風景とでは移動のスピードに差ができる。このことから感じられる遠近感を、運動視差 (Motion Parallax) といい、セルアニメーションなどの風景描写に応用されている。

8 運動知覚

運動の知覚には2種類の機構があると考えられており、それぞれ、像―網膜システム、眼―頭システムという。像―網膜システムは視線を固定した場合に動く対象の網膜上での位置が変化することを運動として知覚するもので、眼―頭システムは動く対象を追視した

場合の眼球や頭の動きから対象の運動を知覚するものである。我々は一般にこの2つの情報を関連づけながら処理することで、運動を知覚していると考えられる。

ただしこの運動も、実際に対象が運動している場合にのみ知覚されとは限らず、「運動しているように見える」という仮現運動の場合もある。仮現運動には自由運動・誘導運動・運動残像・ β 運動・ α 運動・ γ 運動など様々なタイプのものがある。以下、簡単に説明しよう。

自由運動とは、暗い場所で一つの光点を凝視していると、ゆれて見えはじめる現象である。「感覚遮断されると幻覚が見えはじめる」あるいは「入眠幻覚」などの現象とも似て、「人」は刺激情報が少ない場面では、物理的には存在しない架空の現象をつくりだす場合が多い。誘導運動とは、静止した対象が、周囲のものの動きによってその逆向きに運動して見える現象を言う。特に、自分の周囲のものの動きによって自分自身が動きだすように感じられる場合を自己誘導運動といい、「ビツクリハウス」などに応用されている。

運動残像は、滝の水の流れを見つめた後、静止した物体を眺めると物体が上に動き出すように見える現象で、回転する螺旋が止まった後など、運動が停止したあとその運動とは逆方向に感じられる動きはすべて運動残像である。

α 運動はミューラ・ライアの錯視図で2種の矢羽根の図を交互に提示すると、水平の線分が伸縮して見える現象を言い、光刺激が出現・消失する際にそれが膨張・収縮する運動に見える現象を γ 運動と言う。

β 運動は、我々にもっとも馴染み深い現象で、電光掲示板で文字が動く、映画やテレビで現実的な動きが再現される、といった運動の知覚がそれである。これには「視野の持続性」と「ファイ現象」という2つの視覚の要因が作用している。前者は「光(像)の点滅が秒間30回以上まで速くなると、それは点滅ではなく持続して見える」というもので、視覚の情報伝達速度の限界から生じる現象であると考えられている。したがって映画では秒間24コマという画像素材を1コマにつき3回シャッターをあげるという形で秒間72コマにしており、テレビは秒間30コマという画像を1フレームにつき2回(2フィールド)走査するかたちでちらつきを防いでいる。後者は「空間的に位置の異なる2つの光点を一定時間あけて見せると、光が動いて見える」というもので、これには刺激の強さ・刺激間の距離・時間間隔などが適切に計画された場合にきれいな運動に見える。時間間隔のみに注目すれば、一般にそれは60ミリ秒程度と言われており、30ミリ秒を下回ると2つが同時に見えてしまうことがわかつている。

眼はもともと運動の発見器としてスタートしたとも言われ、外界の動くものを知覚するということは、生物一般にとってその生存に関わる重要な問題であった。しかし、現代の「人」にとつてはそのような外界の運動知覚はもちろんだが、虚構の世界の運動知覚である「仮現運動」が非常に大きなウェイトを占めている。我々は日常において多くのものの動きの情報を、現実の運動ではなく、仮現運動(テレビの映像)によって得ていることは銘記すべきであろう。

9 視覚のフレームオブリアレンス

日常的な視覚・文字・静止画・動画などに対して、見る意識の構えは様々に関与しており、それは視覚のフレームオブリアレンスとして、その時々、視覚世界の構築と精巧な読み取りを手助けしている。

日常の視覚では聴覚の場合と同様、頭の中の単語の辞書と文法構造が最も大きなフレームオブリアレンスとして作用しており、すべてを言語的に了解しようとする「カテゴリー態度」が効いている。もちろん視界に「名付けようのないもの」が出現した場合には、このフレームオブリアレンスは効力を失い、緊張感や不快感が高まる状況が発生することになる。

日常の視覚におけるフレームオブリアレンスは、さらに様々な次元に存在する。「人の顔を見る」という状況を例にとると、我々は単にそれを「顔」とカテゴリー化して見る以上に、表情の細かな読みが可能である。しかし顔全体がさかさまに見えている状態ではその判別が難しくなる。すなわち、この場合、顔の目鼻の正立した位置関係がひとつのフレームオブリアレンスであり、それが有効な場合に、表情の違いを敏感に読み取ることが可能だが、フレームオブリアレンスが無効になるさかさまの状態では読み取り能力が落ちることを意味する。同じことは文字などにも言えることで、さかさまでも文章を読む

ことはできるが、内容が頭に入っていない、あるいは誤字や左右反転した文字などを発見する能力が落ちるなど、全体的な情報処理能力は落ちるのである。

「さかさまでは読みにくいのはあたりまえ」だという感想もあるだろうが、「機械」の視覚の場合に、像が正立であれ倒立であれその分析能力に差がないことを考えれば、「人」あるいは「生物」に特有の現象と考えることができる。

次に画像を見る場合についてであるが、一般に「像」として与えられるものには、言語情報における「辞書」・「文法」といったものがないため、それに代わる「何か」が様々なレベルで数多く存在することが考えられる。

ここでは、人の造形的な思考を介して描かれた絵画・漫画と、カメラで自動的に記録された写真やテレビの映像との区別（前者が明らかに「様式」や「文化的な約束事」といった伝え方の枠組みを持つ）をふまえた上で、いくつかの事例を紹介しよう。

例えば漫画を読む（見る）という場合について一つ例をあげると、（日本人の場合）右上から左下へ向かう「方向」が、ひとつのフレームオブリアレンスとしてその読みを拘束しており、右↓左は順方向で「行くもの」として読まれ、左↓右は逆方向で「やってくるもの」として読まれる傾向にある（絵巻物などではそれは典型的である）。これは「縦書き」という文字文化に由来するもので、西洋の場合では逆に左から右が順方向となる。絵画・漫画表現にはこうした様式・約束事が多く存在し、それを読み慣れた者には、作家の意図がスムーズに伝わりやすいと考えられる。現に、いわゆる活字世代と漫画世代では、

漫画を読む（見る）スピードや視線の配分に大きな差が見られる。

また、カメラがとらえた映像を見る場合には、例えば、その映像から逆投影的に読み取れるカメラの切り取った「ワク」・位置（視点）・角度（視線方向）といったものがフレームオブリアレンスとなつて、我々の映像の読みに参画する。映画の理論が誕生して間もないころのアルンハイム（1933）やベラ・バラージュ（1949）もすでに指摘していたように、世界を「ワク」に切り取り、一定の視点から一定の方向を与えるということは、見る者に「世界の見方」というフレームを与えるものであり、そのこと自体がある意味をもつ、あるいは意味を生むものと考えられる。

これらの要素の重要性は、我々の日常的な映像体験からもわかる。例えば、「ワク」に関して言えば、「地面を這う演技が、岩山を登るように見える」などのトリックはその効果を利用している。すなわち、我々は与えられた「ワク」を基準に世界を再構成して見ていることから、撮影時のカメラが90度倒れていてもそのことには気付かないのである（実際には、気付いてもそのようには見えにくい）。また、位置と角度についても、我々はそれがフレームとして機能しなくなると、どこからどう見て撮られたものかわからなくなり、なにか写っているのかすらつかめなくなる。「ワク」・カメラの位置・角度、我々は通常それらを意識して感じてはいないが、映像を見る場合に究めて重要な役割を果たしている。

描かれたものにせよ、カメラでとらえられたものにせよ、そこには（言語コミュニケーションにおける辞書や文法のようなかたちで取り出すことはできないが）様々なフレーム

オブリアレンスが存在して、見る者の読みを支えている。それは経験的に身についてくるものであるが、音楽における「音階スケール」のごとく、その修得に臨界期があるかどうかも定かではなく、したがって「音痴」ならぬ「映痴」についても謎は多い。

最後に慣性系に関わる特殊な問題にも触れておこう。世界に対する我々の関わり方には、アウトサイド・インとインサイド・アウトという2種類のフレームオブリアレンスがあるとされる。この問題は、航空機の操縦における視知覚を考察する場合によく引き合いにだされるものだが、アウトサイド・インとは「航空機の外の世界（地球）が基準座標であり、航空機（自分）はそれに対して傾いている」というように、フレームオブリアレンスを外部において視覚世界を捉えている状態を言い、インサイド・アウトとは「航空機の内部（自分）の上下左右が基準座標であり、外の世界が傾いている」というように、フレームオブリアレンスを内部において視覚世界を見ている状態を言う。

一般に航空機の操縦では航空機自体が慣性系であるために、インサイド・アウトが有効だが、離陸や着陸時にはアウトサイド・インが有効にならざるをえない、この切り替わるタイミングが航空機の操縦で最も危険な瞬間であるとも言われる。自分が傾いているのか外が傾いているのかわらなくなる瞬間というのは、体勢感覚にも視覚にも違和感が生じてあらゆるコントロールに障害が発生する。VRシステムを実用化する過程でもこの問題には注意が必要である。

10 視覚の人工現実

人工現実には仮想現実と遠隔現実の分類があることはすでに述べた。ここでは視覚世界におけるこの問題を聴覚の場合と対応づけながら考えてみよう。

仮想現実世界のもっとも簡単なアイデアは、ディスプレイのサイズを大きくして、我々の現実的な視野を覆ってしまうという発想である。一般に視覚的な臨場感は視角にして20度を超えると大きくなると言われ、約80度から100度でほぼ満足できるサイズになると言われる。要するに非常に大きなスクリーンを用意すれば良いことになる。

しかしこれだけでは「あきらかにこれは映像である」という意識は残るもので、例えば大きなものがこちらに向かってくる映像を見ても「思わず体をよける」というところまではいかない。最大の問題は3次元空間の奥行きをどう再現するかということである。聴覚ではスピーカの多元配置、特にステレオ音場による空間的な音の定位が重要だが、視覚の場合もこれと同様で、両眼視差を応用したステレオグラフィックスの技術がここでの必須要素となる。原理はすでに述べたように、左右の眼の網膜像にあたる映像をそれぞれ左と右の眼に分離して見せることで実現できる。

現在その技術は表示面方式と眼鏡方式とに大別される。表示面方式は、レンティキュラ板やパララックスバリアなどで表示面そのものを加工するもので、肉眼で立体視できるの

が魅力だが、観察者の位置が制限されるなどの問題がある。眼鏡方式では偏光板と偏光眼鏡用いた同時式と、左右交互にシャッターの開閉がおこる眼鏡を用いた継時式があり、特にスクリーンと偏光眼鏡を用いたアトラクションは数多く見受けられるようになった。

さて、遠隔現実の構築では、観察者の手足や頭が動きながら、「向こうの世界」に関わっていくことが前提となる。したがってこの場合は、聴覚におけるヘッドフォンに相当する、HMD (Head Mounted Display) が利用されるのが一般的である。HMDシステムでは、頭の位置・方向・向きの情報を利用して、体や頭の動きに応じた能動的な画像更新が実現される。もちろん右眼と左眼の画像は分離されているから、奥行き感も再現されるかたちで架空の世界を疑似現実として見ることができるようになるのである。

現在さらに進んだディスプレイのアイデアとして、レーザー光によって直接網膜上に像を形成するというシステムが開発中であるという。人の脳そのものにアクセスして視覚像をつくることできない以上、人工的に視覚に介入できる限界は「網膜」であり、この技術に「体勢」の情報をうまくシンクロさせれば、現実と非現実の区別がきわめてつきにくい架空の世界が構築できるであろう。

結び — 「過剰な脳」の独り言 —

なぜかと問われたところで、「脳がよろこぶからだ」としか答えようのないことがある。しかしこれが、人間のあらゆる思索や行為の動機を説明する、最も端的な答えであろう。

人間とサルを分ける最も顕著な違いは、無駄（？）に発達した前頭葉にある。この過剰な脳細胞たちは、自分たちの活性化のために（細胞は活動しなければ衰退するから）、人間に「考える」ことをさせはじめた。要するに生命維持活動には必要のない、自らのためにお互いに信号を交換し合うという余計なことを始めたのである。人間は、考えることによつて、時間的にも空間的にも「今ここ」を離れた事柄を反省したり予見したりするようになり、その想像が描き出す「快適な世界」を実現しようというんなものをつくった。そして、自分たちがつくった世界の複雑さゆえに、20年もの学習期間を経なければ歩き出せないというバカな状況に陥っている。サルはそんな人間を見ておそらくあきれている。歌うことも描くことも物語をつくることも、そして世界の仕組みを探索することも、サルには必要のないことなのだ。すべては、過剰に発達した脳の「よろこび」に起因している*。

かつて、授業のなかで「あなたはなぜ映画を見るのですか」という問いかけをしたことがあった。「好きなタレントが出ているから」「物語の世界を見たいから」「現実逃避」「自分もつくりたいから（ヒントが得られるから）」、「映画と映像について研究したいから」、

様々な回答がなされた。そのままでは面白くないので、さらにディスカッションをしたところ、それらの回答の枝葉を落とすことで、以下の四つの機軸を抽出することができた。

視たい／つくりたい／知りたい／伝えたい —— である。

ここには「脳がよろこぶこと」の理由が見事に見出されているように思える。これらは次のように言い換えてもいいだろう。すなわち、物語を視る（読む）こと／物語をつくること／物語の仕組みを知ること／物語の仕組みを伝えること、——である。「物語」という語を「世界」*に置き換えても良いかもしれないし、あるいは「秩序」と置き換えても良いだろう。サルは物語を理解しないし、つくることもない、当然形而上学的な好奇心は無いし、それを他のサルに伝えようもしない。これらは人間に特有のよろこびなのだ。

この四つの分類は、人間の発達段階に照らしても興味深い。まず視ること聴くこと物語に浸ることが楽しくて、やがて自分でも何かつくってみたくなる。つくっていくうちにその仕組みに興味が行き、最終的には自分が発見した仕組みを人に伝えたくなる。多かれ少なかれ誰でも成人するころには、なんとなく自分の世界観のようなものができあがって、それが成熟すると次の世代へ伝えたくなる。典型的なのは、宗教家や芸術家、あるいは広く研究や教育に携わる人たちである。自分の世界観がクリアであればあるほど「伝えたい」という気持ちは強くなる。

* 過剰な脳活動

もちろん、その過剰な脳活動にも程度はある。脳はやみくもに活動したがつていてるわけではない。信号が多すぎても細胞は破壊されるし、信号のループバックに歯止めがきかなければ癲癇発作のようなことになってしまう。あまりにも多くの新しい情報に対しては、「待った」がかかる。情報は多すぎれば疲れるし、少なければ退屈する。日常的感觉でわかるように、脳はその時々のできるような活性状態を求めている。

* サルとヒト

「サル」と一括するのは実は乱暴で、霊長類の中でもチンパンジーなどの大型類人猿は、鏡による自己確認や写真の認知、またペンを使った落書き行為もする（ニホンサルなど旧世界サルは落書きはしない）。観察できるレベルでの人との区別は、例えば、道具を使って道具を作ること（二次製作）や死体を埋葬する（墓づくり）といった、高度なシンボル操作に関わることができない、ということにある。

さて、脳がよろこぶことの究極には、脳の中で出たり入ったりしている様々な神経伝達物質が関与している。精神疾患の治療などでは脳の状態を良好にするために薬物を使うという場合があるが、通常はそのような外的手段で脳をよろこばす必要はない。厳しい脳血管関門を簡単に通過できる薬物というのは、もともと脳内に存在する物質だということがわかっていて。つまり、脳をよろこばすには、身体状況や各種の感覚刺激、さらに思考のありかたなどをコントロールして、その物質を様々に活性化すればよいのである。おそらく、昔からある様々な「宗教行為」や「癒しの技法」さらには「自己啓発セミナー」といったものも、神経伝達物質のコントロールと無関係ではないだろう。実際、胸の前で手を合わせるだけでも血行が良くなるのだから、内外の刺激が神経伝達物質のふるまいを変化させ得ることは容易に想像できる。

以下、このような物質のコントロールに関与して脳をよろこばせるであろう、見る・つくる・知る・伝える、について順に考察してみたい。

見ること・聴くこと

「見ること、それはもうなにかなのだ」と言ったのは梶井基次郎である。確かに、単なる風景でも、見ているだけで気分が良くなる場合がある。一方で、見ただけで不快感を覚えるような対象があることも事実である。この違いは何か？。記憶である。我々の「見る」や「聴く」は、単純なボトムアップではない。すでに述べたように、そこにはトップダウン

*神経伝達物質

様々な物質が、人間の脳の状態と関わっている。例えば快感・好奇心・学習意欲に関わるドーパミン、注意力・集中力に関わるノルアドレナリン、安心感・幸福感に関わるセロトニンなど数十種類はだが確認されていると言われる。いずれも、不足すると精神活動に支障をきたすが、多すぎても精神分裂・過敏・攻撃・錯乱といった問題につながる。重要なのは適度な活性である。

*自己啓発まがいのお話

「やる気」は、やっかいな存在である。というのも、やる気に関わる側核は、動き始めないと活性があらがない（アセチルコリンの分泌がおこらない）。つまり、やる気になるまで待とうと思っていれば、いつまでたっても「やる気」にはならないのである（逆に休日返上で予定を入れるような「やる気」のある人というのは「がんばった達成感（ドーパミンの快感）がさらに「やる気」を促すことでテンションが上がりっぱなしの状態になっている）。

ンが関与していて、「知覚」と「想像」はともに連想的に記憶を活性化するのである。視ること、それは記憶と関わることなのだ。もちろん人は、楽しい記憶の想起につながる「見る」を求めている。不快な記憶はリンクからはずしたい。人の「見る」は、快適な記憶と結びつくことで快感を覚えるのである。もともと、楽しい記憶と不快な記憶とは、楽しい記憶のほうが長期に保存されることが知られているが（だから想い出は美しい）、それは、気持ちがいい時の方がより結線（記憶のネットワーク）が強化されるからである。

不快な記憶の場合もリンクされれば強化されるのだが、人はそれらをなるべく思い出さないようにして徐々にリンクを切っていく。脳はよくできていて、極端に不快な出来事（強い刺激）の場合は自発的に接合を切って記憶を孤立させてしまうことすらあるのだ（この事実は心的外傷による記憶障害や解離性同一性障害を説明するものでもある）。ソントグ（1979）は、「写真を撮ることは世界をコレクションすることだ」と言った*。美しい記憶の想起を促し、その記憶を強化する刺激、人はそんな風景をコレクションしようと、シャッターを切っているのではないだろうか。

さて、「見る」ことの快感は、単に美しい記憶とのリンクによるものだけではない。視ることによって生じる新たなリンクの形成時にも脳はよろこんでいるようだ。

我々は同じ映画を何度も見て、何度も異なる感動を得る。脳は（記憶は）時と場所とそ
のときの精神状態でネットワークの活性状態（認知的な構え）が異なるから、あらゆる情報体はその都度、重み付けを変えて我々の脳を刺激することになる。旅先で見るテレビが、

朝起きた時は「やる気」がなくて、とりあえず動き出せば、なんとかなる。人はそういう生き物のようである。

「やる気」がでないと言っている学生諸君、とりあえず窓を開けて、光を入れて、手を動かしてみよう。いつのまにか「やる気」がでてくる。「脳はそういう仕組みになっている」。

*イメージのコレクション

我々の欲望の対象は何だろう。確かに人は、いろんな「モノ」を見て、それが欲しくなる。しかし買ったところで、実際には使わないものも多い。欲しかったのはその「イメージ」であり、またそれと連想的に結びつく「夢」なのである。例えば、愛車が盗まれたとしても、それに気付かず「ガレージには愛車がある」と想像できるうちは幸せだ。扉を開けて中身を確かめさえしなければ、実体はなくとも気持ちは満たされたままなのである。

我々の欲望における「実体」の占める割合はこわすか。欲しいのは結局その「イメージ」なのである。

いつもと違って見えるのもそのためである。特に映像のように具体的な情報の場合は、多義的な見方が可能で、新しい「読み」をすることが、また新しい結線を促すことになる。この、リンクが「つながるとき」・「組み換わるとき」の快感も、「視る」ことの一つの動機なのではないだろうか。クリステヴァ(1984)の指摘のとおり、作品は一つの「生産性」であって、生産完成品ではない。その都度新しい読みを促す多義的な「作品」は、我々の脳にとって愉快な存在だと言えよう。脳は視ることによる組み替えをも欲している。

つくること

我々は現実の世界に生きていながら常に新しい「物語」を求めている。「物語」とは別の言い方をすれば「絵空事」であり「異なる秩序」である。小説にせよ音楽にせよ絵画・映画にせよ、新しい情報秩序が出来上がっていくプロセスというのは常に人をワクワクさせるものである。しかし日常的な感覚からもわかるように、楽しいのはつくっているプロセス(それも完成間近の段階)で、出来上がってしまうと興味も薄れてしまう。人は、創造行為における「つくる₁視る」のフィードバックが引き起こす脳内の結線組み換えそのものを楽しんでいくのである(よく部屋の模様替えをするというのも脳内でおこる結線組み換えの投影であろうか)。

つくることは楽しい。自分の脳を駆使して、他にはない自分だけの世界、自分だけの物語、自分だけの情報秩序をつくる。できたものの秩序感が優れていればなおさらである。

しかし、である。人間は、情報秩序を自分の発想でつくったつもりになっているようだが、はたしてそうだろうか。最近思い至った(やっとわかった)事柄なのに、それと同じ記述が10年前のノートに見つかることがある。決してボケているわけではない。10年前にはその言葉の意味の深さに気付いていなかったただけなのだ。深く考えて発した言葉ではないにもかかわらず、言葉は、私の意識より先に結論を言い当てている。つまり「言葉」というシステムは、はじめから、それによって思考すれば出てくるような「答」を内在しているのである。

偶然か盗作かということが話題になることがあるが、情報というものは表現のシステムを確定したとたん、何らかの必然を生じるものである。一定のリズムで、一定の速さで、一定の音域で音符を並べると必然的に音が理想的あるいは合理的な配列を選択することがある。メロディーのように有限の要素、限られた速さと時間で創作されるものでは、複数の作者の作品が似るのは偶然でも盗作でもない、情報の仕組み自体がもつ必然なのである。似ている似ていないで著作権上の揉め事も多いようだが、その作品が優れて音のしくみを反映しているからこそ似てしまうということもあるのではないだろうか。

つくること、それは表現の仕組みやメディアの仕組みを発見することにつながる。そしてさらに言えば、そうした表現の仕組みやメディア自体を作り出した人間の脳の仕組みを発見することにつながるのだ、なぜつくりたくなるのか、それはつくることが通じて自分の脳がわかるからである(明らかに言い過ぎだが、「私の脳」が「書け」と言った)。

*プラトン(B.C.375)は「製作」というものについて、地上の机より「机のアイデア」が先にある、というようなことを言った。また、我々が様々な三角形を見て、それを三角形と認識できるのは、我々が「三角形のアイデア」をすでに持っているからだ、というようなことも言っていた。つまり、ここでの文脈で述べれば、「脳には、はじめからそういう構造があつて、我々はそれを外の世界に投影して形にしているだけだ」ということである。こう言ってしまうと実も蓋もないのだが、別にそれで「芸術」がつまらなくなるわけではない。脳の中は広い。そこからひきずりだせるものはまだまだあるのだ。一時は成りを潜めたチョムスキー(1957)の生成文法論が、脳科学の議論とともに見直されている。つくとも、つくつてよろこぶことも、乳児が自然に文法を探り当てるのと同様に、すべては脳の仕業、脳がそういう構造をもっているからだと言えるのではないだろうか。それを認めてしまえば、さあ、お楽しみはこれからだ。

知ること・発見すること

さて、人間の脳は様々なきっかけで喜びの反応を示す。例えば、人が誰かの言葉に感銘を受ける時、本人は気付いていなくとも、そこには脳の仕組みを反映したよろこびのヒントが隠れている。筆者にとつて四行手前の「お楽しみはこれからだ」というフレーズもそうかもしれない。映画『ジャズシンガー』(1927)のセリフとして見ただけではなく、雑誌のキャッチコピーや日常の会話でも何度も見聞きしている。そして筆者の記憶に強く

ひっかかっているのだ。おそらく、今書いているような文章もそうなのだろう。きつと以前にどこかで聞いたか読んだかした台詞の寄せ集めなのだ。自分で書いている気にはなっているが、所詮こういうものも「脳のお気に入りフレーズ集」なのだと思う。いまのこの文章がぎくしゃくしているから、オリジナルの文章になっている自信があるが、筆が走っているときは要注意だと心得ている。(筆者の論理によれば)「脳が勝手にしゃべる」ような時ほど、既成の文章と構造が似る可能性が高いからだ。

どんなモノにもどんなコトにも人が愉快になれる知の構造が隠れている。ハッと一言葉には脳の結線を最適化するヒントが隠れているのだ。そこに思い至ることができれば、世の中はかなりクリアに見えてくる。

「俳句の会」というのがある。たつた17文字の言葉を相手に、いい大人が集まって、あでもないこうでもないと言い合っている光景は、端から見れば無意味なことのように見えるのだが、本人たちは自分の脳にしくりくる言葉(構造)を求めて楽しんでいるのだ(私はそんな光景をととても愉快に思う)。新聞の「詰め将棋」に熱中している人も、難しい方程式を解いて喜んでいる人も、みんなお互い様、脳へのアプローチの仕方が違うだけで、求めているものは皆同じ「脳の最適構造」なのだ。人が愉快になれるかどうかは、考え方、つまりどの事柄とどの事柄を脳の中で関係付けるかに関わっている。そして、おそらく「ここここを繋いであつちを切断すれば脳の中の信号の流れが快適になって気持ちがいい」というような最適回路があるのだろう。情報システムがもつ仕組みを最も的確に表現でき

*哲学者

おそらく脳科学以前から、人はその思考によつて脳の仕組みを言い当てている。すでに触れたようにサルトルもその哲学的思考によつて、ポトムアップ信号とトップダウン信号がともにイメージ表象を活性化することを言い当てていたのである(もちろん脳科学者がサルトルの発想を直感的に正しいと感じていなければ、そのような検証も進まなかったのかもしれない)。

その意味では、「人間とその世界」についての知見には新しいも古いもない。熟考してピンときた発想であれば、それはおそらく脳の仕組みを上手く反映しているだろうし、同時に我々にとつての世界も上手く説明するはずである。ギリシャ哲学にも、お経のなかにも、近代哲学にも、使っている言葉が異なるだけで、「同じようなこと」が書かれている。

*人間がつくったWWW

インターネットはまさに、人の脳のアイデアを外在化した形式をもつ関連のある情報同士が勝手にリンクしてつながりを強化していく。ネットワーク網を脳に見立てると、そこではあちこちで自己組織化がおこっているのだ。

人の脳を最もよく反映したシステムが、人の世にあるという間に浸透したのは当然の結果と言える。

たとき（脳の中の構造を探り当てた時）「わかった」という達成感が得られる。

「わかった」というときのよろこび。ここに脳がよろこぶヒントがある。「わかる」というのは文字通り「分かる」、もつれていた結線がきれいに分離して、回路が秩序立ってくることが意味している。もちろんこの回路は絶対的なものではないため「わかった気になっている」だけなのだが（だから採め事は絶えない）、要は、脳内で新たに強く新鮮なネットワークができる瞬間が気持ち良いのである。「なるほど、わかった」が繰り返して思考のテンションが上がってくると、ますます気持ちよくなる。ドーパミンのせいだろうか。

伝えること

人間は脳の中の情報を次の世代へ残そうと必死になっている。特に「神」の概念などはその典型で、世界中の人々が自分の信仰の対象を周囲の人々へ、そして次の世代へ伝えようとしている。

リチャード・ドーキンス（1976）は、このような人特有の情報体の伝達欲を説明するのに「Gene（遺伝子）」ならぬ「Meme（文化遺伝子）」という言葉を考えだした。生物の大半は、遺伝子という物質によってのみ次の世代への情報伝達を行っているが、人間はそれ以上に脳の中に形成された情報の痕跡を時間的にも空間的にも伝え・拡大しようとする。サルにも「イモ洗い」程度の文化の伝播はあるようだが、言語・音楽・映像のような高度に編集可能な情報体を用いた伝達は人に特有のものである。

しかし、考えてみれば、遺伝子も（ウイルスも）物質とは言え、四種類（A・T・G・C）の塩基の組み合わせが問題なのだから、コンピュータの中の二種類（0・1）の情報並びとなんら変わりはない。要するに、生物は「情報」を伝えたい（複製したい）のである。「情報」というものは、基本的に「複製されること」・「伝えること」を前提としている。孤立した情報は情報として意味をなさないので。情報は複製されたがっている。

世の中がオープンソース（フリーソフトウェア）へ移行しようとしているのも、積極的に「複製と突然変異」を認めるといふ発想が人間（生物）の直感になじみ易いからだろうし、逆にソフトウェアというものに関して著作権（複製権）の問題が多発するのも、もともと「複製するな」というのが「情報の実態」にそぐわない発想であるからだろう。

音楽にせよ映像にせよ、創作活動に携わる人の大半は、まず聴いてもらうことを視てもらうこと、すなわち「伝えること」を願いつつ活動しているはずである。作者は、それを通して自分の脳内情報の痕跡をあちこちに残したいのである。自分を支えているのは、自分の記憶である。人が、その記憶の痕跡を残したいと願うのは、ごく自然なことだと言えよう。自分がこの世に生きた証として、人は様々な「情報」を伝え遺そうとしているのである。

歌う事・描く事・映す事、上手か下手かはそれほど大きな問題ではない。大事なのは、そこに「伝えたいという気持ち」があるか、「遺したい情報」があるかということである。もちろん、伝えるにはそれなりにテクニクも要るのだが、それはまた、別の話。

*自分を相対化する

世界をどう捉えるかというような問題には、何通りもの答えがある。もともと「正しさ」などというものは相対的なもので、考えている当人にとって、どの視点に立てば世界が最もクリアに見えるかが重要なのである。「地球を中心に太陽を回すこともできるし、太陽を中心に地球を回すこともできる。どちらの記述がより簡単で応用がききやすいかが重要だ」ということと同じである。

すべては「脳」が考えていることである。人間が考えるあらゆる理論は、視点が異なるだけで、正しいとか正しくないとかいう問題ではない。人は、自分の「脳の仕組み」を上手く反映する思考を「正しい」と考えているにすぎない（この記述も例外ではない）。

*サルのイモ洗い

宮崎県の幸島に生息するニホンザルに見られる現象。「イモを海水に浸けて食べるとおいしい」ということがサルの集団の間に伝播・継承されている。

*遺伝子の情報量

人間一人のDNAの情報量を計算してみよう。人のDNAは30億対の塩基の組み合わせから成るので、4の30億乗、つまり2の60億乗の組み合わせが可能。情報量にすると、 $\log_2(4^{3000000000})$ すなわち $60\text{Gbit} \approx 750\text{MB}$ で、ほぼCD-R一枚に記録できる。人間が音楽情報の編集単位としてデザインした記録メディアに、人間のデータがすっぱり入る。何か意味でもあるのだろうか。

*著作権

ガイドをガチガチにしたシステムは流行らずに失敗する場合が多く、複製を（作法的に？）許容したものが方が結果的には成功する。著作権法は、著作者の権利を守るためにあるのだが、現状をふまえると、抜本的なパラダイム転換が必要とも言える。PCでやっていることを思い出そう。やっていることの大半は「複製」なのである。ファイルのコピーができないPCなど何の役にも立たないのだ。

人間にとって「情報の複製」とは何か、そこから考える必要がある。

参考文献

- J・P・サルトル (1955) 想像力の問題 平井啓之訳 人文書院
 J・クリステヴァ (1984) 記号の生成論 中沢新一他訳 せりか書房
 J・モナコ 映画の教科書 (1999) 岩本憲児 (他) 訳 フィルムアート社
 K・プリブラム 甘利俊一 浅田彰 (1985) 脳を考える脳 朝日出版社
 M・シオン (1993) 映画にとって音とは何か 川竹英克・J・ビノン訳 勁草書房
 M・マクルーハン (1967) 人間拡張の原理 後藤和彦・高儀進訳 竹内書店新社
 R・L・ソルソ (1997) 脳は絵をどのように理解するか 鈴木光太郎・小林哲生訳 新曜社
 R・アルンハイム (1963) 美術と視覚 波多野完治・関計夫訳 美術出版社
 R・カイエフ (1973) 遊びと人間 多田道太郎・塚崎幹夫訳 講談社文庫
 R・バルト (1984) 第三の意味 沢崎浩平訳 みすず書房
 甘利俊一 (1970) 情報理論 ダイヤモンド社
 安藤由典 (1978) 楽器の音色を探る 中公新書
 岩本憲児・波多野哲朗編 (1982) 映画理論集成 フィルムアート社
 岡田晋 (1981) 映像学・序説 九州大学出版会
 川添登 (1971) デザインとは何か 角川選書
 川野洋 (1982) 芸術・記号・情報 勁草書房
 神林恒道編 (1987) 現代芸術のトポロジー 勁草書房
 岸田秀 (1985) 希望の原理 朝日出版社
 後藤狷士編 (1979) 芸術と言語 日本文教出版社
 小松和彦 (1985) 異人論 青土社

- 齊藤勇監修・行場次朗編 (1998) 認知心理学重要研究集1 視覚認知 誠信書房
 佐藤良明 (1999) J-POP 進化論 平凡社新書
 澤口俊之 (2000) わがまな脳 筑摩書房
 杉山平一 (1981) 詩のこころ・美のかたち 講談社現代新書
 瀬名秀明 (他) (1998) 「神」に迫るサイエンス 角川書店
 地球大進化プロジェクト編 (2004) 地球大進化1〜6 NHK出版
 角田忠信 (1981) 右脳と左脳 小学館
 中井正一 (1975) 美学入門 朝日新聞社
 中島義明他編 (2004) 心理学辞典 有斐閣
 波多野諄余夫 (1987) 音楽と認知 東京大学出版会
 福井康之 (1984) まなざしの心理学 創元社
 牧田康雄編著 (1976) 現代音響学 オーム社
 正高信男 (2004) 人間の進化史 日本放送出版協会
 松岡正剛 (1996) 知の編集工学 朝日新聞社
 丸山圭三郎 (1984) 文化のフェティシズム 勁草書房
 村上和雄 (2004) 生命の暗号 サンマーク出版
 安田武・多田道太郎 (1982) 日本の美学 ぺりかん社
 山口昌男 (1975) 文化と両義性 岩波書店
 養老猛 (1989) 唯脳論 青土社
 吉武泰水 (1990) 芸術工学概論 九州大学出版会
 吉成真由美 (1986) サイエンスとアートの間に 新書館
 吉本隆明 (1982) 共同幻想論 角川文庫

おわりに

この文章は1998年、筆者が早稲田電子専門学校福岡校在籍中に、「映像音楽情報論」の講義ノートとして書きためたものをベースとしています。音楽と映像の制作に関連のある基礎事項を幅広く紹介することが目的であり、また私的なノートということもあって、首尾一貫しない文章の寄せ集めとなってしまうことは否めません。さらに、表題から期待されるであろう、作詞・作曲・編曲や映像の撮影・編集についての具体的なノウハウについては、それを整理して記述するには至りませんでした。未熟な文章をお許し下さい。いずれ機会があれば全編を再編集をした上で改めて発表したいと考えています。

本文を書くにあたっては、その知見の大半を筆者が学生時代に受講した講義に得ています。映像・画像に関しては直接の恩師である岡田晋先生・瀧山龍三先生、音楽・音響に関しては安藤由典先生・寺西立年先生、また芸術を情報の概念と関連づける視点や、芸術が文化を活性化するという視点は、吉積健先生や波平恵美子先生の講義に啓発されたものです。今は統合によってその名前が無くなってしまった九州芸術工科大学ですが、当時の「文化遺伝子」を少しでも残すことができれば幸いです。

多くの学生諸子との出会いが本文を書く動機を生み、また、彼らと共に学ぶ過程で得ら

れた様々な知見が、本文考察のヒントとなりました。この貴重な数年間を共にした、早稲田電子専門学校・麻生情報ビジネス専門学校・九州産業大学芸術学部スタッフの皆様と学生諸子に、改めて感謝の意を表したいと思います。

最後に、これまで出版に縁のなかった私にこのような機会を与えて下さった編者の瓜生隆弘氏、並びにこうした執筆活動の場を提供して頂いている九州産業大学に心から感謝いたします。そして最後の最後に、こんな部分にまで目を通してくださった読者の「あなた」に感謝いたします。

平成17年3月7日 九州産業大学17号館6階にて 筆者

II 音楽と映像

執筆者 井上貢一 いのうえこういち

1964 福岡県生

1988 九州芸術工科大学大学院芸術工学研究科

情報伝達専攻修士(芸術工学修士)

早稲田電子専門学校福岡校CG科

麻生情報ビジネス専門学校デジタルクリエータ科

他 福岡県内の専門学校講師を歴任

2004 九州産業大学芸術学部デザイン学科 講師

日本映像学会会員

日本デザイン学会会員