

[招待論文] 視覚とメディア処理  
- 画質評価と立体視に関する話題提供 -

大塚 作一<sup>†</sup>

†株式会社 NTT データ 技術開発本部 〒104-0033 東京都中央区新川 1-21-2 茅場町タワービル

E-mail: ootsukask@nttdata.co.jp

**あらまし** 画像の符号化を行ったり CG 画像を制作したりするには、人間の視覚系が実世界をどのように知覚しているかをよく理解することが重要である。本稿では、その一例として、(1)人間の視覚特性を考慮した画質の客観評価法、(2)部分的に劣化した映像の総合画質、(3)遮蔽が存在する場合の視方向、形状知覚、アライメントなど、両眼視差に基づいた立体(3D)の知覚、等に関係するいくつかの知見を述べる。

**キーワード** 画質, 評価, 視知覚, 両眼立体視, 遮蔽, 錯視

[Invited Paper] Visual Perception and Media Processing  
- Topics Concerning Picture Quality Evaluation and Stereopsis -

Sakuichi Ohtsuka<sup>†</sup>

† R&D Headquarters, NTT DATA Corporation

Kayaba-cho Tower Bldg., 1-21-2 Shinkawa, Chuo-ku, Tokyo, 104-0033 Japan

E-mail: ootsukask@nttdata.co.jp

**Abstract** It is important that we know how the human visual system perceives the real visual world in order to design video codecs or create CG images. This paper describes (1) an objective picture quality evaluation method based on visually weighted error, (2) subjective quality evaluation of locally impaired pictures, and (3) some characteristics of human visual system associated with three-dimensional (3-D) perception based on binocular disparity.

**Keywords** Picture Quality, Evaluation, Visual Perception, Stereopsis, Occlusion, Illusion

## 1. まえがき

今年はテレビ放送 50 周年にあたり、映像が巷に溢れる時代となった。「百聞は一見に如かず」の諺にもあるように、各種の視覚情報が簡単に入手できる影響には計り知れないものがある。現在の状況は、わずか半世紀前には想像もつかなかったことであり、最近のインターネットの普及に至ってはテレビジョン普及の速度の比ではない。

ただし、映像メディアを有効に活用するためには実世界から受ける刺激と人工的に作り出された刺激が視覚系に及ぼす違いに十分に注意を払う必要がある。たとえば、1997 年 12 月にテレビ放映されたアニメ番組「ポケットモンスター」視聴中の幼児・児童の多数が光感受性発作を起して医療機関を受診した事件などはその良い例であろう<sup>1)</sup>。

また、テレビ、ゲーム、携帯電話など、生まれながらにして人工的な視覚情報が溢れる環境に育つこれからの世代の人間にとっては、環境の変化に伴って今後どのような事態が発生するのが大問題である。一部では、この問題が週刊誌等でも取り上げられるなど<sup>2)-4)</sup>、身近な話題になりつつある。しかし、まだ一般の理解は十分とは言えないと考えられる。

本稿では、このような大きな問題を正面から取り扱うことは出来ないが、符号化画像の画質の評価や遮蔽

(オクルージョン)が発生した際の人間立体知覚特性について筆者が過去に検討した例に基づいて視覚とメディア処理の関係について話題提供を行うこととした。

## 2. 画像符号化にかかわる基本的な視覚特性

### 2.1. 客観評価と主観評価

符号化後の画質を客観的に計測することは、符号化特性を制御する上で重要であるが、同時に、困難な課題である。客観比較のための指標として、物理誤差に基づく PSNR (Peak Signal to Noise Ratio) が一般に用いられている<sup>5)</sup>。しかし、評価結果が主観的な品質評価とあまり一致しないことも周知の事実である。

これは、評価の観点を忠実性においた場合でも、人間が判断する画質の基準が、「視覚フィルタ」の処理結果に基づいているからである。すなわち、視覚の世界は、決して物理信号と同一ではないことに起因する。また、静止画と動画によって特徴的な視覚特性が異なることにも注意する必要がある。

### 2.2. 静止画と視覚特性

人間は、一般に静止している物体を注視している場合には、中心窩と呼ばれる網膜の狭い領域を使用している。時間的な変化や広い範囲の明るさの変化にはあ

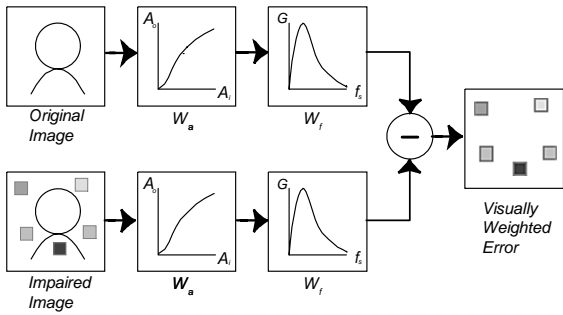


図 1 視感補正に基づく画像の客観評価の模式図

まり敏感ではないが、細部まで識別することが可能である。空間周波数の観点に立てば、バンドパスフィルタを構成している。

劣化画像と原画（参照画像）の差を、視覚を模擬したフィルタを通して客観評価する試みは古くから行われており、一般に WSNR (Weighted SNR) と呼ばれている。図 1 にその概念を示す。輝度方向のフィルタ  $W_a$  (厳密には、単に視覚的非線形性(ウエーバーの法則等)や画像信号の値のみならず、ディスプレイ表面での外光反射の影響等も考慮すべきである)と空間周波数のフィルタ  $W_f$  を用いる。被写体に依存する部分は残るが、概略的に見ると知覚的に異なる種類の劣化であっても、視感補正を行うことによって平均化して扱うことが出来る<sup>6)・8)</sup>。

なお、現在主流の離散的コサイン変換(DCT)符号化<sup>5)</sup>では、小さなブロックに分割して符号化を行うために、ブロックの境界で歪みが目立つことに留意する必要がある。この際、視覚的な劣化はブロックサイズにも大きく依存する<sup>6)</sup>。

また、このような静止物体に対する人間の視覚特性は画像処理におけるマッチングにも応用することが可能であり、ナンバープレート認識等にも応用されている<sup>9)</sup>。

### 2.3. 動画と視覚特性

一方、動いている物体や画像の時間的変化に対しては、人間はむしろ視野の周辺部分を敏感に識別する能力を有している。大面積の変化が目立ちやすく、空間周波数に関してはローパスフィルタと考えることができる。画面のちらつきとして感じられる劣化は、時間周波数で見ると一般にバンドパスフィルタ特性となる。ただし、その周波数は純粋に視覚の時空間周波数特性を計測した場合とは必ずしも一致しないので、個別の検討を要する<sup>10)</sup>。

また、滑らかに移動している被写体の知覚に関しては注意が必要である。このような場合、被写体自体は動いているが、観察者からみると静止画と同様に細部まで観察可能である。また、画面上に静止した劣化が重畳した場合には、被写体とは別の動きをするものとして知覚されるため、大きな劣化として知覚される。

さらに、画質が時間的に変動する場合の総合画質については、近似的に「最悪画質によって決定される」と考えてよく<sup>10)・12)</sup>、QoS (Quality of Service) を考える上で特に注意を要する。図 2 にその概念を示す。全体の画像シーケンスを、その時々画質を判断できる程度に区切り、各々に評点をつけたと仮定する。図 2 において、濃淡が劣化の程度を表すとすると、総合

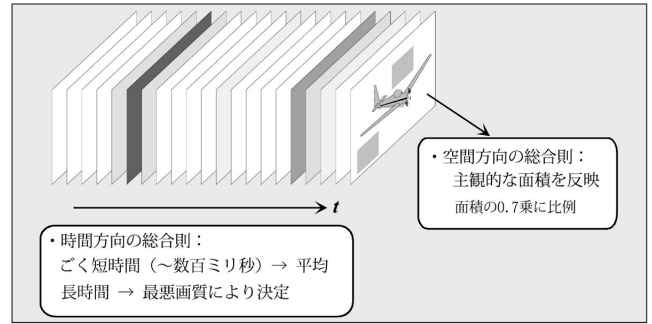


図 2 部分的に劣化した画像の概念図

画質は、濃度の平均ではなく、濃く示した時点の画質で代表されることを意味する。比較的短時間であっても、画質に破綻をきたす箇所があると、その影響は非常に大きいと言える。ただし、視覚系の応答限界よりも短時間で劣化が発生する場合や、継続的であっても高周波成分を多く含む場合には、劣化が視覚の応答特性により平均化され、その影響は緩和される方向にある<sup>10)</sup>。

なお、視覚の時空間周波数特性を加味した動画の客観評価の研究は現在も進められており、今後の進展が期待される。

### 2.4. 画角(視距離)の影響

標準方式のテレビジョン放送では、概ね一定の視距離と考えることができる。しかし、PC のディスプレイ等で画像を観察する場合には、相対的な視距離の変化が大きい。前者では 4~6H (画面高の 4~6 倍) 程度の視距離を前提とすることができる。一方、後者の場合でウィンドウ表示をする際には(画像の高さを基準に) 10H を超える視距離となる小画面の場合もある。

人間の場合、20H 程度までの領域で、画角が狭くなるに従って相対的に細部がよく見える視覚特性を有する<sup>7)</sup>。符号化にとっては厳しい特性であるので、この点に留意する必要がある。

## 3. 両眼立体視と遮蔽領域の知覚

### 3.1. 両眼立体視の基礎

人間は 3 次元空間を知覚するために両眼視差(両眼で知覚される対応点の水平位置のずれに基づいて奥行きが計算される)、輻輳(両眼の角度)、焦点調節、運動視差(物体からの距離に依存して移動に伴う相対的な位置ずれが異なる)、陰影や大きさの恒常性など多くの手がかりを用いていることが知られている<sup>13)</sup>。また、奥行き手がかりは、両眼視差などの両眼によってはじめて得られる手がかりと、運動視差など単眼からも得ることができる手がかりに大別することができる。

両眼立体視より得られる空間情報は、奥行き情報と視方向(2次元的な射影)との二つの成分に分離して考えることができる。奥行き情報を求めるための両眼視差の計算においては、両眼のパララックス(視差)が有効に働いていることは言うまでもない。一方、2次元的な射影を求めるためには、異なった位置から得られた視覚情報を統合するか、あるいは、何れの位置から得られた情報であるかのタグをつけて情報を利用する必要がある。

遮蔽が起らず 3 次元空間に物体が独立して存在する

場合、両眼から得られた視覚情報は、両眼中央の仮想の視点（サイクロプスの眼）からの画像として合成される（視方向の原理）<sup>14)</sup>。図3に示すように、各々の眼で受容された像は、相対的な角度に基づいて平均化される。

従って、遮蔽が無い場合には人間に知覚される画像は、仮想視点にカメラを置いて撮影した写真と同等と考えることができる。これに距離情報を付加すると正しい3次元空間を再構成できることになる。

複数の物体が存在する実世界を観察すると、一般に後方の物体が前方の物体の陰に隠れて見えなくなる遮蔽が発生する。我々の検討により、この場合には、人間が従来の予想とは異なった知覚をしていることが明らかとなった<sup>15)-20)</sup>。

### 3.2. 遮蔽と立体知覚

遮蔽は視点によって変化するので、両眼では遮蔽される領域が異なる。図4にその様子を示す。多少細くなるが、各々の視点からどの範囲が遮蔽されて見えなくなるか確認して頂きたい。

仮に、両眼中央の仮想視点から見た画像を両眼の画像から再現しようとする時、片眼だけに知覚される単眼情報の概ね半分を捨てる必要がある。図4-Aにおいて  $d-e$  と  $f-g$  の部分がこれにあたる。この情報を捨てることによって、矛盾なく空間を再構成できる。

しかし、人間は実際にはこの単眼情報を捨てることをせず、知覚に反映させている（なお、ここでは両眼の視差が十分に小さく、二重像や視野闘争などが起らない場合を想定している）。図5に示す例（図5-Cと図5-D）を比較することにより、人間の知覚がカメラ画像とどのように異なるかが分かる。たとえば、中央の横線上にあるドット（●）の数について注目すると、左右の眼では各々3個知覚される場合（内1個は単眼領域）統合されて人間には4個知覚されるにも拘らず、カメラ画像には2個しか存在しない（その他にもいく

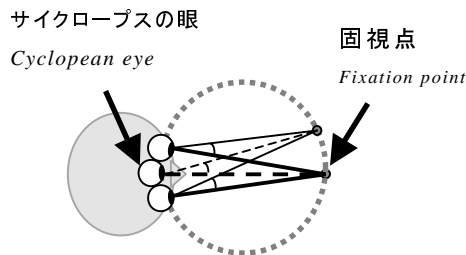


図3 サイクロプスの眼

つかの相違点があるが、紙面の関係で説明を省略する。2枚の画像を重ね合わせて比較して頂きたい<sup>19)</sup>。

視覚系が情報を捨てなかった代償はなんであるか。カメラ画像（幾何学的に正しく射影された画像）とは異なり両眼で捕えた情報を全て知覚する代償として、人間の視覚系が得た情報には、水平方向の視方向に歪みを含んでいると考えられる。換言すると、物体のアスペクト比が正しく再現されず、歪んで知覚されているはずである。

## 4. 遮蔽と立体錯視

### 4.1. Poggendorff 錯視と Kanizsa 錯視

Poggendorff 錯視と Kanizsa 錯視はよく知られた幾何学的錯視で、共に絵画的な遮蔽情報（単眼で知覚することが出来る陰影情報）を含んでいる（図6参照）。前者は、後方に隠れた斜線が上下にずれて知覚される錯視であり、後者は、後方に隠れた四角形が横方向に縮んで縦長の長方形として知覚される錯視である。これらの錯視が発生するメカニズムについては諸説があるが、これらの図形を立体で観察すると錯視量が減少することが以前から知られており<sup>21)</sup>、両者が同一の要因によって生じるのではないかという指摘もあった<sup>22)</sup>。

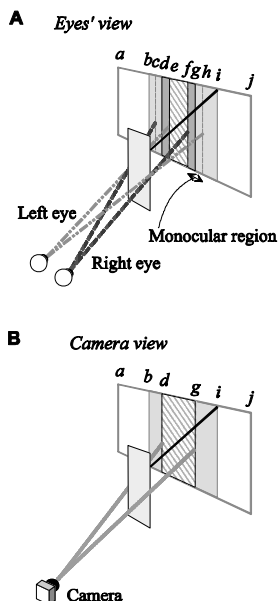


図4 視点と遮蔽領域

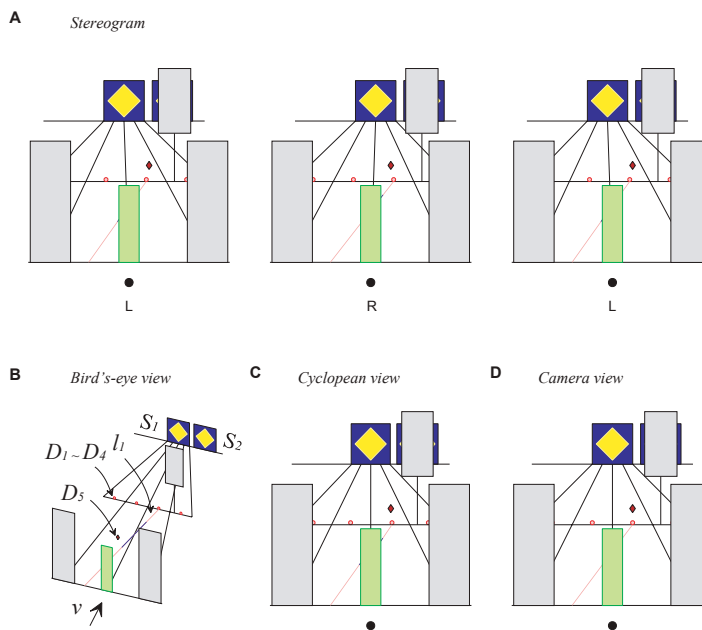


図5 人間の知覚とカメラ画像の違い

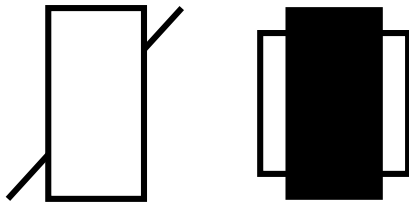


図 6 Poggendorff 錯視と Kanizsa 錯視

#### 4.2. 歪み補正機構の仮説

2次元画像ではなく実際に立体視で遮蔽が起る環境では、3.2 で述べた視方向の歪みを補正しアライメントや形状知覚を正しく知覚するための補正機構を有しているとの仮説を立てることができる<sup>15)-20)</sup>。このとき、補正の手がかりが視差情報ではなく、絵画的な遮蔽の情報(すなわち Tジャンクションの存在)によって起ると仮定すれば上記の現象をうまく説明することが可能である。

図 5-C の例に示した様に実際の知覚に基づいて描いた 2次元画像を観察すると、中央下部の Poggendorff 錯視に対応する部分や右上部の Kanizsa 錯視に相当する部分で錯視が感じられないことが分かる。

このような現象は、人間の視覚系が情報を取捨選択する際に実世界では希にしか発生しない問題を犠牲にすることによって相対的に効率的な情報を取得できるように進化してきたことを示唆する。そのことは、時として我々が常識的な理論に基づいて予測する情報取得とかなり異なった結果を生み出していることを意味する。

実は、画家はこのような状況を Poggendorff 錯視として発見される約 250 年も前から理解していたことが明らかになっている。Rubens のキリスト降架 (The Descent from the Cross) がその証拠である<sup>23)</sup>。「カメラ画像が『真』を写すか?」という問いに対して、人間の感性からみると明らかに異なるという証としてとることができる。

#### 5. むすび

視覚は映像メディアの処理と密接な関連を有している。本稿では、画質評価や両眼立体視に関する人間の視覚特性についていくつかの知見を述べた。人間の視覚は、光学的に取得される物理情報とはかなり異なった情報処理を行っている一端を理解していただければ幸いである。なお、紙面の関係でこのような立体視覚特性をディスプレイ技術に応用する点については触れることが出来なかったが、その点については別稿を参照頂きたい<sup>24)</sup>。

映像メディアが広く利用される時代において、メディア工学に従事する者として、今後、人間の視覚特性を十分理解することの重要性を再認識して行きたい。

#### 文 献

[1] 例えば、厚生省：“平成 9 年度厚生科学特別研究「光感受性発作に関する臨床的研究」”，<http://www1.mhlw.go.jp/shingi/s1226-3.html> (Dec. 1997)

- [2] 山下：“シリーズ<五感&警告する>第 5 弾 視覚編 仮想現実が「視覚世界」を変えていく”，ヨミウリウイークリー，*62*，38，pp. 90-93，(Sept. 7, 2003)
- [3] 山下：“同上「トリックアート」の不思議と視覚依存社会の危険性”，ヨミウリウイークリー，*62*，39，pp. 78-81，(Sept. 14, 2003)
- [4] 山下：“同上 闇夜の世界が「五感」と「身体」を拓く瞬間”，ヨミウリウイークリー，*62*，40，pp. 92-95，(Sept. 21, 2003)
- [5] 例えば、酒井，吉田：映像情報符号化，オーム社 (2001)
- [6] 大塚，井上，渡辺：“輝度・空間周波数特性を考慮した画質評価法”，信学論 A，*J-70A*，11，pp. 1659-1666 (1987)
- [7] 大塚：“客観評価技術の動向”，テレビ学誌，*45*，3，pp. 293-298 (1991)
- [8] 半谷，和田，宮内：“静止画の画質を評価するための空間周波数領域上の重み付け関数の実験的導出”，テレビ学誌，*46*，3，pp. 295-299 (1992)
- [9] 高橋，田中，鈴木，塩，大塚：“対数輝度変換と空間周波数強調を適用したマッチとフィルタによるナンバープレート認識法”，映情学技報，*27*，35，pp. 9-12，ME2003-104 (June 2003)
- [10] 大塚：“部分的に劣化した画像の品質に関する一検討”，テレビ学技報，*12*，38，pp. 13-18，VVI88-36 (Oct. 1988)
- [11] S. Ohtsuka, M. Kosugi: “Quality evaluation of locally impaired pictures”, Proc. of the SID, *Vol. 32/1*, pp. 19-24 (1991)
- [12] 藁科，稲積，吉田：“最適フレームレート推定に基づく動画のトランスコード”，映情学技報，*26*，81，pp. 17-20，ME2002-88 (Dec. 2002)
- [13] 例えば，原島監修：“3次元画像と人間の科学”，オーム (2000)
- [14] Ono, H.: “Binocular visual directions of an object when seen as single or double”, In Regan, D. (Ed.), Vision and visual dysfunction, *9*, pp. 1-18, Macmillan, London (1991)
- [15] 大塚ほか：“両眼立体視に基づく 2次元画像生成方法”，特願平 2-122928 (1990)
- [16] 大塚，矢野：“立体画像の知覚はずみと Poggendorff 錯視”，テレビ学技報，*18*，60，pp. 25-30，VAI94-50/ IPCV94-60 (Oct. 1994)
- [17] 大塚：“遮蔽のある立体観視条件下での視方向の知覚”，信学技報，pp. 31-36，PRU95-84/ HIP95-12 (July 1995)
- [18] Ohtsuka, S. et al.: “Kanizsa’s amodal figure made with da Vinci stereopsis”, ECVF 99 B034, Perception, *28*, Supplement, p. 118 (1999)
- [19] Ohtsuka, S., Ono, H.: “Correct 3-D but illusory 2-D percepts linked to binocularly unpaired regions”, Proc. of SPIE, *Vol. 4864*, pp. 167-174 (Aug. 2002)
- [20] Ono, H. et al.: “Leonardo’s Constraint: Two opaque objects cannot be seen in the same direction”, Journal of Experimental Psychology: General, *132*，2，pp. 253-265 (2003)
- [21] Gyoba, J.: “The Poggendorff illusion under stereopsis”, Tohoku Psychologica Folia, *37*, pp. 94-101 (1978)
- [22] Zanuttini, L.: “A new explanation for the Poggendorff illusion”, Perception and Psychophysics, *20*，1，pp. 29-32 (1976)
- [23] Topper, D.: “The Poggendorff illusion in Descent from the Cross by Rubens”, Perception, *13*，6，pp. 655-658 (1984)
- [24] 大塚，陶山，高田，石樽：“両眼立体視に基づく奥行き知覚と立体表示 - 奥行き融像立体表示 (DFD 表示) や仮想画枠 (VFM) を例として - ”，信学技報，pp. 37-42，EID2000-51 (Sept. 2000)