

■2 群 (画像・音・言語)－5 編 (画像符号化)

9 章 符号化画質

■2群 - 5編 - 9章

9-1 符号化雑音の種類

(執筆者：杉本 修) [2013年4月 受領]

ディジタル画像信号の情報量は膨大であり、これを何の処理を施さず蓄積・伝送することは実用の観点から現実的ではない。よって、予測符号化の適用により時間・空間的に分布する冗長な信号成分の情報量を削減し、更に視覚的に検知されにくい信号成分については、周波数成分に対する量子化の適用などによる非可逆的な情報削減を行うことで、画像の蓄積・伝送に適するサイズ・帯域を実現している。このうち、後者の非可逆的な処理は圧縮前の原画像には存在しない信号成分を付加することとなり、いわゆる符号化雑音の原因となる。

符号化雑音は、視覚上目立たないよう配慮されているとはいえ、圧縮率が高くなった場合には検知されやすいという特性をもつ。また、符号化雑音の見え方は一通りではなく、原画像のもつテクスチャや動きの複雑さなどの特性によって種々の劣化が知覚されることが知られている。本節では、圧縮符号化により生じる雑音のうち典型的なものを紹介する。

9-1-1 ブロックノイズ

ブロックノイズは、圧縮符号化の制御単位となっている画素ブロックの境界で輝度値の不連続が生じ、その結果、ブロック外周部が際立ち矩形のアーティファクトとして知覚される空間的な劣化要素である。DCT符号化で各画素ブロックに適用される量子化処理においては、視覚感度の低い高周波数成分を削減し低周波数成分を保存するように量子化マトリクスが設定されている。しかし、圧縮率が高くなると情報量の削減のため順次低い周波数成分も削減されることになり、これによりブロック境界に沿った領域で信号値に大きな不連続が発生するようになり境界が検知されるようになる。また、画面間の予測を適用しているフレームにおいて隣接ブロック間で異なる動きベクトルを用いる場合、残差画像がブロック境界で不連続になることも、その後の量子化処理とあわせて境界が知覚されやすくなる原因と考えられる。また、予測符号化では、特に圧縮率の高い場合にブロックひずみをもつ画像領域を参照することによるひずみの伝搬の影響も指摘されており、こうした原因が複合して発生すると考えられる。

図 9・1 にブロックひずみの発生例を示す。MPEG-2 適用時のブロックは、H.264 のそれに比べると境界部がシャープであり境界部に沿った直線上のアーティファクトは目立って見える。



(a) 原画像

(b) MPEG-2 適用時

(c) H.264 適用時

図 9・1 ブロックノイズの発生例

一方、H.264 適用時には、後述するデブロッキングフィルタの効果で境界部が先鋭に見えることはないものの、ブロックの内部がぼやけている領域では大きな矩形の領域が知覚されることとなる。

9-1-2 リンギング、モスキートノイズ

リンギングは、粗い量子化により高周波数成分の打ち切られた場合に、原画像のエッジ周辺にいわゆる Gibbs の現象が発生し、高振幅のアーティファクトが観察される現象である。図 9・2 にリンギングの発生例を示す。同図では、植物の穂の周辺部にもやもやした雑音が発生しているのが確認できる。エッジに沿って蚊の集団が飛んでいるような見え方をするため、モスキートノイズとも呼ばれる。放送・映像通信向けのフォーマットでは、MPEG-2 符号化を適用する際に多く観測される劣化である。

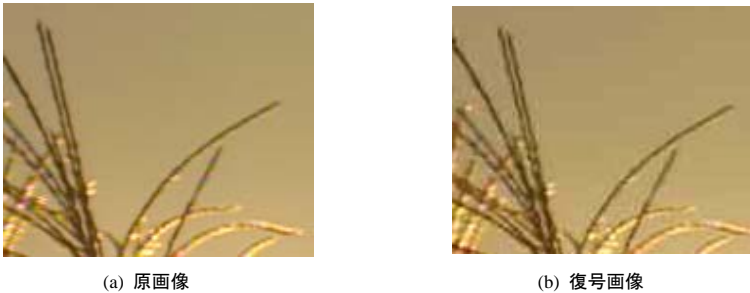


図 9・2 リンギングの発生例

9-1-3 ぼやけ

ぼやけは、図 9・1 からも観察されるように H.264 を適用した際によく観測される劣化であり、原画像に存在していたテクスチャの解像感が失われる現象である。ぼやけの発生は量子化による高域成分の消失に起因するが、もうひとつの原因として、H.264 で採用されたデブロッキングフィルタの影響が考えられる。デブロッキングフィルタは、符号化ループ内で適用される処理であり、予測画像に存在するブロックひずみの伝搬防止を目的にブロック境界に適用される低域通過特性をもつ画像処理である。フィルタ適用の有無やその強度は、ブロックひずみの強度に応じて設定されるため、圧縮率が高い場合解像感が失われる傾向がある。

9-1-4 フリッカ

フリッカとは、映像の明るさが短時間のうちに急激に変化することにより、ちらつき感が知覚される現象である。人間の視覚は、空間方向に比べて時間方向の劣化の知覚感度が低い傾向にある。そこで、動画像符号化のレート制御ではピクチャタイプ (H.264 ではスライスタイプ) ごとに異なる圧縮率を設定し、それらを時間方向に組み合わせることにより劣化を知覚させることなく情報量の削減を行っている。

一般に、フレーム間予測で多く参照されることとなる I ピクチャにおいては劣化の伝搬を防ぐため圧縮率を低く設定する一方、P、B ピクチャでは I ピクチャに比べて高い圧縮率を設

定する。ここで、Iピクチャと隣接するP、Bピクチャとの画質差が大きい場合にフリッカが知覚されやすくなる。これは、P、Bピクチャは高い圧縮率が設定されるためフレーム内の高域の周波数成分が失われているのに対し、Iピクチャは高域成分を比較的多く保持するため、Iピクチャ挿入の都度この高域成分の差が画面の輝度変化として現れ、それがちらつきとして知覚されるためである。

9-1-5 偽輪郭

偽輪郭とは、滑らかな階調変化をもつ信号が量子化され、その滑らかさが失われることにより粗い階調変化となり、原画像には見られなかった階段状のアーティファクトが発生する現象である。ブロックひずみの発生原理と同様に、周波数成分の量子化により高周波数成分が失われることにより信号の連続性が失われ、階調変化の再現が不十分になることが主な原因であると考えられる。偽輪郭の発生例を図9・3に示す。原画像では、画面中央にかかる雲が右側に向かって薄くなっているが、復号画像では、右下方向に等高線状の輪郭が発生していることがわかる。



(a) 原画像



(b) 復号画像

図9・3 偽輪郭の発生例

■2群 - 5編 - 9章

9-2 画質の客観評価

(執筆者：杉本 修) [2013年4月 受領]

客観評価とは、圧縮符号化における劣化程度を信号解析により定量的に表現することを意味する。この定量的評価が「客観」と形容されるのは、後述する主観評価と対をなす評価手法であるためである。客観評価の特徴として、数値計算により一意に評価値を求めることが可能であるため、主観評価に比べると導出が比較的容易であり、また、同一の画像に対する評価値の再現性が保証されている点があげられる。一方で、客観評価を主観評価の代替として考えた場合、人間の知覚心理特性を完全に反映させることが困難であることから、場合によっては主観評価値からの乖離が大きくなる（主観評価値との相関が低下する）点が課題とされている。

以下に符号化品質評価の目的でよく用いられる客観評価尺度を示す。

9-2-1 SNR, PSNR

圧縮符号化にともない発生する雑音の評価においても、信号対雑音比（Signal to Noise Ratio : SNR）はよく用いられる。原画像の信号分散を s^2 、雑音電力を e^2 とすると、SNRは、

$$SNR = 10 \log_{10} \frac{s^2}{e^2} \quad (9 \cdot 1)$$

により定義される。ここで、

$$e^2 = \frac{1}{N} \sum_i (x[i] - y[i])^2 \quad (9 \cdot 2)$$

であり、 $x[i], y[i]$ はそれぞれ原画像及び復号画像の i 番目の画素位置における信号値、 N は計算対象の画素数であるとする。符号化品質評価の観点では、 $x[i], y[i]$ は主観画質に強く寄与する輝度成分の信号値とするのが通例である。上式は原画像と復号画像の差分二乗平均であり、MSE (Mean Squared Error) と呼ばれる。

また、上式において原画像の信号分散ではなくピーク-ピークの輝度変化と MSE の比に基づいた PSNR (Peak SNR) もよく用いられる。輝度信号が 8 ビットで表現される場合、輝度信号は 0~255 の範囲で表現されるため、 $s = 255$ とし、

$$PSNR = 10 \log_{10} \frac{255^2}{e^2} \quad (9 \cdot 3)$$

により定義される。動画像に対しては各フレームに対して PSNR が得られるため、その PSNR 平均値をその動画像の評価値とする。

式(9・1)及び式(9・3)において、信号分散または最大振幅と雑音の電力比の対数を求めているのは、信号のダイナミックレンジが広いためデシベル (dB) 単位で表現することを意図したものであるが、人間の視覚がウェバー-フェヒナー則に従うことから、対数化された MSE が主観画質によりよく一致するという効果も同時にもたらしている。特に、PSNR が高い、すなわち劣化程度の低い画像に対しては、PSNR は主観画質とよく相関することが知られている。対して、PSNR が著しく低い画像や画面内で品質に偏りがある（局所的に著しい劣化がある）場合、主観画質との乖離が生じる場合があることが知られている。また、PSNR は基

本的にはフレーム単位の品質を示すものであり、動画像に対しては単純な加算平均によりその画像の評価値を表すため、時間的な品質変動の大きい動画像に対しては主観画質との乖離を示す場合がある。しかし、MSEがMPEG-2, H.264などの主要な動画像符号化方式におけるレート-ひずみ最適化のひずみの指標に用いられていることから、符号化効率を図る用途としてもPSNRの有効性は高く、また複雑な計算を要しないことから、PSNRは客観評価尺度として現状最もよく使われている。

9-2-2 SSIM

上述のとおり、PSNRは主観評価の代替としては万能とはいき切れない。よって、主観画質との相関の観点で、PSNRの性能を上回る客観評価法が求められている。SSIM (Structural SIMilarity) はこうした背景からWangらによって提案された指標で、PSNRが符号化にともない生じた雑音成分の知覚感度に基づく指標であるのに対し、画像構造の類似度が人間の画質劣化の知覚に寄与するとの仮説に基づいて定義されている。画像構造の類似度には、ピアソンの相関係数を用いており、これに原画像及び復号画像の輝度・コントラストの比較項を乗じ評価値を構成している。原画像 $x[i]$ 、復号画像 $y[i]$ の画像構造の類似度項を $s(x, y)$ 、輝度及びコントラストの比較項をそれぞれ $l(x, y)$, $c(x, y)$, $s(x, y)$ とすると、SSIMは、

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \quad (9 \cdot 4)$$

により与えられる。 α, β, γ は、正の定数であり、文献では $\alpha = \beta = \gamma = 1$ とされている。また、 $l(x, y)$, $c(x, y)$, $s(x, y)$ は、

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (9 \cdot 5)$$

である。ここで、 C_1 及び C_2 は L を信号の最大振幅 (8ビット表現であれば $L = 255$)、 K_1, K_2 を $K_1 \ll 1, K_2 \ll 1$ なる定数とすると $C_1 = (K_1 L)^2$, $C_2 = (K_2 L)^2$ であり、 C_3 は計算の簡略化のため $C_3 = C_2/2$ とされる。なお、文献においては、 $K_1 = 0.01, K_2 = 0.03$ が好適であるとされている。また、 $\mu_x, \sigma_x, \sigma_{xy}$ は 11×11 画素の正規化された円対称ガウス関数 $w[i]$ により重みづけられた原画像 $x[i]$ の平均、標準偏差及び $x[i], y[i]$ の共分散であり、

$$\mu_x = \sum_i w[i] \cdot x[i], \quad \sigma_x^2 = \sum_i w[i] (x[i] - \mu_x)^2, \quad \sigma_{xy} = \sum_i w[i] (x[i] - \mu_x)(y[i] - \mu_y) \quad (9 \cdot 6)$$

と定義されている。なお、 μ_y, σ_y は復号画像 $y[i]$ に対して上記と同様の定義を適用したものである。最終的に、式(9.4)及び式(9.5)を統合することにより、SSIMは以下のように定義される。

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (9 \cdot 7)$$

μ_x, σ_x などの計算にあたって、有限の領域の重み係数 $w[i]$ が適用されているため、式(9.7)の評価値は、実際にはある小領域ごとに定義されることになる。この小領域に属する原画像・復号画像を $x_j[i], y_j[i]$ とし、この小領域ごとの評価値を $SSIM(x_j, y_j)$ とすると、画面全体の評価値は各領域の評価値の平均により与えられ、これをMSSIM (Mean SSIM) と呼ぶ。

$$MSSIM(x, y) = \frac{1}{M} \sum_{j=1}^M SSIM(x_j, y_j) \quad (9 \cdot 8)$$

通常、SSIMとは、上式に基づく評価値のことを意味する。SSIMは主観画質との相関の観点でPSNRより優れる場合があると報告されている。Wangらは、これは画像構造の類似度が人間の画質劣化の知覚程度と相関するという自身らの仮説が成立するためであると主張するが、知覚心理分野においてこの関係を明示する知見はいまだ得られていない。むしろ、小領域ごとに重みを適用する手法や、極めて品質が高い/低い領域において評価値が飽和するような設計を適用した点が、PSNRに対する優位性に寄与しているものと考えられる。

9-2-3 客観評価方式のフレームワーク

主観評価における一重刺激・二重刺激の分類と同様に、リファレンス画像の有無によって3種のフレームワークが定義されている³⁾。この三つのフレームワークを図9・4に示す。

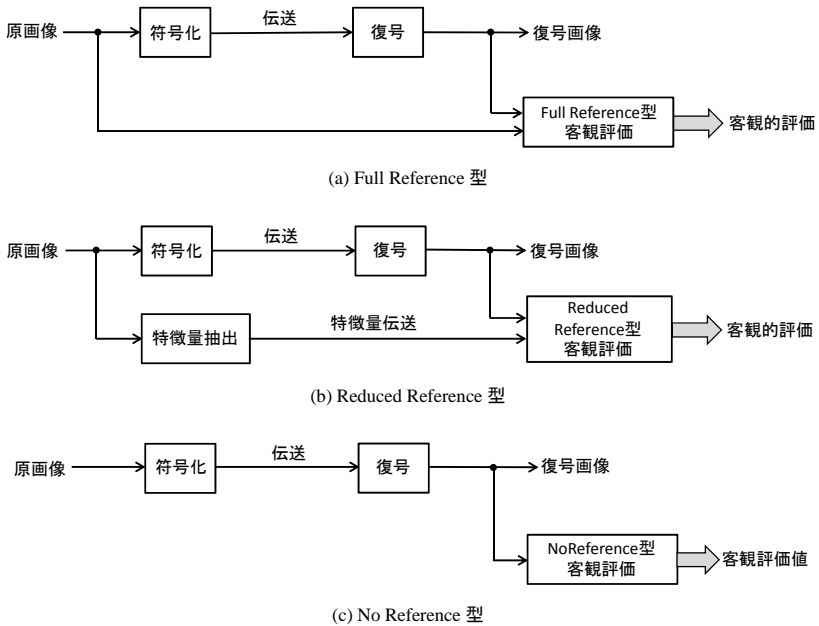


図9・4 客観評価方式のフレームワーク

Full Reference (FR) 型は、原画像・復号画像の両方を参照するフレームワークである。符号化前後の画像比較により圧縮及び伝送により生じた誤差を正確に把握することが可能であり、図中の3フレームワークの中で最も高い評価精度を実現する。上述のPSNR, SSIMはこのFR型の一つに分類される。ただし、原画像と復号画像を同時に参照するため、遠隔地または受信地点での品質モニタリングには適さず、試験環境におけるコーデックの性能評価などに用いられる。

No Reference (NR) 型は、復号画像のみを参照するフレームワークである。原画像の情報

を必要とせず、シングルエンド型の構成をとることが可能となるため、インサービスの品質モニタリングに適している。一方で、復号画像のみを参照するため、原画像が本来有するテクスチャや動きの特性と符号化劣化により生じたアーティファクトの判別が困難となり、それゆえ評価精度の面でFR型評価には及ばないという制約がある。

Reduced Reference (RR)型は、復号側の構成は上述のNR型と同じ構成であるが、原画像から抽出した画像特徴と合わせて評価する点がNR型とは異なる。この構成は映像伝送のインサービス監視において、映像回線とは別に設けた低速のデータ回線で原画像の映像特徴を復号側に送信することにより、原画像を参照できないNR型の精度の問題を補うことを意図したものである。データ回線の伝送帯域は数10k～数100kbps程度が想定されているため、限られた情報量で画像特徴を表現することが求められる。

9-2-4 客観評価方式の国際標準

本節では、客観評価技術の例として、ITUにより標準化された客観画質評価方式について概観する。ITUで画質評価技術の標準化を行う際は、画質評価技術の専門家グループVQEG(Video Quality Experts Group)に複数機関からの提案方式の性能評価を行う方式評価テストの実施を依頼し、そこで所定の水準の性能を達成した方式を勧告として採用している。所定の水準とは、他の提案方式に対する精度の優位性のほか、PSNRに対する優位性も含まれている。表9-1に、各映像フォーマットとフレームワークの組合せに対する標準方式の一覧を示す。

表9-1 各フレームワーク、フォーマットに対する客観評価の標準方式

Table with 4 columns: Reference Type, Multimedia (QCIF/CIF/VGA, 3~30fps), SDTV (525/60, 625/50), and HDTV (1080/50i, 60i, 25p, 30p, 720/50p, 60p). Rows include Full Reference, Reduced Reference, and No Reference.

表9-1の各方式は、符号化劣化のほか伝送エラー劣化(パケットロスなどの伝送路障害による復号失敗にともなう再生画像の劣化)の評価にも対応しているが、J.144は符号化劣化のみを対象としている。J.144もVQEGによる方式評価テストの段階では伝送エラー劣化の評価をスコープに含んでいたが、性能評価の結果PSNRの性能を超える方式が出現しなかったため、スコープを符号化劣化評価のみに変更した経緯がある。NR型の標準が存在しないのも同様の理由で、過去にMultimedia, HDTVの各フォーマットに対しては提案があったものの、PSNRの性能には届かなかったとの理由でこれまで勧告化が見送られている。NR型評価に関しては、リファレンス画像を参照しないという制約があるため、幅広い圧縮符号化及び伝送エラー条件においてPSNRと同等の性能を達成することに対する技術的な困難さはFR/RR型に比べて高いと考えられる。しかし、前述のとおりNR型はその構成の自由度の高さから多様な品質モニタリングへの応用が期待されており、こうした課題を克服した標準方式の確立が望まれている。

■2群 - 5編 - 9章

9-3 画質の主観評価

(執筆者：杉本 修) [2013年4月 受領]

客観評価が圧縮符号化により生じた雑音成分を所与の計算基準に基づき定量的に表すのに対し、主観評価は、評価対象の画像を見た人間が自身の感覚に基づき評価値を決定する手法である。主観評価は人間による評価であるため、その評価値は客観評価値よりも実際のアプリケーションにおけるサービス品質により高く相関すると考えられる。ただし、主観評価値は人間の知覚心理特性や個人の評価基準の差異に起因する評価値のばらつきがあり、同一の評価画像に対する結果の再現性を保証するためには相応の数の被験者による評価を行い、更に適切な統計処理の適用が求められるなど、客観評価に比べて評価の実施にあたっての要件は厳しいと言える。

主観評価法は画像提示の方法により二重刺激法と一重刺激法に大別される。二重刺激法では、評価対象となる画像(評価画像)に加えて品質の基準となる画像(リファレンス画像)を対にして提示する手法であり、両者の品質差を評価尺度として導出する。なお、符号化画質の評価の目的では、評価画像には符号化処理適用後の復号画像を、リファレンス画像には符号化前の原画像を通常使用する。対して、一重刺激法は評価画像(復号画像)のみを提示し、その品質を評価尺度とする手法である。両者を比較すると、二重刺激法は評価対象となる復号画像に対応する原画像を毎回提示する必要がある点が一重刺激法とは異なる。このように、復号画像と原画像を対にして提示することにより、主観評価値に両者の相対的な品質差が反映されるため、一重刺激法に比べて被験者間のばらつきの少ない安定した主観評価値を求めることが可能となる。一方、二重刺激法においては、同一の原画像に対して複数の符号化条件を適用するような場合であっても都度その原画像を提示する必要があり、一重刺激法に比べて映像提示の総数が増える、すなわち被験者の負担が大きくなり実験設計の難易度が上がるという課題もある。よって、実験設計の難易度及び得られる結果の精度に対する要件を勘案して適切な主観評価法を選択する必要がある。

主観評価実験を実施するにあたっての視聴環境、映像の提示順、評価尺度の定義などの具体的な実験手順は、ITUにより勧告化されている。以下に、符号化画像の画質評価によく使われる主観評価法を示す。

9-3-1 DSCQS 法

DSCQS (Double Stimulus Continuous Quality Scale ; 二重刺激連続品質尺度) 法は、ITU-R 勧告 BT.500-13¹⁰⁾ で定義されている二重刺激法の一つである。勧告には 2 種類の実施形態 (Variant I 及び Variant II) が示されているが、本節では、符号化品質の評価によく用いられる Variant II について記す。DSCQS 法における画像提示順を図 9・5 に示す。ここで、リファレンス画像と評価画像のうち、先に表示するものを画像 A、後に表示するものを画像 B としている。同図のとおり、A, B, A, B と各画像がグレー画像を挟みながら 2 回ずつ繰り返し提示される手順となっている。リファレンス画像、評価画像と画像 A, B の対応は、画像番号(実験において、評価画像を一意に識別するための番号)の提示ごとにランダムに入れ替えることができる。すなわち、DSCQS 法は主観評価において被験者に画像 A, B のどちらがリファ

レンズであるかを明示しないことを意図している。図中のT1~T4は画像の提示時間であり、動画画像の場合 T1=T3=10〔秒〕程度が望ましいとされている。T2, T4は、画像A, Bの間に挿入するグレー画像の表示時間であり、T2=3〔秒〕及びT4=5~11〔秒〕とされている。T4は、評点の記入の時間を確保する意味で、T2より長い時間が設定されている。ただし、評点の記入は2回目の画像A, Bの繰り返しの後から開始することが許容されているため、T4の値はT1~T3の時間を鑑みて設定する必要がある。

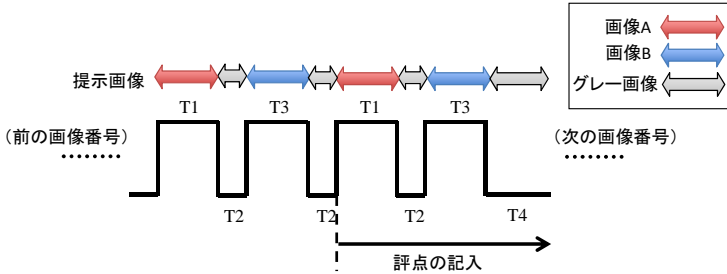


図 9・5 DSCQS 法における画像提示順

被験者はリファレンス画像と評価画像のそれぞれの品質を 0~100 の範囲の評点で評価する。評価の記入は、図 9・6 のような様式の採点用紙を用いる。評価用紙には 5 等分された縦線の評価軸が表 9・2 に示す 5 段階の品質尺度（英語表記は勧告に記されている評価語）とともに示されており、被験者はこれらの評価語を参照しながら、自身の評価に最も近い評価軸上の位置に横線を引いて評価の記入を行う。評価は 5 択ではなく、評価軸の下端からの距離と評価軸の全長の割合を 0~100 の評点に変換する。最後に、リファレンス画像と評価画像の評点の差分を求め、当該被験者の評価値とする。一般に、DSCQS 値とは、この各被験者の評価値の平均のことを指す。

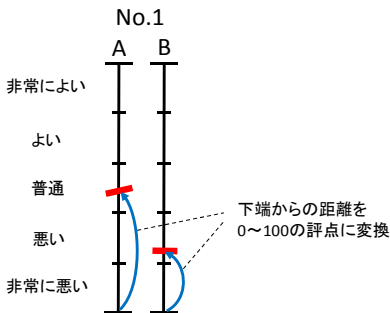


図 9・6 DSCQS 法における評価の記入

表 9・2 DSCQS 法で用いられる品質尺度

非常によい	Excellent
よい	Good
普通	Fair
悪い	Poor
非常に悪い	Bad

9-3-2 DSIS 法

DSIS (Double Stimulus Impairment Scale ; 二重刺激劣化尺度) 法は、DSCQS 法と同じく勧告 BT.500-13 に定義されている二重刺激法の一つであり、EBU 法とも呼ばれている。DSIS

法についても画像提示の方法によって Variant I, II の 2 形態があるが、ここでは Variant II について記す。

画像提示の方法は、図 9・5 に示した DSCQS 法の提示順と同じく、リファレンス画像と評価画像を順に 2 回ずつ提示する手法をとる。ただし、DSIS 法では必ずリファレンス画像、評価画像の順に提示することが定められている。その後、リファレンス画像に対する評価画像の品質を以下に示す 5 段階の劣化尺度により評価する。

表 9・3 DSIS 法で用いられる劣化尺度

5.	劣化がわからない	imperceptible
4.	劣化がわかるが気にならない	perceptible, but not annoying
3.	劣化が少し気になる	slightly annoying
2.	劣化が気になる	annoying
1.	劣化が非常に気になる	very annoying

DSCQS 法では被験者は二つの画像それぞれについて評価を行うのに対し、DSIS 法は一つの評価画像に対して一つの評価値のみを回答することになる点が異なる。また、DSIS 法においては、各評価語の間を選ぶような評価は認められておらず、5 段階の評価語のいずれかを選択する点も DSCQS 法とは異なる。一般に、DSIS 法はリファレンス画像を明示し、かつリファレンスとの間のわずかな品質差を問う品質尺度を用いているため、評価画像との品質差が小さい場合の評価に効果的であると考えられる。

9-3-3 ACR 法

ACR (Absolute Category Rating) 法は、ITU-T 勧告 P.910¹¹⁾ で定められる一重刺激法である。ACR 法における画像提示順を図 9・7 に示す。一重刺激のため評価画像のみを提示する点、評価画像の提示を繰り返さない点が前節に記した二重刺激法とは異なっている。画像の提示時

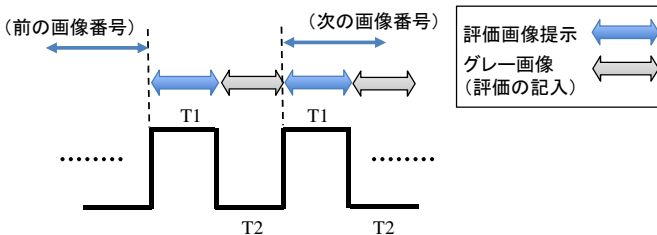


図 9・7 ACR 法における画像提示順

表 9・4 ACR 法で用いられる品質尺度

5.	非常によい	Excellent
4.	よい	Good
3.	普通	Fair
2.	悪い	Poor
1.	非常に悪い	Bad

間は、 $T_1 = 10$ [秒]、 T_2 は 10 秒以内とされている。被験者は、図 9・7 の T_2 (グレー画像表示) の期間に表 9・2 に示した 5 段階の品質尺度に基づき評価を記入する (DSIS と同じく、評価語間の中間の評価は認められない)。この品質尺度は、DSCQS 法で用いているものと同じであるが、評価語に 1~5 の評点を示す数値を添えて表示する点のみが異なる (表 9・4)。各被験者が与えた評価値の平均が MOS (Mean Opinion Score) 尺度となる。

9-3-4 ACR-HR 法

ACR-HR (Absolute Category Rating with Hidden Reference) 法は、画像提示や評価の方法は ACR 法に準ずるが、実験中に評価画像に加え、対応するリファレンス画像も提示するよう設計された手法である。実験においてリファレンス画像の存在は被験者には明かされず、評価画像に混じってランダムな順序で提示されるため、Hidden Reference と呼ばれる。

主観評価尺度 DMOS (Difference MOS) は、評価画像と対応するリファレンス画像の MOS の差分に基づき求められる。評価画像及びリファレンス画像の MOS をそれぞれ MOS (PVS)、MOS (REF) とするとき、

$$DMOS = MOS(REF) - MOS(PVS) + 5$$

により与えられる。MOS の差分に 5 を加えているのは、表 9・4 の評価語と評点の関係 (評点が高いほど品質が高い) に合わせるためである。DSCQS 法と比較すると、リファレンス画像を評価画像に都度対応させて提示する必要がないため、主観評価実験に要する時間を大幅に低減できる点が利点としてあげられる。ただし、ACR-HR 法ではリファレンス画像の評点が不正確な場合、対応するすべての評価画像の DMOS の精度に影響を与えるため、不正な主観評価データを除外するスクリーニング処理などをより慎重に行う必要がある。

■参考文献

- 1) N. Hara, A. Ichigaya, M. Kurozumi, Y. Nishida, and Y. Ohtsuka, "Flicker reduction in MPEG-2 video by post-processing," IEEE Trans. on Consumer Electronics vol.55, pp.210-217, 2005.
- 2) Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Trans. on Image Processing, vol.13, no.4, pp.600-612, Apr. 2004.
- 3) ITU-T Recommendation J.143, "User requirements for objective perceptual video quality measurements in digital cable television," May 2005.
- 4) ITU-T Recommendation J.247, "Objective perceptual multimedia video quality measurement in the presence of a full reference," Aug. 2008.
- 5) ITU-T Recommendation J.246, "Perceptual visual quality measurement techniques for multimedia services over digital cable television networks in the presence of a reduced bandwidth reference," Aug. 2008.
- 6) ITU-T Recommendation J.144, "Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference," Mar. 2004.
- 7) ITU-T Recommendation J.249, "Perceptual video quality measurement techniques for digital cable television in the presence of a reduced reference," Jan. 2010.
- 8) ITU-T Recommendation J.341, "Objective perceptual multimedia video quality measurement in the presence of a full reference," Jan. 2011.
- 9) ITU-T Recommendation J.342, "Objective multimedia video quality measurement of HDTV for digital cable television in the presence of a reduced reference signal," Apr. 2011.
- 10) Recommendation ITU-R BT.500-13, "Methodology for the subjective assessment of the quality of television pictures," Jan. 2012.
- 11) ITU-T Recommendation P.910, "Subjective video quality assessment methods for multimedia applications," Apr. 2008.