

第8章 目視による主観評価と 定量的な客観評価の方法

デジタル・ビデオ画像の 評価の実際

川田 亮一
Ryoichi Kawada

デジタル・ビデオ信号には、アナログ信号の場合とは異なったノイズが乗ります。この章では、このデジタル・ビデオ特有のノイズの種類とその評価方法について説明します。

● 主観評価と客観評価がある

デジタル・ビデオの画質の評価には、**主観評価**と**客観評価**があります。主観評価とは、人間の目で実際に見たときの評価値で、ITU(国際電気通信連合)で標準の評価法が定められています。ITU-R BT.500が相当します。

客観評価とは、人間ではなく機械で自動的に評価値を計算することです。なるべく人間の主観に近い値を出すように、いま盛んに研究開発が進められています。

主観評価の実際

● **テレビ用かマルチメディア用かで評価の条件が違う**
ひとことでデジタル・ビデオといっても、普通のテレビで見る場合もあれば、パソコンで映す場合もあります。さらに、携帯電話で動画を見ることもできます。

見る環境も、それぞれでかなり違うでしょう。テレビなら家の中で、数m離れて見るでしょうし、携帯電話なら場所は問いませんがせいぜい30cmぐらいの距離から見ることになります。

このように従来のテレビと最近のマルチメディアの映像では、主観評価の条件が大きく異なります。ITUで勧告されている主観評価実験方式も、**従来はテレビ、それもアナログ伝送信号を対象として考案されたものでした。**

しかし現在では、**デジタル圧縮の性質やマルチメディア視聴環境を考慮したものができつつあります。**これらについて以下で説明します。

● 二つの主観評価法…2重刺激と単一刺激

国際電気通信連合無線通信部門(ITU-R)の勧告

BT.500⁽¹⁾では、テレビの主観評価実験の条件を定めています。これには枠組みとして大きく分けて2種類あります。

評価したい画像と比較対象画像を比較しながら評価値を出す「2重刺激」方式である**DSCQS**(Double-Stimulus Continuous Quality-Scale, 2重刺激連続品質評価尺度)法や**DSIS**(Double-Stimulus Impairment scale, 2重刺激劣化尺度)法と、評価したい画像だけを提示して評価値を出す「絶対評価」方式である**SSCQE**(Single Stimulus Continuous Quality Evaluation, 単一刺激連続品質評価)法です。

前者は、原画と比べて処理画がどのくらい劣化しているかを厳密に評価するものであるのに対して、後者は処理画だけでの評価なので、例えば家庭でテレビを見ている人が、その画像に対してどのように感じるかということの評価するための枠組みであるといえます。

言い換えると、SSCQEは画像そのものの評価用であり、DSCQSは処理系の評価用であるといういい方もできます。

このほか、DSISとSSCQEの特徴を両方取り入れた新しい評価方式として、**SDSCE**(Simultaneous Double Stimulus for Continuous Evaluation, 同時二重刺激連続評価)法もあります。以下に詳しく解説します。

● DSCQS

▶ 2種類の画像を交互に比較しながら評価する

BT.500では、DSCQSについて、主観評価実験の評価者の人数(非専門家の評価者を15人以上)、モニタの明るさやコントラスト、モニタの大きさ、モニタから評価者までの距離、画像の提示方法(順番、時間)などについて、細かく規定しています。

図8-1にしたがって説明します。この方式は2種類の画像(仮にA, Bとします)を比較しながら評価します。まず、画像Aをモニタに表示します。時間は10秒間です。その後、一面灰色の画像を3秒間表示し、今度は画像Bを10秒間表示します。そしてまた一面

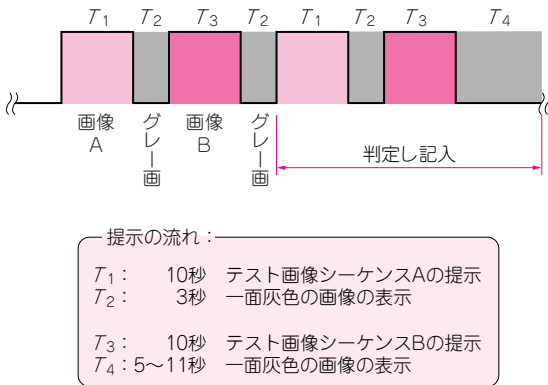


図8-1 DSCQSの画像提示方法

灰色の画像を3秒間表示します。ここまで評価者は、ただ観察するだけで、評価シートへの記入は行いません。

次に、以上の表示をもう一度繰り返します。つまり、Aを10秒間表示し、灰色画像を3秒間、Bを10秒間、最後に灰色画像を5秒から11秒間表示します。この間に各評価者はA、Bそれぞれの評価値をシートへ記入します。実際には、目盛り上へのマーキングになります。

評価シートは図8-2のように5段階表現になっていますが、目盛りのどの部分にマークをつけてもよく、実際には連続値による評価になります。これが**連続品質尺度**と呼ばれる理由です。

▶ **精度を高めるためのくふう**

この評価実験では精度を高めるため、いろいろな絵柄の画像をさまざまな組み合わせで提示します。図8-2では、27回目から31回目までの組み合わせの評価を記入する部分が表示されています。例えば、画像圧縮コーデックの画質評価を行う場合、AとBのどちらかに原画、どちらかにコーデックの出力画が提示されますが、どちらを原画にするかは各回によってランダムに変えられます。

このようにして得られた連続評価値は、差分をとられ、それが0から100(%)に正規化されて、平均化されます。これが評価対象の劣化度となります。

このテストに先立っては、評価者への事前の説明や、評価作業になれるための練習セッションが行われます。これらの準備作業は、評価データを信頼性の高いものにするために重要です。テスト時間は、評価者の集中力が持続するよう30分以内と勧告されています。

▶ **ほかのITU勧告での伝送品質の指標としても使われる**

以上のようにして得られた劣化尺度は、ほかのITU勧告における伝送品質の要求条件などの勧告で指標として指定されるなど、重要なものとなっています。

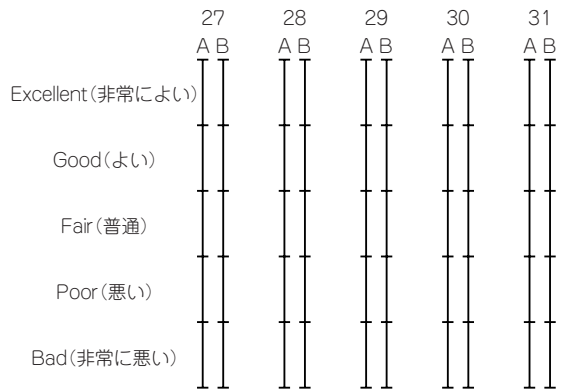


図8-2 評価シートの一部

す。例えばITU-R勧告BT.800では、テレビ素材伝送のコーデックの要件として、圧縮画像と原画の品質差がDSCQSで12%以下であることと規定しています。

● **DSIS**

▶ **あらかじめどちらが原画かわかっている**

DSIS法はDSCQS法と似ていますが、まず先に原画を見せ、次に評価対象画像を提示します。評価者にはどちらが原画かがわかっているという点が大きな違いです。つまり、**原画に対して評価対象画像がどのくらい異なっているか**、ということ測定するわけです。

例えば被写体の輪郭をはっきりさせるエッジ強調、またはノイズを低減する平滑化フィルタがかかった画像が評価対象の場合、DSCQSでは処理画(評価対象画像)のほうが良い点がつくかもしれません。しかしDSISでは、原画に対してどの程度違っているかを見るわけですから、そのようなことは起こりません。処理画だけを5段階評価します。

▶ **評価のための画像提示方法**

図8-3にDSIS法の画像提示方法を示します。図中の方法Ⅱは、方法Ⅰに比べて原画と処理画の差が小さい場合に使用します。

▶ **評価の各段階**

DSISにおける評価の各段階は、次のとおりです。

- 5 : imperceptible (劣化がわからない)
- 4 : perceptible, but not annoying (劣化がわかるが気にならない)
- 3 : slightly annoying (劣化が気になるが、邪魔にならない)
- 2 : annoying (劣化が邪魔になる)
- 1 : very annoying (劣化が非常に邪魔になる)

● **SSCQE**

▶ **原画と見比べることはしない**

MPEGなどの画像圧縮では、絵柄によって圧縮の