4-2 客観画質評価法の標準化動向

正会員 西田 幸博[†]

キーワード: ITU, VQEG, 画質, 客観評価, 標準化

1. ま え が き

映像品質の評価法に関する ITU (国際電気通信連合) での標準化は, ITU-R SG 11 (テレビ放送), ITU-T SG 9 (ケーブルテレビなどテレビ・音声の伝送), ITU-T SG 12 (ネットワーク端末間伝送性能) において, 各研究委員会 (SG) が所掌する分野ごとに行われている. ITU のほか IEEE (Institute of Electrical and Electronic Engineers) や ANSI (American National Standards Institute) などでも評価法に関する標準化活動が行われている.

デジタル符号化技術の放送への導入に伴い, 映像や音声の客観評価法のニーズが増大し, ITU においても客観評価法の標準化の気運が高まってきた. 音質の客観評価法については,すでに ITU-R 勧告が制定されている¹⁾. 画質の客観評価法については, VQEG (Video Quality Experts Group) が ITU 勧告の作成に向けて活動中である. 本稿では, 特に VQEG の活動について述べる.

2. VQEG

1997 年 10 月, ITU-R SG 11, ITU-T SG 9, ITU-T SG 12 から専門家が集まり, 「映像品質の主観的および客観的評価に関する専門家会議」がイタリアで開催された. この専門家グループを VQEG と呼ぶようになった. VQEG は ITU としての公式なグループではないが, その活動は ITU から強く支持されている. この第 1 回会合では, 画質の客観評価法の

表 1 方式提案機関

CPqD-TELEBRAS (Brazil)
Institut für Nachrichtentechnik (Germany)
KDD (Japan)
KPN Research (Netherlands)/Swisscom CIT (Switzerland)
NASA (USA)
NHK/Mitsubishi Electric Corp. (Japan)
NTIA/ITS (USA)
Swiss Federal Institute of Technology (Switzerland)
TAPESTRIES (European ACTS project)
Tektronix/Sarnoff (USA)

ITU 勧告を検討していくための活動方針が決定された. 第 1 段階として, 放送応用を念頭に置いた符号化方式やビットレートを中心にテストを行うこととなった. ITU の各 SG 会合や Email ベースでの議論を経て, 1998 年 5 月に第 2 回 VQEG 会合がアメリカで開催された. この会合において, 方式提案募集から提案方式の実験, 性能評価に至る具体的な活動内容と手順が決定された.

1998 年 6 月の客観評価法提案募集に対して, 日本, 北南米, ヨーロッパの 10 機関 (グループ) から応募があった. 現在, 提案方式の性能を評価するための実験が行われており, 1999 年 9 月の第 3 回会合でテスト結果が報告される予定である. 現在は競争的フェーズであるが, 複数の提案方式を統合することで大きな性能向上が見込めると判断されれば共同のフェーズへ移行することもあり得る. このテスト結果に基づき, ITU の各 SG はそれぞれの応用分野に応じて ITU 勧告を作成することになる.

3. VQEG 性能評価テスト²⁾³⁾

提案方式は図 1 の手順に基づいて評価される. 符号化などの映像処理は,

ビットレート 768 k~50 Mb/s の範囲で, 符号化方式として MPEG-2 や H.263 を用い, 多段接続やコンポーネント/コンポジット変換なども想定した 16 種類の条件を設定した. テスト画像には, 525/60 と 625/50 それぞれ 10 種類のコンポーネント素材を用意した. そして, 原画像 20 個と処理画像 320 個から成るテスト画像データベース (1 画像は約 8 秒間) を作成した. 映像処理条件やテスト画像の選定に当たっては, 偏ったテストとならないよう一般性が考慮されている. また, 原画像と処理画像を用いて評価値を求めるため, 原画像に対する処理画像の時間的なアライメント, ゲインやオフセット変化の補正が行われている. 評価装置としては, このような補正機能も含まれている必要があるが, ここでは方式提案の対象外とされている. 客観評価法の提案は, いくつかのプラットフォームを決めて実行形式のソフトウェアを提出する形で行われた. テスト画像データベースを用いて, 各提案方式による客観評価実験を行うと共に主観評価実験を行い, 客観評価結果が主観評価結果とどの程度合致しているかという観点を中心に, 方式の性能が評価される.

[†] NHK 放送技術研究所

“Recent Standardization Activities of Objective Picture Quality Assessment Methods” by Yukihiro Nishida (NHK Science and Technical Research Laboratories, Tokyo)

図1 客観評価法のテスト手順

表3 主観評価実験条件

項目	条件
評価法	DSCQS (Double Stimulus Continuous Quality Scale)
評定者	非専門家 1実施機関当たり15～18人 HQとLQで異なる評定者 教示の統一
観視条件	ITU-R勧告BT. 500-8およびBT. 1129-2の実験室環境
ディスプレイ	19～20インチ スタジオ用CRT形 解像度や調整値などの報告義務付け
視距離	画面高の5倍
セッション	HQとLQでそれぞれ3セッション 1セッション当たり35テスト(ダミーを含む)
評価画像	コンポーネント信号(D1-VTR) シーン長8秒間 原画像を基準画像とする
その他	セッション順序、提示順序のランダム化

表2 主観評価実験実施機関

525/60	625/50
Berkom (Germany)	CCETT (France)
CRC (Canada)	CSELT (Italy)
FUB (Italy)	DCITA (Australia)
NHK STRL (Japan)	RAI (Italy)

3.1 主観評価実験

前述の16種類の映像処理条件をビットレートを基に2つのグループ(HQとLQ)に分けて評価実験を行う。ここで、中間的な2つの条件を両グループに共通に含めており、HQとLQはそれぞれ9種類の条件から構成されている。表2に主観評価実験の実施機関を示すが、1つの実施機関では525/60あるいは625/50についてHQとLQ両方の評価実験を行う。

提案方式の性能は主観評価実験結果と比較して分析されるため、主観評価結果には偏りなく安定したものであることが求められる。そのため、主観評価実験計画については特に詳細かつ慎重に審議が行われた。評価結果は、実施機関の間の相違やHQとLQに分けたことの影響などについても分析される。

3.2 客観評価実験

各方式提案機関に対して前述の20個の原画像と320個の処理画像が供給され、すべての処理画像について評価

図2 精度

結果を提出するよう求められた。一方、あらかじめ提出されている実行モジュール(ソフトウェア)を用いて、方式提案機関とは利害関係のない第三者機関がその出力結果の確認を行っている。

3.3 性能評価

性能評価は、客観評価実験結果を主観評価実験結果と比較することで行われる。比較に当たっては、まず最初に客観評価結果と主観評価結果との非線形回帰を行い、最も相関が高くなる回帰関数で非線形変換した客観評価結果を用いて、「精度」、「単調性」、「一貫性」の3つの判断基準から分析される。客観評価法として求められる性能は明確に規定されるには至っていないが、主観評価実験結果自体に含まれる実施機関や評定者の間でのばらつき度

合いが参考にされる。また、提案方式の複雑さの観点からも必要に応じて評価されることになる。

(1) 精度

客観評価結果との平均的な一致度を測るもので、ピアソン相関係数(Pearson linear correlation coefficient)などが用いられる。図2では(a)の方が精度が高い。

(2) 単調性

主観評価値の大小関係と同じ大小関係が得られているかどうかを測るもので、スピアマン順位相関係数(Spearman rank correlation coefficient)などが用いられる。図3では(a)の方が単調性に優れている。

(3) 一貫性

画像に依存して精度が異なることは適当でない。図4において、(a)は大

図 3 単調性

図 4 一貫性

部分は主観評価値とよく一致しているが、一部の結果が大きく乖離している。一方、(b)は全体的に同程度のば

らつきである。はずれ(outlier)の割合や尖度(kurtosis)などが用いられる。

4. む す び

客観画質評価法の標準化に向けて、VQEGでは世界各国から専門家が集まった共同活動として大規模なテストが行われている。ITUではVQEGでのテスト結果を基に客観評価法の勧告を作成することになるが、アプリケーションに応じた要求条件が充分に考慮される必要がある。

VQEGでの現在のテストは基準画像と処理画像の両方を用いる double-ended 方式であるが、処理画像だけから画質を測定する single-ended 方式のニーズも高く、今後の検討課題となることが予想される。

(1999年7月2日受付)

〔文 献〕

- 1) ITU-R Recommendation BS. 1387, "Method for Objective Measurement of Perceived Audio Quality" (1999)
- 2) VQEG Subjective Test Plan (Jan. 1999)
- 3) VQEG Objective Test Plan (Sep. 1998)
- 4) VQEG Official Web Site
<<http://www.crc.ca/vqeg/>>

にしだ ゆきひろ
西田 幸博

1983年、慶應義塾大学工学部電気工学科卒業。1985年、同大学大学院工学研究科電気工学専攻修士過程修了。同年、NHKに入局。1988年より、同放送技術研究所にて、テレビジョン信号の符号化・伝送方式の研究開発に従事。正会員。