

計量心理学的測定法によるユーザレベル QoS の多次元的评价の研究

代表研究者 伊 藤 嘉 浩 名古屋工業大学工学研究科講師

共同研究者 田 坂 修 二 名古屋工業大学工学研究科教授

1 研究調査の目的

アクセス網の整備と端末の高性能化により、インターネット上で音声やビデオなどの連続メディアを同時に伝送するマルチメディアアプリケーションが普及しつつある。マルチメディア伝送においては、その品質を劣化させる要因が多数考えられる。したがって、インターネット上でのマルチメディア伝送におけるサービス品質(QoS)を評価する場合は、複数の QoS パラメータを検討する必要がある。

一般に、ネットワークの機能は階層的に表現されるので、QoS も階層的に定義される。例えば、インターネットは六つのレベル(物理レベル、ノードレベル、ネットワークレベル、エンド・ツー・エンドレベル、アプリケーションレベル、ユーザレベル)に分類できる。

音声・ビデオ伝送においては、ユーザが最終的に判断する QoS (ユーザレベル QoS) が重要となる。したがって、音声・ビデオ伝送を行う場合、最終的に高いユーザレベル QoS が得られるような QoS 制御を考えなければならない。

ユーザレベル QoS はユーザが判断するものであるため、これを直接制御することはできない。そこで、QoS の制御が可能な下位のレベルでの制御により、ユーザレベル QoS を高く保つようにすることが必要である。そのためには、ユーザレベル QoS と下位のレベルの QoS 間の関係を明確にしておかなければならない。QoS 間の関係を明らかにすることを、QoS マッピングと呼ぶ。ユーザレベル QoS とのマッピングを行うためには、ユーザレベル QoS の定量的な評価が不可欠である。

音声・ビデオ伝送において、ユーザがその品質の良し悪しを総合的に判断する際、音声の品質やビデオの品質、コンテンツの内容など、多くの要因が評価された上で、最終的な判断が下されている。したがって、これらの判断の要因をユーザレベル QoS パラメータとして考え、複数のパラメータでユーザレベル QoS を多次元的に評価することが必要である。更に、この判断の要因となるユーザレベル QoS パラメータに影響を与える物理的な要因を調べることで、ネットワークの設計、制御における有用なガイドラインを示すことができる。

本研究では、ユーザレベル QoS、特に、IP ネットワーク上での音声・ビデオ伝送における主観品質を、SD (Semantic Differential) 法により多次元的に評価する。この多次元的な評価に加えて、メディアの最終的な良し悪しを表す総合的な品質も評価する。そして、この総合的な品質と SD 法で求めた複数のユーザレベル QoS パラメータとの関係を QoS マッピングにより明らかにし、総合的な判断における主要因を調査する。更に、ユーザレベル QoS と下位レベルの QoS、特にアプリケーションレベル QoS との関係を調査する。本研究では、重回帰分析による QoS マッピングを行う。

2 研究調査の方法

本研究では、SD 法によりユーザレベル QoS の多次元評価を行い、総合的な主観品質との関係を求める。まず、SD 法の概要を示す。また、SD 法で得られた結果を処理するための系列カテゴリー法と主成分分析手法も述べる。

2.1 SD 法

SD 法は計量心理学的測定法の一つであり、多数の形容詞対を用いて対象を評価するものである。SD 法の特徴は、多数の形容詞対を用いることにより、評価対象を多くの視点から簡単に評価できることである。

SD 法では、評価に用いる形容詞対をいかに選定するかが重要となる。一般には予備アンケートなどにより、

分析に適した数十種類程度の形容詞対が選択される。

選択された各形容詞対に対して、評定尺度法により各評価対象（ここでは刺激と呼ぶ）の主観評価値が求められる。この方法は、被験者の判断基準に従って、被験者に刺激をカテゴリーに分類させるものである。すなわち、形容詞対間の主観的な感じ方をいくつかのカテゴリーに分類し、被験者は、すべての形容詞対に対して、刺激をいずれかのカテゴリーに分類する。

一般的に、SD 法による多次元評価では、各カテゴリーに予め尺度値が付与される。そして、各形容詞対におけるカテゴリーの心理的な境界が等しいという仮定の下で、刺激が分類されたカテゴリーの尺度値の平均をとり、その値をその形容詞対における刺激の評価尺度値とすることが多い。しかしながら、SD 法で用いられる様々な形容詞対に対してこのような仮定が成り立つ保証はない。したがって、各カテゴリーに付与された値の平均をとったものは順序尺度（大小関係のみに意味がある尺度）とみなす方が自然であり、これをユーザレベル QoS パラメータとして考えることは必ずしも適切ではない。

本研究では、形容詞対ごとに、以下に示す系列カテゴリー法を適用し、形容詞対ごとの距離尺度（数値の差のみに意味がある尺度）を求める。そして、求められた距離尺度をユーザレベル QoS パラメータとして考える。

2.2 系列カテゴリー法

系列カテゴリー法とは、評定尺度法で得られた結果に対しカテゴリー判断の法則を適用して距離尺度を得る手法である。評定尺度法では、カテゴリーの心理的な境界が等間隔であるという保証がない場合、この方法により得られた評価値は順序尺度にすぎない。順序尺度に対して、多くの統計的処理を行うことは不適切である。そこで、系列カテゴリー法では、カテゴリー判断の法則を用いて、評定尺度法によって得られた結果を、ほとんどの統計処理が可能な距離尺度に変換する。

2.3 主成分分析

SD 法では、形容詞対の数に等しい次元数で刺激を評価するため、このままでは、次元数が大きく評価が困難となる。そこで、主成分分析や因子分析で次元数を縮小した後、評価が行われる。本研究では、各形容詞対に対して得られた尺度を変数として主成分分析を行い、必要な次元数の空間で刺激を評価するものとする。主成分分析を行う際、最初に評価に用いるための主成分の数を決定する必要がある。本研究では、主成分に対する各主成分の累積寄与率から必要な次元数を決定し、この最小限必要な次元数までの各主成分をユーザレベル QoS パラメータとして扱う。

2.4 実験方法

ネットワークシミュレータ上で、音声・ビデオ伝送をシミュレートし、伝送された音声・ビデオを対象としてユーザレベル QoS の評価を行う。評価に用いる音声・ビデオには、異なる六つのコンテンツを扱い、ネットワークシミュレータソフトウェアとしては、ns2 を用いた。コンテンツとしては、特徴が大きく異なる三つのジャンルを選び、各ジャンルから二種類ずつコンテンツを採用した。以下に、本実験で用いた六種類のコンテンツを示す。

- ・ニュース番組の一シーン：二種類（以下、ニュース 1（N1）、ニュース 2（N2）と記す）
- ・テニスの試合中継の一シーン：二種類（以下、テニス 1（T1）、テニス 2（T2））
- ・アニメーションの一シーン：二種類（以下、アニメ 1（A1）、アニメ 2（A2））

本実験では、負荷を加えたネットワーク上で音声・ビデオ伝送を行うことで、アプリケーションレベル QoS を変動させる。ns2 上で、サーバからクライアントに対して、各コンテンツの音声・ビデオを送信する。サーバの属するネットワークとクライアントの属するネットワークをルータを介して接続する。負荷生成器が負荷トラフィックを生成する一方で、サーバは音声・ビデオストリームの送信を行う。クライアントは音声・ビデオストリームを受信し、これを出力する。クライアントが出力した音声・ビデオストリームを記録し、記録された合計 18 個の音声・ビデオストリームをユーザレベル QoS 評価用の刺激とした。

2.5 形容詞対の選択

初めに、SD 法で使用する形容詞対を以下のように選定する。まず、SD 法を用いた過去の様々な文献より 90 の形容詞対を収集し、予備実験により計 20 組の形容詞対を選んだ。選定した 20 組の形容詞対を示す。

雑音がないーある，画像が明るいー暗い，音声が強力ー弱々しい，画像が滑らかなー滑らかでない，音声
が澄んだー濁った，画像がきれいー汚い，音声が静かなー騒々しい，画像が止まらないー止まる，音声と画像のず
れがないーある，画像が鮮明なーぼやけた，音声豊かなー貧弱な，画像が崩れないー崩れる，音声途切れな
いー途切れる，画像にぎやかなー静かな，音声快いー不快な，画像が安定したー不安定な，音声がかすれな
いーかすれる，画像に迫力があるー物足りない，音声がきれいー汚い，画像が自然なー不自然な

2.6 SD 法による多次元評価

被験者は、前節で得られたすべての形容詞対に対し、18個の刺激を五段階の評定尺度法で評価する。被験者は、
20代の男女58名である。

SD 法で得られた結果に対して系列カテゴリ法を適用し、形容詞対ごとに各刺激の心理尺度値を求める。得
られた心理尺度がユーザレベル QoS パラメータとなる。求められたユーザレベル QoS パラメータを変数として
主成分分析を行い、最小限必要な主成分数を求め、得られた各主成分を新たにユーザレベル QoS パラメータと
する。

2.7 総合品質評価

SD 法による多次元的な評価とは別に、音声・ビデオ伝送における総合的な主観品質を求め、SD 法で得られ
た各主成分との関係を調べる。総合的な主観品質は SD 法の被験者の内の24名に対して、系列カテゴリ法に
よって測定する。被験者には、SD 法による評価で用いたものと同じ18個の刺激を提示し、評定尺度法で、総合
品質を判断させる。得られた結果に対し、カテゴリ判断の法則を適用し距離尺度を求め、これを音声・ビデオ
伝送の総合的な主観品質を表すユーザレベル QoS パラメータとする。

3 研究調査の結果

3.1 ユーザレベル QoS の多次元評価結果

20個の形容詞対に対し、評定尺度法を適用した。その結果、形容詞対「画像が止まらないー止まる」において、
被験者全員が二つの刺激を、最も「画像が止まる」側にあるカテゴリに分類した。したがって、この形容詞対
においては、カテゴリ判断の法則によって、この二つの刺激の距離尺度を求めることができない。そこで、こ
の形容詞対「画像が止まらないー止まる」を評価用の尺度として扱わない。

残りの19個の形容詞対をそれぞれ、形容詞対 J ($J=1, \dots, 19$) と呼び、系列カテゴリ法で求められた形容詞
対 j の距離尺度を心理尺度 j とする。表 1 に j に対応する形容詞対を示す。

表 1 形容詞対 J

j	形容詞対	j	形容詞対
1	雑音がないーある	11	画像が崩れないー崩れる
2	画像が明るいー暗い	12	音声途切れぬー途切れる
3	音声が強力ー弱々しい	13	画像にぎやかなー静かな
4	画像が滑らかなー滑らかでない	14	音声快いー不快な
5	音声澄んだー濁った	15	画像安定したー不安定な
6	画像がきれいー汚い	16	音声がかすれぬーかすれる
7	音声が静かなー騒々しい	17	画像に迫力があるー物足りない
8	音声と画像のずれがないーある	18	音声がきれいー汚い
9	画像鮮明なーぼやけた	19	画像が自然なー不自然な
20	声豊かなー貧弱な		

例として、形容詞対 1～10における各刺激の心理尺度値を、ネットワーク負荷量ごとに、図 1～図 3 に示す。
図 1～図 3 より、多くの形容詞対において、ネットワーク負荷量が多くなるにつれ心理尺度値が低くなっている
が、その減少の度合いは、形容詞対やコンテンツによって異なっていることがわかる。例えば、形容詞対 4 の“画
像が滑らかなー滑らかでない”では、ネットワーク負荷量が多くなるにつれて、対応する心理尺度値は大きく減

少している。一方、形容詞対 2 の“画像が明るいー暗い”においては、対応する心理尺度値の減少があまり見られない。

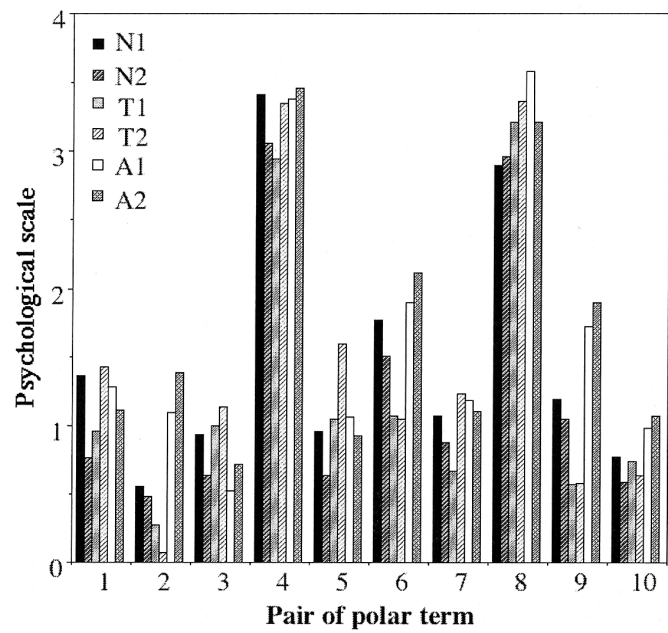


図1 心理尺度 j (負荷が 3.0Mb/s の時)

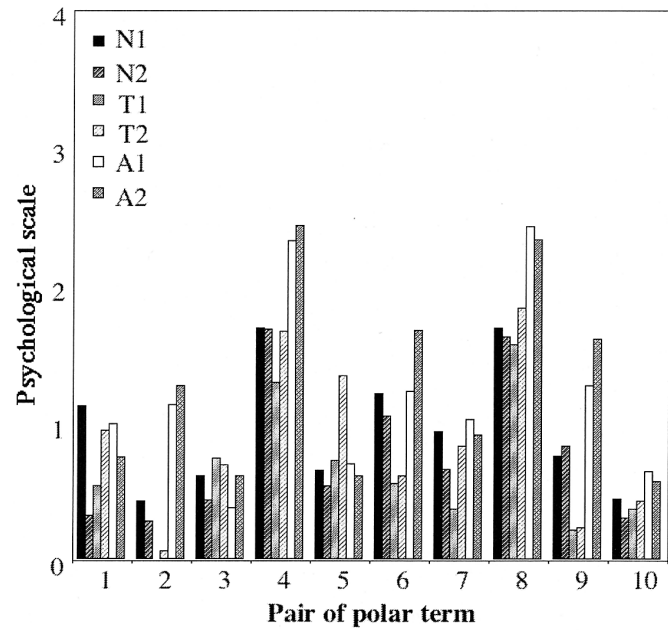


図2 心理尺度 j (負荷が 3.4Mb/s の時)

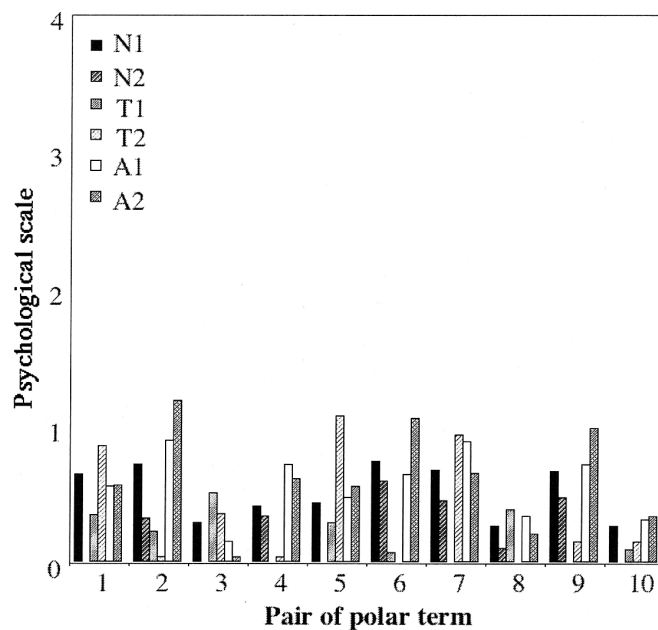


図3 心理尺度 j (負荷が 3.8Mb/s の時)

得られた心理尺度を変数として主成分分析を行い、19次元の評価空間の圧縮を行った。主成分分析の結果、主成分数に対する各主成分の累積寄与率は第三主成分までで、93.6%となっている。そこで、第三主成分までをユーザレベル QoS パラメータとして考えるものとする。各主成分に対する各心理尺度の主成分負荷量を図4、図5に示す。

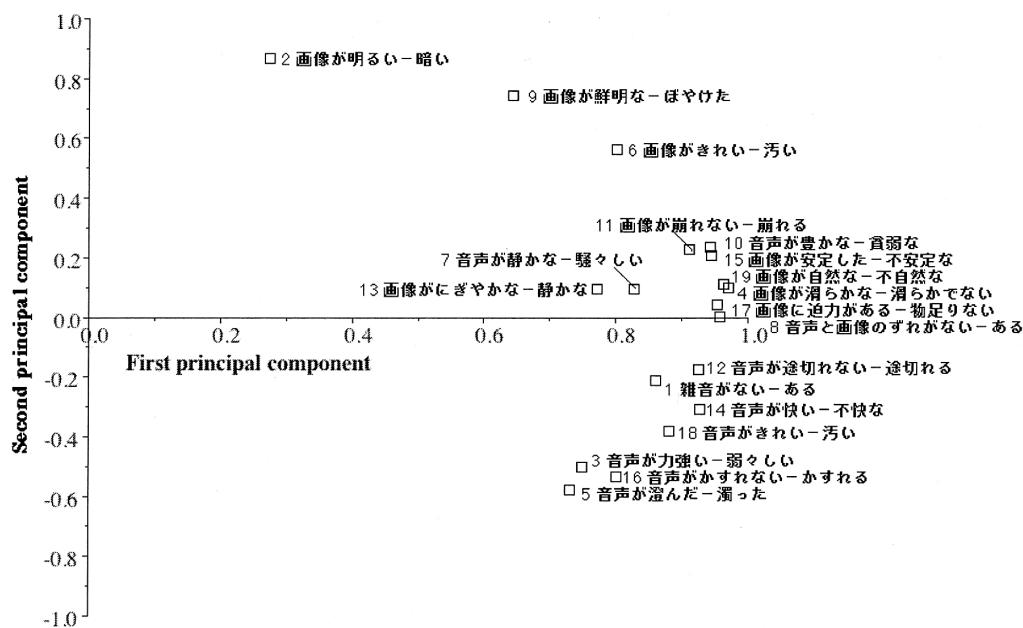


図4 各心理尺度の主成分負荷量 (第一, 第二主成分)

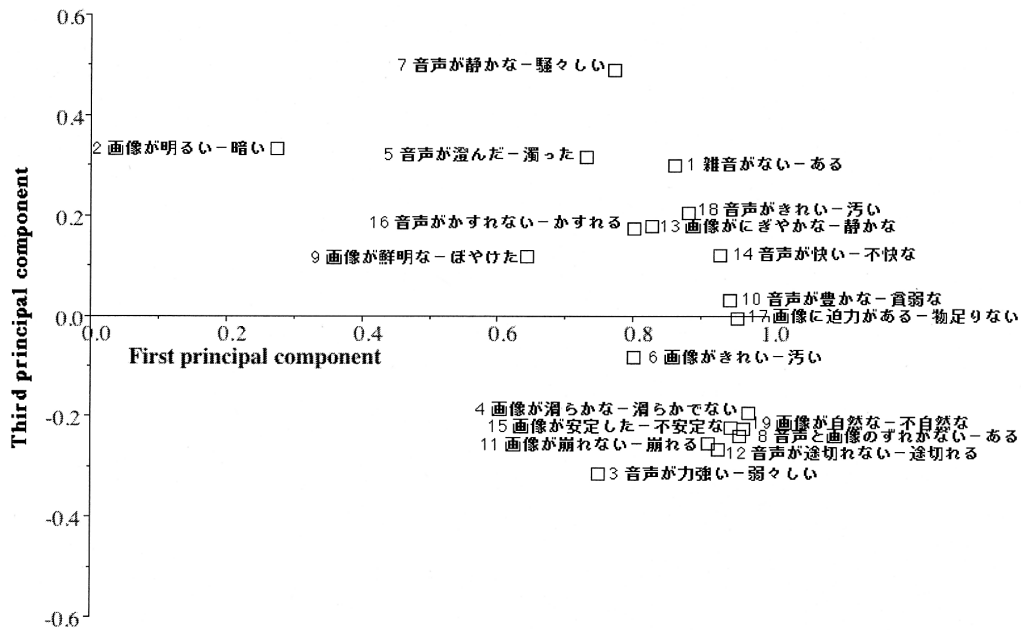


図5 各心理尺度の主成分負荷量（第一，第三主成分）

図4より，第一主成分は，ビデオや音声の滑らかさを表す成分であると考えられる。また，第二主成分は，ビデオの明るさ・鮮明さに関する成分であると推測できる。一方，図5より，第三主成分は，音声の静かさに関する成分であると考えられる。

各刺激に対する各主成分の主成分得点を図6，図7に示す。図中において，各刺激はコンテンツ平均ネットワーク負荷量の形で表されている。

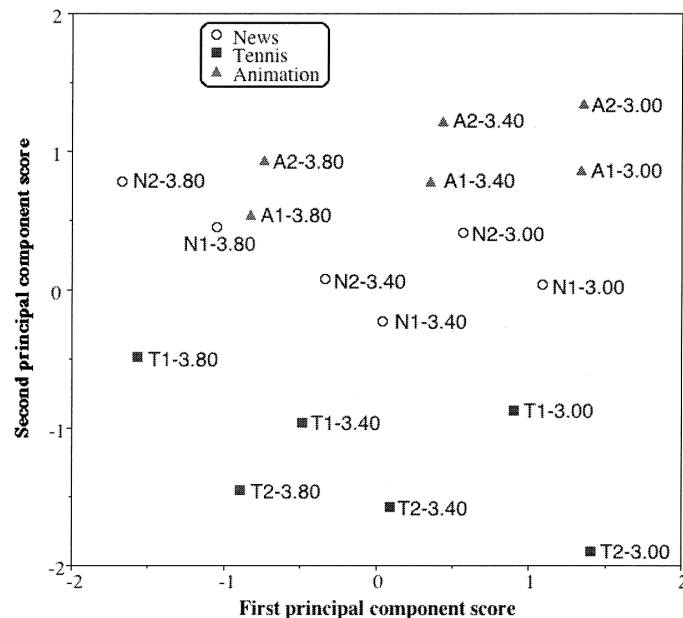


図6 各刺激の主成分得点（第一，第二主成分）

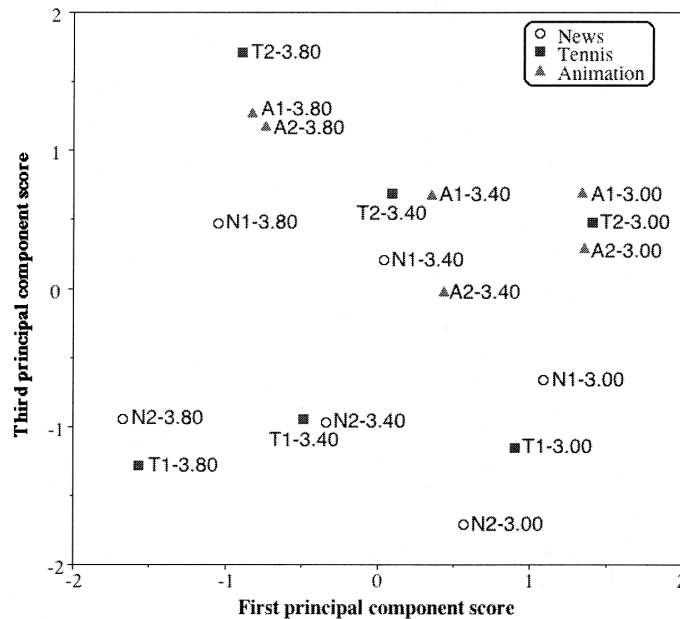


図7 各刺激の主成分得点（第一，第三主成分）

図6より，第一主成分の主成分得点は，コンテンツに関わらず，ネットワーク負荷量が大きくなるに従って減少していることがわかる。これに対して，第二主成分の主成分得点はネットワーク負荷量に関わらず，おおよそ，アニメ2，アニメ1，ニュース2，ニュース1，テニス1，テニス2の順番で小さくなっていることがわかる。これより，第一主成分，すなわち，音声やビデオの滑らかさに関係する尺度は，ネットワークの品質によって影響を受ける尺度であり，第二主成分の表すビデオの明るさ・鮮明さと関連する尺度は，コンテンツに含まれる内容に依存する尺度であることが予想できる。各コンテンツのシーンを確認すると，アニメは，手書きの線を着色したものであり，画像がはっきりしている。一方，テニスは，屋外のシーンを継ぎ接ぎしたものであり，他のコンテンツよりも画像が不鮮明である。これらのことから，第二主成分がビデオの明るさ・鮮明さを表す尺度であることが確認できる。また，図7より，第三主成分の主成分得点もネットワーク負荷量との関連は見られない。その値は，おおよそ，テニス2，アニメ1，アニメ2，ニュース1，テニス1，ニュース2の順番で小さくなっており，特にテニス1とニュース2の値が低くなっている。各コンテンツに含まれるシーンを確認すると，テニス2は選手がボールを打ち合うシーンであるが，観客は静まりかえり，ボールの打ち合う音のみが確認できるものであった。また，アニメ1，アニメ2はそれぞれ，登場人物の会話以外の音声は含まれていなかった。ニュース1，ニュース2はそれぞれ，アナウンサーが記事を読み上げているものであるが，ニュース1は社会記事を扱ったものであり，ニュース2は明るい内容の記事を取り上げている。テニス1もボールを打ち合うシーンであるが，選手の掛け声と，観客のざわめきが随所に現れている。また，ニュース2は明るい内容の記事を扱ったものであり，騒々しい感があった。この結果から，第三主成分も第二主成分と同様にコンテンツ自身に依存する尺度であり，そのコンテンツに含まれる音声の静かなものであるかどうかを示すものであると考えられる。

3.2 総合的な主観品質の測定結果

24名の被験者に対し，総合的な品質に関する主観評価値を評定尺度法により求めた。得られた結果に対し，カテゴリ判断の法則を適用し，総合的な品質における主観評価値を表す距離尺度を求めた。この距離尺度をユーザレベル QoS パラメータの一つとして考え，総合品質心理尺度と呼ぶ。また，本論文では，便宜上，総合品質心理尺度の最小値を原点とする。得られた総合品質心理尺度を図8に示す。

図8より，すべてのコンテンツにおいて，平均ネットワーク負荷量が増加するに従い，総合品質心理尺度値は減少しているが，その減少の割合は，コンテンツによって異なることが分かる。コンテンツがアニメとニュースの場合，平均ネットワーク負荷量の増加による総合品質心理尺度の減少の割合にほとんど違いは見られない。しかし，コンテンツがテニスの場合，平均ネットワーク負荷量の増加による総合品質心理尺度の減少の割合は，他のコンテンツよりも大きくなっていることがわかる。

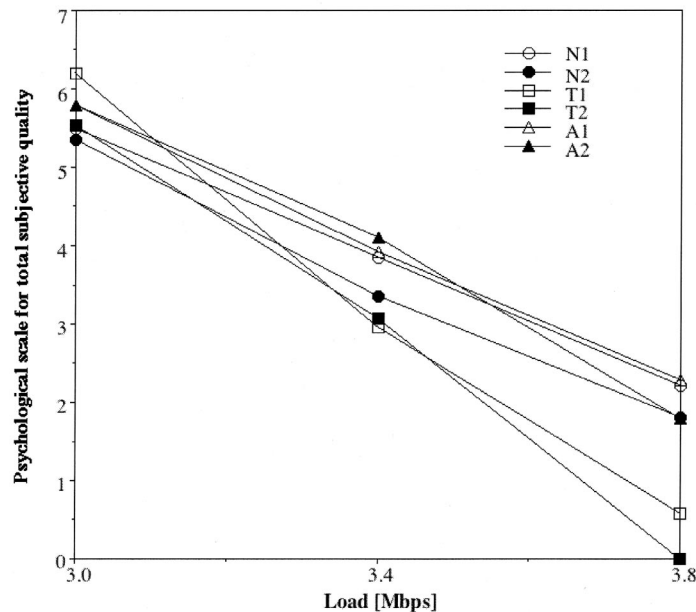


図8 平均ネットワーク負荷に対する総合品質心理尺度

3.3 ユーザレベル QoS 間のマッピング結果

総合品質心理尺度とユーザレベル QoS の各主成分との関係を明らかにするため、総合品質心理尺度を従属変数、各主成分を独立変数として重回帰分析を行った。得られた回帰式の各重回帰係数に統計的検定を適用したところ、有意水準 5 % で、第一主成分及び第三主成分の係数は有意であったが、第二主成分の係数は有意ではなかった。すなわち、第二主成分は音声・ビデオ伝送における評価基準の一つではあるが、今回の実験においては、総合的な主観品質への大きな寄与は見られなかった。そこで、独立変数を第一主成分、第三主成分の二つとして重回帰分析を行った。得られた重回帰式を次に示す。

$$\hat{I} = 3.559 + 1.738U_1 - 4.995 \times 10^{-1}U_3 \quad (1)$$

ここで、 $\hat{I}$ は総合品質心理尺度の推定値であり、 U_1 、 U_3 はそれぞれ、第一主成分、第三主成分である、この式の自由度調整済み寄与率は 0.957 である。以上の結果から、総合的な主観品質は、コンテンツの内容にも影響を受ける部分があることになる。

式(1)において、第三主成分の係数が負になっている。すなわち、第三主成分の値が大きくなると、総合的な主観品質は劣化する。この理由を以下に述べる。ここまで、第三主成分は、「音声の静かさ」を表すことを述べた。このことから、音声静かなコンテンツの場合は、音声品質劣化を検知しやすくなるため、総合的な品質が低くなってしまったものと考えられる。

以上の結果から、音声・ビデオ伝送の総合的な評価において、通信品質に影響を受ける第一主成分が判断の主要因であり、第三主成分で表されるコンテンツの内容も判断の要因であることがわかった。また、コンテンツに関しては、個々のコンテンツが音の静かなものであるかどうか総合的な品質に影響を与えていると言える。

3.3 アプリケーションレベルとユーザレベル間の QoS マッピング結果

得られたユーザレベル QoS パラメータの各主成分とアプリケーションレベル QoS とのマッピングを行った。ここでは、総合品質への寄与を確認できたユーザレベル QoS パラメータの第一主成分と第三主成分を扱う。アプリケーションレベル QoS パラメータとしては、メディア同期に関する以下の九つの尺度を用いた。

まず、音声及びビデオのメディア内同期の QoS パラメータとして、メディア出力の滑らかさを表す出力間隔の変動係 C_a 、 C_v を用いた。更に、音声、ビデオのメディア内同期の平均二乗誤差 E_a 、 E_v を検討した。これらは、MU の生成間隔と出力間隔の差の二乗平均として定義される。次に、音声、ビデオの平均 MU レート R_a 、 R_v を用いた。平均 MU レートは、受信側で 1 秒間に出力された平均 MU 数である。また、総生成 MU 数に対する

失われた MU 数の比率である MU 損失率も用いた。音声、ビデオの MU 損失率は、それぞれ、 L_a , L_v で表す。メディア間同期品質を評価するために、メディア間同期平均二乗誤差 E_{im} を用いた。メディア間同期平均二乗誤差は、ビデオの MU の出力時刻とこれに対応する音声の MU の出力時刻の差から、それぞれのタイプスタンプの差を引いた値の二乗平均として定義される。

シミュレーション上での音声・ビデオ伝送時に、九つのアプリケーションレベル QoS パラメータ値の測定も行った。測定結果より、多くのアプリケーションレベル QoS パラメータ間で高い相関があることがわかった。そこで、九つのアプリケーションレベル QoS パラメータを変数として主成分分析を行ったところ、第二主成分までを用いれば、九つのアプリケーションレベル QoS パラメータの97.9%の情報を表現できることがわかった。主成分分析の結果より、アプリケーションレベル QoS パラメータを以下の二つのカテゴリーに分類した。

a) E_v 以外

b) E_v

多重共線性を考慮して、重回帰分析における独立変数として、先に分類した二つのカテゴリーからアプリケーションレベル QoS パラメータを一つずつ選択する。したがって、考えられる独立変数の組み合わせは8種類となる。このすべての独立変数の組み合わせに対し、従属変数をユーザレベル QoS パラメータの各主成分として重回帰分析を行った。

重回帰分析の結果、ユーザレベル QoS パラメータの第一主成分は、アプリケーションレベル QoS パラメータから高い精度で推定できるが、第三主成分を推定することはできないことがわかった。これは、第三主成分が、コンテンツの内容に依存する成分であり、ネットワーク負荷による品質の劣化に影響されないためと考えられる。

重回帰分析の結果を基に、アプリケーションレベル QoS からユーザレベル QoS パラメータの第一主成分を推定する重回帰式を選ぶ。得られた重回帰式は次のようになった。

$$\hat{U}_1 = 3.173 + 1.595C_a - 2.499 \times 10^{-4} E_v \quad (2)$$

ここで、 $\hat{U}_1$ は第一主成分の推定値である。この重回帰式の自由度調整済み寄与率は0.896であり、アプリケーションレベル QoS から高い精度で第一主成分値を推定できることがわかる。ユーザレベル QoS パラメータの第三主成分がコンテンツに含まれる内容に関連のある尺度であることを確認するため、図XXにおける各コンテンツの第三主成分値の大きさに基づき、以下に定義したダミー変数 C を独立変数とし、第三主成分を従属変数とした単回帰分析を行った。

$$C = \begin{cases} 0 & (\text{Contents is } N2) \\ 1 & (\text{Contents is } T1) \\ 2 & (\text{Contents is } N1) \\ 3 & (\text{Contents is } A2) \\ 4 & (\text{Contents is } T2) \\ 5 & (\text{Contents is } A1) \end{cases} \quad (3)$$

第三主成分の推定値を $\hat{U}_3$ で表すと、得られた回帰式は次のようになる。

$$\hat{U}_3 = -1.229 + 0.492C \quad (4)$$

自由度調整済み寄与率は0.855であり、おおその推定ができています。したがって、第三主成分がコンテンツ自身の内容に関連する尺度であることが確認できる。

以上の結果より、音声・ビデオ伝送における総合的な主観品質は、先に求めたユーザレベル QoS パラメータの第一主成分と第三主成分で表現でき、特に、第一主成分のみで、総合的な主観品質の大部分を表すことができた。また、第一主成分は、ビデオや音声の滑らかさに関連の深い尺度であり、二つのアプリケーションレベル QoS パラメータ C_a , E_v により、推定が可能であった。したがって、 C_a , E_v の値がわかれば、音声・ビデオ伝送における総合的な主観品質を概ね推定できる。第三主成分は、コンテンツの内容に依存する尺度であり、音声の静かさに関するものであった。コンテンツの内容が既知である場合、その内容から、第三主成分を推定することができれば、第一主成分の推定値と、第三主成分のものをを用いることで、より高い精度で総合的な主観品質を推定可能である。

3.4 結論

本論文では、SD 法により、音声・ビデオ伝送におけるユーザレベル QoS の多次元評価を行った。また、系列カテゴリー法により総合的な主観品質を求め、各主成分との関係を調査した、その結果、今回の実験環境においては、総合的な主観品質は、二つの主成分と関連があり、一つはネットワークの品質に影響を受ける成分、もう一つはコンテンツの内容に依存する成分であった。特に、前者によって、総合的な主観品質の大部分を表すことができることがわかった。また、各主成分とアプリケーションレベル QoS との QoS マッピングを行った結果、第一主成分は、メディア同期に関するアプリケーションレベル QoS パラメータの二つから推定でき、第三主成分は、コンテンツの内容から推定できることがわかった。

〈発 表 資 料〉

題 名	掲 載 誌 ・ 学 会 名 等	発 表 年 月
IP ネットワーク上での音声・ビデオ伝送におけるユーザレベルQoS の多次元評価	電子情報通信学会和文論文誌B	2005年3月