

# 視聴覚寄与を考慮した高品位映像のための視聴覚感性モデルの構築

## Construction of Audio-Visual Sensibility Model for High-quality Movies in Consideration of Audio-Visual Contribution

精密工学専攻 13 号 大野 瑠衣  
Rui Ono

### 1. はじめに

音は人の生活を快適や不快にするように、生活に欠かせないものであり、人に対し心理および生理的に影響を与える効果がある。一方、人は五感の複合刺激を受けることで、単一で受けた場合とは異なる印象を持つことが知られている<sup>(1)</sup>。音の印象も視覚などの他刺激により変化するといわれ<sup>(2)</sup>、複合刺激に関する研究が広く行われている。例えば、映画に臨場感や緊迫感を与える音楽を使用することで、シーンの立体感が増すことが明らかになっている<sup>(3)</sup>。これより、動画と音の組み合わせによる映像の更なる好感度を生み出す可能性が示唆される。

そこで本研究では、映像の音に着目し、主観評価の SD 法と一対比較法を用いることで、好感度の高い映像を生み出す映像効果を調査する。その際、映像に対する視覚および聴覚の影響度を調べることで、それらの関連性を把握する。また、生体情報を用いた客観評価<sup>(4,5)</sup>を行うことにより主観評価との相互関係を確認し、映像に対する適切な評価方法を提案する。そして、評価を行う各映像について、視聴覚に関する感性モデルの構築を目指す。なお、本研究では一般的に撮影頻度の高いシーンとして、パーティでの会話を撮影した「会話シーン(S1)」、海の光景を撮影した「風景シーン(S2)」、花火大会の打ち上げ花火を撮影した「花火シーン(S3)」、およびスポーツの速い動きを撮影した「スポーツシーン(S4)」の 4 シーンを使用する。なお、被験者は全評価を通して共通した 20 代男性 16 名で、十分なインフォームドコンセントを実施している。

### 2. 視覚または聴覚刺激による主観映像評価

映像の動画または音に効果を施し、印象変化を把握する。また、動画と音の一方および双方の再生による評価を行い、印象変化を比較する<sup>(6,7)</sup>。

#### 2.1 主観評価方法

##### 2.1.1 評価形容詞対

本章では、7 段階の SD 法相対評価による主観評価を行う。S1 の評価で用いる形容詞対を Table 1 に示す。他シーンの評価では類似した形容詞対を用いるが、S1 の特徴を表す形容詞対「面白いー面白くない」、「賑やかなーさびしい」などに代わり、各シーンの特徴を表す形容詞対として、S2 では「冬のようなー夏のような」、「安らぎのあるー不安な」、S3 では「豪快なー繊細な」、「印象に残るー印象の薄い」、S4 では「白熱したー淡々とした」、「機敏なー遅鈍な」などの形容詞対を用いる。

##### 2.1.2 評価映像効果

本評価で用いる動画効果は、撮影時の動画(Normal)に対し、明るさ(Bri)、ぼかし(Bra)、明るさとぼかし(BB)、彩度(Sa)、コントラスト(Con)の増加である。一方、音効果は、カメラ内蔵のマイクによる録音(ME)に対し、計測用高性能マイクによる録音(IC)、ME にシーンごとに異なる 3 つの中心周波数のハイパスフィルタ(Hi)およびローパスフィルタ(Lo)の適用である。

Table 1 Result of factor analysis on the conversation scene

Adjective pairs	Factor loadings		
	Comfort	Power	Clearness
Consistent – Strange	0.90	0.18	0.21
Comfortable – Uncomfortable	0.86	0.25	0.22
Desirable – Undesirable	0.86	0.29	0.15
Natural – Artistic	0.79	0.16	0.24
:	:	:	:
Tense – Relaxed	0.01	0.77	0.08
Energetic – Insipid	0.27	0.74	0.17
Real – Virtual	0.55	0.64	0.03
:	:	:	:
Light – Heavy	0.24	-0.12	0.71
Bright – Dark	0.14	0.43	0.58
Cheerful – Cheerless	0.15	0.05	0.68
Contribution ratio	29.7%	19.8%	10.6%
Cumulative contribution ratio	29.7%	49.5%	60.1%

動画効果に音を再生する評価では ME、音効果に動画を再生する評価では Normal を再生する。また、SD 法による相対評価の基準映像には、動画に Normal、音に ME(Normal\_ME)を用いる。

#### 2.2 シーン毎の主観評価

各シーンの主観評価結果について、バリマックス回転による因子分析を行い、シーン毎の印象変化を把握し、比較する。

##### 2.2.1 会話シーンの主観評価

効果を施した動画のみを再生した際の主観評価による因子分析から、Table 1 に示すように快適、迫力、および明瞭因子が得られた。基準音を追加した際や、音に効果を施した際も同様の因子が抽出された。

Hi により、快適性と明瞭性が上昇し、Lo により全因子が低下する。さらに、Bri により全因子が上昇し、ぼかしの増加で迫力性が低下、明るさの増加では明瞭性が上昇した。また、Normal 動画の追加により Hi の快適性は ME と変わらず、迫力性は低下した。ME 音の追加により BB の迫力が増加した。

##### 2.2.2 風景シーンの主観評価

全条件において、因子分析より、「違和感のない」、「好ましい」などから快適因子、「力強い」、「印象に残る」などから迫力因子、「冬のような」、「冷たい」などから緊張因子、「明るい」、「軽快な」などから軽快因子が得られた。

IC で快適性の上昇と緊張性、軽快さの低下、Hi で迫力性の低下、Lo で迫力性以外の低下が見られた。さらに、Bri により快適性が微増、緊張性が低下し、Bra、BB で軽快さ以外の因子が低下、Sa で快適性、緊張性の低下と迫力性、軽快さの微増が確認された。また、Normal 動画の追加による傾向の変化は見られず、迫力性は低下した。ME 音の追加により、Bri の緊張性低下と明るさによる軽快さの印象が明白になった。

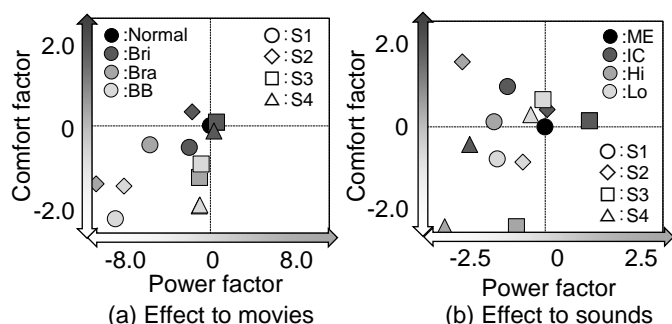


Fig. 1 Transition of impressions for visual or auditory sense between 4 scenes

### 2.2.3 花火シーンの主観評価

動画効果の評価において、因子分析より、「快適な」、「好ましい」などから快適因子、「迫力のある」、「印象に残る」などから迫力因子、「はっきりとした」、「鮮明な」などから明瞭因子、「明るい」から明瞭因子が得られた。一方、音効果の評価において、快適、迫力、明瞭因子に加え、「落ち着いた」、「自然な」から安心因子が得られた。

IC では快適性の上昇と明瞭性の低下、Hi では迫力性の低下と明瞭性の上昇、Lo では快適性、迫力性の低下が見られた。Bri, Con により迫力性が増し、Bra, BB により全因子の印象が低下した。また、動画の追加により、Lo は ME に類似の印象を与えた。一方、音の追加による印象の傾向変化は見られなかった。

### 2.2.4 スポーツシーンの主観評価

動画効果の評価において、因子分析より、「快適な」、「自然な」などから快適因子、「迫力のある」、「白熱した」などから迫力因子、「明るい」、「軽快な」などから軽快因子が得られた。一方、音効果の評価において、快適、迫力因子に加え、「鮮明な」、「明確な」などから明瞭因子が得られた。

IC と Hi は類似した傾向であり、全因子で低下が見られた。Lo では明瞭性が低下したが、他因子は ME と類似した印象を与えた。Bri, BB では迫力性、Bra, BB では快適性が低下し、Sa では軽快さは低下したが、他因子は上昇した。また、動画の追加により IC の快適性、明瞭性が Lo に似た傾向を示したが、音の追加による変化はなかった。

### 2.3 各シーンの印象変化比較

各シーンで確認できた動画または音の追加による印象変化の有無から動画と音の影響度を検討する。S1 では双方に変化がありどちらの影響も受けていると考えられるが、S2 では音追加時のみ、S3, S4 では動画追加時のみ変化が生じ、それぞれ音および動画の影響が大きいと考えられる。また、Fig. 1 に動画または音のみに効果を施した際の全シーンの印象推移を示す。動画は全シーン類似した推移だが、音はシーンにより大きく異なることがわかる。なお、推移の表示には全シーンに共通する効果のみを用い、Hi および Lo はある 1 つの中心周波数による印象を示す。

## 3. 視覚および聴覚刺激による主観映像評価と視聴覚感性モデルの構築

映像の動画および音の双方を変更した際の各シーンの印象変化を確認し、好感度の高い映像を検討する。

### 3.1 主観評価方法

#### 3.1.1 評価形容詞対

本章では 2 章での SD 法に加え、4 つの形容詞を用いた一対比較法による主観評価を行う。使用する形容詞は全シーンに共通した「好ましい」、「迫力がある」、「一致した」の 3 つに加え、S1 に「会話が聞き取りやすい」、S2 に「緊張感がある」、S3 に「印象に残る」、S4 に「臨場感がある」を用いる。

### 3.1.2 評価映像効果

本評価で用いる動画効果は、全シーンに共通した Normal, Bra, BB, に加え、S2 および S4 では Sa, S3 では Con を用いる。一方、音効果は、全シーンにおいて ME, IC, Hi, Lo を用いる。基準映像には動画 Normal と音 ME の映像 (Normal\_ME) を用いる。

### 3.2 シーン毎の主観評価

各シーンでの因子分析から映像の印象傾向を確認する。また、一対比較法による評価から好感度が高い映像を検証する。

#### 3.2.1 会話シーンの主観評価

SD 法による因子分析より、「好ましい」、「快適な」などから快適因子、「迫力のある」、「躍動感のある」などから迫力因子、「こもらない」、「歯切れの良い」などから明瞭因子を得た。Fig. 2(a) より、快適性と迫力性は動画効果、明瞭性は音効果に寄与している。

また、Fig. 3(a) に示す一対比較法による主観評価では、Bri\_IC が最も好ましく、迫力性の結果と傾向が類似していることから、迫力が好感度に繋がると考えられる。

#### 3.2.2 風景シーンの主観評価

SD 法による因子分析より、「自然な」、「好ましい」などから快適因子、「迫力のある」から迫力因子、「冷たい」、「硬い」などから緊張因子、「明るい」、「軽快な」などから軽快因子を得た。Fig. 2(b) より、快適性は動画効果、迫力性は音効果に寄与している。緊張性、軽快さには Hi が大きく寄与するが、それ以外の音では動画効果が寄与している。

Fig. 3(b) に示す一対比較法による主観評価では、Sa\_ME が最も好ましく、一致性の結果と傾向が類似していることから、音と動画の一致性が好感度に繋がると考えられる。

#### 3.2.3 花火シーン的主観評価

SD 法による因子分析より、「好ましい」、「快適な」などから快適因子、「迫力のある」、「力強い」などから迫力因子、「澄んだ」、「はっきりとした」などから明瞭因子を得た。Fig. 2(c) より、快適性は Bra, BB を除き音効果、迫力性は音効果に寄与しているが、明瞭性は双方の効果に寄与している。

また、Fig. 3(c) に示す一対比較法による主観評価では、Con\_ME が最も好ましく、一致性の結果と傾向が類似していることから、音と動画の一致性が好感度に繋がると考えられる。

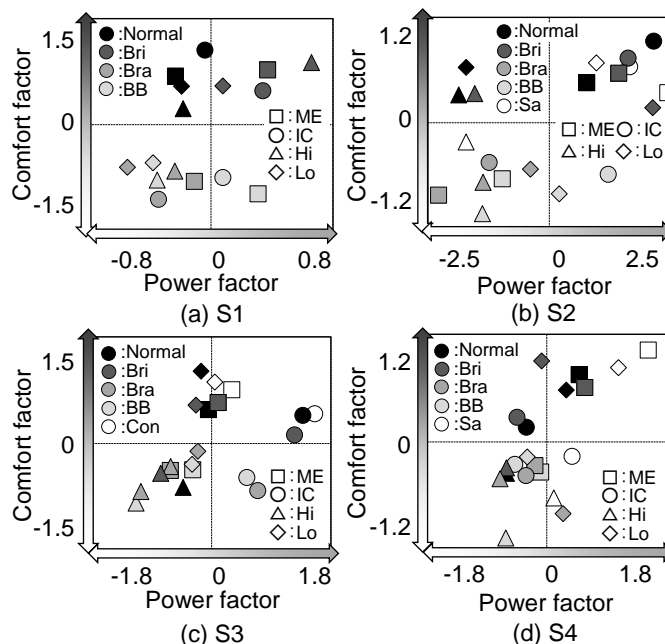


Fig. 2 Power and comfort factor scores

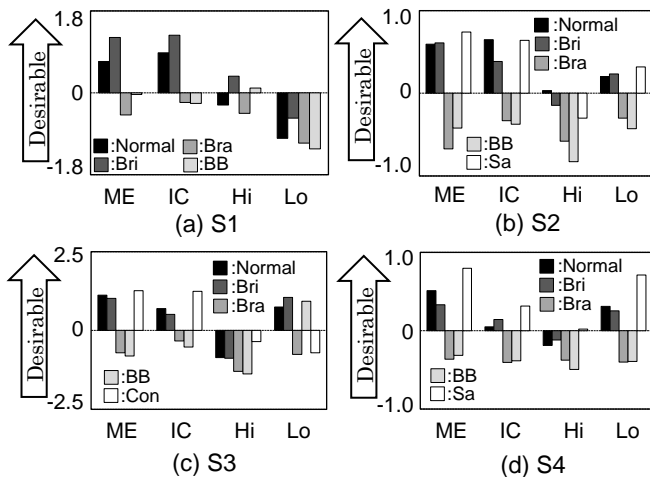


Fig. 3 Comparison of the various movies by paired comparison method

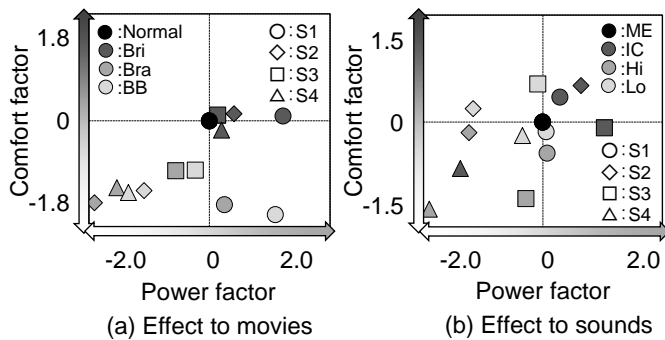


Fig. 4 Transition of impressions for visual and auditory sense between 4 scenes

### 3.2.4 スポーツシーンの主観評価

SD 法による主観評価では、「自然な」、「好ましい」などから快適因子、「白熱した」、「迫力のある」などから迫力因子、「はっきりとした」、「鮮明な」などから明瞭因子を得た。Fig. 2(d)より、快適性はBraを除き音効果、明瞭性は動画効果の影響を受け、迫力性は双方の効果に寄与している。

Fig. 3(d)に示す一対比較法による主観評価では、Sa\_MEが最も好ましく、躍動感の有無で傾向が類似していることから、映像の躍動性が好感度に繋がると考えられる。

### 3.3 各シーンの印象変化比較

双方の効果による主観評価から各シーンの動画と音の影響度を確認する。S1は因子により双方の効果に影響を与え、S2はHiに強く寄与していることから音効果、S3とS4はBraに強く寄与していることから動画効果の影響が大きいと考えられ、2章と同様の結果を得た。また、Fig. 4(a)に動画効果、Fig. 4(b)に音効果による全シーンの印象推移を示す。動画効果では、全シーンにおいて類似の推移傾向であり、Briが有効な効果であることがわかる。音効果では、Hiは一律に低下するが、その他の音はシーンに依存することがわかる。

### 3.4 視聴覚感性モデルの構築

重回帰分析よりモデル式を算出し、シーンの比較および動画と音の関連度を検討する<sup>(8)</sup>。

#### 3.4.1 モデル式の構築条件

主観評価から得られた各因子得点を用い、シーンごとのモデル式を算出する。本分析において、目的変数には3.2節のSD法評価から抽出された各シーンの因子得点を用いる。また、動画に対する説明変数には、動画効果のパラメータとしてBri, Braに加え、S2とS4ではSa, S3ではConを、音に対する説明変数には、音効果の心理音響パラメータとしてラウドネス(L)、シャープネス(S)、ラフネス(R)を用いる。

### 3.4.2 視聴覚感性モデル

各シーンにおいて、重回帰分析より算出した式を下記に示す。本分析では、目的変数に対する優位確率を表す各説明変数のP値が算出され、P値が小さいほど目的変数への関連性が高いと考えられる。下記式では、P値が小さい値を示す説明変数に下線を表記する。

S1について、

$$\begin{aligned} Comfort &= 0.06*L - 0.87*S - 0.28*R \\ &\quad + 0.00*Bri - 0.10*\underline{Bra} - 0.14 \\ Power &= -0.07*L + 0.86*S + 0.44*R \\ &\quad + 0.00*Bri - 0.02*\underline{Bra} + 1.25 \\ Clearness &= -0.25*L + 5.57*\underline{S} - 0.24*R \\ &\quad - 0.00*Bri + 0.01*\underline{Bra} - 1.93 \end{aligned}$$

S2について、

$$\begin{aligned} Comfort &= 0.13*L - 0.48*\underline{S} - 6.70*R \\ &\quad - 0.01*Bri - 0.08*\underline{Bra} - 0.02*Sa + 6.38 \\ Power &= 0.28*L - 0.70*S - 8.89*R \\ &\quad + 0.01*Bri - 0.02*\underline{Bra} + 0.03*Sa + 6.89 \\ Tense &= -0.47*\underline{L} + 0.64*\underline{S} + 7.13*\underline{R} \\ &\quad - 0.01*Bri - 0.00*\underline{Bra} - 0.02*Sa - 3.62 \\ Light &= -0.05*L + 1.25*\underline{S} + 5.03*R \\ &\quad + 0.02*\underline{Bri} - 0.04*\underline{Bra} + 0.01*Sa - 6.37 \end{aligned}$$

S3について、

$$\begin{aligned} Comfort &= 0.01*L - 3.04*\underline{S} + 2.31*\underline{R} \\ &\quad - 0.01*Bri - 0.05*\underline{Bra} + 0.01*Con - 1.51 \\ Power &= -0.015*\underline{L} - 2.315*\underline{S} - 1.682*\underline{R} \\ &\quad - 0.01*Bri - 0.03*\underline{Bra} + 0.01*Con + 3.73 \\ Clearness &= -0.01*L + 2.68*\underline{S} - 0.51*R \\ &\quad - 0.01*Bri - 0.06*\underline{Bra} + 0.01*Con - 2.02 \end{aligned}$$

S4について、

$$\begin{aligned} Comfort &= -0.04*L - 0.09*S + 6.92*R \\ &\quad + 0.00*Bri - 0.04*\underline{Bra} - 0.00*Sa - 0.27 \\ Power &= 0.04*L + 0.14*S + 0.96*R \\ &\quad - 0.00*Bri - 0.01*\underline{Bra} + 0.04*\underline{Sa} - 1.46 \\ Clearness &= -0.17*L + 0.10*\underline{S} + 15.6*R \\ &\quad + 0.01*Bri - 0.04*\underline{Bra} + 0.02*Sa + 2.05 \end{aligned}$$

上記より、S1では快適因子Comfortおよび迫力因子PowerでBra, 明瞭因子ClearnessでSの影響を受けていることがわかる。これより、好感度には動画の影響が大きいと考えられる。S2では、快適因子でSとBra, 緊張因子TenseでL, S, RおよびBra, 軽快因子LightでS, BriおよびBraの影響を大きく受け、動画の寄与が大きいことがわかる。S3では、快適因子でS, R, Bra, 迫力因子でL, S, RおよびBra, 明瞭因子でS, Braの影響を受け、動画の寄与が大きいことがわかる。S4では、快適因子でBra, 迫力因子でSa, 明瞭因子でSおよびBraの影響を動画の寄与が大きいことがわかる。以上より、主観評価に基づき視聴覚感性モデルが構築できた。

## 4. 生体情報による客観映像評価

動画と音の双方に効果を施した映像について、客観評価により印象変化を把握する。

### 4.1 客観評価手法と評価条件

#### 4.1.1 生体情報による客観評価手法

生体情報を用いた客観評価手法として、脳血流量と心電の計測を行う。脳血流量計測では、測定脳血流内における酸化ヘモグロビン(Oxy-Hb)濃度の変化量を非侵襲的に測定する近赤外分光法(NIRS: Near Infrared Spectroscopy)を用いる。その際、視覚刺激を捉える視覚野が存在する後頭葉、および聴覚刺激を捉える聴覚野が存在する側頭葉に加え、情動を司るとされる前頭前野が存在する前頭葉の活動が映像の好感度に関連すると考え、各部の脳血流量を測定することで脳の

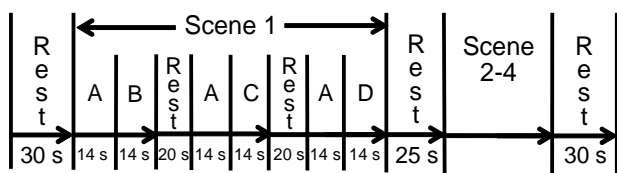


Fig. 5 Time protocol

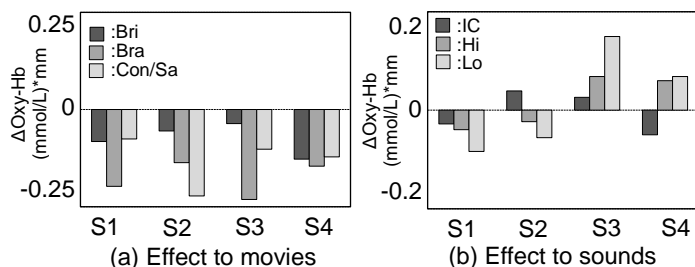


Fig. 6 Evaluation results of NIRS in frontal lobe

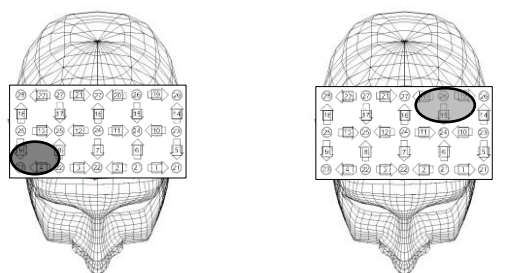


Fig. 7 Effective area in frontal lobe

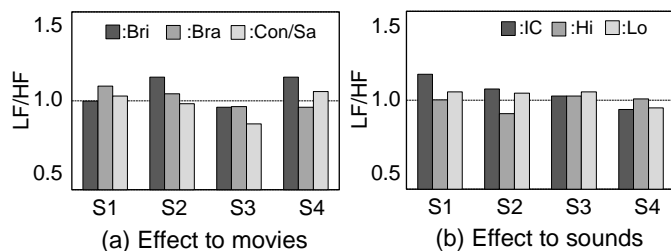


Fig. 8 Evaluation results of ECG

活性状態を評価する。心電計測では、心電波形から LF 成分 (0.05~0.15 Hz) と HF 成分 (0.15~0.40 Hz) を用い交感神経と副交感神経の優位さの検討に有効な指標 LF/HF を算出する。

#### 4.1.2 客観評価条件

評価に用いる動画効果は、Normal, Bri, Bra であり、音効果は、ME, IC, Hi, Lo である。実験のタイムプロトコルを Fig. 5 に示す。本実験では動画および音効果として 2 度に分けて実験を行う。図中の A には基準となる Normal または ME を、B, C, D には上記の効果を実施した動画および音を再生する。なお、脳血流量計測において、動画効果に対する実験では前頭葉および後頭葉、音効果に対する実験では前頭葉および側頭葉を測定対象領域とする。

#### 4.2 生体情報計測による客観評価

脳血流量計測において、Fig. 6(a) に動画効果、Fig. 6(b) に音効果に対する NIRS による前頭葉の脳血流量変化を示す。本研究では、Oxy-Hb の増加により脳が活性化しと考えられるため、前頭葉の Oxy-Hb 増加が映像好感度の高さと一致すると判断する。これより、全シーンにおいて、主観評価による印象変化で快適性が上昇した動画および音効果では Oxy-Hb の増加が確認できる。さらに、前頭葉領域のなかでも、動画効果に対しては Fig. 7(a) に示す前頭葉の上部中央、音効果に対しては Fig. 7(b) に示す前頭葉下部右側において、主観

評価と脳血流量計測による客観評価で類似の結果を示すことが確認できる。したがって、脳血流量計測に基づく映像好感度の評価にはこれらの部位が有効であると考えられる。

心電計測において、Fig. 8(a) に動画効果、Fig. 8(b) に音効果に対する結果を示す。本研究では、効果を施した B~D の LF/HF を基準 A の LF/HF で正規化した変化比を算出する。LF/HF は値が大きいほどストレス負荷が大きいことを示すが、この算出法から得られる LF/HF 変化値は値が大きいほど基準 A に対する LF/HF が低下し、ストレス負荷が小さくなるため、LF/HF 変化値の大きさが主観評価における映像好感度の上昇と一致すると判断する。これより、音効果では S1 で IC, S3 で Lo のストレス負荷が減少し、2 章の主観評価と類似した結果を得た。動画効果は主観評価との一致が得られないため、心電計測による客観評価は音に対して有効であると考えられる。

## 5. 研究成果

- (1) 動画または音効果による主観評価より、動画効果または音効果のみを再生した場合と映像に動画効果または音効果を施した場合では異なる印象を与えることがわかった。
- (2) 動画および音効果による主観評価より、各シーンにおける映像の印象変化に違いが生じることを明らかにした。動画効果ではシーンに共通の印象変化が生じたことから、シーンに異なる印象を与えるのは音効果であることがわかった。
- (3) 主観評価において SD 法の因子分析から得られた因子得点より、重回帰分析を用いて視聴覚感性モデルを構築した。これより、シーンごとに動画と音の影響度が異なることを把握した。
- (4) 脳血流と心電を用いた客観的な映像評価と主観評価との相関から、映像評価における生体計測の有用性を把握した。また、脳血流量計測では、映像評価に有効な部位を提案した。

## 参考文献

- (1) 丸山欣哉, 佐々木隆之, 視覚と聴覚間の相互作用諸効果, 日本音響学会誌, Vol.52, No.1, (1996), pp.250-259.
- (2) 岩宮眞一郎, 映像メディアにおける ((音)) の価値, 日本騒音制御工学研究発表会講演論文集 (秋), (2015), pp.77-80.
- (3) 佐竹秀一, 萩原啓, 立体映像・音響刺激が脳血行動態、心拍および主観評価に与える影響, 日本人間工学会, Vol.51, No.Supplement, (2015), pp.S338-S339.
- (4) 堀田浩史, 久保木将之, 山口雅夫, 戸井武司, 生体情報に基づく複合刺激下における音質評価, 日本音響学会講演論文集, (2009), pp.1023-1024.
- (5) 林秀彦, 國藤進, 宮原誠, 高品位映像の評価—脳波を指標とする客観評価法—, 映像情報メディア学会誌, Vol.56, No.6, (2002), pp.954-962.
- (6) 大野瑠衣, 有光哲彦, 戸井武司, 視覚と聴覚の複合刺激下での会話シーンの印象変化, 日本音響学会講演論文集 (春), (2015), pp.1465-1466.
- (7) 大野瑠衣, 有光哲彦, 三木裕一郎, 奥村武, 宮崎将哉, 戸井武司, 視聴覚調和を考慮したシーンの違いによる映像の印象変化, 日本音響学会講演論文集 (春), (2016), 2-4-5.
- (8) 大野瑠衣, 有光哲彦, 三木裕一郎, 奥村武, 宮崎将哉, 戸井武司, 高品位映像のための視聴覚寄与を考慮した視聴覚感性モデルの構築, 日本音響学会講演論文集 (春), (2016), 2-4-6.