

五度圏上の移行に基づく和音感情モデルに基づく 感情表現エージェントとの共聴体験効果

Agent Expressing Emotions Based on the Circle of Fifths on the Co-listening Experience

石田 真子^{1*} 竹村 響¹ 西野 歩真¹ 米澤 朋子¹
Mako Ishida¹ Hibiki Takemura¹ Ayuma Nishino¹ Tomoko Yonezawa¹

¹ 関西大学大学院総合情報学研究科

¹ Graduate School of Informatics, Kansai University

Abstract: 本研究は、和音に対する感情を表現する擬人化エージェントとの和音進行共聴を通して、ユーザの和音に伴う情動を促進することを目指す。エージェントの和音感情はこれまでに検討した和音の機能や移行に伴う感情生成モデルを再検討し、五度圏上の 1) 移動距離、2) 終着和音の調性、3) 回転方向、をパラメタとして感情値を算出する方式を、知覚実験結果に基づき導出した。これにより、和音遷移に対する快不快感や緊張・落ち着きをモデル化し、エージェントの顔表情に表した。検証の結果、提案モデルに基づく全ての感情表現が、エージェントへの共感性や体験への没入感、エージェントの人間らしさ、存在感、好ましさを高めることが示された。

1 はじめに

音楽が引き起こす感動は、時に人の判断や行動を変容させる。音楽による情動喚起は旋律、和声、リズム、音色といった複数の音楽要素が複合的に作用して生じるとされる [28]。中でも和音は、聴き手に対して明暗や安定・不安定の印象を与える要素であり [16]、このような和音に基づく情動的知覚は乳児期から既に備わっているとされる [31]。また、明るく感じられる和音は長三和音と分類され幸福感や陽気さなどの明るい情動を、暗く感じられる和音は短三和音と分類され悲しみなどの陰性的な情動を喚起するとされる [10, 17]。そして、不安定の印象は緊張 [3, 5]・落ち着き [19] の情動をもたらす。さらに、これらの情動は和音単体の響きだけでなく、和音間の移行も影響するとされる [2, 19]。

また、音楽体験における情動は音楽要素だけでなく、他者の存在や振る舞いと相互作用によっても変化する。複数人で音楽を聴取する共聴場面では、参加者同士が表情や身体動作、視線などを通じて感情を共有し合う音楽コミュニケーションが生じる [25, 14]。この過程を通じて、他者も自分と同じように情動を抱いていると感じる共感が生まれ、結果として音楽体験の質や楽しさが増幅される [15, 26]。

このような音楽コミュニケーションを人と仮想エージェント・ロボットとの間で実現する試みが行われてい

る。これまでに、リズムに基づく感情表現 [12, 13] や、和音の機能に基づく感情表現 [29] を行うエージェントにより、ユーザがエージェントに対して共感を抱く感覚が高まる効果が示されてきた。しかし、既存の和音機能に基づく感情表現 [29] は、その時点での和音単体の機能に対応した静的な感情表現に留まっており、和音間の移行（和音進行）による動的な影響や、連続的に推移する感情状態までは十分に扱えていない。そこで本研究では、和音進行に伴う感情変化を構造的に捉えるために、五度圏に着目した。五度圏は和音間の親和性や調性関係を幾何学的に表現できるため、1) 五度圏上の移動距離、2) 終着和音の調性、3) 回転方向、の 3 要素を用いることで、連続的な和音進行による感情の揺らぎをモデル化できると考えられる。

具体的には、和音共聴エージェントの内的状態として、五度圏上での移動角度や位置関係に基づき、和音の明暗と不安定さに基づく感情状態を算出する和音感情モデル、およびそれを反映する和音共聴エージェントを提案する。前回の和音と現在の和音から明暗度と不安定度を算出する回帰式を、予備実験 [30] によりあらかじめ導き、和音の明暗度に応じて快不快感情、不安定度に応じて緊張・落ち着き感情を、エージェントの表情や呼吸動作（ため息）で表出する。本研究では、このモデルに基づく和音共聴エージェントシステムを構築し、エージェントによる和音進行に連動した感情表出が、ユーザのエージェントへの共感性に与える影響や、和音に対する情動強化の効果を検証する。

*連絡先：関西大学

〒 569-1095 大阪府高槻市霊仙寺町 2-1-1

E-mail: k753488@kansai-u.ac.jp

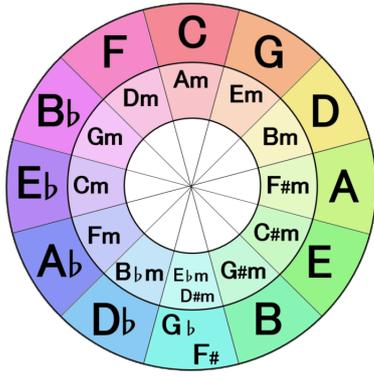


図 1: 五度圏

2 関連研究

2.1 和音の情動

音楽理論において、和音は 1) 和音単体の構成音の組み合わせ [16], 2) 和音間の移行 [2, 19, 18], によって異なる情動を喚起するとされる。

1) の和音構成音による情動として、和音の長短に基づく明暗の印象が喜びや悲しみなどの快・不快感 [16, 10, 17] を、安定・不安定の印象が緊張や落ち着きをもたらす [3, 5, 4, 19], この感覚は乳児期から既に備わるとされる [31]。

また、2) の和音移行による感情的影響について、和音の安定・不安定の印象は絶対的なものではなく、前後の和音の組み合わせにより相対的に変化する。音楽理論ではこの連続的知覚の指標として五度圏が用いられる。五度圏は、全 12 調を完全 5 度/4 度順に円環状に配置し、調や和音間の親近性を距離として可視化する (図 1)。五度圏上で近接する調への遷移は自然に知覚される一方、遠い調への移動は不自然さや緊張感を高める傾向があるとされる [18]。

そこで我々は、五度圏の移動角度と絶対位置を活用し、和音情動のモデル化を試みた [30]。結果、和音間の 3 要素: 1) 五度圏上の移動距離 (五度圏距離), 2) 終着和音の調性, 3) 回転方向が、明暗と安定性の印象に影響を与えることが示唆された。具体的には、五度圏距離が長いほど不安定に、終着和音が短調の場合は暗い印象を与え、長調で移動距離が短い場合は安定的で明るい印象を与える傾向が示された。また、回転方向については、反時計回りの方が時計回りより安定した印象を与えることも示された。

本研究では、これら 1) - 3) の 3 要素を変数として和音の明暗と安定性を算出する和音知覚モデルを、エージェントの和音を感じる心として応用する。

2.2 エージェントとの音楽的共感

人間はエージェントやロボットを社会的存在とみなし、そのふるまいから心の状態を推測する [22]。例えば、共感的にふるまう対話エージェントに対して、ユーザが共感するだけでなく、エージェントから共感されていると感じる喜びも生じるとされる [20]。

また、ユーザとエージェントがゲームや映画、物語などの同一のメディアを共有しながらコミュニケーションをする場合、メディアに対してエージェントが示す感情表現は、ユーザにエージェントの心的状態を想定させ、結果としてユーザ自身の感情にも影響を与える [6, 24, 27, 1]。

特に、音楽メディアを共有する場合、エージェントやロボットの身体的・感情的表現が、ユーザの音楽体験に対する情動を高めたり、エージェントやロボットへの親しみを強化する可能性が示されている。例えば、ロボットとの音楽共聴において、ロボットが音楽に同調した動作を示すことでユーザにロボットの楽しさを感じさせ音楽の印象を強めたり [11], ロボットがジェスチャーで観客の感情を演奏者に伝えて音楽が変化すると観客の感情が強化されたり演奏者との一体感を感じさせた [9]。

また、音楽が励起させるエージェントの感情をどのようにモデル化するかや、内部感情に応じた表現について、リズムに励起される共聴エージェントの感情が表出されると、ユーザのリズムの感じ方が変化したり共感する感覚を高める効果は確認した [12, 13] が、和音は、興奮やノリだけではなく喜びや悲しみなどポジティブ度の異なる感情の表現色彩が増えるとも考えられる。本研究では和音に着目し、和音の響きやその移行によって喚起される感情をエージェントの内部状態としてモデル化し、それを表情や動作により表出することで、和音とエージェント双方への共感を促進する手法を提案する。

3 システム

3.1 システム概要

本研究では、和音に喚起される情動を表現する和音共聴エージェントシステムを提案する。本エージェントは、五度圏における和音移行の移動距離、調性、回転方向といった特徴量に基づいて快・不快および緊張・落ち着きを生成し、それらを外観を通じた視覚的表現として提示する。

提案システム (図 2) は、和音刺激を生成する和音再生部 [29], 和音感情を生成する和音感情生成部, 生成感情を外観で表現する外観表現部から構成される。

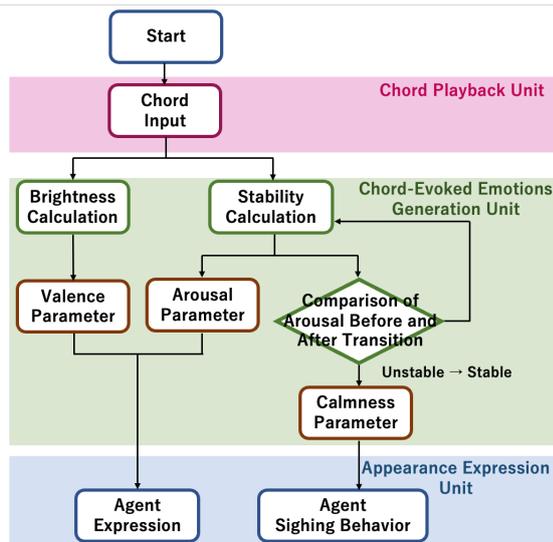


図 2: システムフロー

3.2 和音再生部

和音は、ピッチの安定性や適度な複雑さの観点からピアノ音色の3和音を採用した。また、五度圏上の全和音(図1)から自由に構成した和音進行(計16音:4音/小節×4小節)を再生でき、テンポも可変(初期値50BPM)で各和音は1拍(四分音符)の間保持される。和音再生部はPureData Vanilla¹で実装した。再生和音の例を表1に示す。

3.3 和音感情生成部

和音感情生成部は、ラッセルの感情円環モデル[23]に基づき、和音の明暗度(*Brightness*)から快感度(*Valence*)を、不安定度(*Instability*)から覚醒度(*Arousal*)を算出する。

まず、和音の明暗度と不安定度に関して、五度圏の移動角度と和音の明暗・不安定さの印象との関係を調査した先行研究[30]に基づく線形回帰式(数式1: $R^2 = 0.177$, 数式3: $R^2 = 0.182$)により算出する。これらの回帰モデルの決定係数は高いとは言えないものの、いずれも分散分析の結果において $p < .001$ を示しており、五度圏上の移動角度と和音の印象との間に統計的に有意な関係が確認されている。そこで、本研究では、これらの線形回帰式をエージェントの明暗と安定不安定を感じる指標として適用した。和音再生部より受信した和音データから、前回発音された和音(V_{base})と今回発音された和音(V_{target})における、五度圏上の移動最短距離(D :最長6)、回転方向(R :反時計回り0, 時計周り1)、調性変化(V :長調1, 短調0)、の

¹<https://puredatajapan.info>

パラメタに対して係数を割り当て、下記の線形回帰式にそれぞれ代入し、明暗度と不安定度を算出する。

$$Stability = 5.256 - 1.301 \times D + 0.238 \times V_{base} + 0.246 \times R + 1.066 \times V_{target} \quad (1)$$

$$Instability = (\max(Stability) + \min(Stability)) - Stability \quad (2)$$

$$Brightness = 3.658 - 0.207 \times D + 0.363 \times V_{base} - 0.066 \times R + 1.605 \times V_{target} \quad (3)$$

算出された明暗度・不安定度は0-1に正規化し、それぞれエージェントの快・不快感を生成する快感度と緊張を生成する覚醒度として適用する(表2)。また、不安定から安定の遷移(和音解決)で生じる一時的な解決感[19]も、緊張に対する落ち着きとしてモデル化する。これは、基準和音(V_{base})から評価和音(V_{target})への覚醒度が0.13以上低下する場合かつ、評価和音の覚醒度が0.4未満の場合に $Calmness = 1$ ($Arousal = 0$)として設定され、次節で述べるため息表現後に0にリセットされる。これら快感度、覚醒度、落ち着き値の3値を和音励起感情モデルの出力として外観表現部に送信する。

3.4 外観表現部

外観表現部は、和音感情生成部の快感度・覚醒度・落ち着き値に応じ、エージェントの顔表情とため息のふるまいを2次的に表示する。

この外観を動かすアニメーションの表現(図3)には、1)和音の明暗に対する快不快感に応じた顔表情と安定・不安定に対する緊張(覚醒)に応じた顔表情、2)落ち着きに対するため息のふるまいを実装した。

1)の表情は、Facial Action Coding SystemにおけるAction Unit[7](AUs)での顔パーツの操作がエージェントの表情表出における自然さをもたらすとされる[21]ことから、エージェントの連続的な快感と覚醒の感情パラメタを、対応するAUsにマッピングした(図3)。これらの感情は感情値算出(和音が切り替わる)タイミングと同時に更新される。一方で、2)の落ち着きでは、1)の連続的AUsマッピングの一部とは独立して、和音解決による緊張が解きほぐされて息をつく様子をため息として表現する。具体的には、落ち着き値が1に設定された際に離散的にため息のふるまいがトリガーされる(図3)

エージェントの外観はProcessing²で実装し、Open Sound Control 通信³を用いて和音再生部からの和音

²<https://processing.org>

³<https://opensoundcontrol.stanford.edu/>

表 1: 再生和音の例

| 小節 | 1 | | | | 2 | | | | 3 | | | | 4 | | | |
|-------------|-----|-----|-----|-----|-----|-----|---------------|-----|--------------|--------------|----------------------|--------------|-----|-----|-----|-----|
| 拍 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 和音例 | C | F | G | C | Am | Dm | E | Am | Fm | Bm | E \flat | A | F | G | C | C |
| 五度圏 回転位置 | - | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 |
| 構成音 | CEG | FAC | GBD | CEG | ACE | DFA | EG \sharp B | ACE | FA \flat C | BDF \sharp | E \flat GB \flat | AC \flat E | FAC | GBD | CEG | CEG |

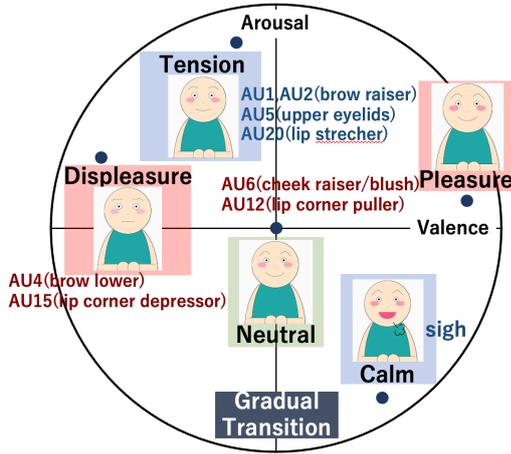


図 3: Action Unit に対応したエージェントの感情表出

データに連携させた。

4 実験

4.1 実験目的

本稿では、五度圏上の移動角度に基づいて快感度と覚醒度を変化させ表情を変える和音共聴エージェントの内部モデルの妥当性や、感情表現がユーザの共感促進や和音による情動強化の効果を検証する。

4.2 実験仮説

和音励起感情モデルに沿ったエージェントの感情表現は、

- H1 和音と一致した印象を与える。
- H2 人間らしい印象を与える。
- H3 他者と音楽を聴く感覚を与える。
- H4 他者と共感し合う感覚を与える。
- H5 和音の示す感情を理解させる。
- H6 音楽の魅力を高める。

H7 エージェントの好感度を高める。

4.3 実験条件

提案する和音感情モデルの感情変化に対する、快・不快表現の変化方向（要因 A, A1: 正変化, A2: 逆変化）、緊張表現の変化方向（要因 B, B1: 正変化, B2: 逆変化）、落ち着き表現の有無（要因 C, C1: あり, C2: なし）の 3 要因 8 条件の被験者内実験計画とした。

要因 A にランダム変化を導入しない理由は、和音と表情が偶然に一致する区間が生じ、モデルの有効性を厳密に評価できない可能性があるためである。そこで、本実験ではあえてモデルと真逆の反応を示す逆変化を比較対象として設定した。

4.4 実験参加者

本実験は、参加に際する中止の自由などを含む説明に対し同意した、日本語を母国語とする、20 歳–28 歳（平均年齢：22.225 歳，標準偏差：2.094）の 40 名（男性 34 名，女性 6 名，うち過去 3 年以上の楽器演奏経験者 12 名）が参加した。

4.5 実験刺激

和音刺激には、適度な複雑さと明暗を表現可能な 3 和音を用いた。また、和音刺激の調性は、幼少時の音楽から用いられるハ長調とした。あらかじめ 4 小節分の和音進行で、1 小節あたり 4 種の和音が進行する合計 16 音（4 小節）の刺激を準備した。使用した和音の明暗度、不安定度とこれら対応する感情値（快感度、覚醒度と落ち着き値）を表 2 に示す。

短い和音刺激に一般的な楽曲に使用される和音進行のみを用いると、通常は大きな感情遷移が起きにくい。また、音楽的な要因に加え、エージェントの表情に大きな変化が生じにくい。そこで、実験刺激中にエージェントの緊張方向および快・不快方向の全ての種類の表情（図 3）を提示できるような和音進行を作成した。

表 2: 実験に使用した和音刺激に対応する感情値

| No. | Chord | Brightness | Valence | Instability | Arousal | Calmness |
|-----|-------|------------|---------|-------------|---------|----------|
| 1 | C | 0.000 | 0.000 | 0.000 | 0.160 | 0 |
| 2 | F | 5.419 | 0.937 | 0.055 | 0.167 | 0 |
| 3 | G | 5.146 | 0.853 | 1.110 | 0.302 | 0 |
| 4 | C | 5.419 | 0.937 | 0.055 | 0.000 | 1 |
| 5 | Am | 3.334 | 0.300 | -0.426 | 0.105 | 0 |
| 6 | Dm | 3.451 | 0.336 | 1.359 | 0.334 | 0 |
| 7 | E | 4.783 | 0.743 | 5.251 | 0.832 | 0 |
| 8 | Am | 3.814 | 0.447 | 5.024 | 0.803 | 0 |
| 9 | Fm | 2.830 | 0.147 | 5.262 | 0.834 | 0 |
| 10 | Bm | 2.350 | 0.000 | 3.961 | 0.667 | 0 |
| 11 | Eb | 4.369 | 0.616 | 3.950 | 0.666 | 0 |
| 12 | A | 4.318 | 0.601 | 6.560 | 1.000 | 0 |
| 13 | F | 4.798 | 0.747 | 3.958 | 0.667 | 0 |
| 14 | G | 5.146 | 0.853 | 1.110 | 0.000 | 1 |
| 15 | C | 5.419 | 0.937 | 0.055 | 0.000 | 1 |
| 16 | C | 5.626 | 1.000 | -1.246 | 0.000 | 1 |

4.6 実験手順

本実験はオンラインで実施し、実験説明および評価画面は GoogleForm の Web サイトを用いて教示した。

実験刺激である音楽とエージェントの実験刺激映像は、各条件の評価フォーム内に提示した。エージェントの表示サイズは、実験参加者（ノート PC、もしくはデスクトップ PC 保持者限定）各自が所持する PC のサイズに依存して変化するが GoogleForm 上で提示可能な最大サイズで提示した（14 インチ PC の場合 4.7 インチ程度）。これにより、タブレットやスマートフォンでの小型端末での動画視聴を防いだ。

また、参加者の聴覚的な不快感や疲労による実験への悪影響を回避するため、全員各自の所持するイヤホンを使用し、各自が聞きやすい音量に調整させ、設定後は変更しないよう指示した。このことにより参加者ごとに音量は異なる可能性が高いが、主観的に快適でかつ明瞭に聴き取れる音量での聴取を優先した。実験前にインフォームドコンセントとして実験内容や個人情報扱いを説明し、それらに同意した者のみ実験に参加した。

実験参加者は実験の 8 試行に取り組み、各条件を体験した後に、エージェントとの共聴体験について下記の評価項目に回答を求めた。各実験条件提示順はラテン方格法 [8] に基づいてカウンタバランスを考慮して決定した。

4.7 評価項目

各仮説に対応する評価項目を準備した（表 3）。Visual Analogue Scale 法を用い、0-99 点（0: 全くそう思わない、99: とてもそう思う）とし、直感的評価を行うための数値設定スライダーとして、操作中は点数を表示せず操作後に点数を表示する仕様のスライダーシステムを準備した。

表 3: 評価項目

| No. | 評価項目 | 仮説 |
|-----|-------------------------------------|----|
| Q1 | エージェントの表情は音楽の雰囲気（音楽から得られる感情）に合っていた。 | H1 |
| Q2 | エージェントの表現に人間らしさを感じた。 | H2 |
| Q3 | エージェントと一緒に音楽を聴いているように感じた。 | H3 |
| Q4 | エージェントの感情にあなたが共感した。 | H4 |
| Q5 | エージェントから共感されているように感じた。 | H4 |
| Q6 | 音楽の中に、喜び-悲しみの感情の移り変わりを感じた。 | H5 |
| Q7 | 音楽の中に、緊張-落ち着きの感情の移り変わりを感じた。 | H5 |
| Q8 | 音楽に引き込まれるような感覚があった。 | H6 |
| Q9 | エージェントを好ましく感じた。 | H7 |

4.8 実験結果

実験により得られた各評価結果に対して繰り返しのある分散分析を有意水準 $\alpha = 0.05$ で実施した。多重比較は Ryan's method に従った。図 4 に各評価項目の平均値およびその標準偏差を、表 4, 5, 6 に分散分析の結果を示す。

4.8.1 要因 A,B,C それぞれの主効果

全ての評価項目において要因 A の主効果が確認され、いずれも $A1 > A2$ が示された。また、Q6, Q7 を除く全ての項目で効果量偏 η^2 が 0.3 を上回り、他に有意差が確認された項目の効果量と相対的に比較して大きな値を示した。このことから、快・不快感情遷移が全ての仮説を支持する強力な要因であることが示唆された。

次に要因 B の主効果は Q1, Q4, Q9 で確認され、全て $B1 > B2$ が示された。つまり、これらの評価項目において要因 A の主効果も確認されたことも踏まえると、快・不快感情遷移と緊張表情遷移のそれぞれが、提案モデルに対応した表現を示すことによって、エージェントが和音に沿って正しい感情を表出しているように感じさせること、エージェントに共感する感覚を促進すること、エージェントへの好感を高めることが示唆された。

最後に、要因 C の主効果は、全評価項目の中で Q7 にのみ確認され、 $C1 > C2$ が示された。つまり、落ち着き表現がある方が和音進行内の緊張-落ち着きの変化の理解を有意に深めることが示唆された。

4.8.2 2 要因間交互作用

全ての評価項目において、要因 BC 間に交互作用が確認された。まず、Q6 を除く全ての項目で、C1 において $B1 > B2$, B1 において $C1 > C2$ が示された。エージェントの落ち着き表現と正方向の緊張表情遷移の両方が、Q6 以外の全ての評価を高めた。このことから、エージェントの覚醒軸に関わる緊張と落ち着きの両表

表 4: 主効果に関する ANOVA の結果

| Evaluation Items | FactorA | | | | FactorB | | | | FactorC | | | |
|------------------|---------|--------|-------------|----------------------|---------|-------|-------------|----------------------|---------|--------|-------------|----------------------|
| | F | p | Effect Size | Multiple Comparisons | F | p | Effect Size | Multiple Comparisons | F | p | Effect Size | Multiple Comparisons |
| Q1 | 47.044 | <.001* | 0.553 | A1>A2 | 7.647 | .009* | 0.168 | B1>B2 | 0.000 | 1.000 | 0.000 | - |
| Q2 | 26.931 | <.001* | 0.415 | A1>A2 | 3.662 | .063 | 0.088 | - | 0.253 | .618 | 0.007 | - |
| Q3 | 27.939 | <.001* | 0.424 | A1>A2 | 0.861 | .359 | 0.022 | - | 1.442 | .237 | 0.037 | - |
| Q4 | 48.336 | <.001* | 0.560 | A1>A2 | 11.912 | .001* | 0.239 | B1>B2 | 0.088 | .768 | 0.002 | - |
| Q5 | 21.742 | <.001* | 0.364 | A1>A2 | 1.501 | .228 | 0.038 | - | 0.099 | .755 | 0.003 | - |
| Q6 | 7.385 | .010* | 0.163 | A1>A2 | 0.001 | .980 | 0.000 | - | 0.035 | .852 | 0.001 | - |
| Q7 | 4.345 | .044* | 0.103 | A1>A2 | 1.523 | .225 | 0.039 | - | 22.362 | <.001* | 0.370 | C1>C2 |
| Q8 | 24.138 | <.001* | 0.388 | A1>A2 | 2.669 | .111 | 0.066 | - | 1.685 | .202 | 0.042 | - |
| Q9 | 45.269 | <.001* | 0.544 | A1>A2 | 7.046 | .012* | 0.156 | B1>B2 | 3.476 | .070 | 0.084 | - |

*: $p < .05$

現相乗効果的に機能し、仮説 H1-H4, H5 の覚醒軸に関わる評価側面と H6, H7 を支持することが示唆された。そして、Q1, Q2, Q4-Q6 では、いずれも B2 において C2>C1 が確認され、緊張表情遷移がモデルと逆変化する場合、落ち着き表現がない方が高評価となった。これは、緊張表現が和音の感情変化と連携していない状況で、落ち着き表現を加えると印象に矛盾が生じ、実験参加者を混乱させた可能性が考えられる。つまり矛盾がないよう組み合わせる (B1C1 または B2C2) ことで評価が高くなるといえる。

4.8.3 3 要因間交互作用

要因 ABC 間の交互作用は、Q4, Q6, Q7, Q8 で確認された。これらの評価項目は共通して、[A1B1] の条件で C1>C2, [A1C1] の条件で、B1>B2 が確認された。また、Q4, Q7, Q8 においては、[B1C1] の条件下で A1>A2 も確認された。これらのことから、3 要因全てが連動する [A1B1C1] の条件でエージェントへ共感する感覚 (Q4), 和音が示す喜び-悲しみ (Q6), 緊張-落ち着き (Q7) の感情理解, 体験への没入感 (Q8) を高めることがわかった。

また Q7 では、[A2B1][A2B2] それぞれの条件下で C1>C2 が確認され、両方の表情遷移が提案モデルに連携していない条件下においてさえも、落ち着き表現がある方が有意に評価が高かった。これは、落ち着きを表現するため息の表現が、他の要因 (A, B) から独立し、単独で「落ち着き」のサインとして強く機能したことを示唆している

さらに、Q4, Q6, Q8 は、[B2C1] や [B2C2] の条件下でも、A1>A2 が確認された。つまり、緊張表情遷移が提案モデルに対して逆変化の場合でも、提案モデルに沿った快・不快表情遷移が独立して、エージェントへ共感する感覚、喜び-悲しみの感情理解、没入感の効果に対して有効に働くことが示された。

一方、Q4, Q6 は、[A1B2] の条件下では、C2>C1 となり、快不快表現が正しく変化する時に、緊張表現が

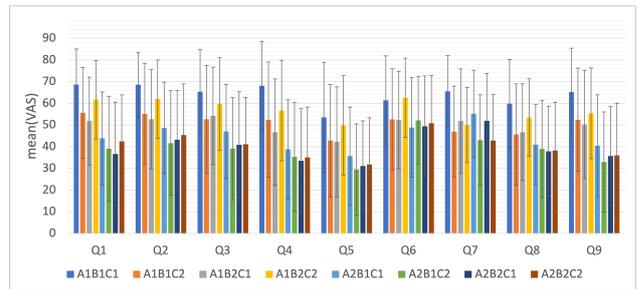


図 4: VAS の平均値と標準偏差

矛盾すれば、落ち着き表現も伴わない方が共感や喜び-悲しみの感情を理解しやすい可能性が示された。

5 考察

本研究では、五度圏での和音位置と移動距離に基づくエージェントの和音感情表現が、エージェントとの共感し合う感覚の促進や和音による情動強化に与える効果を検証した。実験の結果を踏まえ以下に考察する。

感情的基盤としての快・不快感表現の優位性: まず、全評価項目において、快・不快感表情遷移 (要因 A) の主効果が確認された。これは、提案モデルに基づく快・不快感表現が共聴体験の質を担保する根本的な基盤として有効に機能したことが示唆された。和音に対応した快・不快の感情的理解が、エージェントとの共聴時の共感的な印象形成の基盤であると考えられる。

複合的な感情表現の整合性と認知的混乱: 覚醒度に関わる緊張表情遷移 (要因 B) と落ち着き表現 (要因 C) の間の交互作用 ([B1C1]: 緊張が正変化・ため息ありで相乗的に評価が高まる一方で、[B2C1]: 緊張が逆変化・ため息あり < [B2C2]: 緊張が逆・ため息なし, 図 5) からはエージェントの感情設計における表現の整合性の重要性が示唆された。これは、エージェントが和音感情に反した表現 [B2] を示すことよりも、感情チャンネル間に矛盾を含む表現 [B2C1] を行う方がユーザの違和感

表 5: 2 要因間交互作用に関する ANOVA の結果

| Evaluation Items | FactorA × B | | | FactorA × C | | | FactorB × C | | | | |
|------------------|-------------|------|-------------|-------------|------|-------------|-------------|--------|-------------|--------------------|----------------------|
| | F | p | Effect Size | F | p | Effect Size | F | p | Effect Size | Simple Main Effect | Multiple Comparisons |
| Q1 | 1.327 | .257 | .034 | .333 | .567 | .009 | 15.955 | <.001* | .296 | B(C1),C(B1),C(B2) | B1>B2,C1>C2,C2>C1 |
| Q2 | 3.526 | .068 | .085 | .137 | .714 | .004 | 12.876 | <.001* | .253 | B(C1),C(B1),C(B2) | B1>B2,C1>C2,C2>C1 |
| Q3 | 0.422 | .520 | .011 | .214 | .646 | .006 | 13.387 | <.001* | .261 | B(C1),C(B1) | B1>B2,C1>C2 |
| Q4 | 2.765 | .105 | .068 | .189 | .666 | .005 | 10.915 | .002* | .223 | B(C1),C(B1),C(B2) | B1>B2,C1>C2,C2>C1 |
| Q5 | 0.014 | .906 | .000 | .164 | .688 | .004 | 12.350 | .001* | .245 | B(C1),C(B1),C(B2) | B1>B2,C1>C2,C2>C1 |
| Q6 | 0.292 | .592 | .008 | .003 | .960 | .000 | 5.369 | .026* | .124 | C(B2) | C2>C1 |
| Q7 | 2.53 | .120 | .062 | .262 | .611 | .007 | 7.118 | .011* | .158 | B(C1),C(B1) | B1>B2,C1>C2 |
| Q8 | 0.526 | .473 | .014 | .621 | .436 | .016 | 12.526 | .001* | .248 | B(C1),C(B1) | B1>B2,C1>C2 |
| Q9 | 2.358 | .133 | .058 | .068 | .795 | .002 | 9.559 | .004* | .201 | B(C1),C(B1) | B1>B2,C1>C2 |

* : $p < .05$

表 6: 3 要因間交互作用に関する ANOVA の結果

| Evaluation Items | FactorA × B × C | | | | | |
|------------------|-----------------|------|-------------|-------------------------------------------------------------|-----------------------------|---|
| | F | p | Effect Size | Simple Main Effect | Multiple Comparisons | |
| Q1 | 1.987 | .167 | .050 | - | - | - |
| Q2 | 3.545 | .067 | .085 | - | - | - |
| Q3 | 0.755 | .390 | .019 | - | - | - |
| Q4 | 7.718 | .008 | .169 | A(B1C1),A(B1C2),A(B2C1),A(B2C2), B(A1C1),C(A1B1),C(A1B2) | A1>A2, B1>B2,C1>C2,C2>C1 | |
| Q5 | 1.517 | .226 | .038 | - | - | - |
| Q6 | 6.006 | .019 | .136 | A(B2C1),A(B2C2),B(A1C1), C(A1B1),C(A1B2) | A1>A2,B1>B2, C1>C2,C2>C1 | |
| Q7 | 5.602 | .023 | .128 | A(B1C1),B(A1C1) C(A1B1),C(A2B1),C(A2B2) | A1>A2,B1>B2 C1>C2 | |
| Q8 | 8.801 | .005 | .188 | A(B1C1),A(B2C1),A(B2C2), B(A1C1),C(A1B1) | A1>A2, B1>B2,C1>C2 | |
| Q9 | 1.856 | .181 | .047 | - | - | - |

* : $p < .05$

や混乱を招き、体験の質を損なったことが推測される。つまり、単にエージェントの表現チャンネルを増やせばよいわけではなく、複数の感情チャンネルが五度圏理論に照らして互いに論理的に整合していることが、エージェントの人間らしさ (Q2) やエージェントとの共感し合う感覚 (Q4, Q5) の獲得に不可欠であると考えられる。

高次な共感体験を生む感情軸の統合: エージェントへの共感する感覚 (Q4) や、和音の示す感情理解 (Q6, Q7), 没入感 (Q8) などの体験評価において 3 要因全ての交互作用が確認され、特に、Q4, Q7, Q8 では全ての要因が整合的に連動する条件 [A1B1C1] で最も高い評価が得られた (図 4)。つまり、エージェントへの共感や、和音の示す感情理解、共聴体験への没入といった高次の体験を生み出すには、快・不快と覚醒の感情軸が、互いに矛盾なく統合的に表現される必要があると考えられる。音楽が喚起する情動は、快・不快感だけでなく、緊張・落ち着きなど、複数の感情が複雑に絡み合って生じる [28]。本結果は、提案する和音共聴エージェントが、この整合的に統合された感情状態を再現できている可能性を支持するものであり、この再現により、ユーザはエージェントを「音楽を深く理

解・共有している存在」として受容したことが示唆される。

ため息の効果: さらに、落ち着き表現として用いたため息 (要因 C) は、Q7 (緊張-落ち着きの理解) において、主効果が確認され、効果量は偏 $\eta^2 = 0.37$ を示した。これは Q7 において同じく有意差が確認された要因 A の効果量 (偏 $\eta^2 = 0.103$) を大きく上回った。また、他の感情表現が間違った表現 [A2B2] の状況下でさえも評価を高めた。これは、ため息が他の連続的な表情変化とは異なり、和音の解決による落ち着きを示す解釈容易なシグナルとして機能したと考えられ、緊張と落ち着きの和音感情理解を促進する上では非常に有効な表現であったことがわかる。

6 おわりに

本研究では、エージェントとの和音に応じた感情表現がエージェントや和音への共感を促進する効果について検証した。

その結果、快・不快感と緊張・落ち着きの表現はそれぞれ、和音の示す感情と一致した印象を高め、エージェ

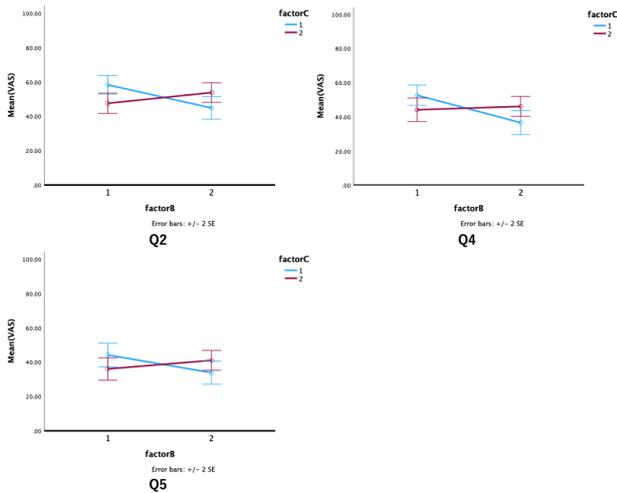


図 5: Q2 と Q4 と Q5 の交互作用グラフ

ントの人間らしさや好感度，他者としての印象，エージェントと共感し合う感覚を高めた。さらに，和音の示す感情の理解を促進し，共聴体験への没入感を高めた。これらの効果は，単一の感情表現だけでは完全には生み出せず，3要素全てを統合することで相補的にユーザの共聴感覚やエージェントへの共感する感覚を高めることがわかった。このように，エージェントの和音に対応した感情表現が，ユーザの和音による情動を強化した。

今後は本研究の限界に対し，表現するエージェントの感情種類の妥当性を調査すると共に，他の種類の和音に対する知覚を調査することで，多様な和音進行に対応したエージェントの反応の実装を目指す。

謝辞

本研究の一部は，科研費 24K02977, 23K11278, 22K19792, および一般財団法人カワイサウンド技術・音楽振興財団サウンド技術振興部門における「音楽共感型エージェントによるリトミック支援メタバース技術の開発」として助成を受けた。

参考文献

- [1] Timothy Bickmore and Amanda Gruber. Relational agents in clinical psychiatry. *Harvard review of psychiatry*, Vol. 18, No. 2, pp. 119–130, 2010.
- [2] Emmanuel Bigand, Barbara Tillmann, Brigitte Poulin, David A. D’Adamo, and Fabrice Madurell. The effect of harmonic context on

chord perception. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 25, No. 1, pp. 184–197, 1999.

- [3] Norman D. Cook. *Tone of Music*. Oxford University Press, Oxford, 2002.
- [4] Norman D. Cook. Calculation of the acoustical properties of triadic harmonies. *The Journal of the Acoustical Society of America*, Vol. 142, No. 6, pp. 3748–3755, 2017.
- [5] Norman D. Cook and Takashi X. Fujisawa. The psychophysics of harmony perception: Harmony is a three-tone phenomenon. *Empirical Musicology Review*, Vol. 1, No. 2, pp. 106–126, 2006.
- [6] Sandra Costa, Alberto Brunete, Byung-Chull Bae, and Nikolaos Mavridis. Emotional storytelling using virtual and robotic agents. *International Journal of Humanoid Robotics*, Vol. 15, No. 03, p. 1850006, 2018.
- [7] Paul Ekman and Wallace V Friesen. *Facial action coding system*. Consulting Psychologists Press, 1978.
- [8] Leonhard Euler. De quadratis magicis. *Commentationes arithmeticae collectae*, pp. 593–602, 1849.
- [9] Marcello Giardina, Salvatore Tramonte, Vito Gentile, Samuele Vinanzi, Antonio Chella, Salvatore Sorce, and Rosario Sorbello. Conveying audience emotions through humanoid robot gestures to an orchestra during a live musical exhibition. In *Complex, Intelligent, and Software Intensive Systems: Proceedings of the 11th International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS-2017)*, pp. 249–261. Springer, 2018.
- [10] Kate Hevner. The affective character of the major and minor modes in music. *The American Journal of Psychology*, Vol. 47, No. 1, pp. 103–118, 1935.
- [11] Guy Hoffman, Shira Bauman, and Keinan Vannun. Robotic experience companionship in music listening and video watching. *Personal and Ubiquitous Computing*, Vol. 20, No. 1, pp. 51–63, 2016.

- [12] Mako Ishida, Hibiki Takemura, and Tomoko Yonezawa. Empathetic co-listening agent based on exciting/positive emotions and internal-time. *Human Interface Society*, Vol. 25, No. 4, pp. 283–298, 2023. (in Japanese).
- [13] Mako Ishida, Hibiki Takemura, and Tomoko Yonezawa. Confused expressions of co-listening agent based on deviations from expected rhythm. *Human Interface Society*, Vol. 27, No. 3, pp. 213–224, 2025. (in Japanese).
- [14] Patrik N. Juslin. From mimesis to catharsis: Expression, perception, and induction of emotion in music. pp. 85–115. Oxford University Press, 2005.
- [15] Patrik N Juslin, Simon Liljeström, Daniel Västfjäll, and Lars-Olov Lundqvist. *How does music evoke emotions? Exploring the underlying mechanisms*. Oxford University Press, 2010.
- [16] Patrik N Juslin and John A Sloboda. *Music and emotion*. D. DEUTSCH (Org.), 2001.
- [17] Marianne P. Kastner and Robert G. Crowder. Perception of the major/minor distinction: Developmental considerations. *Psychomusicology: A Journal of Research in Music Perception*, Vol. 9, No. 2, pp. 90–97, 1990.
- [18] Carol L. Krumhansl. *Cognitive foundations of musical pitch*. Oxford University Press, 1990.
- [19] Leonard B. Meyer. *Emotion and Meaning in Music*. University of Chicago Press, 1956.
- [20] Ana Paiva, Iolanda Leite, Hana Boukricha, and Ipke Wachsmuth. Empathy in virtual agents and robots: A survey. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, Vol. 7, No. 3, pp. 11:1–11:40, 2017.
- [21] Catherine Pelachaud. Studies on gesture and facial expression for embodied conversational agents. In Justine Cassell, Joseph Sullivan, Scott Prevost, and Elizabeth Churchill, editors, *Embodied Conversational Agents*, pp. 208–235. MIT Press, 2000.
- [22] Byron Reeves and Clifford Nass. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press, 1996.
- [23] James A Russell. A circumplex model of affect. *Journal of personality and social psychology*, Vol. 39, No. 6, pp. 1161–1178, 1980.
- [24] Jeongwoo Ryu, Kyusik Kim, Dongseok Heo, Hyungwoo Song, Changhoon Oh, and Bongwon Suh. Cinema multiverse lounge: Enhancing film appreciation via multi-agent conversations. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 2025.
- [25] Claude E. Shannon and Warren Weaver. コミュニケーションの数学的理論：情報理論の基礎. 明治図書出版, 1969. Translated from: *The Mathematical Theory of Communication (1949)*, (in Japanese).
- [26] Takanori Shibata and Kazuyoshi Wada. Robot therapy: a new approach for mental healthcare of the elderly—a mini-review. *Gerontology*, Vol. 57, No. 4, pp. 378–386, 2011.
- [27] Masahiro Shiomi, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. Interactive humanoid robots for a science museum. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*, pp. 305–312. ACM, 2006.
- [28] John A Sloboda and Patrik N Juslin. Psychological perspectives on music and emotion. *Music and emotion: Theory and research*, pp. 71–104, 2001.
- [29] Hibiki Takemura, Mako Ishida, and Tomoko Yonezawa. Appropriateness of facial expressions of chord empathetic agents on user’s sense of music. *IEICE; technical reports*, Vol. 121, No. 349, pp. 13–18, 2022. (in Japanese).
- [30] Hibiki Takemura, Mako Ishida, and Tomoko Yonezawa. Evaluation of chord impressions based on music-theoretic parameters and a study on systematization for artificial emotion generation. *IPSJ SIG technical reports*, Vol. 2023, No. 2, pp. 1–16, 2023. (in Japanese).
- [31] Laurel J Trainor, Christine D Tsang, and Vivian HW Cheung. Preference for sensory consonance in 2-and 4-month-old infants. *Music Perception*, Vol. 20, No. 2, pp. 187–194, 2002.