

人の身体性に着目した音階の生成過程に関する考察

中山 功一[†] 下原 勝憲[†]

本論文では、人の身体性に着目した音階の生成過程について考察する。「同一周波数成分を多く含む音を快適と感じる人の聴覚（身体性）」と、「倍音を多く含むという特徴を持つ楽器（環境）」との相互作用により、現在の12音階が創発的に生成されるという仮説を立て、音階発生過程をマルチエージェント・シミュレーションを用いて検証する。

Generation process of musical scale influenced by human being

KOICHI NAKAYAMA[†] and KATSUNORI SHIMOHARA[†]

This paper describes the generation process of music scales. We assume that the conventional music scale is generated by interactions between human's auditory organs and musical instruments. We simulated the scales generation process by a multi-agent system. As a result, a music scale similar to the conventional one emerged.

1. はじめに

現代音楽の多くでは、1オクターブがほぼ対数的に12半音に分けられた純正律（Pure Temperament）、または平均律（Equal Temperament）が用いられる。一部地域の民族音楽では、1オクターブを5等分や7等分した音階が用いられており、日本でも非12音平均律による作曲がこころみられている^{1),2)}が、一般的に人が心地よいと感じるのは12半音に分けられた音階（12音階）であるといわれている。歴史上、ピタゴラス音階から純正律、そして平均律へと使用される音階は変化してきたことは知られているが、そもそもなぜ、この12音階が用いられるのであろうか？

古代ギリシャにおいて音の協和性を数学的に探求したピタゴラスは、弦楽器を用いた実験から、二つの弦の長さの比が単純な整数比となるときに、二つの音の響きが心地よいことを発見したと言われている。弦楽器では倍音成分を多く含むため、二つの弦の長さの比が単純な整数比となる場合、二つの弦から発する音に同一周波数成分が多く含まれる。一方、人は、同一周波数成分を多く含む複数の音を聞くと心地よいと感じる。このため、人は弦の長さの比が単純な整数比となるときに音の響きが心地よいと感じると考えられている。これらのように、音階を科学的・解析的に理解す

る研究はなされているのに対し、工学的・構成論的な方法論による研究、すなわち、音階の創発的な生成にむけた研究は、あまり検討されていない。

人が身体的に同一周波数成分を多く含む音を快適と感じる感性（身体性）を共有し、弦楽器のように音を発生させる機器（環境）が倍音を多く含むという特徴を持つ場合、この身体性と環境との相互作用により現在の12音階が成立したとは考えられないであろうか？本論文では、発音機器としての楽器の特徴と、受音機器としての身体性の特徴、および心地よいと感じられる音ほどより多くの頻度で発せられるようになるメカニズムを、マルチエージェント・システムを用いてモデル化する。このモデルを用いたシミュレーションにより、音階が創発的に生成され、文化的に共有される過程を検証する。また、12音階の生成にいたる過程や、12音階が生成されるための条件について考察する。

本論文の構成は以下のとおりである。2章で、本論文で用いる実験モデルについて述べる。3章で、実験について述べ、結果を考察する。4章で、まとめと今後の課題について述べる。

2. 実験モデル

2.1 モデル化

本論文では、音階の創発的な生成過程をマルチエージェント・システムを用いてモデル化し、実験的に検証する。人が発音や受音に用いる身体や機器類を単弦楽器モデルとして、発音用楽器および受音用楽器を用

[†] ATR ネットワーク情報学研究所
ATR Network Informatics Laboratories

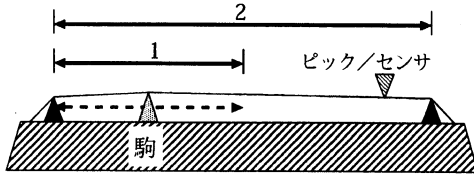


図1 発音および受音用楽器の概念

いた発音行為、および受音行為をエージェントとしてモデル化する。また、人がより心地よいと感じる音をより多くの頻度で発音するようになるメカニズムを、エージェントの学習アルゴリズムとしてモデル化する。

- ・ エージェント：楽器などを用いた発音行為（音の発生）、および聴覚などを用いた受音行為（音の知覚）を、それぞれエージェントとしてモデル化する。エージェントは、発音および受音に関係する遺伝子を持つ。

各エージェントは、単位時間ごとに自らの遺伝子に応じて発音する。このとき、ランダムに選択された他のエージェントが、発せられた音を受音する。

- ・ 発音・受音用単弦楽器：エージェントが発音および受音に用いる機器として、それぞれ図1に示す長さ2の1本の弦からなる単弦楽器を想定する。発音用楽器と受音用楽器が平行においてあり、エージェントは、それぞれの楽器を用いて発音および受音する。

発音用の楽器には弦を弾くピックが、受音用の楽器には弦の振動を感知するセンサがそれぞれ駒の右寄りにあり、それぞれの弦を押さえる駒が弦の左寄りにある。駒は、それぞれ左から $L_{Out} \in \{0 \sim 1\}$ 、 $L_{In} \in \{0 \sim 1\}$ の範囲で移動可能であり、駒の位置は、各エージェントが持つ遺伝子により決定される。この弦は、駒の位置により決定される基本周波数（固有振動数）を持つ。楽器の出す音の基本周波数は、弦の持つ固有振動数と等しいものとする。ここでは、400Hz から 800Hz までの1オクターブの音を出すモデルとして、基本周波数 $F_1(L)$ は、 L_{Out}, L_{In} に応じて式(1)で示される値とする。

$$F_1(L_{Out}/L_{In}) = \frac{800}{2 - L_{Out}/L_{In}} \quad (1)$$

ピックにより弾かれた発音用楽器の弦は、基本周波数 $F_1(L_{Out})$ と、その倍音の周波数成分 $F_n(L_{Out}) \{= n \times F_1(L_{Out})\}$ (n は自然数) を含む音を発する。

- ・ 共鳴：弦は、弦の固有振動数と同一周波数成分を多く含む音を受けると共鳴する。受音用楽器の弦は、発音用楽器から発せられた音の周波数成分 $F_n(L_{Out})$ と、受音用楽器の持つ基本周波数の m 倍 (m は自然数) の周波数成分 $F_m(L_{In})$ が近い場合、式(2)に従い共鳴する。ただし、音階生成の観点から、発音エージェントと受音エージェントの基本周波数がほぼ等しく同一音と判断できる場合 ($|F_1(L_{Out}) - F_1(L_{In})| < 10$) は除くものとする。

$$R(L_{Out}, L_{In}) = \begin{cases} MAX_{n,m} \\ \{10 - |(F_n(L_{Out})) - (F_m(L_{In}))|\} \\ \text{(if } |(F_n(L_{Out})) - (F_m(L_{In}))| < 10 \\ \wedge |(F_1(L_{Out})) - (F_1(L_{In}))| > 10) \\ 0(\text{else}) \end{cases} \quad (2)$$

($MAX_{n,m} \{X(n, m)\}$ は、

n, m に関して $X(n, m)$ の最大値を返す関数)

- ・ 学習アルゴリズム：エージェントは、より心地よい音をより多くの頻度で発音するように駒の位置 L を学習する。本論文では、同一周波数成分を多く含む音を心地よいと感じるという前提に基づき、時刻 t におけるエージェントの評価値 $E(a, t)$ を、式(2)で示された共鳴の度合いに応じて式(3)に従い決定する。より高い評価を得たエージェントが分裂し、低い評価を得たエージェントが消滅する。

$$E(a, t+1) = E(a, t) + R(L_{Out}, L_{In}) \quad (3)$$

エージェント間相互作用のモデルとして、局所性を考慮できる動的離隔モデル^{3),4)}を用いる。実空間では、より近くの個体間ほどより多くの頻度で相互作用が発生する。動的離隔モデルは、そのような相互作用の局所性や密度依存性をモデル化したものである。動的離隔モデルでは、各エージェントはコロニーと呼ぶグループごとに離隔される(図2)。異なるコロニーに存在するエージェントとは相互作用しない。同一コロニー内に存在するエージェント数が増加し限界数以上になる場合、そのコロニーに存在するエージェントは、さらに半数ずつのコロニーに離隔され、エージェント数が0になる場合、そのコロニーを消滅させる(図3)。

2.2 シミュレーション

本論文の実験で用いるシミュレーションの概要を NS

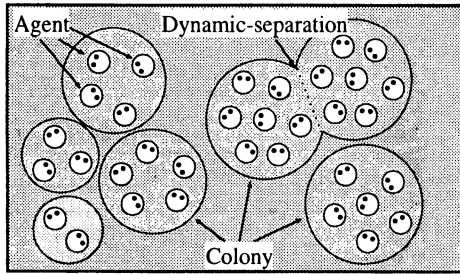


図2 動的離隔モデルの概念

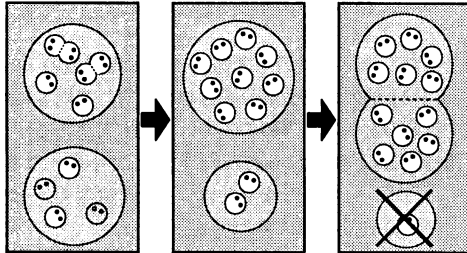


図3 動的離隔によるエージェント集団の変化

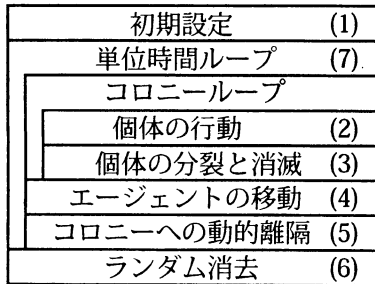


図4 NSチャートで示すシミュレーションの概要

チャートで図4に示し、詳細を以下に述べる。

- (1) 初期設定：仮想環境内に、1000個のコロニーを作成する。各コロニー内に、100個のエージェント a を作成する。エージェント a は、初期評価値 10 と、無作為に決定された駒の位置を示す遺伝子 $L(a)$ を持つ。
- (2) 行動：全エージェントが、それぞれの遺伝子に従い駒の位置を決定し発音する。エージェントの発音ごとに、同一コロニー内からランダムに選択されたエージェントがその音を受音する。発音および受音エージェントは、式(2)で示される共鳴の程度に応じて評価値を得る。このとき得た評価値は式(3)に従い各エージェントに累積される。
- (3) エージェントの分裂と消滅：エージェントは、エー

ジェントの評価値が初期値の倍以上になると分裂する。このとき、エージェントは、分裂前のエージェントの評価値を半分ずつ持つ2つのエージェントに分裂し、それぞれ遺伝子を引き継ぐ。ただし、遺伝子は突然変異確率 $P_{mut} = 0.01$ で変異する。また、エージェントは、エージェントの評価値が0以下になると消滅する。

- (4) エージェントの移動：エージェントは、移動確率 $P_{mig} = 0.01$ に従い、ランダムに選択されたコロニーに移動する。
- (5) コロニーへの動的離隔：コロニーへの動的離隔は、コロニー内のエージェント数がコロニーの限界エージェント数 100 を超えると起きる。このとき、1個のコロニー内に存在するエージェントは、2個のコロニーに離隔される。ただし、2個のコロニーに存在するエージェント数の差は1以下とする。また、コロニー内エージェント数が1以下になるとコロニーは消滅する。
- (6) ランダム消去：コロニー数が初期コロニー数より増えた場合、コロニー数が初期コロニー数と同数になるまでコロニーを無作為に消去する。
- (7) 単位時間ループ：(2)~(6)を繰り返す。

3. 実験

3.1 単純なモデルによる音階の生成

本節では、まず本論文の実験モデルにおける音階の生成過程を明らかにするため、単純なモデルを用いた音階の生成について考える。発音用楽器が発する音には、 $n \in \{2, 3, 4, 5\}$ 倍の倍音成分を均等に含むものとする。

3.1.1 実験結果

実験結果として、エージェントの遺伝子(駒の位置)の値を 0.01 ごとに分類した個体数比率を図5に、単位時間ごとの獲得した評価値の平均を図6に示す。

図6から、進化的学習の結果、エージェントはより多く共鳴する音を獲得し、高い評価値を得たことが分かる。図5から、エージェントの遺伝子が3つの値に収束したことが分かる。これら3つの値が示す音の周波数比は、452Hzを基準としたときほぼ3:4:5となり、452Hzをドとすると、それぞれド・ファ・ラに相当する。また、このときのファとラの周波数は、平均律におけるファ・ラより絶対律におけるファ・ラに近い値となった。

3.1.2 考察

本節の実験では、いずれも事前に明示的な音階は与えていないが、エージェントの自発的な学習から、現

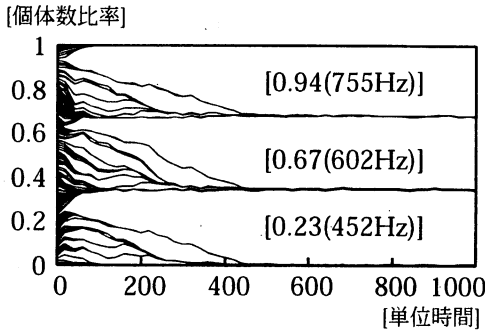


図5 駒の位置 L ごとの個体数比率

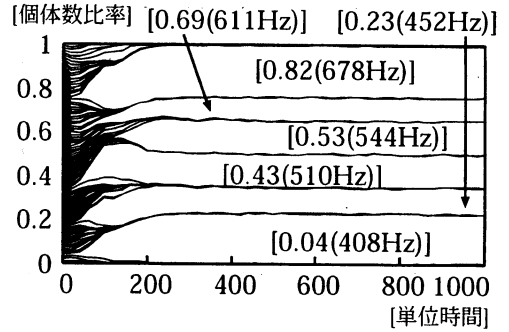


図7 駒の位置 L ごとの個体数比率

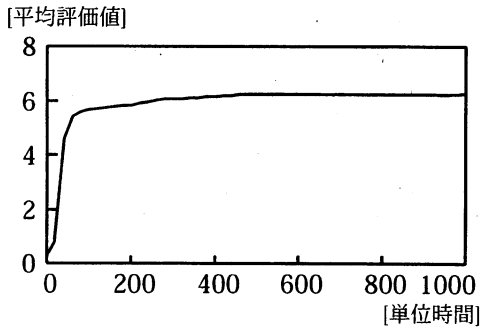


図6 単位時間ごとに獲得した評価値の平均

在の12音階においてド・ファ・ラに相当する音からなる音階を創発的に獲得した。これは、人類が用いてきた楽器などの発音機器が「発せられる音に倍音成分が含まれる」という条件を満たし、人の聴覚などの受音機器が「同一周波数成分を含む音を心地よいと感じる」という条件を満たす場合には、「より心地よいと感じる音の発生頻度を増加させる」という学習メカニズムにより、元始の音階が創発的に生成される可能性を示すものである。

本節のモデルでは、3つの音からなる音階は現れたが、それ以上の音からなる音階は現れなかった。これは、本モデルに組み込んだ5倍音までしか用いない条件は、現在の12音階の生成には十分ではないためであると考えられる。

本節のような音階の構成論的なアプローチを用いることで、どのような条件ではどのような音階が生成されるのかが検証可能となり、既に存在する音階の解析的なアプローチでは得ることが難しかった音階に対するより深い理解や新たな知見が得られると期待される。

3.2 実際の環境により近いモデルによる音階の生成

前節の実験では、発音用楽器は、基本周波数の $n \in \{2, 3, 4, 5\}$ 倍の周波数成分を均等に含む音を発するも

のとした。また、受音用楽器は、基本周波数の $m \in \{2, 3, 4, 5\}$ 倍の周波数成分を含む音に対して均等に共鳴するものとした。実際の弦では、基本周波数に比べて倍音成分の音量は高音ほど減衰して小さくなる。また、必ずしも5倍音までを含むわけではなく、それ以上の倍音成分を含む場合がある。

本節では、現在の音階が生成されたと考えられる環境により近いモデルにより、より複雑な音階の生成について議論する。発音用楽器が含む倍音成分を基本周波数の $n \in \{2, 3, \dots, 10\}$ 倍とし、 n 倍音の音量を基本周波数における音の 0.9^n 倍とする。さらに、音階生成の観点から、発音エージェントと受音エージェントの基本周波数がほぼ等しく同一音と判断できる場合 ($|F_1(L_{Out}) - F_1(L_{In})| < 10$)、負の評価として -10 を与えるものとする。すなわち、式(4)に従い共鳴する。

$$R(L_{Out}, L_{In}) = \begin{cases} \text{MAX}_{n,m} \{ \{10 - |F_n(L_{Out}) - F_m(L_{In})| \} \times 0.9^{n+m} \} \\ \quad (if |F_n(L_{Out}) - F_m(L_{In})| < 10 \\ \quad \wedge |F_1(L_{Out}) - F_1(L_{In})| > 10) \\ -10 (if |F_1(L_{Out}) - F_1(L_{In})| < 10) \\ 0 (else) \end{cases} \quad (4)$$

$(\text{MAX}_{n,m} \{X(n, m)\})$ は、 n, m に関して $X(n, m)$ の最大値を返す関数)

3.2.1 実験結果

実験結果として、エージェントの遺伝子(駒の位置)の値を0.01ごとに分類した個体数比率を図7に、単位時間ごとの平均評価値を図8に示す。

図8から、進化的学習の結果、エージェントはより多く共鳴する音を獲得し、高い評価値を得たことが分かる。図7から、エージェントの遺伝子が6つの値に収束したことが分かる。これら6つの値が示す音の

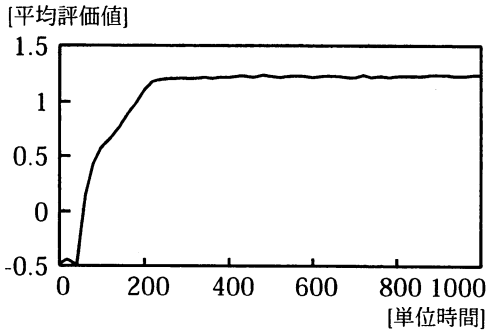


図 8 単位時間ごとに獲得した評価値の平均

周波数比は、408Hz を基準としたとき、ほぼ 36 : 40 : 45 : 48 : 54 : 60 となり、408Hz をソとすると、それぞれソ・ラ・シ・ド・レ・ミに相当する。また、このときそれぞれの音の周波数は、平均律における値より絶対律における値に近いものであった。

3.2.2 考察

本節の実験では、より現実に近いと考えられる環境のモデルとして、10 倍音までを含み、高倍音ほど音量が減衰し、同一音を心地よいと感じない場合に、現在の 12 音階により近いソ・ラ・シ・ド・レ・ミに相当する 6 つの音からなる音階を創発的に獲得した。これは、「発音用楽器が 10 倍音までを含む」、「高倍音の音量が減衰して小さくなる」、「同一音を発音した場合に不快と感じる」といった条件が満たされる場合には、現在の 12 音階に近いより多くの音からなる音階が生成される可能性を示すものである。また、この結果は、どのような成分の音を発するかという楽器（発音機器）の特徴や、どのような音を心地よいと感じるかという人の身体性（受音機器）に応じて、生成される音階は異なる可能性を示している。

本手法を用いることで、楽器の特徴に応じたその楽器のための音階や、聴覚などの個人が持つ身体性の特徴に応じた音階など、カスタムメイドな音階生成・提示システムなどへの展開が期待される。

3.3 現実の環境とは異なるモデルによる音階の生成

前節の実験では、現実により近いと考えられるモデルにより、現在の 12 音階に含まれる 6 音からなる音階を獲得した。本節では、現実の環境とは異なると考えられるモデルを用いて、新しい音階の生成について述べる。

本節の実験では、20 倍音 ($n \in \{2, 3, \dots, 20\}$ 、約 4.3 オクターブ上) まで減衰なく含む音を発生する発音機器と、それらの全ての音を知覚可能な受音機器をモデル化する。また、前節と同様に、発音エージェン

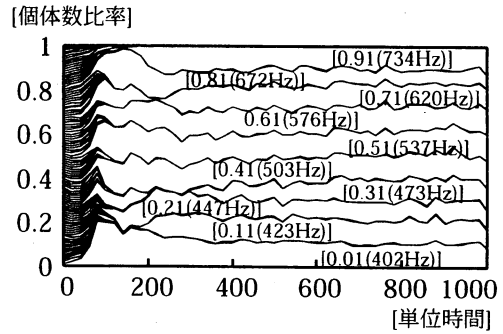


図 9 駒の位置 L ごとの個体数比率

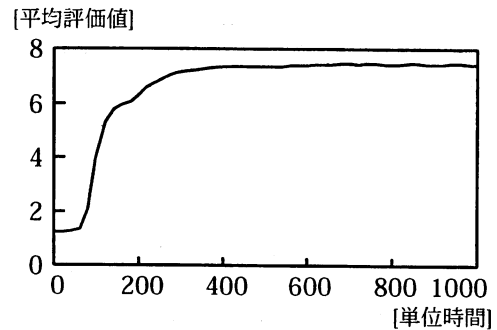


図 10 単位時間ごとに獲得した評価値の平均

トと受音エージェントの基本周波数がほぼ等しく同一音と判断できる場合 ($|F_1(L_{Out}) - F_1(L_{In})| < 10$) には、負の評価として -10 を与えるものとする。すなわち、式 (5) に従い共鳴する。

$$R(L_{Out}, L_{In}) = \begin{cases} \text{MAX}_{n,m} \\ \left\{ \begin{array}{l} 10 - |(F_n(L_{Out})) - (F_m(L_{In}))| \\ (if |(F_n(L_{Out})) - (F_m(L_{In}))| < 10 \\ \wedge |(F_1(L_{Out})) - (F_1(L_{In}))| > 10) \\ -10 (if |(F_1(L_{Out})) - (F_1(L_{In}))| < 10) \\ 0 (else) \end{array} \right. \quad (5) \end{cases}$$

($\text{MAX}_{n,m}\{X(n, m)\}$) は、

n, m に関して $X(n, m)$ の最大値を返す関数)

3.3.1 実験結果

実験結果として、エージェントの遺伝子 (駒の位置) の値を 0.01 ごとに分類した個体数比率を図 9 に、単位時間ごとの平均評価値を図 10 に示す。

図 10 から、進化的学習の結果、エージェントはより多く共鳴する音を獲得し、高い評価値を得たことが分かる。

図 9 から、エージェントの遺伝子が 9 つの値に収束

したことが分かる。これら9つの値が示す音の周波数比は、10等分平均律に近い値となり、現在の12音階や、主に用いられている他の音階とは異なる音階が生成された。

3.3.2 考察

本節の結果は、もし倍数の大きな音まで減衰なく発生する楽器が古くから存在し、それらの音を知覚できる人々からなる集団では、現在とは異なる音階が生成されていた可能性を意味する。すなわち、「人間の聴覚により知覚できる倍音成分には、ある程度の限界がある」、「人間の周りにあった音を発する機器類には、ある程度までの倍音成分しか含まない」といった要因が共通であったため、現在の12音階が文化的に共有されたと考えられる。逆に、一部の民族で用いられる12音階とは異なる音階が生成され、共有される条件を検討することで、その民族の身体的特徴や楽器の特徴が推測できる可能性がある点も興味深い。

4. おわりに

本論文では、「楽器など音を発する機器に倍音成分を含む」という特徴と、「人が身体的に同一周波数成分を多く含む音を心地よいと感じる」という特徴がある場合、より心地よい音をより多くの頻度で発生させることにより、音階が創発的・自然発生的に生成され、共有されることをマルチエージェント・シミュレーションにより示した。

その結果、含まれる倍音成分が少ない場合にはド・ファ・ラに相当するの3つの音からなる単純な音階が生成されることを示した。また、含まれる倍音成分がある程度多い、倍数が大きいほど音量が小さい、といった現実により近いと考えられる環境モデルにより、ソ・ラ・シ・ド・レ・ミに相当する6つの音からなる現在の12音階に近い音階が生成された。さらに、現実とは異なると考えられる環境のモデルにより、現実ではあまり用いられない10等分平均律に近い音階が生成された。これらの結果は、倍音成分を含む楽器により奏でられる音楽において、現在の12音階が創発的に生成され、文化的に共有される過程について知見を与えるものである。また、12音階が生成されるための条件について検討する方法論を明らかにしたものである。

本論文で示した方法は、2つの音の共鳴に注目した局所的なエージェントの相互作用により、事前に明示的に与えていない音階を創発的に生成するものである。今回の実験では、動的離隔モデルでモデル化した相互作用の局所性による影響が見られなかった。今後

の課題として、生成される音階に対し、局所性の強さが与える影響の検証があげられる。また、音楽を、同一周波数成分を多く含む音が「空間的に調和した和音 (code)」や「時間的に調和した旋律 (melody)」の組み合わせであるとして捉えることで、本論文で示した方法の応用により、音階の生成だけでなく、新しい音楽の創発的な生成法(作曲法)への展開が期待される。

*本研究は独立行政法人情報通信研究機構の研究委託「人間情報コミュニケーションの研究開発」により実施したものである。

参考文献

- 1) 別宮貞雄：十九平均律の可能性 - 音律再考，音楽芸術，第44巻，8号(1986)。
- 2) 小方 厚：非12音平均律音楽のこころみ，情報処理学会研究報告，2004-MUS-54，pp. 47-50(2004)。
- 3) 中山功一，松井博和，野村由司彦：動的離隔型GAの提案，情報処理学会論文誌：数理モデル化と応用，進化的計算特集号，Vol.43，No.SIG 10 (TOM 7)，pp. 95-109 (2002)。
- 4) K. Nakayama, K. Shimohara, and O. Katai: Dynamically Separating GA: A New Method of Achieving the System-level Optimality in MAS, The journal of the three dimensional images, Vol. 16, No. 4, pp. 177-183 (2002)。