

Stock to Music: 多変量株価時系列データの音楽変換

Stock to Music: Transformation of Multivariate Stock Time Series into Music Data

平松 祐紀^{*1} 中川 慧^{*2} 高野 海斗^{*2} 中村 栄太^{*3}
Yuki Hiramatsu Kei Nakagawa Kaito Takano Eita Nakamura^{*1}三井住友海上あいおい生命 Mitsui Sumitomo Aioi Life Insurance Co., Ltd. ^{*2}野村アセットマネジメント株式会社 Nomura Asset Management Co., Ltd. ^{*3}京都大学 Kyoto University

We propose a method for transforming multivariate stock time series data into music. Typically, investors rely on visual information, such as stock charts, to make investment decisions based on stock time series. However, watching numerous stock data simultaneously and for long time is highly demanding and challenging. Therefore, we focus on representing essential information for investment decisions (trends and sudden changes) through sound information, or more precisely, music, which does not depend on visual interpretation. We expect that decision-making through musical transformation can reduce cognitive load and be accessible to a broader range of users. We conduct experiments to determine whether the essential information can be successfully acquired from the music generated by the proposed method.

1. はじめに

トレーディングには、人間の判断によって執行を行う裁量取引と機械的に投資判断を行うシステム取引の二種類が存在する [Jansen 20]。裁量取引では、トレーダーが直接市場の動向を分析し、投資判断を下す。対して、システム取引では、あらかじめ設定されたルールないしモデルに基づき、自動的に、機械的に取引を行う。様々な資産において、短期的な取引は機械化されているものの、機関投資家における大口取引の執行や中長期の投資戦略において取引を執行する際などは裁量取引が行われている。裁量取引を行うトレーダーは、刻一刻と変化する株価時系列をもとに投資意思決定を行うが、その際にはチャート形式などの視覚情報を頼りにすることが多い [Taylor 92, Edwards 18]。また、金融市場は非常に複雑であり、かつ、相互依存関係が存在するために [Raddant 21]、関連する複数の情報 (複数の資産の価格やニュースなど) を同時に監視する必要がある。

しかしながら、投資家がこれら複数の情報を同時にかつ長時間にわたって視覚的に追跡するのは、負荷が高く困難である。このような課題に対処することは、金融機関をはじめとする企業が健康経営^{*1}を推進することにつながり、人的資本経営の観点からも重要である。さらに、人間の注意や認知リソースが限られて (注意力の限界)、重要な情報を見落とすリスクもある [Hirshleifer 09]。特に複数の作業時に、視覚的タスクが重なるとパフォーマンスが低下することが知られている [Wickens 02]。したがって、複数の情報を視覚的に追跡する場合、他の作業を同時に効率的にはできないという問題もある。

本研究では上記の課題に対処するために、投資家が投資判断

に必要な情報を、視覚に依存しないデータ表現方法である音情報、特に音楽として表現する。具体的には、株価指数の動きと、株価指数に含まれる各銘柄の情報 (出来高など) を一つの音楽で表現することで、投資家が他の作業をしながらでも市場の動きを追跡できるようにする。これまでに、金利や株価といった複数の金融時系列データを音に変換する手法が提案されているが、各時系列データの動きを音高の変化で表し、異なる音色で再生するアプローチが多い [Janata 04]。この場合、同時に聞き分けられる音色の数に限りがあるため、一つの音楽で表現できるデータの数が限定的になるという問題がある。そこで、本研究では異なる銘柄のデータを異なる音域で表現することでより多くのデータを同時に追跡できるようにする。

視覚情報だけに依存する方法と比較し、音楽情報を利用する方法は、負荷を軽減し、特に視覚情報に頼ることができないユーザーを含む、多くのユーザーに有益であると考えられる。さらに、他の作業へ視覚情報資源を投入することで作業効率化につながると考えられる。評価実験では、実際に生成した音楽を他の作業をしながら聞くことで、投資判断に必要な情報が取得可能かどうかを検証する。

2. 提案手法

2.1 問題設定

$N + 1$ 個の株価時系列データ $\mathbf{S} = (\mathbf{S}^0, \mathbf{S}^1, \dots, \mathbf{S}^N)$ に対し、コード進行と N 個のメロディーパートからなる音楽を生成する。ここで、 \mathbf{S}^0 は株価指数、 $\mathbf{S}^1, \dots, \mathbf{S}^N$ は個別銘柄のデータを表す。本研究で扱う多変量株価時系列データ

$$\mathbf{S}^n = \{o_t^n, h_t^n, l_t^n, c_t^n, v_t^n\}_{t=1}^T$$

は、時点 $t - 1$ から t までの期間における始値 o_t^n 、高値 h_t^n 、安値 l_t^n 、終値 c_t^n 、出来高 v_t^n からなる。時点 $t - 1$ と t の間の時間幅を 1 ティックと呼び、日足データを用いる場合、1 ティックは 1 日となる。なお、生成する音楽は、4/4 拍子で長さは $2T$ 小節とし、テンポは一定とする。また、音高はハ長調の音階 (C, D, E, F, G, A, B) 上に制限する。

2.2 株価データの特徴量

本手法では、株価指数の動きと、指数に含まれる個別銘柄のデータ (出来高など) を 1 つの音楽で表現する。まず、株価指

連絡先: 平松 祐紀, 三井住友海上あいおい生命

〒 101-8011; 東京都千代田区神田駿河台 3 丁目 9 番地

E-mail: yuki.hiramatsu.bb@gmail.com

本稿の内容は筆者らが所属する組織を代表するものではなく、本稿の全ての誤りは、筆者らの責に属するものである。

*1 同様の問題意識から経済産業省は、従業員の健康管理を経営的な視点で考え、戦略的に取り組んでいる法人を評価する取り組みを行っている https://www.meti.go.jp/policy/mono_info_service/healthcare/kenkoukeiei_yuryouhouzin.html

数 S^0 については、以下の2つの指標を音楽として表現する。

1. 上昇/下落: 株価指数の終値 c_t^0 が前ティックと比較して上昇したか下落したかを表す。

$$\text{sign}_t^0 = \text{sign}(c_t^0 - c_{t-1}^0)$$

ここで、 $\text{sign}(\cdot)$ は引数が正または0のときに1、負のときに0を返す符号関数である。

2. ボラティリティ (標準偏差) の大きさ: 株価指数の対数収益率 $r_t^0 = \log(c_t^0/c_{t-1}^0)$ の時間幅 w 期間の標準偏差である。

$$\sigma(r_{t-w+1:t}^0) = \sqrt{\frac{A}{w} \sum_{s=0}^{w-1} (r_{t-s}^0 - \mu(r_{t-w+1:t}^0))^2}$$

ここで、 $\mu(\cdot)$ は平均を表し、 A は年率換算するための定数である。

次に個別銘柄 S^1, \dots, S^N については、以下の3つの指標を音楽として表現する。

3. 出来高の多さ: 個別銘柄の出来高 v_t^n である。
4. ボラティリティ (標準偏差) の大きさ: 株価指数と同様に、株価の対数収益率 $r_t^n = \log(c_t^n/c_{t-1}^n)$ の時間幅 w 期間の標準偏差である。

$$\sigma(r_{t-w+1:t}^n) = \sqrt{\frac{A}{w} \sum_{s=0}^{w-1} (r_{t-s}^n - \mu(r_{t-w+1:t}^n))^2}$$

ここで、 $\mu(\cdot)$ は平均を表し、 A は年率換算するための定数である。

5. 相対力指数 (RSI) の大きさ: 銘柄の価格変動の勢いを表す指標 [Wilder 78] であり、以下の式で計算される。

$$\text{RSI}_t^n = 100 - \frac{100}{1 + \text{RS}_{t-w+1:t}^n}$$

ここで、 $\text{RS}_{t-w+1:t}^n = \frac{\text{平均上昇幅}}{\text{平均下降幅}}$ であり、平均上昇幅と平均下降幅は時間幅 w 期間における上昇幅と下降幅の平均値をそれぞれ示す。一般に、RSIが70を上回ると買われすぎ、30を下回ると売られすぎと判断される。

2.3 音楽の生成

本手法では、1ティックのデータ $S_t^n = (o_t^n, h_t^n, l_t^n, c_t^n, v_t^n)$ に対して、1小節の音楽を逐次生成し、2回ずつ繰り返し再生する。株価指数の動きからコード進行と基本リズムパターンの系列を生成した上で、個別銘柄ごとにメロディーパートを生成することで、パート間の調和を図る (図1)。

2.3.1 コード進行の生成

まず、コード進行を株価指数の上昇/下落をもとに生成する。各ティックごとに、上昇 ($\text{sign}_t^0 = 1$) のときは、ハ長調の長三和音 (C major, F major, G major) から、下落 ($\text{sign}_t^0 = 0$) のときは、短三和音 (A minor, D minor, E minor) からひとつ選択する。コードの遷移は、学習に用いる音楽データから最尤推定で推定した上記6和音の遷移確率行列 (一次マルコフモデル) を、長三和音と短三和音のいずれかに制限してからサンプリングすることで決定する。

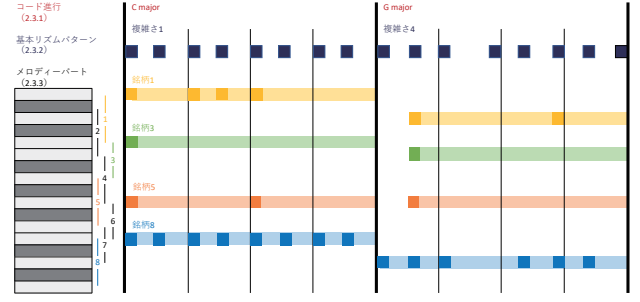


図1: 音楽の生成方法

2.3.2 基本リズムパターンの生成

個別銘柄ごとのメロディーパートを生成するために、まず、小節ごとに基本リズムパターンを生成する。基本リズムパターンは、メロディーパートに含まれる音符の発音時刻の可能な候補を与えるもので、株価指数のボラティリティが高いほど、リズムが複雑になるように設計する。まず、各小節 m において、対応するティックのボラティリティの値を $\max(1, \lceil \sigma(r_{t-w+1:t}^0) / 5 \rceil)$ により1から8の整数値 x_m として表す。

次に、各小節におけるリズムは、音符の発音時刻を表す、1小節を16分割した拍節位置 $b \in \{0, 1, \dots, 15\}$ の集合として表す (例えば、 $b=0$ は小節の第1拍目、 $b=4$ は第2拍目を表す)。小節 m における基本リズムパターンは8個の発音位置 $\mathbf{b}_m = \{b_{mk}\}_{k=1}^8$ ($0 \leq b_{m1} < \dots < b_{m8} \leq 15$) により表す。可能な基本リズムパターンの種類は $\binom{16}{8} = 12870$ 個ある。リズムパターン $\mathbf{b} = (b_1, \dots, b_8)$ の複雑度 $C(\mathbf{b})$ は、 \mathbf{b} により生み出される音の長さ (音価) の相対頻度分布のパープレキシティー $S(\mathbf{b})$ と、拍節位置の拍の強度に応じた拍節複雑度 $M(\mathbf{b})$ (拍上に発音が多いほど大きな値を取る) の和により定義する。具体的には、音価 $\nu_k = (b_{k+1} - b_k + 16) \bmod 16$ ($\nu_k \in \{1, \dots, 16\}$) の集合から求めた相対頻度分布を $\mathbf{f} = (f_1, \dots, f_{16})$ とすると、 $S(\mathbf{b}) = \exp(-\sum_{\nu=1}^{16} f_\nu \ln f_\nu)$ である (右辺の括弧内は Shannon エントロピーを表す)。また、有理数 $b_k/16$ を既約分解したときの分母の値を $\rho_k \in \{1, 2, 4, 8, 16\}$ とおくと、拍節複雑度は $M(\mathbf{b}) = \sum_{k=1}^8 \rho_k$ で定義する。例えば、発音時刻の列 b_k が等間隔に並ぶ時 (即ち、音価が全て8分音符になる時) に $S(\mathbf{b})$ が最小となり、さらにこれらの発音位置が小節の8分割した拍上に乗る時に $M(\mathbf{b})$ が最小になる。全てのリズムパターンの中での $S(\mathbf{b})$ と $M(\mathbf{b})$ の最大値をそれぞれ S_{\max} と M_{\max} とするとき、複雑度 $C(\mathbf{b})$ は次の式で定義する。

$$C(\mathbf{b}) = \frac{S(\mathbf{b})/S_{\max} + M(\mathbf{b})/M_{\max}}{2}$$

全てのリズムパターン \mathbf{b} を $C(\mathbf{b})$ の昇順にソートして、その添字の集合を $\{1\}, \{2, \dots, 920\}, \{921, \dots, 1839\}, \dots$ のように15分割する。この奇数番目の分割を8段階の複雑性を持つリズムパターンの集合 $B(1), \dots, B(8)$ とする。8段階のボラティリティの系列 x_m が与えられた時、基本リズムパターン \mathbf{b}_m を $\mathbf{b}_m \in B(x_m)$ を満たすように選択する。この際、隣り合う小節のリズムがなるべく滑らかにつながるようにするため、両者の間の Hamming 距離 $d(\mathbf{b}_m, \mathbf{b}_{m+1}) = \sum_{k=1}^8 (1 - \delta_{b_{mk}, b_{(m+1)k}})$ が小さくなるように選ぶこととする。具体的には、 \mathbf{b}_{m+1} は $B(x_{m+1})$ からランダムに選び、 \mathbf{b}_m が与えられた時には、 \mathbf{b}_{m+1} の候補を $B(x_{m+1})$ からランダムに100個選び、その中で $d(\mathbf{b}_m, \mathbf{b}_{m+1})$ が最小なものを \mathbf{b}_{m+1} として帰納的に選択する。これにより、小節間でリズムが唐突に変化することを避けながら、生成の度

ごとに多様な生成結果が得られるように工夫を行っている。

2.3.3 メロディーパートの生成

株価指数の動きから決定したコード進行と基本リズムパターンに基づき、個別銘柄数 N だけメロディーパートを生成する。各メロディーパートの 1. 音高、2. 音量、3. リズム、4. 音色の 4 要素を、個別銘柄データから決定する。

- 音高:** 音高は、銘柄名のアルファベット順とコードから決定する。生成された音楽を聞いたとき、どの銘柄に重要な変化があったのか把握できるようにするため、アルファベット順で一番最後の銘柄 $n = N$ には (C2, D2, E2)、 $n = N - 1$ には (D2, E2, F2) のように、3つの音高を重複させながら割り当てる。各小節において発音する音高は1つであり、割り当てられた3つの音高に対してコード内の音高に重みをつけてサンプリングすることで決定する。例えばコードが C major のとき、 $n = N$ に割り当てられた3つの音高 (C2, D2, E2) に対し、(10, 1, 10) の重みをつけてサンプリングする。

- 音量:** 音量は、個別銘柄の出来高の多さによって決定する。出来高の多さは時間幅 w 期間の Z スコアで表す。

$$Z(v_t^n) = \frac{v_t^n - \mu(v_{t-w+1:t}^n)}{\sigma(v_{t-w+1:t}^n)}$$

ここで、 $\mu(\cdot)$ と $\sigma(\cdot)$ は平均と標準偏差を表す。MIDI のペロシティ (0 から 127 の整数) を $\max(0, [15Z(v_t^n)])$ で設定することで、音量を調整する。

- リズム:** リズムは、個別銘柄のボラティリティの大きさによって決定する。まず、ボラティリティが大きいほど発音回数が多くなるように、各小節における発音回数を $\lceil \sigma(r_{t-w+1:t}^n) / 5 \rceil$ で決定する。次に発音時刻の集合 (リズム) を、基本リズムパターンの部分集合から選択する。具体的には、基本リズムパターン $\mathbf{b} = \{b_k\}_{k=1}^8$ から1つの発音時刻を除いた部分リズムパターン \mathbf{b}^7 、さらに1つの発音時刻を除いた部分リズムパターン \mathbf{b}^6 のように、発音回数が7回から0回となる8つの部分リズムパターン $\mathbf{b}^7 \supset \mathbf{b}^6 \supset \dots \supset \mathbf{b}^0$ をサンプリングによって決定し、各小節におけるリズムを選択する。
- 音色:** 音色は RSI の水準によって決定する。具体的には、買われすぎ ($RSI_t^n > 70$)、通常 ($30 \leq RSI_t^n \leq 70$)、売られすぎ ($RSI_t^n < 30$) の3状態を異なる音色で再生する。通常の状態については柔らかく心地よい音を選択し、買われすぎと売られすぎの状態については鋭く固い音で、両者の違いが判別しやすい音を選択する。

3. 評価実験

3.1 実験設定

実験では、ダウ工業株 30 種平均の日足データを使用し、図 2 に示すように、30 銘柄のデータを表示した。コード進行の遷移確率行列の学習には、日本のポピュラー音楽の約 2500 曲のデータセットを用いた [Nakamura 23]。ボラティリティは過去 20 ティックの標準偏差を $A = 250$ として年率換算したもの、出来高の Z スコアは過去 20 ティックの平均と標準偏差を使って計算したもの、RSI は過去 14 ティックについて計算したものを採用した。RSI が示す買われすぎ、通常、売られすぎの各

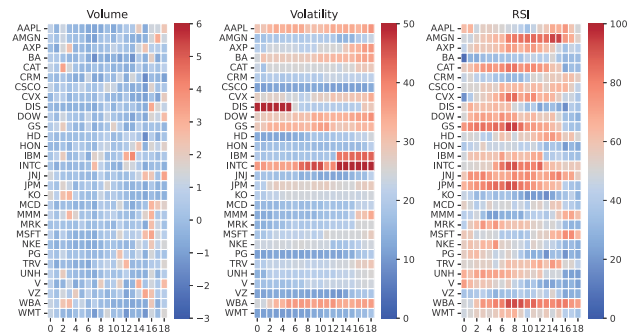


図 2: 実験環境の例

表 1: 実験 1 の正解率

上昇/下落	標準偏差	銘柄の違い
9/10	9/10	6/10

状態に対する音色はそれぞれ、GM 規格の 84:chiff、95:halo、81:square とした。

提案手法により生成した音楽を聞くことで、他の作業をしながらでも投資判断に必要な情報を取得可能かどうかを評価するため、運用業務に携わる 10 人を対象に 2 つの実験を行った。

3.1.1 実験 1: 株価指数の情報取得

まず実験 1 として、生成した音楽から株価指数の、1. 上昇/下落、2. ボラティリティの大きさ、3. 銘柄の違いの 3 点を認識できるかそれぞれ検証した。具体的には、株価データから音楽への変換方法を説明した上で、下記 3 つの質問をした。

1. 上昇系列と下落系列から生成した音楽のうち、上昇トレンドはどちらか。
2. ボラティリティの大きさが異なる 2 系列から生成した音楽のうち、ボラティリティが大きいのはどちらか。
3. 高い音高と低い音高を割り当てた銘柄のうち、出来高が多いのはどちらか。

実験 1 の評価指標は、各質問の正解率とした。

3.1.2 実験 2: 個別銘柄の情報取得

次に実験 2 として、他の作業をしながらでも、必要な情報、すなわち個別銘柄の出来高、ボラティリティ、RSI の動きを追跡可能かどうかを検証した。具体的には、株価指数に含まれるすべての個別銘柄の出来高、ボラティリティ、RSI のデータを 5 秒おきに更新して表示し、被験者にはそれぞれのデータが一定の閾値を超えたところを記録するタスクを行ってもらった。同時に、読書などの別の作業を行ってもらい、2 つの作業の出来栄を自己評価してもらった。提案手法の有効性を検証するため、比較として視覚情報のみの場合と、各データが閾値を超えた時にアラーム音を鳴らした場合についても実験した。アラーム音は、出来高の Z スコアが 3 を超えたときに 1 拍目、ヒストリカルボラティリティが 30 を超えたときに 2 拍目、RSI が 70 を上回ったときに 3 拍目、30 を下回ったときに 4 拍目に鳴らした。実験 2 の評価指標は、とてもできた: 2、ややできた: 1、あまりできなかった: -1、全くできなかった: -2 として計算した平均スコアとした。

3.2 実験結果

実験 1 の結果を表 1 に示す。上昇/下落の判定とボラティリティの大きさの判定は正解率が高いが、銘柄の認識は比較的

表 2: 実験 2 の評価スコア

	出来高	標準偏差	RSI	平均	別の作業
視覚のみ	-0.9	-0.1	-0.8	-0.6	-0.1
アラーム	-0.7	0.2	-0.3	-0.3	-0.3
提案手法	0.2	-0.2	-0.3	-0.1	0.5

答率が低かった。音高の認識は音感という能力が必要な比較的難しいタスクであることや、各銘柄に対し 3 つの音高を重複させながら割り当てたことで銘柄の順序関係が複雑になったことが原因として考えられる。

実験 2 の結果を表 2 に示す。株価データの追跡タスクと別の作業のスコアは、どちらも提案手法が最も高く、生成した音楽を聞くことで、株価の重要な変化に気がつきやすくなり、別の作業をより効率的に進められるようになることが確認できた。しかし、株価追跡タスクのスコアはマイナスであり、改善の余地が大きい。また、実験 2 において「アラーム音と音楽を聞きながら作業して、心地よかったのはどちらか」と質問したところ、音楽を選択したのは 10 人中 9 人であり、音情報のうち音楽を利用することの有効性が確かめられた。

4. おわりに

本研究では、投資家が他の作業をしながらでも株価データを追跡できるよう、株価指数の動きと個別銘柄のデータを 1 つの音楽で表現する手法を提案した。実験により音情報、中でも音楽として表現することで株価の重要な変化に気がつきやすくなり、別の作業をより効率的に行うことができることが確かめられた。今後は、メロディーの音高とリズムを学習ベースのアルゴリズムで生成する手法を検討し、リアルタイムデータへの応用を行う予定である。

参考文献

- [Edwards 18] Edwards, R. D., Magee, J., and Bassetti, W. C.: *Technical Analysis of Stock Trends*, CRC Press (2018)
- [Hirshleifer 09] Hirshleifer, D., Lim, S. S., and Teoh, S. H.: Driven to Distraction: Extraneous Events and Underreaction to Earnings News, *The Journal of Finance*, Vol. 64, No. 5, pp. 2289–2325 (2009)
- [Janata 04] Janata, P. and Childs, E.: Marketbuzz: Sonification of Real-Time Financial Data, in *International Conference on Auditory Display (ICAD)* (2004)
- [Jansen 20] Jansen, S.: *Machine Learning for Algorithmic Trading: Predictive Models to Extract Signals from Market and Alternative Data for Systematic Trading Strategies with Python*, Packt Publishing Ltd (2020)
- [Nakamura 23] Nakamura, E.: Computational Analysis of Selection and Mutation Probabilities in the Evolution of Chord Progressions, in *Proc. International Symposium on Computer Music Multidisciplinary Research (CMMR)*, pp. 462–473 (2023)
- [Raddant 21] Raddant, M. and Kenett, D. Y.: Interconnectedness in the Global Financial Market, *Journal of*

International Money and Finance, Vol. 110, p. 102280 (2021)

[Taylor 92] Taylor, M. P. and Allen, H.: The Use of Technical Analysis in the Foreign Exchange Market, *Journal of International Money and Finance*, Vol. 11, No. 3, pp. 304–314 (1992)

[Wickens 02] Wickens, C. D.: Multiple Resources and Performance Prediction, *Theoretical Issues in Ergonomics Science*, Vol. 3, No. 2, pp. 159–177 (2002)

[Wilder 78] Wilder, J. W.: *New Concepts in Technical Trading Systems*, Hunter Publishing Company (1978)