

音楽の配布形態の差による音響的特徴の分析

Analysis of Acoustic Characteristics in Different Music Distribution Media

神谷 達夫, 北口 千華

Tatsuo Kamitani, Chihana Kitaguchi

要旨

音楽の配布形態には、いわゆるアナログレコードや CD などのメディアが存在する。インターネット上等での評判によると、同一の楽曲であっても、配布形態やデジタルリマスターと呼ばれる修正により、異なった感じを受けるとされていることがある。ただし、定量化された分析は少なく、「このような感じがする」という定性的な意見が散見される。

本論文は、同一の楽曲がアナログレコードや CD によって違いがあるのかを定量的に調査することを目的としている。対象としたのは、イングヴェイ・マルムスティーンの曲である。この曲は、各国で販売されているアナログレコードや CD を容易に入手可能であり、インターネット上でその差が議論されていたため、研究の題材とした。

特徴の抽出には、一定間隔毎のスペクトルを図に表現し、画像の類似度で評価する方法を採用した。類似度判定結果は、定性的な楽曲の判定と一致していることが確認できた。

キーワード: 楽曲の特徴、CD、アナログレコード、イングヴェイ・マルムスティーン

Keywords: Characteristics of the music, CD, vinyl record, Yngwie Malmsteen

1. はじめに

デジタル録音技術が民生用の音楽配布に用いられるようになって 40 年以上が経過した。デジタル録音の特徴は、複製時の劣化が無いため、音響特性を劣化させずに編集できることにある。一方、アナログ録音による音楽配布の場合は、編集段階や配布用メディアに記録する段階での音質の変化が存在する。また、アナログ録音により録音されたマスターテープをデジタル録音化し、配布する際に、いわゆる”デジタルリマスター”と呼ばれるような音質の修正のなされる場合がある。

聴感上の感想としては、アナログ録音の版によって音質にどのような違いがあるかや、デジタル録音化されて配布された録音がどのように聞こえるかが議論されることがある。また、リマスターやリミックスなどの処理によってどのように音質が変化したかが話題に上がることもある。聴感上の感想として議論される音質は、指している内容が人によって異なる場合がある。音量や音圧の大きさを音質としている場合もあれば、高音域や低音域がどの程度よく聞こえるかや、それぞれの楽器の音がどれくらいハッキリ聞こえるかを音質としている場合もある。

本論文では、各国版で音質が違うと言われることがあり、アナログ録音配布形態である LP レコードとデジタル録音配布形態である CD が各種入手可能なイングヴェイ・マルムスティーンのアルバムである *Marching Out* に収録されている曲を実験材料にし、配布形態の差による音質の定量化を検討した。

このアルバムは、各国版によって音質が違うとされており、中でも日本版はアナログレコード、CD 共に「音がモコモコと奥まって」おり、音質が悪いとされている。そのため、評価の低い日本版のものを基準とした、各種形態の音質に対する聴感上の感想がインターネット上に多く存在する。特に、アナログレコード、CD ともに入手が容易である日本版、米国版、西ドイツ版の 3 種を比較した感想や、リマスターによって音質を修正した CD に対する感想が多く見られる。CD の音質に対する評価は「西独盤（ただし位相揺れ有り）＞米盤（若干ハイ上がり）＞日本盤（一番悪い）」が通説となっている^{注2) 注3)}。また、アナログレコードに関しては、米国版及び西ドイツ版は音のこもりがなく、西ドイツ版はリマスターが施された CD よりも日本版との音質の違いを実感できるとされている^{注4) 注5)}。

音楽の特徴を信号処理等の工学的手法を用いて分析する方法は、各種知られている。旋律パターンを周波数離散形音響パターン（FM 検波出力）と捉える方法では、パターン間の類似関係をパターンの概形から細かい変化まで帯域通過フィルタリングによるパターンの時間的変化の階層下で評価し、フィルタ処理部、線形予測部、類似度評価部からなるシステムで 2 つのパターン間の類似度を検出している。その結果、試験音列パターンに対して、輪郭の類似や半音程度のパターンの相違も明確に検出されている¹⁾。

また、長唄三味線の旋律プロセスを分類する研究では、長唄のリズムパターンを抽出及び分類するとともに、抽出された各パターンについて各曲の例の高音の動きを比較している。比較は、曲データを数値データに変換し、コンピュータ上で同じまたは類似した箇所を抜き出したものを、人手で判定している²⁾。

旋律の類似性を推定する研究では、調の推定と作曲者の識別実験が試みられている。入力データに写像処理をし、音高別頻度を計測し、終止高音レベルから抽出した音程列パターンを基に曲の特徴ベクトルを抽出しており、特徴ベクトルと標準パターンの照合により調が推定されている。一方、作曲者の識別ではデジタル像として捉えた旋律情報を一つの信号とみなし信号処理手法を適用して旋律の類似性について検討しており、サンプリングした楽曲に FFT 処理と相関処理をし、楽曲のスペクトル分布の概要を自己相関関数によって視覚的に捉え、作曲者の識別を試みている³⁾。

歌謡曲の旋律部分を五線譜から読み取り楽曲の旋律を分析する研究では、時系列データの予測の可能性が周波数空間で“べき指数”の値と関連することに注目し、旋律を予測の可能性の尺度で分析している。旋律パターンをフーリエ解析し、スペクトル密度関数のべき型一次近似による伸長パターンを定義している。その結果、伸長パターンが楽曲の分類や作曲家の識別などに有効な情報を与えることが確認されている⁴⁾。

旋律進行の視覚的表現の研究では、フーリエ解析による音色の分析手法を旋律構造の分析に適用して、旋律を動的に分析している。フーリエ間区間における部分音列に対して抽出したパワースペクトルを旋律進行とともに順次計測し、三次元表示すると、旋律輪郭以外に一定区間におけるエネルギー・スペクトルパターンやエネルギー・スペクトル一次近似パターンの類似が旋律の類似性（刺激構造の類似性）と関連があるとの知見が得られた⁵⁾。

上記のように、音楽の分析に信号処理技術が用いられることが多数報告されている。特に、フーリエ変換を用いてスペクトル解析により音楽の特徴を分析する方法は、人の耳が位相情報を単独で処理できない構造を持っている⁶⁾ため、有効であると考えられる。

ところが、フーリエ変換によってスペクトル化された結果を処理する方法は、人手を介する必要がある。例えば、通信機やソナー、測定器では、スペクトルを表示し、それを人間が判断する例が現在でも存在する。

一方、近年は画像の類似度を調べる技術が発達し、信号処理結果を画像化し、その結果を画像として分析する手法が使用されるようになってきた。本論文では、フーリエ変換の結果を画像化し、その画像を画像の特徴量抽出によって分析する手法を検討した。

2. 音声信号の画像化

2.1 ランニングフーリエ変換

スペクトル情報の取り出しには、連続した信号を区間に区切ってその区間ごとにフーリエ変換しスペクトルを求める方法を採用した。本論文では、区間の重なりを考慮しなかった。処理する N 個のデータを持つ音声信号を $x(t_i)$ $i = \{1, 2, \dots, N\}$ とする。フーリエ変換の幅を w とし、区間の重なりが無い場合、フーリエ変換の回数は次の式(1)となる。

$$M = \left\lfloor \frac{N}{w} \right\rfloor \quad (1)$$

つまり、信号 $x(t_i)$ は、 M 個の区間に分割されることになる(式(2))。

$$x(t) = \{x_1(t), x_2(t), \dots, x_M(t)\} \quad (2)$$

信号は、式(3)のように M 個のフーリエ変換された結果となる。

$$X_j(\omega_k) = \mathcal{F}(x_j(t_k)) \quad (3)$$

ただし、 i, j は次の式(4)となる。

$$\begin{aligned} j &= \{1, 2, \dots, M\} \\ k &= \{1, 2, \dots, w\} \end{aligned} \quad (4)$$

フーリエ変換の区間の前後で不連続になるため、フーリエ変換時には、窓関数としてハミング窓を用いた。画像はスペクトルを表示することとしたため、パワースペクトルを求めた。パワースペクトルは、 w の半分の区間が繰り返されるため、画像生成には半分の区間を用いた。

したがって、生成された画像は、画像の縦座標を上から j 、横座標を左から k とすると、式(5)で示された画像となる。ただし、座標 j, k の濃度を $Y_{j,k}$ とし、濃度は最大値を 255、最小値を 0 として整数値に正規化している。

$$\begin{aligned} Y_{j,k} &= |X_j(\omega_k)|^2 \\ j &= \{1, 2, \dots, M\} \\ k &= \left\{1, 2, \dots, \frac{w}{2}\right\} \end{aligned} \quad (5)$$

2.2 画像の特徴量

画像の特徴量と画像の類似性を計算するために、MATLAB 2022b の `indexImages` 関数と `retrieveImages` 関数を用いた^{注1)}。まず、使用する全ての画像の特徴量を `indexImages` 関数によって求め、対象とする画像との一致度を `retrieveImages` 関数で求める。`IndexImages` 関数で作った特徴量には、比較対象の画像も含まれるため、最も似ている画像は、比較対象そのものになる。したがって、`indexImages` 関数によって作成された特徴量と最も近い画像は、`retrieveImages` 関数で比較された画像そのものとなるため、次に近いと判定された画像が最も近いと判定された画像となる。したがって、本論文では、`retrieveImages` 関数で 2 番目に近いと判定された画像を最も近い画像とし、3 番目に近いと判定された画像を次に近い画像であると定義した。

3. 音声信号データの取得

3.1 CD 音声の取り出し

対象とする楽曲の CD から、いわゆるリップニングにより音声データを取り出した。取り出しには

Exact Audio Copy(Ver1.6)を用いた。取り出し時にエラーが無いかを Exact Audio Copy により確認している。

3.2 アナログレコード音声の取り出し

アナログレコードは、その再生環境により音質が変動する。したがって、できるだけ普及している再生環境で再生することが重要である。特に、再生音に影響を与えるカートリッジには、放送局用に開発された DENON 社製 DL-103 を用いた。イコライザー回路には、図 1 のように示す回路を用いた。この回路の両側に増幅器を配置し、必要な増幅率を得ている。増幅器は DC サーボ付き DC アンプである。カートリッジの特性は、図 2 に示す。

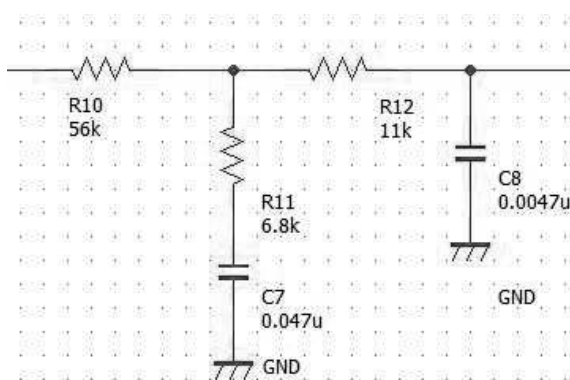


図 1 イコライザー回路

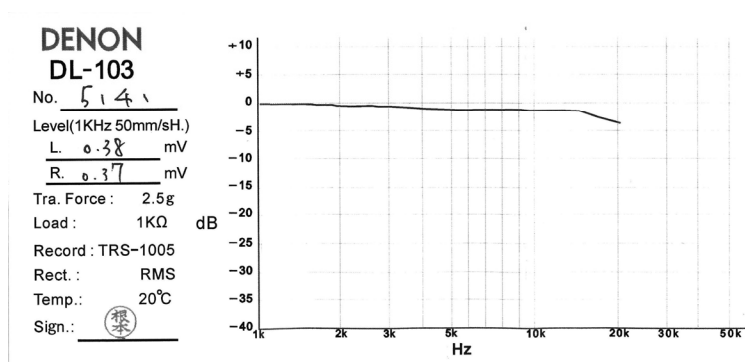


図 2 カートリッジの周波数特性

イコライザー回路を通した信号は、TEAC 社製 Portacapture X8 によりデジタル化した(図 3)。使用したレコードプレーヤは、DENON 社製 DP-57M である。デジタル化時のサンプリング周波数は 192[kHz]で、32bit 浮動小数点フォーマットで記録した。

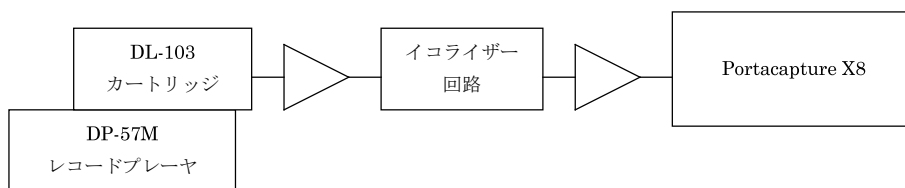


図 3 デジタイズ時の機器配置

4 画像化した音声の特徴

4.1 使用した音声データ

本論文では、音声データの特徴抽出のために、音声信号を画像化する手法を採用した。音声信号を一定時間毎に区切り、そのスペクトルを時系列に上から順に並べた。CD のサンプリング周波数は 44.1[kHz]であるため、アナログレコードのデータも 44.1[kHz]サンプリングに変換した。この変換には、Audacity(Ver.2.1)を用いた。高速フーリエ変換を用いて計算を高速化するために、フーリエ変換の幅(式(5)の w)を 4096 サンプルとした。したがって、フーリエ変換の間隔は 0.0929[s]となり、約 0.1 秒ごとにスペクトルを求めていることになる。サンプリング周波数であるため、図の右端が 22.05[kHz]となる。

使用した楽曲は、イングヴェイ・マルムスティーンの *Marching Out* に収録されている *Don't Let It End* である。今回使用したのは、CD は 1986 年に発売された日本での初版(日本版 CD)、デジタルリマスターを施し 1998 年に日本で発売されたもの(日本 98 年版 CD)、日本で 2007 年に制作されたマスターを使用したもの(日本 07 年版 CD)、米国版、西ドイツ版であり、アナログレコード(LP)は日本版、米国版、西ドイツ版である。それぞれの音声を図化したものを図 4 から図 11 に掲載する。計算には、MATLAB2022b を用い、計算結果を PNG 画像ファイルに書き出した。

4.2 聴感上の特徴と画像の特徴

日本版 CD と日本版 LP を聞くとどちらも高音が詰まったような感じに聞こえる。図 4 と図 9 は、他の画像に対して図の右側に白い部分が少ない。これは、他のものよりも高音域が少ないことを示しており、聴感上感じる感覚と一致していると考えられる。

一方、図 5,6 は、日本国内でリマスターを施して再度販売された CD の音声データから作成した画像である。図 5 は、日本 98 年版 CD で、図 6 は日本 07 年版 CD の画像である。図 4,5,6 を比較すると、日本 98 年版 CD と日本 07 年版 CD は、日本版 CD と比較して高音域が大きくなっていることが分かる。特に日本 07 年版 CD は、高音域が非常に大きくなって強調されていることが分かる。また、

日本版 CD は日本版 LP と近い傾向となっている。

一方、それ以外の CD を比べると、日本 98 年版 CD と米国版 CD、西ドイツ版 CD の差は分かりにくい。日本版 CD は不評であったが、米国版 CD と西ドイツ版 CD は日本でも日本版より良いと評価されており、日本 98 年版 CD は海外版 CD に合わせて改良したと考えられる。

日本 07 年版 CD は、他のものと大きく異なっており、高音域を多く含んでいる。これは、音楽の聴取環境が異なり、高音域が大きな方が望まれる時代になったため、編集作業により高音域を大きくしたものと推測される。しかし、日本 07 年版 CD は、定性的には高音域を大きくされすぎているためか、やや不自然に感じられる。

著者らが聴感上で最も良いと感じたのは、西ドイツ版 LP である。この版は、図 8 のように、米国版 LP や日本版 LP に比べて、高音域まで記録されている。聴感上も歌手の声や楽器の高音域が良く聞こえるように感じられる。

4.3 画像の特徴量

2 章で述べたように、画像化したデータを MATLAB 2022b の `indexImages` 関数と `retrieveImages` 関数を用いて、特徴の近いものを計算により判定した。最も近いと判定されたものを表 1、次に似ていると判定されたものを表 2 に示す。表の上段は、MATLAB 2022b の `retrieveImages` 関数の比較対象とした画像を示している。つまり、上段の画像に近いと判定されたものが下段である。

表 1 から、聴感上も近いと判断できたものと近い結果になっている。CD に関しては、日本 98 年版 CD が米国版 CD と近く、米国版 CD は西ドイツ版 CD と近く、西ドイツ版 CD が日本 98 年版 CD に近い。このことから、日本版 CD と日本 07 年版 CD を除く CD 版は、それぞれが似たような特徴を持っていることが分かる。

日本版 CD と最も近いと判定されたのは日本版 LP であり、日本版 LP と最も近いと判定されたのは日本版 CD であった。また、日本版 CD と日本版 LP に 2 番目に似ていると判断されたのは日本 98 年版 CD であった。これらのことから、日本版 CD と日本版 LP は相互に似ていることが分かる。

特徴的なのは、日本 07 年版 CD である。日本 07 年版 CD は、他の版から近いとされた版が存在しない。高音域が多く含まれているため、高音域が大きくなっている西ドイツ版 LP や西ドイツ版 CD と近いことになっているが、西ドイツ版 LP は米国版 CD や西ドイツ版 CD に、西ドイツ版 CD は日本 98 年版 CD や米国版 CD に近い。このことは、日本 07 年版 CD の特徴が孤立していることを示している。

4.4 画像の特徴量の比較結果と聴感上の特徴の対比

日本版 LP と日本版 CD は共に「音がこもっている」という理由で低く評価されている。日本版 LP と日本版 CD の画像の特徴は相互に近く、聴感上同じように聞こえることをこの結果は支持している。

西ドイツ版 LP は、米国版 CD や西ドイツ版 CD と似ていると判別された。西ドイツ版 LP が特に評価が高く、米国版 CD や西ドイツ版 CD が概ね好評であることから、聴感上の定性的な特徴が、画像化して特徴量を求めた結果と一致していると考えられる。

また、特徴が孤立していた日本 07 年版 CD は、国内オリジナル・アナログ・マスターを基にした 24bit リマスター音源とされており、大きく編集操作が入っているものと思われる。このことは、聴感上も感じることができる。インターネット上では、日本 07 年版 CD について「ギターの輪郭が明確になったかもしれない」「かなり音質的に向上している」「ヴォーカルとドラムがクリアになった」と音質の変化を評価する感想が見られる^{注6)}。一方、日本 07 年版 CD について「音が大きくなっただけ」「モコモコ感が修正されていない」とする感想も見られる^{注7)}。このように聴感上の感想において様々な意見があることは、画像の特徴が孤立していることと関連していると思われる。

表 1 最も似ていると判定されたもの

日本版 CD	日本 98 年版 CD	日本 07 年版 CD	米国版 CD	西ドイツ版 CD	日本版 LP	米国版 LP	西ドイツ版 LP
日本版 LP	米国版 CD	西ドイツ版 LP	西ドイツ版 CD	日本 98 年版 CD	日本版 CD	西ドイツ版 CD	米国版 CD

表 2 次に似ていると判定されたもの

日本版 CD	日本 98 年版 CD	日本 07 年版 CD	アメリカ版 CD	西ドイツ版 CD	日本版 LP	米国版 LP	西ドイツ版 LP
日本 98 年版 CD	米国版 LP	西ドイツ版 CD	西ドイツ版 LP	米国版 CD	日本 98 年版 CD	米国版 CD	西ドイツ版 CD

6. まとめ

本論文では、聴感上定性的に述べられている同一楽曲の配布形態による特徴の差が、定量的に確認できるかを確認した。この確認の方法に、本論文は、音声データを図化して、その図を絵の特徴として分析する方法を提案した。

本論文では、同一の楽曲がアナログレコードや CD によって違いがあるのかを題材とした。実験には、各国版のアナログレコードや CD を容易に入手可能であり、配布形態により異なる特徴があるとされているイングヴェイ・マルムスティーンの曲を用いた。

特徴の抽出には、一定間隔毎のスペクトルを図に表現し、画像の類似度で評価する方法を採用した。その図を MATLAB 2022b の `indexImages` と `retrieveImages` 関数によって分析した。分析の結果、類似度判定は、定性的な楽曲の判定と一致していることが確認できた。

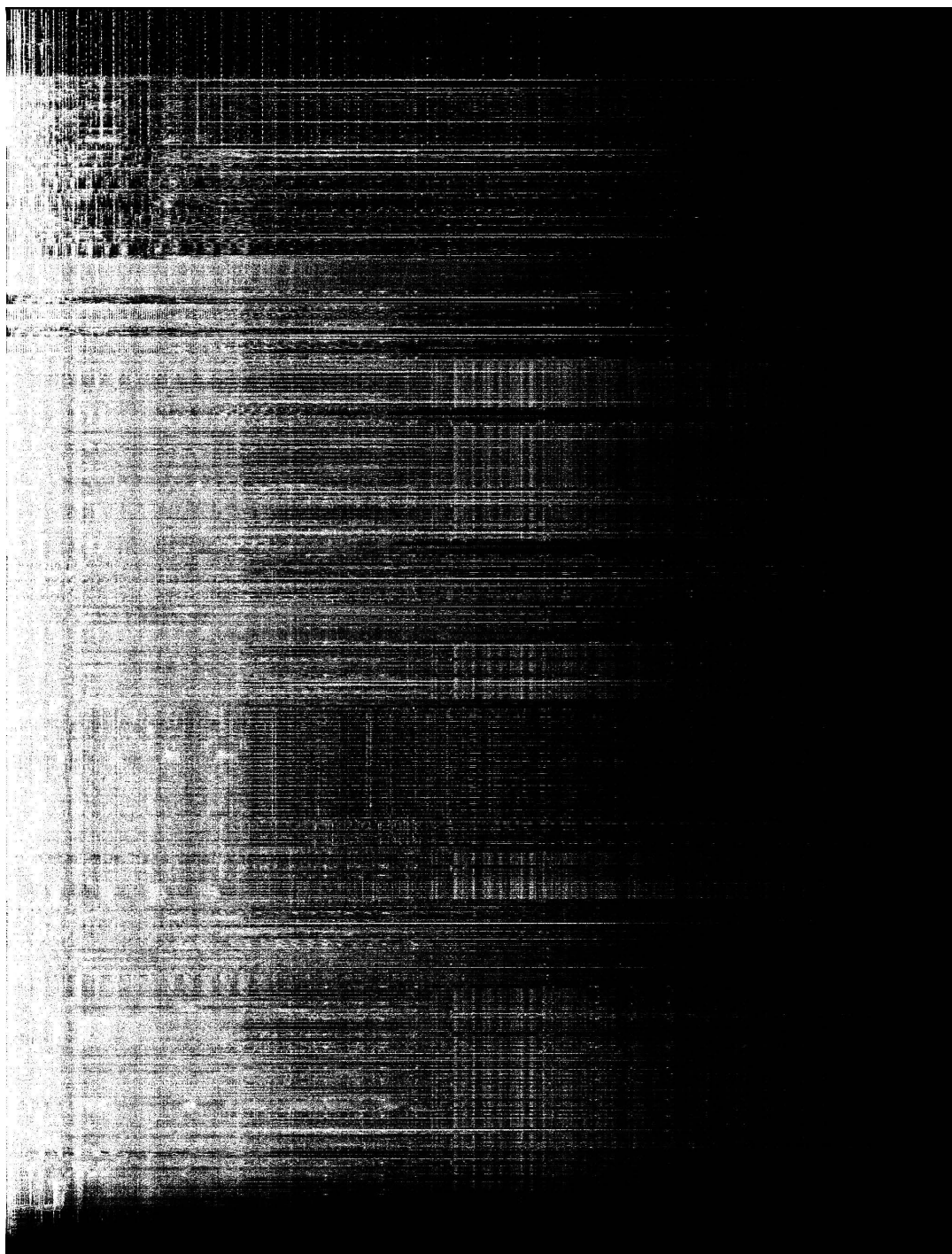


图 4 日本版 CD

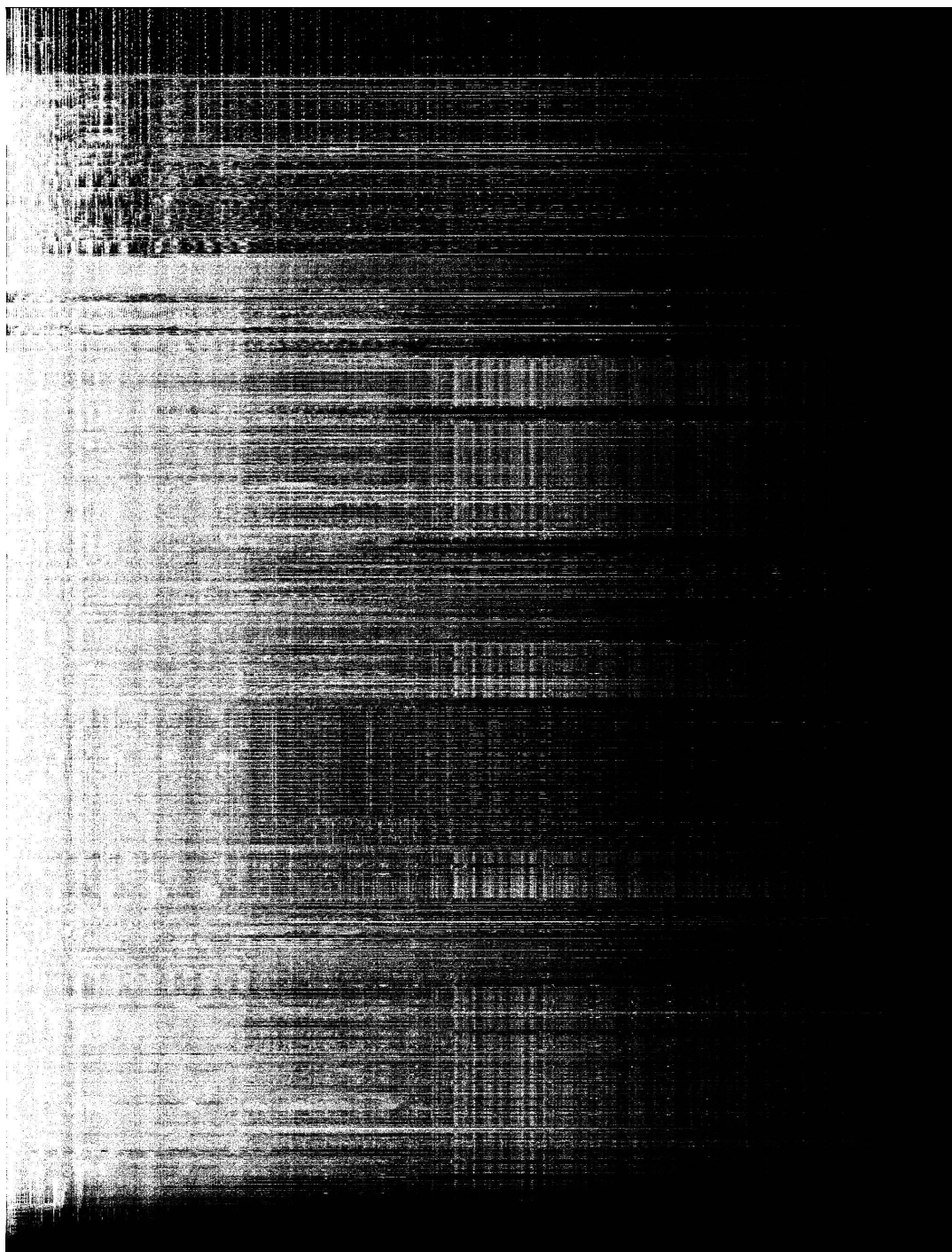


图 5 日本 98 年版 CD

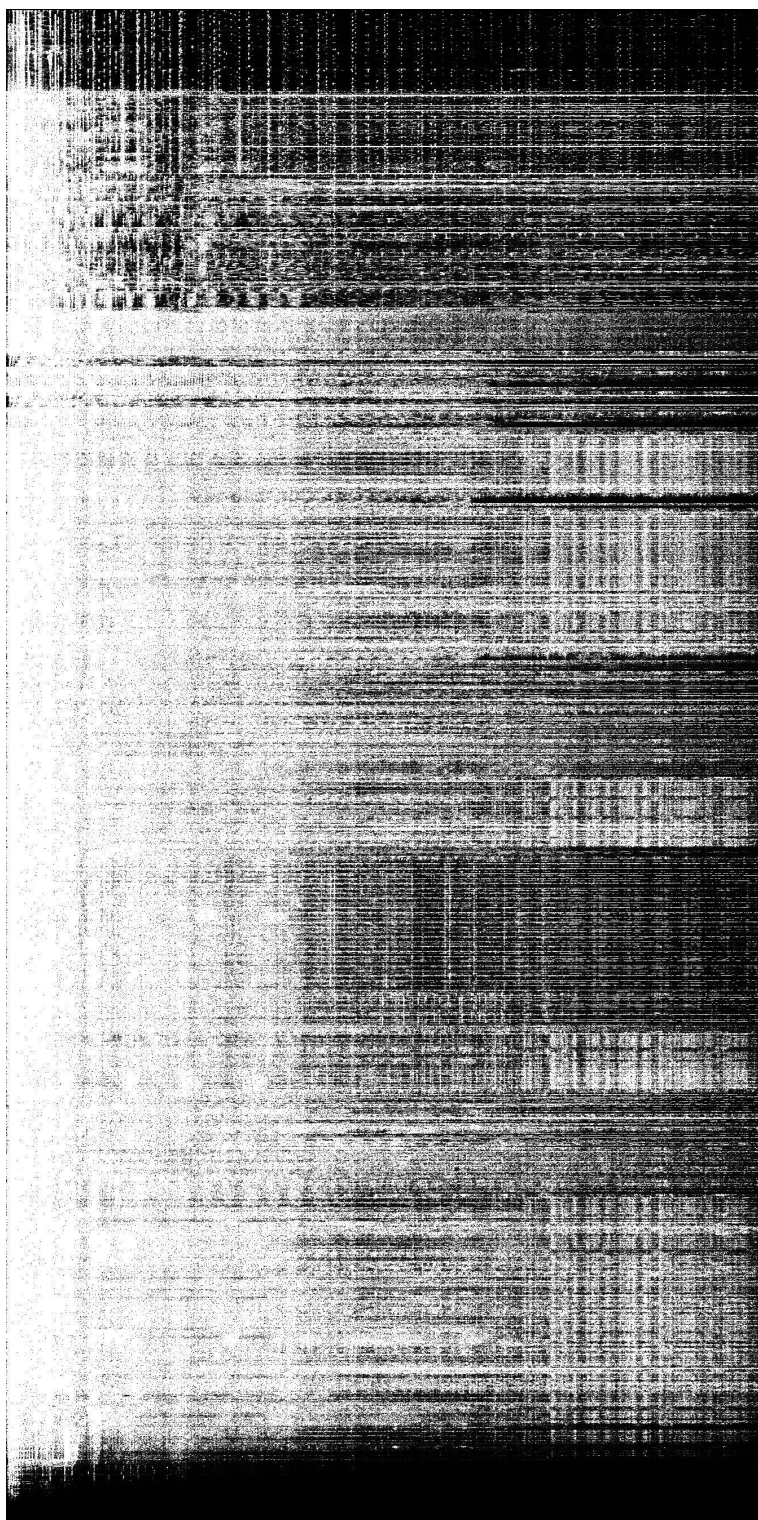


图 6 日本 07 年版 CD

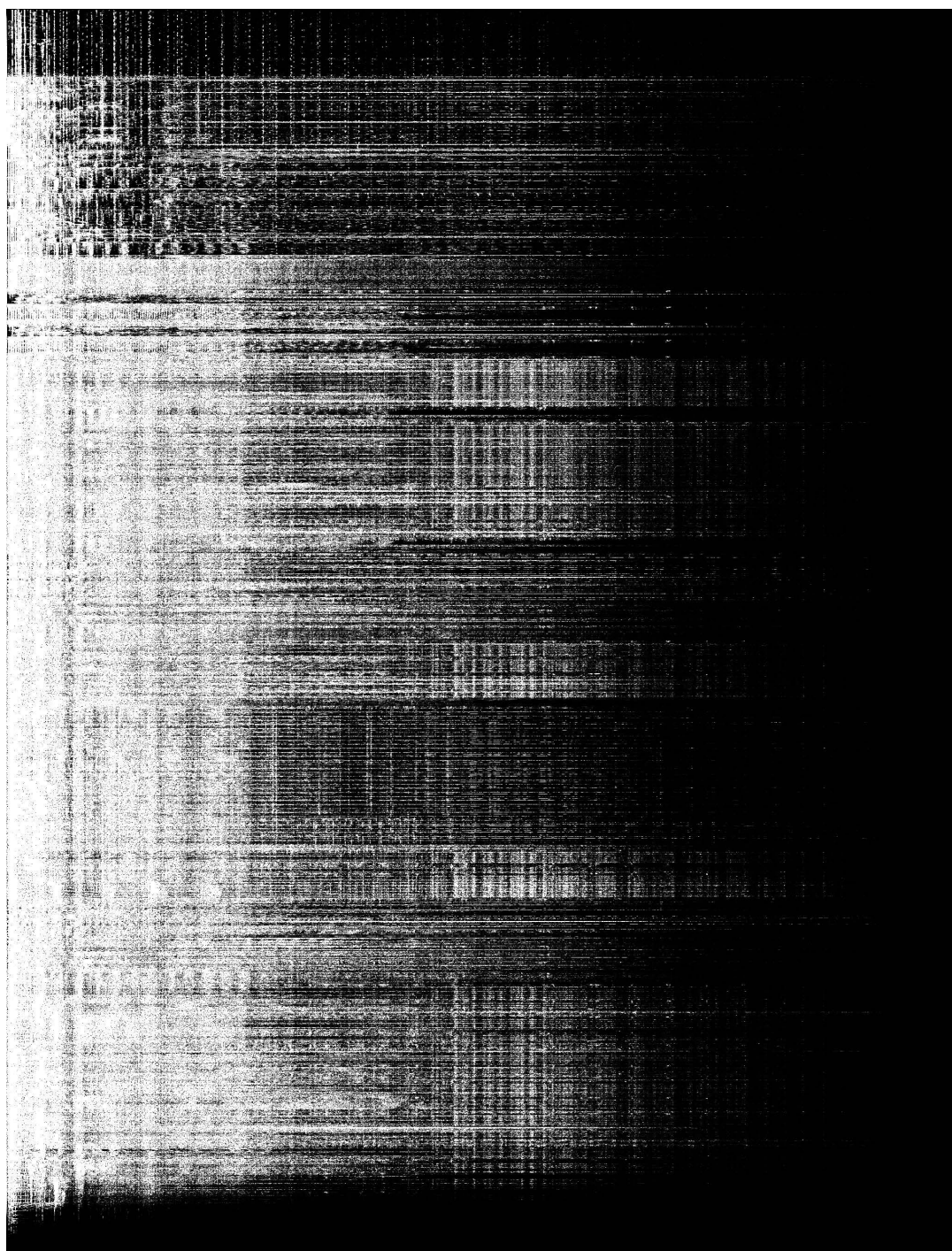


图 7 米国版 CD

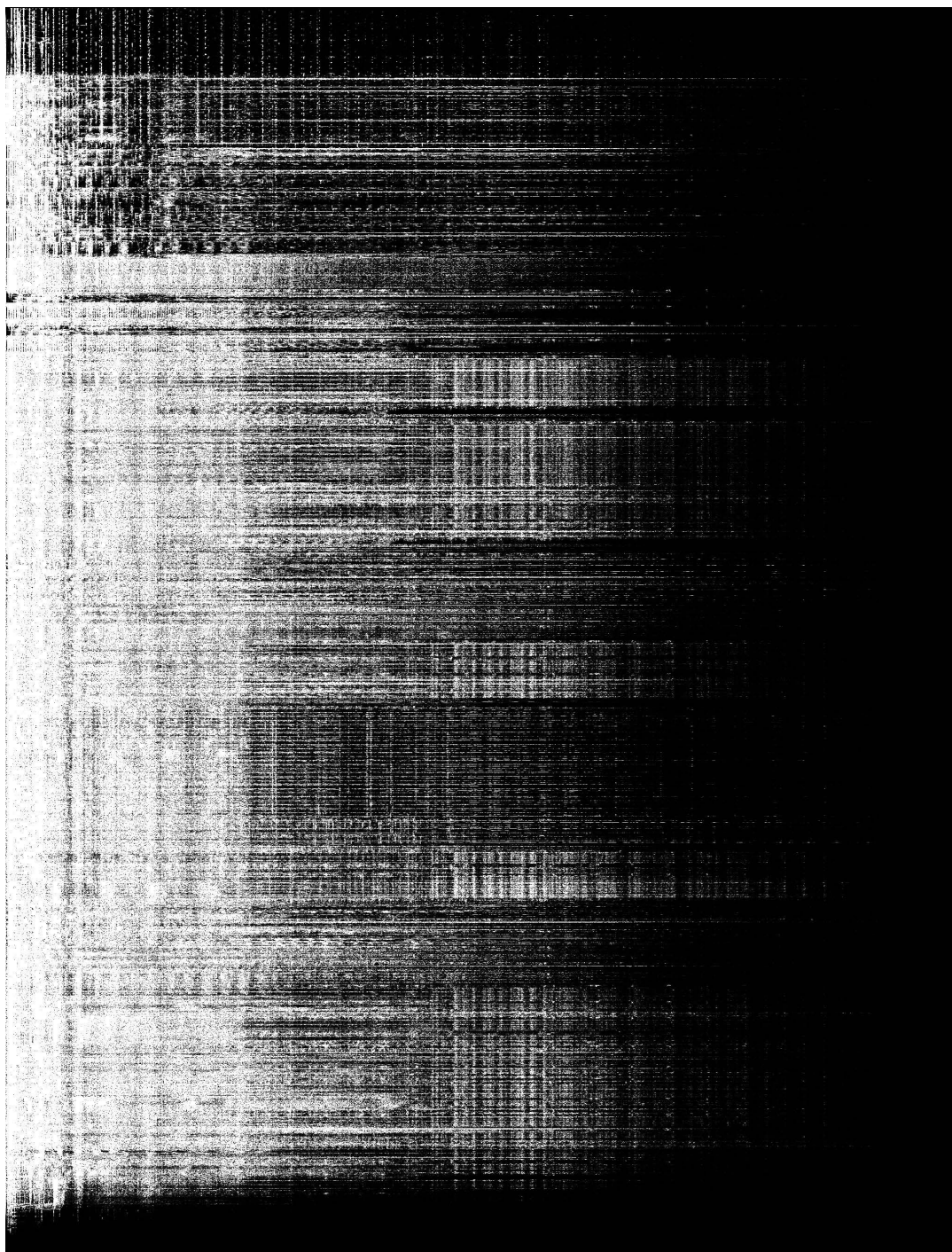


図 8 西ドイツ版 CD

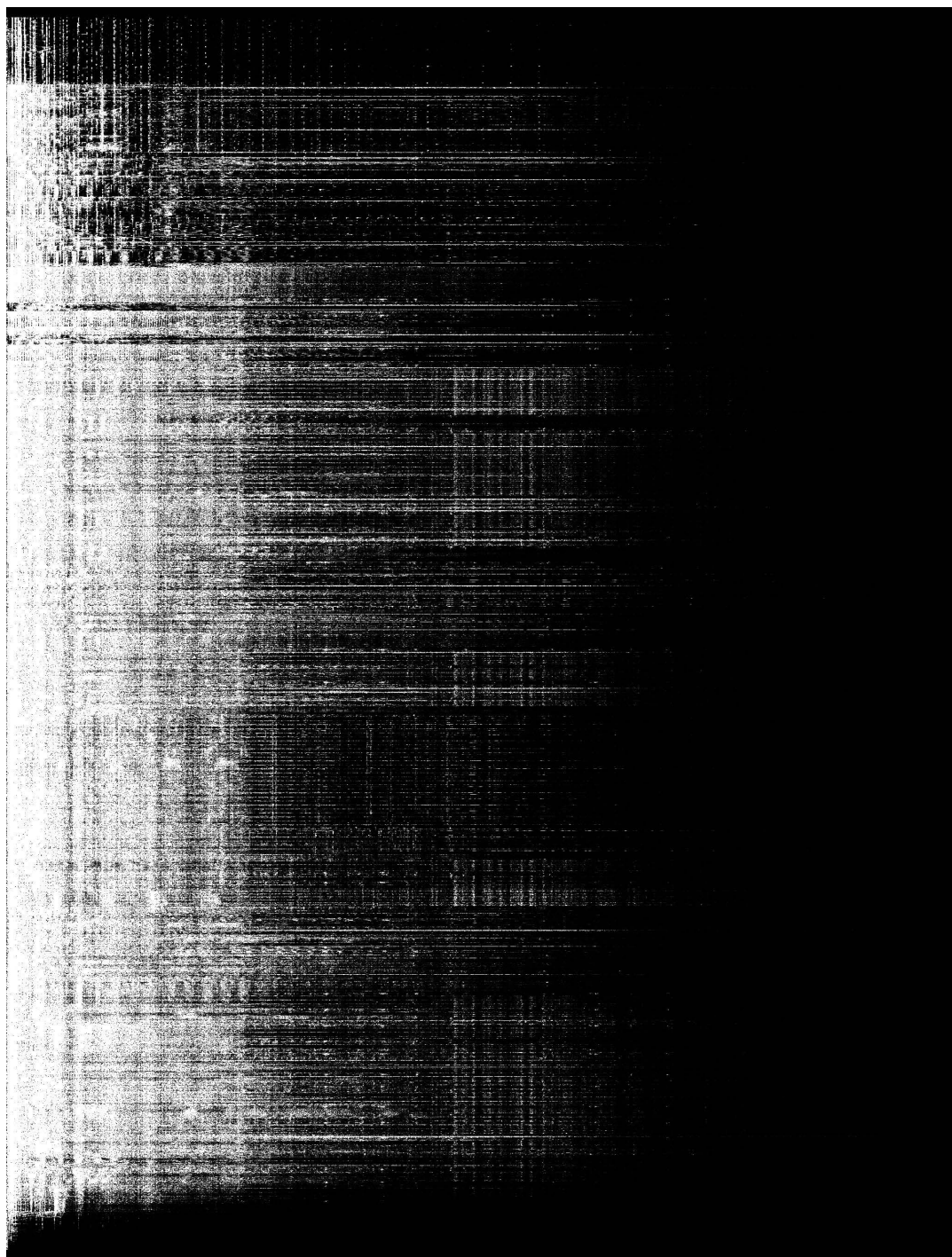


图9 日本版LP

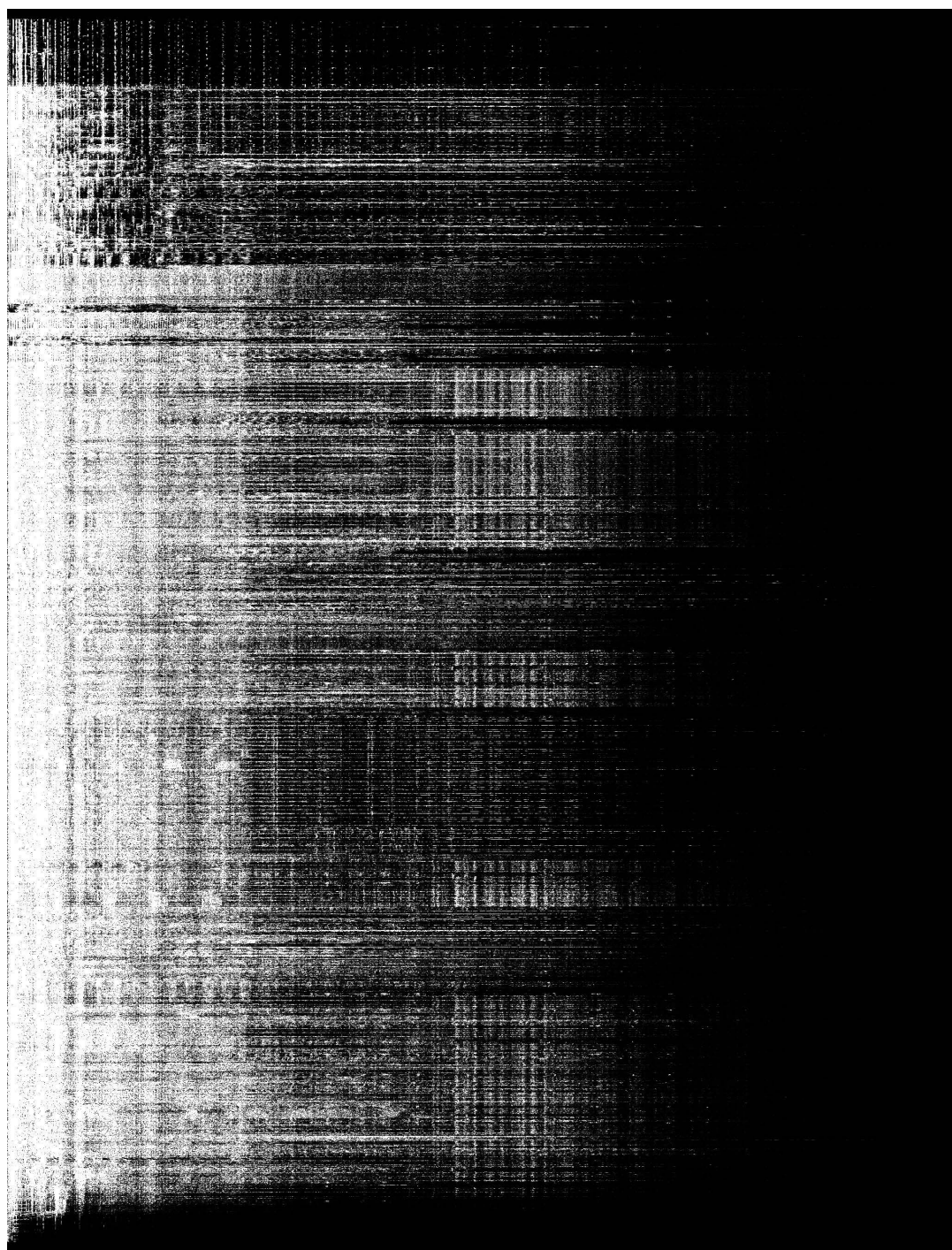


图 10 米国版 LP

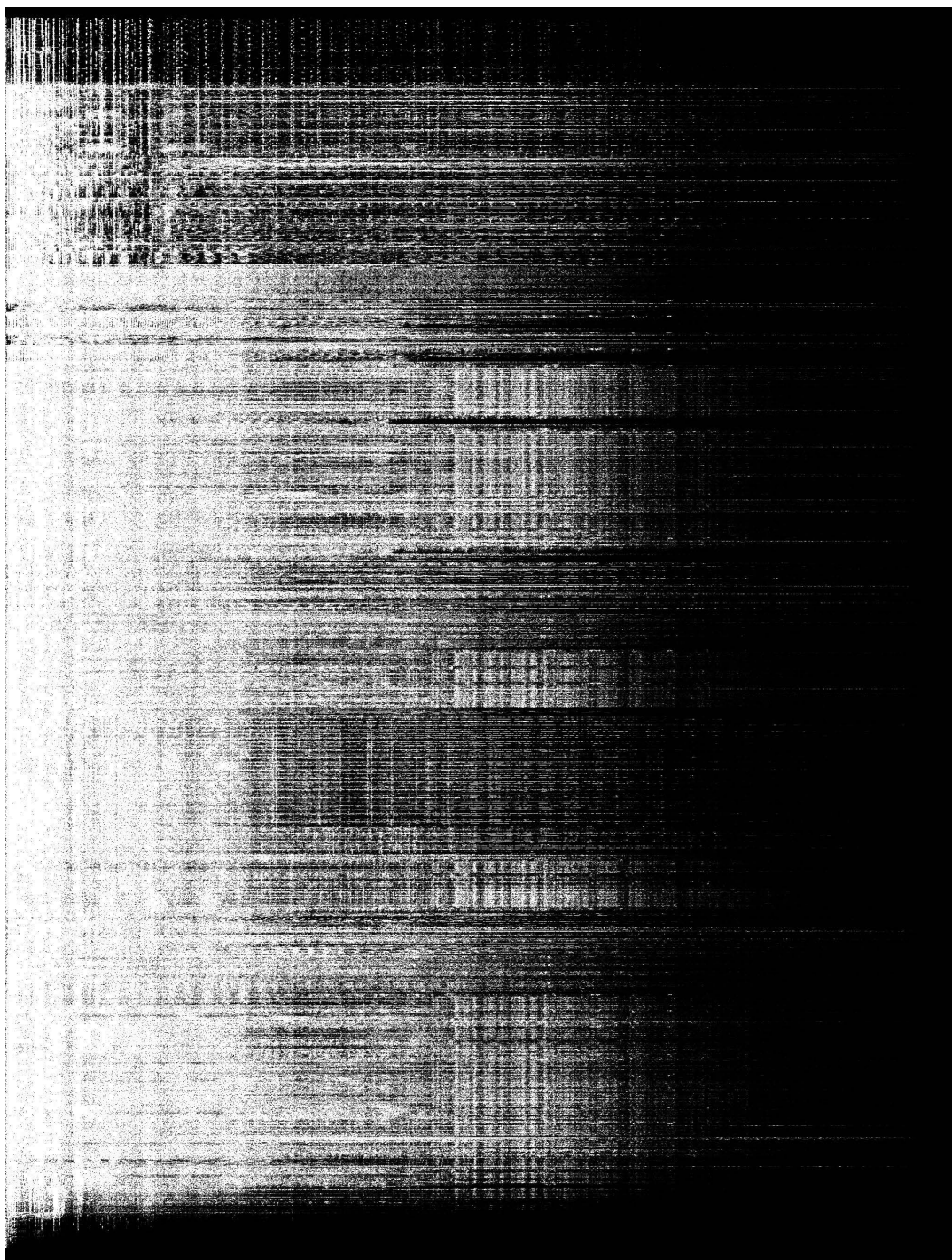


図 11 西ドイツ版 LP

《参考文献》

- (1) 秋山 好一, 松田 稔, 中野 稔, メロディパターンの類似性判別—聴感上の類似と相違—, 音楽音響研究会資料, Vol.14, No.4, pp.7-12 (1995)
- (2) 矢向 正人, 長唄三味線の旋律プロセスの分析, 音楽音響研究会資料, Vol.11, No.6, pp.1-8 (1992)
- (3) 秋山 好一, 松田 稔, 中野 稔, 楽曲分析 単旋律の類似性に関する検討, 音楽音響研究会資料, Vol.5, No.6, pp.1-8 (1987)
- (4) 秋山 好一, 松田 稔, 中野 稔, 楽曲の旋律分析—周波数揺らぎの伸長パターン—, 音楽音響研究会資料, Vol.9, No.2, pp.11-18 (1990)
- (5) 秋山 好一, 松田 稔, 中野 稔, 旋律進行の視覚的表現, 音楽音響研究会資料, Vol.7, No.6, pp.1-7 (1988)
- (6) 大串 健吾, 音のピッチ知覚, コロナ社, pp.10-50 (2016)

《注》

- (1) <https://jp.mathworks.com/help/vision/ref/retrieveimages.html> (2023 年 1 月 23 日閲覧)
- (2) <https://lavender.5ch.net/test/read.cgi/hrhm/1573466419/> (2023 年 1 月 25 日閲覧)
- (3) <https://lavender.5ch.net/test/read.cgi/hrhm/1673054335/150> (2023 年 1 月 25 日閲覧)
- (4) <http://blog.fantomas.kill.jp/?eid=928638> (2023 年 1 月 25 日閲覧)
- (5) <http://fantomas.jugem.jp/?eid=929185#comments> (2023 年 1 月 25 日閲覧)
- (6) <https://www.amazon.co.jp/dp/B005S1Y6TM> (2023 年 1 月 25 日閲覧)
- (7) <http://narick.jugem.jp/?eid=779> (2023 年 1 月 25 日閲覧)