

修士論文

非負値行列因子分解を用いた
ピアノ音の個体差分析に関する研究

北陸先端科学技術大学
情報科学研究科情報科学専攻

小林 慶祐

2014年3月

修 士 論 文

**非負値行列因子分解を用いた
ピアノ音の個体差分析に関する研究**

指導教員 赤木正人 教授

審査委員主査 赤木正人 教授
審査委員 鵜木祐史 准教授
審査委員 党建武 教授

北陸先端科学技術大学院大学
情報科学研究科情報科学専攻

1210023 小林 慶祐

提出年月: 2014年2月

概要

近年の計算機の普及、発展に伴い、Desktop Music (DTM) による音楽制作が、個人でも簡単に行えるようになってきた。DAW (Digital Audio Workstation) の発達により、イコライザーやコンプレッサーといった、エフェクタ群をハードウェアで持つ必要がなくなり、ほとんどが計算機上で行えるようになった。また、YAMAHA の歌声合成ソフト Vocaloid によって、歌唱自体も人が行う必要がなくなり、純粹に作曲活動に専念出来る環境が整っている。DTM では、MIDI (Musical Instruments Digital Interface) 音源に代表されるソフトウェア音源が一般的に用いられているが、それらの多くは楽器音の中でも、典型的な音しか表現出来ない。そのため、個体差を含む実楽器音を使用する場合、大規模な実楽器音の単音データベースを所有するか、実際に演奏する必要がある。単音データベースは楽器の種類によっては非常に少なく、高価であるため、容易に入手しづらい。また、演奏する場合には、演奏者の技量に左右され、楽器の演奏が未経験であれば、演奏することすらままならない。従って、個体差を含む実楽器音を用いた音楽制作は困難になっている。

この楽器音の個体差を扱うにあたり、楽器の生成機構をモデル化する研究が行われてきた。特に発音機構がとても複雑なピアノに関する研究は数多く行われてきた。しかし、提案されているモデルを用いても、ピアノ音を完全に再現は出来ない。また、制御パラメータが多く、実験的にしかパラメータの値を決める事が出来ないため、音色再現は非常に困難である。山家らはピアノ音の個体差と演奏時の印象評価を対応づけるにあたり、調波構造と時間波形それぞれに個体差が含まれると指摘している。しかし、山家らが指摘しているのはピアノ音の立ち上がりの箇所であり、他の箇所については触れられていない。また、Lee らが指摘するように、調波構造は時間的に変動をするものであるため、一部分のみを分析するのでは不十分である。

本研究は上記の問題点を解決するために、ピアノ音の調波構造とその時間変動を同時に扱いながら個体差の分析を行う。この分析にあたり、非負値行列因子分解 (Non-negative Matrix Factorization; NMF) を分析手法として用いる。NMF はサウンドスペクトログラムを、 K 個の調波構造を表す基底行列と、それらに対応する時間変動パターンを表すアクティベーション行列の2つの行列に分解することができる。また、基底行列の一部をピアノ音の共通成分で固定する事が出来れば、ピアノ音の共通成分と、それ以外の成分、つまり入力ピアノ音固有の成分が表れると考えられる。NMF で分析するにあたり、ピアノの発音機構を考慮した結果、入力はパワースペクトログラムではなく、対数パワースペクトログラムを入力として用いる。

まず、NMF の基底数 K を決定するために、 K を 1~5 まで変化させて NMF で分析を行った。その結果、信号対歪み率 (SDR) は基底数が 3 以上になると、約 11 dB 程度で横ばいになる傾向を示した。

基底数が 4 以上の場合、ピアノの発音機構と対応しない分離結果や、冗長な分離結果が生じた。一方、基底数が 3 の場合は、定常部、立ち上がり部、減衰部の 3 種類に分類され

た。定常部は響板，ふた等の反響と，立ち上がり部は打鍵，打弦時に発生する雑音と，減衰部は弦の2段階減衰と対応していると考えられる。ピアノ音を表現出来る基底数を求めた後，具体的な個体差を示すために，基底行列の一部をピアノ音の共通成分で固定して再度分析を行った。その結果，個体差成分として響板などによる反響時の調波構造，弦の2段階減衰の後半の緩やかな減衰部分が，グランドピアノ音，アップライトピアノ音それぞれ表れた。特に弦の2段階減衰は共通成分で最初の急峻な減衰が表れ，グランドピアノ音とアップライトピアノ音，それぞれの発音機構の影響が表れていた。これらに対し，MIDI音源では，打源位置による第6～8倍音のピーク値の影響や，2段階減衰，ダンパの影響などが反映されていないことも，分析結果の比較によりわかった。

本研究ではピアノ音の個体差分析として対数パワースペクトログラムを用いてNMFで分析を行った。この分析では，要因が直列でつながっているものを分析出来るので，他の楽器でも要因が直列的であれば，同様に個体差を導出出来ると考えられる。また，この得られた個体差成分を別の音源のものと入れ替え，対数パワースペクトログラムを再合成，そしてそこから，時間波形に戻せば，音色変換が行えることが考えられる。今回MIDI音源では弦の2段階減衰とダンパの影響が表現出来ていなかった。そこで，NMFで分析してそのときに得られた減衰部を表すアクティベーションを実音源のアクティベーションに入れ替えて，スペクトログラム，そして時間波形を再合成すれば，より実音源に近いMIDI音源に変換出来ると思われる。このように，今回得られた個体差の情報を用いて，音色変換や，演奏の表情付けなどに転用出来るかと期待される。

目次

第1章 序論	1
1.1 音楽制作の現状	1
1.2 本研究の背景	1
1.2.1 ピアノ音の個体差	1
1.2.2 従来手法の問題点	2
1.3 本研究の目的	2
1.4 本論文の構成	3
第2章 ピアノの発音原理	4
2.1 ピアノの種類	4
2.2 ピアノ音の発音機構	4
2.3 インハーモニシティ	9
2.4 まとめ	9
第3章 ピアノ音の個体差分析手法	11
3.1 ピアノ音の個体差分析の流れ	11
3.2 非負値行列因子分解	11
3.2.1 距離尺度	12
3.2.2 更新則	12
3.3 高域強調処理	14
3.4 対数パワースペクトログラムを用いた非負値行列因子分解による分析	14
3.5 値のスケール変換	15
3.6 まとめ	16
第4章 非負値行列因子分解を用いたピアノ音の個体差分析	17
4.1 基底数 K の決定	17
4.1.1 実験目的	17
4.1.2 実験条件	17
4.1.3 分析結果と考察	22
4.2 基底行列の部分固定を用いた非負値行列因子分解によるピアノ音の個体差分析	30
4.2.1 実験目的	30

4.2.2	実験条件	31
4.2.3	分析結果と考察	31
4.3	まとめ	39
第5章	総合検討	40
5.1	グランドピアノ音間での個体差の違い	40
5.2	アップライトピアノ音間での個体差の違い	43
5.3	各ピアノ間での個体差の違い	46
5.3.1	アクティベーションベクトルの差異	46
5.3.2	定常部ベクトルの差異	46
5.3.3	高周波数領域の影響	50
5.4	まとめ	51
第6章	結論	53
6.1	本研究のまとめ	53
6.2	今後の展望	53
	謝辞	55
	参考文献	58
	付録	58
	研究業績	126

目次

2.1	グランドピアノ	5
2.2	アップライトピアノ	5
2.3	グランドピアノのアクション	6
2.4	アップライトピアノのアクション	6
2.5	ピアノ音の発音機構	7
2.6	響板の各周波数における相対的な強さのパターン	8
2.7	パラメータ B と第 n 倍音の周波数値の変化	10
3.1	ピアノ音分析のブロック図	11
3.2	各距離尺度における 2 変数の近さ度合い	13
3.3	高域強調無しで NMF による分析をしたときの基底行列	15
4.1	分析窓幅 512 点での分析結果	19
4.2	分析窓幅 2048 点での分析結果	20
4.3	分析窓幅 4096 点での分析結果	21
4.4	基底数と SDR の関係	24
4.5	$K = 1$ の場合の分離結果 (RWC1)	25
4.6	$K = 2$ の場合の分離結果 (RWC1)	26
4.7	$K = 3$ の場合の分離結果 (RWC1)	27
4.8	$K = 4$ の場合の分離結果 (RWC1)	28
4.9	$K = 5$ の場合の分離結果 (RWC1)	29
4.10	基底固定を用いる NMF の概念図	30
4.11	グランドピアノ音の共通成分	33
4.12	アップライトピアノ音の共通成分	33
4.13	グランドピアノ音の分析結果 (GP2)	34
4.14	グランドピアノ音の分析結果 (RWC3)	35
4.15	基底行列の一部を固定して NMF を適応した結果 (UP1)	36
4.16	基底行列の一部を固定して NMF を適応した結果 (UP2)	37
4.17	基底行列の一部を固定して NMF を適応した結果 (MAPS)	38
5.1	アップライトピアノ音の非固定部の定常部ベクトル	41
5.2	グランドピアノ音の減衰部ベクトルのアクティベーション	42

5.3	アップライトピアノ音の非固定部の定常部ベクトル	44
5.4	アップライトピアノ音の非固定部のアタックベクトル	45
5.5	各ピアノ音の減衰部ベクトルのアクティベーション	47
5.6	各ピアノ音の共通成分の減衰部ベクトルのアクティベーション	48
5.7	各ピアノ音の定常部ベクトル	49
5.8	アップライトピアノ音の基底行列 (UP3)	51
1	YAMAHA GRAND C3	60
2	GP1 のマイク設置	60
3	GP2 のマイク設置	62
4	アップライトピアノのマイク設置	63
5	$K = 1$ の場合の分離結果 (MIDI)	66
6	$K = 2$ の場合の分離結果 (MIDI)	67
7	$K = 3$ の場合の分離結果 (MIDI)	68
8	$K = 4$ の場合の分離結果 (MIDI)	69
9	$K = 5$ の場合の分離結果 (MIDI)	70
10	$K = 1$ の場合の分離結果 (UP1)	71
11	$K = 2$ の場合の分離結果 (UP1)	72
12	$K = 3$ の場合の分離結果 (UP1)	73
13	$K = 4$ の場合の分離結果 (UP1)	74
14	$K = 5$ の場合の分離結果 (UP1)	75
15	$K = 1$ の場合の分離結果 (UP2)	76
16	$K = 2$ の場合の分離結果 (UP2)	77
17	$K = 3$ の場合の分離結果 (UP2)	78
18	$K = 4$ の場合の分離結果 (UP2)	79
19	$K = 5$ の場合の分離結果 (UP2)	80
20	$K = 1$ の場合の分離結果 (UP3)	81
21	$K = 2$ の場合の分離結果 (UP3)	82
22	$K = 3$ の場合の分離結果 (UP3)	83
23	$K = 4$ の場合の分離結果 (UP3)	84
24	$K = 5$ の場合の分離結果 (UP3)	85
25	$K = 1$ の場合の分離結果 (UP4)	86
26	$K = 2$ の場合の分離結果 (UP4)	87
27	$K = 3$ の場合の分離結果 (UP4)	88
28	$K = 4$ の場合の分離結果 (UP4)	89
29	$K = 5$ の場合の分離結果 (UP4)	90
30	$K = 1$ の場合の分離結果 (MAPS)	91
31	$K = 2$ の場合の分離結果 (MAPS)	92
32	$K = 3$ の場合の分離結果 (MAPS)	93

33	$K = 4$ の場合の分離結果 (MAPS)	94
34	$K = 5$ の場合の分離結果 (MAPS)	95
35	$K = 1$ の場合の分離結果 (GP1)	96
36	$K = 2$ の場合の分離結果 (GP1)	97
37	$K = 3$ の場合の分離結果 (GP1)	98
38	$K = 4$ の場合の分離結果 (GP1)	99
39	$K = 5$ の場合の分離結果 (GP1)	100
40	$K = 1$ の場合の分離結果 (GP2)	101
41	$K = 2$ の場合の分離結果 (GP2)	102
42	$K = 3$ の場合の分離結果 (GP2)	103
43	$K = 4$ の場合の分離結果 (GP2)	104
44	$K = 5$ の場合の分離結果 (GP2)	105
45	$K = 1$ の場合の分離結果 (RWC1)	106
46	$K = 2$ の場合の分離結果 (RWC1)	107
47	$K = 3$ の場合の分離結果 (RWC1)	108
48	$K = 4$ の場合の分離結果 (RWC1)	109
49	$K = 5$ の場合の分離結果 (RWC1)	110
50	$K = 1$ の場合の分離結果 (RWC3)	111
51	$K = 2$ の場合の分離結果 (RWC3)	112
52	$K = 3$ の場合の分離結果 (RWC3)	113
53	$K = 4$ の場合の分離結果 (RWC3)	114
54	$K = 5$ の場合の分離結果 (RWC3)	115
55	基底行列の一部を固定して NMF を適応した結果 (UP1)	117
56	基底行列の一部を固定して NMF を適応した結果 (UP2)	118
57	基底行列の一部を固定して NMF を適応した結果 (UP3)	119
58	基底行列の一部を固定して NMF を適応した結果 (UP4)	120
59	基底行列の一部を固定して NMF を適応した結果 (MAPS)	121
60	基底行列の一部を固定して NMF を適応した結果 (GP1)	122
61	基底行列の一部を固定して NMF を適応した結果 (GP2)	123
62	基底行列の一部を固定して NMF を適応した結果 (RWC1)	124
63	基底行列の一部を固定して NMF を適応した結果 (RWC3)	125

表 目 次

4.1	基底数 $K = 3$ と $K = 4$ の各基底ベクトル間での相関係数	22
4.2	基底数 $K = 3$ と $K = 5$ の各基底ベクトル間での相関係数	23
5.1	グランドピアノ音の減衰の傾き	43
5.2	アップライトピアノ音の減衰の傾き	43
5.3	各ピアノ間の相関係数の平均値	50
1	収録機材一覧	61
2	収録音の設定	64
3	GP1 収録時のマイク位置	64
4	GP2 収録時のマイク位置	65
5	アップライトピアノ音収録時のマイク位置	65

第1章 序論

1.1 音楽制作の現状

近年コンピュータの発展に伴い、Desktop Music (DTM) による音楽制作が個人でも簡単に行えるようになってきた。特に DAW (Digital Audio Workstation) の高性能化のおかげで、イコライザーやコンプレッサーなどといった、エフェクタをハードウェアで用意する必要がない。また、Vocaloid [1] に代表されるような歌声合成ソフトウェアを用いることにより、楽器音のみの楽曲だけでなく、歌唱曲の制作も行える。

DTM で使われる音源としては、MIDI (Musical Instruments Digital Interface) 音源に代表されるソフトウェア音源が一般的に用いられている。しかし、それらの多くは楽器音の中でも、典型的な音しか表現出来ない。ソフトウェア音源の代わりに個性差を含む実楽器音を使用する場合、大規模な実楽器音の単音データベースを所有するか、実際に演奏する必要がある。楽器の単音データベースの数は少なく、高価であるため、容易に入手しづらい。また、演奏する場合には、演奏者の技量に左右され、楽器の演奏が未経験であれば、演奏することすらままならない。従って、個性差を含む実楽器音を用いた音楽制作は困難になっている。

1.2 本研究の背景

1.2.1 ピアノ音の個性差

まずはじめに、楽器音の個性差を以下のように定義する。

定義 楽器音の個性差とは材質、構造、などに起因する楽器音の差異

この楽器音の個性差には、音高、音量、演奏者、演奏法に関しては同一の条件である。つまり、外的要因による影響がない、純粋にその楽器固有の持つ成分こそが、楽器音の個性差である。

本研究では楽器音のなかでもピアノに焦点をあてて個性差の分析を行う。ピアノは後述する発音機構が非常に複雑であり、現在でもこの発音機構等の解析の研究が行われている。そのため、この複雑な機構から生じる個性差を捉えることを試みる。

1.2.2 従来手法の問題点

ピアノ音の個性差を扱うにあたり、楽器の生成機構のモデル化に関する研究が従来から行われてきた [2-7]. しかし、これらのモデルを用いてもピアノ音を完全に再現出来ず、再現しているのはピアノを構成している一部の影響である。また、これらのモデルは制御パラメータが多く、実験的にしかパラメータの値を決める事が出来ないため、実楽器音の音色再現は難しい。

楽音の個性差を扱う研究として音色変換に関するものがある。時間波形に対してテンプレートフィルタリングを行うもの [8] や、スペクトルを変形して音色を変化させる方法 [9] などがある。しかし、これらの手法は変換こそ出来るが、個性差が何なのかについては明記されていない。

個性差そのものについては、山家ら [10] はピアノ音の個性差が調波構造及び時間波形それぞれにあると指摘している。したがって、テンプレートフィルタリングを用いた方法は時間波形のみを、スペクトルを変形させる手法に関しては調波構造のみをターゲットにしているのので、山家らが指摘している調波構造と時間波形それぞれの関係性はわからない。

また、山家らが指摘している箇所は、ピアノ音の立ち上がり部分であり、音の定常的な部分や減衰特性などについては触れられていない。そもそも、Lee *et al.* [11] が指摘するように、調波構造は時間的に変動するので、山家らの分析のように、ある一時点における時間波形とその調波構造だけでは不十分と言える。従って、ピアノ音の個性差を分析するには、調波構造とその長期の時間変動を同時に扱う必要がある。

1.3 本研究の目的

本研究の目的はピアノ音の調波構造とその時間変動を同時に分析し、ピアノ音の個性差を明らかにすることである。ピアノ音の個性差分析にあたり、本研究では非負値行列因子分解 (Non-negative Matrix Factorization; NMF) [12] を用いる。NMF は元々画像処理に用いられる手法だが、近年音響信号に適用され、音源分離や自動採譜などで非常に優れた効果を発揮している [9, 13-16]. NMF は入力音のサウンドスペクトログラムを、 K 個の調波構造パターンと、その時間変動の2つに分離する事が出来る。つまり、周波数情報である調波構造と、時間変動の両方を同時に扱う事が出来る手法である。また、 K 個のパターンのうち、 R ($R < K$) 個をピアノの共通成分でうまく固定することにより、ピアノ音を、 R 個の共通成分パターン、 $K - R$ 個の個性差パターンで表現することが可能と考えられる。つまり、この $K - R$ 個の成分が、入力音の個性差そのものと考えられる。

この個性差を分析することにより、音楽の演奏表現 [17] や音色変換 [18] などにおける指標としての利用や、物理モデルにおいて、個性差を含む音色を制御する場合の指標などに利用できることが期待される。

1.4 本論文の構成

本論文は6章で構成される。

第2章 本研究の分析対象であるピアノの発音原理について述べる。ピアノの種類による構造上の違い，そこから考えられるピアノ音に対する影響等について説明する。また，ピアノ独特の調波構造である非調波性 (インハーモニシティ) についても触れ，同様に予測される影響についても考察する。

第3章 ピアノ音の個体差を分析する手法について説明を行う。また，そこで用いる技術である NMF の基本原理についての紹介も行う。

第4章 第3章で示した分析手法を基に行った，ピアノ音の個体差分析実験について記載する。

第5章 ピアノの発音原理と，先ほどの分析結果を踏まえて，ピアノ音の個体差についての考察を行う。

第6章 本研究で得られた結果についてまとめる。

第2章 ピアノの発音原理

本章では、ピアノの発音原理について述べ、そこから考えられる個体差として表れる違いについて考察を行う。

2.1 ピアノの種類

ピアノは大きく分けて2種類存在する。グランドピアノとアップライトピアノである。グランドピアノは図 2.1 に示すように大型である。グランドピアノ内でも、奥行きが 270 cm もあるフルコンサートから、奥行き 150 cm 程度のミニグランドまで様々である。

アップライトピアノは図 2.2 に示す通り、グランドピアノに比べ小型であり、グランドピアノが置けないような部屋、及び家庭用ピアノとして普及している。

これら2種類のピアノの大きな違いは2つである。1つはピアノそのものの大きさ。もう一つは打弦機構である”アクション”の違いである。グランドピアノ、アップライトピアノそれぞれのアクションを図 2.3, 2.4 に示す。図から分かる通り、グランドピアノはハンマーが弦の下から打つものであるのに対し、アップライトは縦に張られた弦を横からハンマーが打つ機構になっている。グランドピアノでは、ハンマーが下から弦を叩くので、打弦後は重力によって打弦前の位置まで戻される。しかし、アップライトピアノでは、ハンマーが横から弦を叩いた後、元の位置に戻るために、バネの力を利用している。ヤマハ楽器では、この2つの機構の違いによって、グランドピアノはアップライトピアノの2倍の早さで打弦できるという [19]。つまり、アップライトピアノの方が打弦時の接触時間が長いということがわかる。弦との接触時間が長いということは、それだけ弦の振動を抑えているという事と等価であるため、アップライトピアノはグランドピアノよりも弦振動の減衰が早いと考えられる。

2.2 ピアノ音の発音機構

ピアノの発音機構の概形図を図 2.5 に示す。この発音機構は、大きく分けて4つの段階がある。(1) 演奏者がピアノの鍵盤を叩く。(2) その動きがアクションと呼ばれる打弦機構に伝わり、ハンマーが弦を叩く。(3) ハンマーの打弦により弦が振動をはじめ。(4) 弦の振動が駒を通して響板へ伝わる。これらのプロセスを通じて、ピアノの音が発せられる [2]。



図 2.1: グランドピアノ

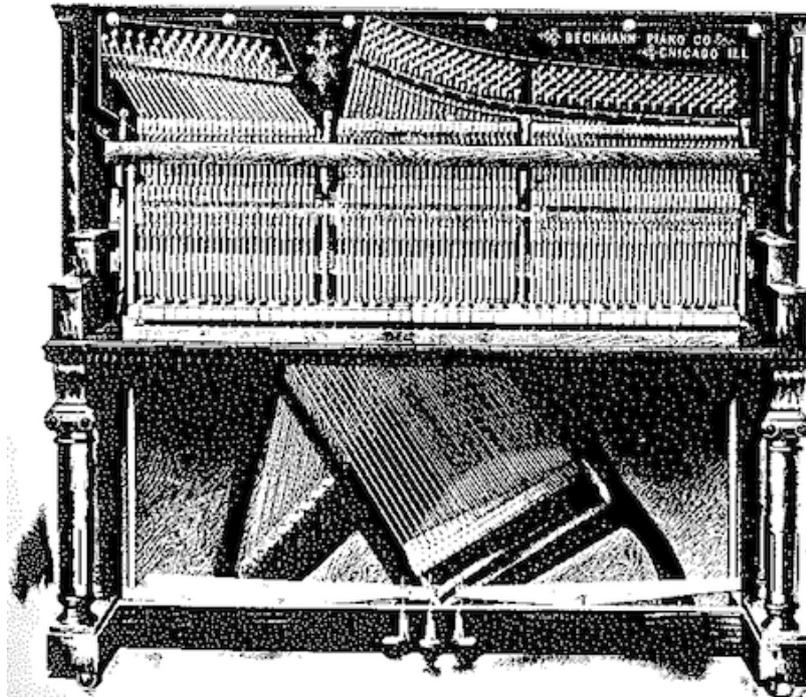


図 2.2: アップライトピアノ

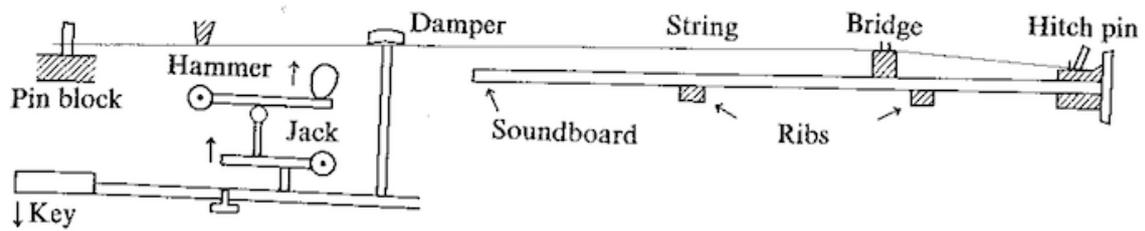


図 2.5: ピアノ音の発音機構

(1)の打鍵時には、(2)のハンマーによる打弦と同等の音量を発する、鍵盤が底にぶつかったときの雑音(下部雑音)が発生する。この下部雑音がピアノ音の音色知覚に影響を与えているという報告がある [20]。後述する弦の振動にも、ピアノ特有の非調波性があり、弦の違いによる音への影響も確認されている。駒及び響板についても音に影響があり、橋本ら [4] は、駒がある事により、打弦による基本周波数以下の低周波数成分が小さくなるという結果を報告している。また、響板については Fletcher らの著書 [7] にて、図 2.6 のように各周波数によって、響板の共鳴する箇所が異なる事が報告されている。つまり、これらの影響も個体差として表れてくる事が予想される。

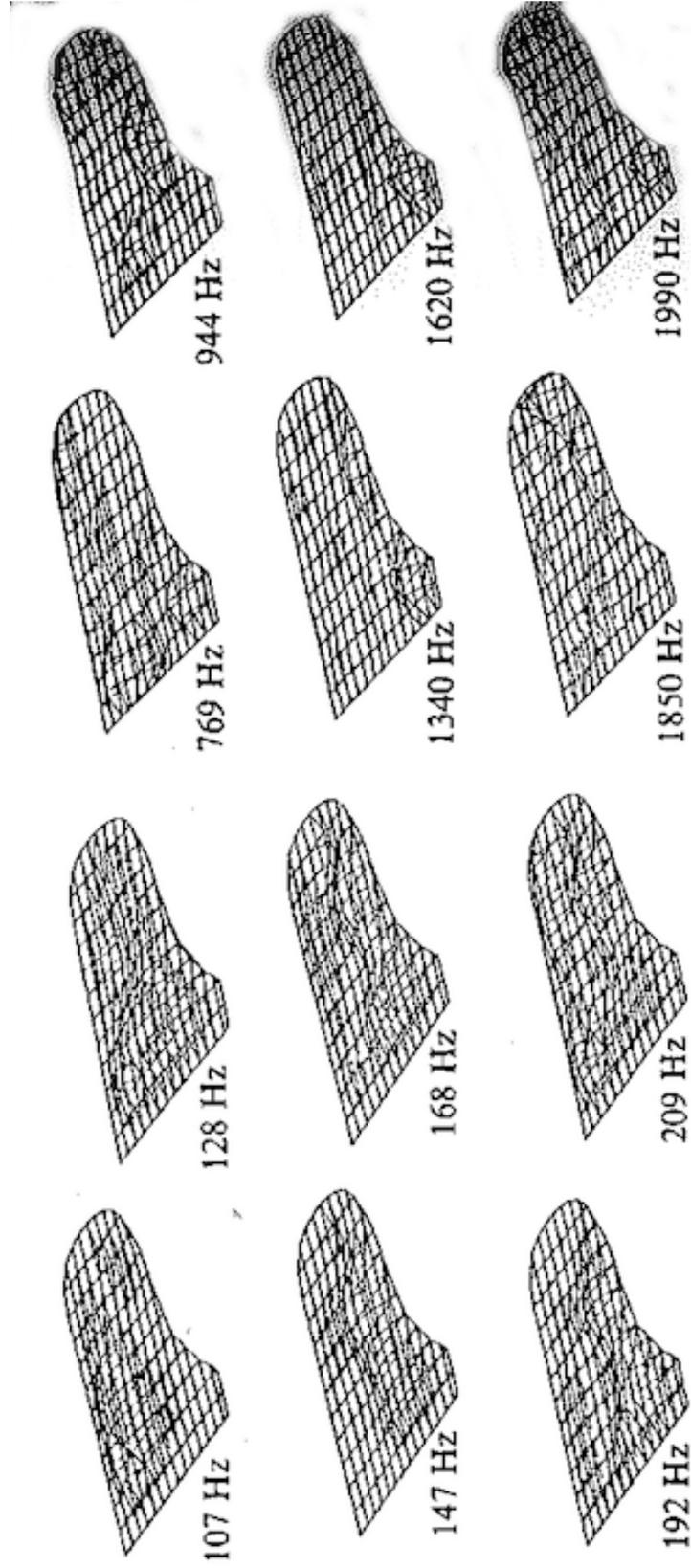


図 2.6: 響板の各周波数における相対的な強さのパターン

2.3 インハーモニシティ

ピアノの中・高音部では、各音高に対して2~3本の弦が張られている。それぞれの弦は微妙に調律がずらされている。3本弦の場合であれば、真ん中の弦を中心に、高音側の弦は真ん中の弦よりわずかに高く、低音側の弦は真ん中の弦よりわずかに低く調律されている。この微妙な調律の違いにより非調波性(インハーモニシティ)が生じる。

このインハーモニシティは、弦の自由振動時の基本周波数を F_0 とした場合、第 n 倍音 (n は正の整数) を f_n としたとき、式 (2.1) によって周波数の値を計算する事が出来る [7,21].

$$f_n = nF_0\sqrt{1+Bn^2} \quad (2.1)$$

弦の自由振動 F_0 は式 (2.2) で計算され、式 (2.1) の f_1 とほぼ同じ値である。

$$F_0 = \frac{1}{2L}\sqrt{\frac{T}{\mu}} \quad (2.2)$$

ただし、 L は弦の長さ、 T は張力、 μ は単位長さ辺りの質量(線密度)を表す。 B はインハーモニシティ係数と呼ばれ式 (2.3) によって計算される。

$$B = \frac{\pi^3 Ed^4}{64TL^2} \quad (2.3)$$

E はヤング率、 d は弦の直径、 T, L は先ほどと同じである。この B は一般に $[10^{-5}, 10^{-2}]$ の値をとる [22]。この B の値によって図 2.7 に示すように、高調波でも特に第 10 倍音以降の周波数の値が変わってくる。このインハーモニシティの違いは個体差として表れてくるはずである。

2.4 まとめ

本章では、ピアノの種類による構造の違い、並びにピアノの発音機構について説明を行った。ピアノの発音機構から個体差として考えられる要因は (1) 打鍵時のノイズ、(2) ハンマーが打弦時の打撃音、(3) 弦の振動、(4) 駒・響板の影響の4種類である。特に、弦の振動においてはインハーモニシティが見られるため、これらの要因毎に分離して個体差がどのようにあるのかを見ていく必要がある。

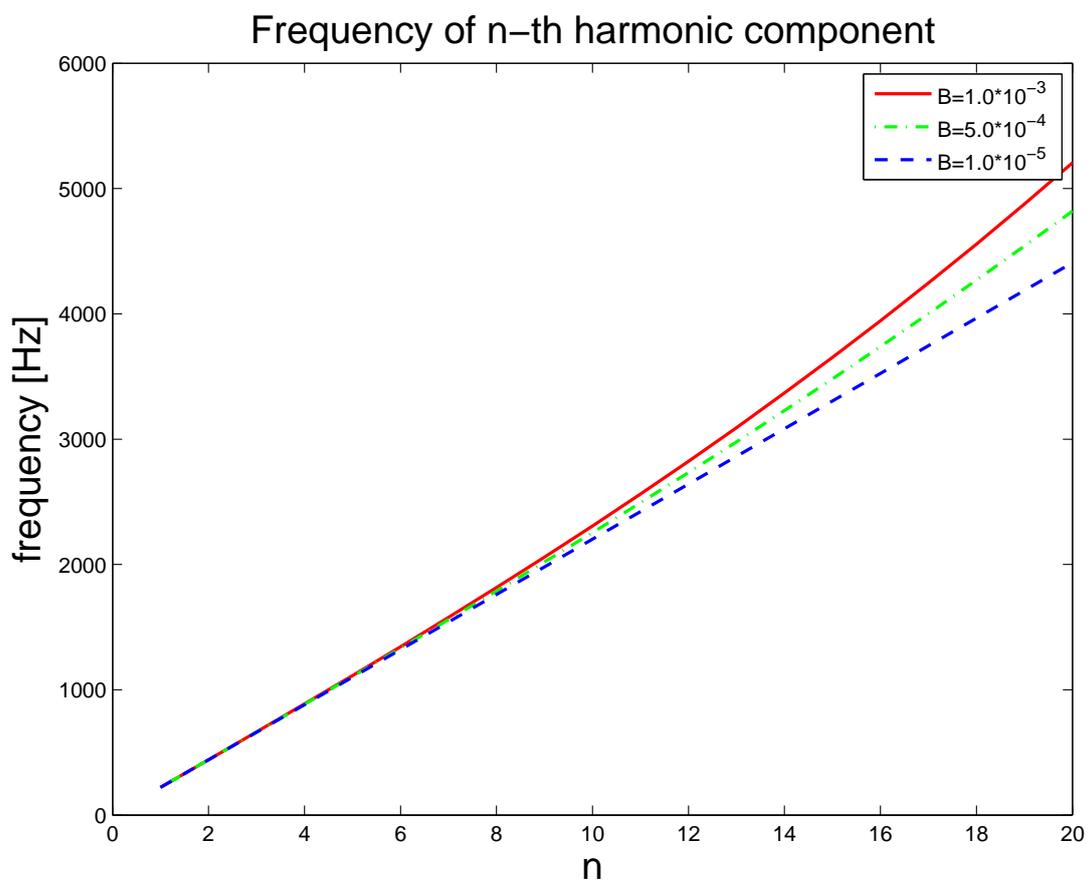


図 2.7: パラメータ B と第 n 倍音の周波数値の変化

第3章 ピアノ音の個体差分析手法

この章では、ピアノ音の個体差を分析する手続きについて述べる。各節にて個々の処理及び、用いる技術の基本原理について説明を行う。

3.1 ピアノ音の個体差分析の流れ

図 3.1 にピアノ音の個体差分析の流れを示す。NMF の入力として対数パワースペクトログラムを使うので、NMF に適応させるためにも様々な前処理を施してある。それぞれの処理については次節で説明する。

3.2 非負値行列因子分解

NMF とは、ある非負値行列 $Y (\in \mathbb{R}^{\Omega \times T})$ を異なる 2 つの非負値行列 $U (\in \mathbb{R}^{\Omega \times K})$ と $V (\in \mathbb{R}^{K \times T})$ の積に近似分解する手法であり、式 (3.1) によって定義される。

$$Y_{\omega,t} \approx \hat{Y}_{\omega,t} = \sum_k^K U_{\omega,k} V_{k,t} \quad (3.1)$$

ここで、 Ω は周波数ビン数、 T はフレーム数、 K は NMF の基底数を表す。また、 ω, t はそれぞれ周波数ビン番号、フレーム番号を示す。音響信号に NMF を適応する場合、パワー (振幅) スペクトログラムが行列 Y に相当する。行列 U は基底行列と呼ばれ、 U の各列ベクトルは調波構造を表す。また、行列 V はアクティベーション行列と呼ばれ、各行ベクトルは、 U の各調波構造の時間変動を表現する。

このように NMF は調波構造と、それぞれに対応する時間変動の両者を同時に見る事が出来る。



図 3.1: ピアノ音分析のブロック図

3.2.1 距離尺度

NMF は非負値行列 Y を2つの異なる非負値行列 U, V に分解する。単に分解するだけでは、解が定まらない。そこで、解が収束するようにある一定のルールに基づいて U, V の最適化を行う。そのルールとしては、ユークリッド距離 (式 (3.2)) や一般化 KL ダイバージェンス (式 (3.3)), 板倉齋藤距離 (式 (3.4)) [23, 24] などが主として用いられている。

$$D_{\text{Euc}}(x|y) = (x - y)^2 \quad (3.2)$$

$$D_{\text{KL}}(x|y) = (x - y) + y \log \frac{y}{x} \quad (3.3)$$

$$D_{\text{IS}}(x|y) = \frac{y}{x} - \log \frac{y}{x} - 1 \quad (3.4)$$

各距離尺度における、2つの変数の近さ度合いを表した図が図 3.2 である。

3.2.2 更新則

U, V の最適化は先ほどの距離尺度を基に、乗法更新則と言われる方法で値の更新を行う。ユークリッド距離での更新則は式 (3.6) のようになる。

$$U \leftarrow U \cdot \frac{YV^t}{UVV^t} \quad (3.5)$$

$$V \leftarrow V \cdot \frac{U^tY}{U^tUV} \quad (3.6)$$

$*$ は転置, \cdot はアダマール積を表す。導出方法としては補助関数法と呼ばれる、目的関数の上限関数を設計し、その上限関数を最小化することにより、目的関数を最小化する手法である [25, 26].

以下にユークリッド距離に置ける場合の更新則の導出方法を記述する。

$$\begin{aligned} D_{\text{Euc}}(Y|UV) &= \|Y - UV\|_F^2 \\ &= \sum_{\omega,t} |Y_{\omega,t} - \sum_k U_{\omega,k} V_{k,t}|^2 \\ &= \sum_{\omega,t} (|Y_{\omega,t}|^2 - 2|Y_{\omega,t}| \sum_k U_{\omega,k} V_{k,t} + |\sum_k U_{\omega,k} V_{k,t}|^2) \end{aligned} \quad (3.7)$$

この右辺の第3項に対して Jensen の不等式をたてる。Jensen の不等式とは、ある凸関数 $f(x_i)$ に対して

$$f\left(\sum_i \lambda_i x_i\right) \leq \sum_i \lambda_i f(x_i) \quad (3.8)$$

D(y|x) : Degree of proximity between x and y

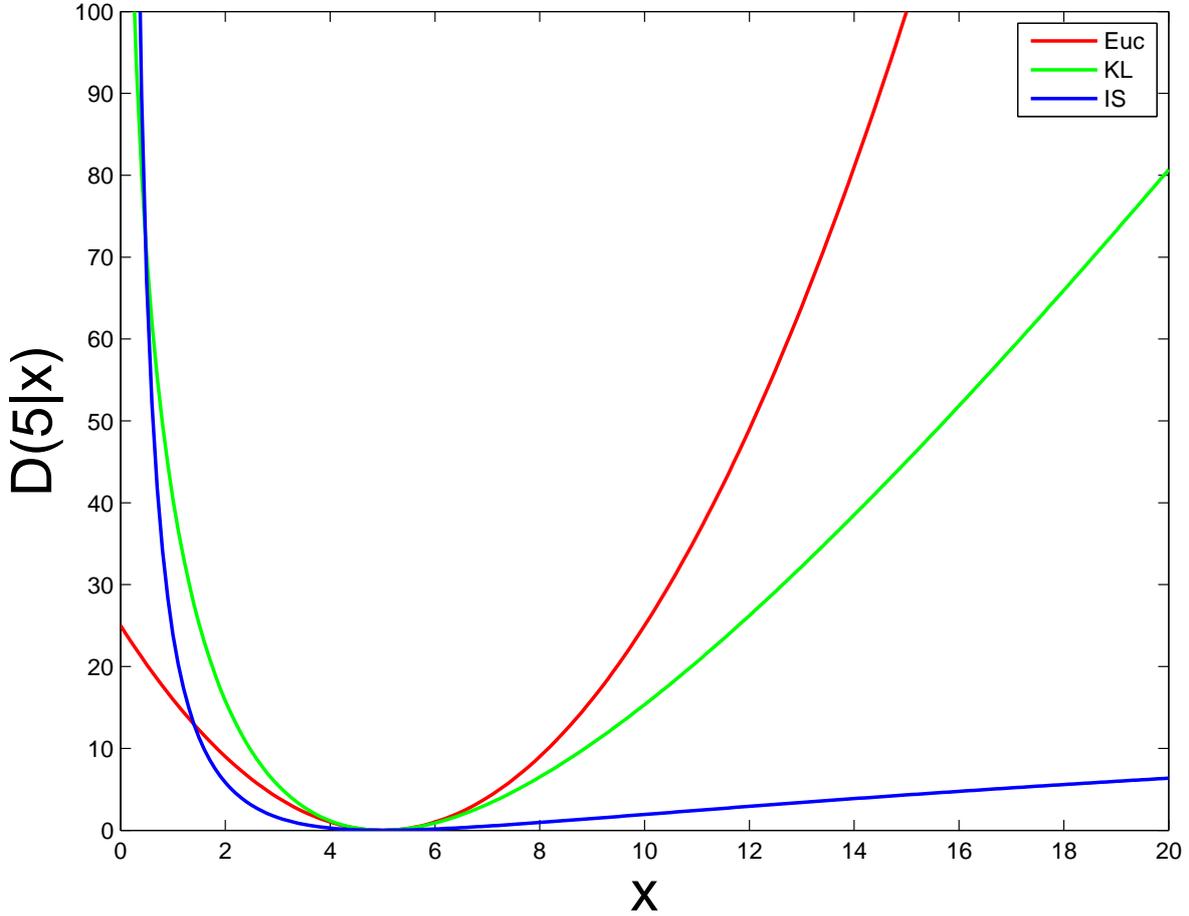


図 3.2: 各距離尺度における 2 変数の近さ度合い

という大小関係を示す不等式である。2 次関数であれば式 (3.9) のようになる。

$$\left(\sum_i x_i\right)^2 = \left(\sum_i \lambda_i \frac{x_i}{\lambda_i}\right)^2 \leq \sum_i \lambda_i \left(\frac{x_i}{\lambda_i}\right)^2 = \sum_i \frac{x_i^2}{\lambda_i} \quad (3.9)$$

つまり、第 3 項に対して Jensen の不等式を立てると式のようなになる。

$$\left|\sum_k U_{\omega,k} V_{k,t}\right|^2 \leq \sum_k \frac{|U_{\omega,k} V_{k,t}|^2}{\lambda_{k,\omega,t}} \quad (3.10)$$

よって、式 (3.7) は式 (3.11) のように書き換えられる。

$$D_{Euc}(Y|UV) = \sum_{\omega,t} (|Y_{\omega,t}|^2 - 2Y_{\omega,t} \sum_k U_{\omega,k} V_{k,t}) + \sum_k \frac{|U_{\omega,k} V_{k,t}|^2}{\lambda_{k,\omega,t}} \quad (3.11)$$

この式 (3.11) を $U_{\omega,k}$ と $V_{k,t}$ について解くと

$$U_{\omega,k} = \frac{\sum_t Y_{\omega,t} V_{k,t}}{\sum_t \frac{V_{k,t}^2}{\lambda_{k,\omega,t}}} \quad (3.12)$$

$$V_{\omega,k} = \frac{\sum_t Y_{\omega,t} U_{k,t}}{\sum_t \frac{U_{k,t}^2}{\lambda_{k,\omega,t}}} \quad (3.13)$$

となる。この λ は全ての λ_i 和が 1 となる係数であれば何でも良いので、式 (3.14) のようにした。

$$\lambda_{k,\omega,t} = \frac{U_{\omega,k} V_{k,t}}{\sum_{k'} U_{\omega,k'} V_{k',t}} \quad (3.14)$$

よって、式 (3.13) に式 (3.14) を代入すると、式 (3.16) のようになり、更新式が導出出来る。

$$U_{\omega,k} = \frac{\sum_t Y_{\omega,t} V_{k,t}}{\sum_t V_{k,t} \sum_{k'} U_{\omega,k'} V_{k',t}} \quad (3.15)$$

$$V_{\omega,k} = \frac{\sum_t Y_{\omega,t} U_{k,t}}{\sum_t U_{k,t} \sum_{k'} U_{\omega,k'} V_{k',t}} \quad (3.16)$$

U, V の更新に先立ち、各行列の初期値は乱数を用いる。この初期値を元に更新を行うので、NMF は初期値依存のアルゴリズムであるので、局所最適解に陥る危険性がある。

3.3 高域強調処理

本研究の予備実験より、図 3.3 に示す通り、分析対象の信号に対し、高域強調処理を施さない状態では、低周波数領域の値が大きく、高周波数領域の値は結果にほとんど影響しなかった

そこで、周波数による値の偏りを無くすために高域強調処理を行った。分析対象の全音源に対し、スペクトル傾斜を回帰直線近似で求めたところ、平均 6.6 dB/oct の傾きがあったため、この傾きを打ち消す高域通過フィルタとして式 (3.17) に示すフィルタを高域強調処理に用いた。

$$H(z) = 1 - 0.97z^{-1} \quad (3.17)$$

3.4 対数パワースペクトログラムを用いた非負値行列因子分解による分析

ピアノ音の個体差分析にあたり、NMF の入力信号 Y として、対数パワースペクトログラムを用いる。

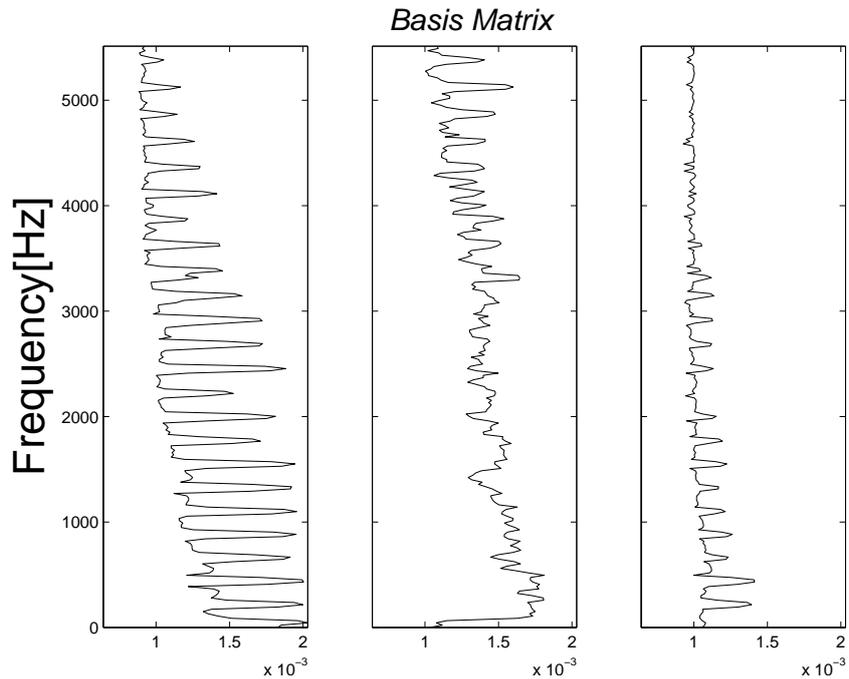


図 3.3: 高域強調無しでNMF による分析をしたときの基底行列

NMF は、各要因の和で表現する手法である。そのため、 X を要因 A, B, C, \dots の和で表現するにあたり、要因が並列的な関係なら式 (3.18) に示す通り、そのままの和で十分表現出来る。

$$X = A + B + C + \dots + N \quad (3.18)$$

しかし、各要因が従属的な関係なら、式 (3.19) に示す通り対数をとらない限り、各要因の和で表現する事が出来ない。

$$\begin{aligned} X &= \log A + \log B + \log C + \dots + \log N \\ &= \log(A \times B \times C \times D) \end{aligned} \quad (3.19)$$

ピアノの発音原理は従属的な流れであることから、音の要素も従属的なフィルタの影響を受けると考えた。そこで、従属的な要因を分離するために今回はパワースペクトログラムの対数をとった。

3.5 値のスケール変換

NMF は原理上、非負値しか扱う事が出来ない。従って、対数をとったものに対して、全て非負値になるようにスケール変換を行う必要がある。そこで、パワースペクトログラムの値が0 のところを、0 を除いた最小値で置き換えて対数をとる。これにより、値の範

囲が有限になるので，対数パワースペクトログラムの最小値を全体から減算することで，値を全て非負値にすることができる。

3.6 まとめ

これまで各節にて紹介した処理，すなわち，高域強調処理，対数パワースペクトログラムの非負値化，を用いてNMFにてピアノ音の個体差分析を行う．分析結果等については次章にて説明する．

第4章 非負値行列因子分解を用いたピアノ音の個体差分析

4.1 基底数 K の決定

4.1.1 実験目的

この実験では、NMF の基底数を決める事を目的に行った。NMF は基底数 K 個のパターンに分離、表現する手法である。しかし、基底数 K は一般的には決め打ちである。NMF の基底数を自動的に求めるアルゴリズムはいくつか提案されてはいる [27,28] もの、ある1つの楽器音を分析するための方法ではない。そのため、基底数を変化させて、分析に適切な基底数を決める必要がある。NMF の基底数を決定するにあたり、 K の値を1から5まで変化させて、それぞれの分離結果及び、信号対歪み比の値を用いて決めた。

4.1.2 実験条件

ピアノ音の個体差の分析にあたり、分析対象のピアノ音として、グランドピアノ音4つとアップライトピアノ音5つを用いた。グランドピアノ音のうち2種はRWC研究用音楽データベース:楽器音 (以下 RWC-DB とする) [29] からグランドピアノ音源2つ、MIDI Aligned Piano Sound (以下 MAPS-DB とする) [30] からアップライトピアノ音源1つを用いた。RWC-DB には、グランドピアノ音源が3つ収録されているが、うち一つが音が鳴り終わる前に切れてしまっているものがあったため、2つのみの使用した。また、MAPS-DB には、物理モデル音源やソフトウェア音源があるが、今回は実音源を対象としているので用いなかった。残りのグランドピアノ音2つとアップライトピアノ音4つは著者が収録したものをを用いる。

今後、便宜上 RWC-DB の2つの音源をそれぞれ RWC1, RWC3, MAPS-DB の音源を MAPS と称する。また、収録したグランドピアノ音2つを GP1, GP2, アップライトピアノ音4つを UP1, UP2, UP3, UP4 と名付ける。

分析対象のピアノ音の音高はいずれも A3 (220 Hz) とし、音量はメゾフォルテとした。全ての音源で量子化ビット数は16, サンプリング周波数は44.1 kHz である。パワースペクトログラムは短時間フーリエ変換 (Short Time Fourier Transform; STFT) を用い、分析窓はハニング窓を使用して算出した。分析窓幅は2048点 (約50 ms), 窓シフト幅は128点 (約3 ms) である。

分析窓幅によって、周波数解像度が変化することは周知の事実である。今回窓幅を 2048 点としたことには理由がある。図 4.1 4.3 に分析窓幅の異なる場合の NMF での分析結果を示す。

図 4.1 は分析窓幅が 512 点で分析したときの結果であり、図 4.2 は、2048 点で分析したときの結果、そして、図 4.3 は 4096 点で分析を行ったときの結果である。また、各図において (a) は基底行列を、(b) はアクティベーション行列を表す。

これらから分かる事として、512 点では、基底ベクトルである周波数情報がわからず、調波構造らしい構造すら見えてこない。調波構造は 2048 点のものと、4096 点のものは確認される。このとき、調波構造の概形についてはほとんど変わらない事から、4096 点のものではなく、2048 点のものでも表現出来ていると考えられる。また、4096 点でのアクティベーションベクトルについて着目すると、第 3 ベクトルの減衰部分がピーク付近の値がほぼ同じ値になっており、2048 点での結果のような急峻なピークが見られなかった。4096 点以上での分析だと、時間変動がぼやてけしまう結果が生じたため、周波数解像度と時間変動のバランスをとった 2048 点での STFT を行った。

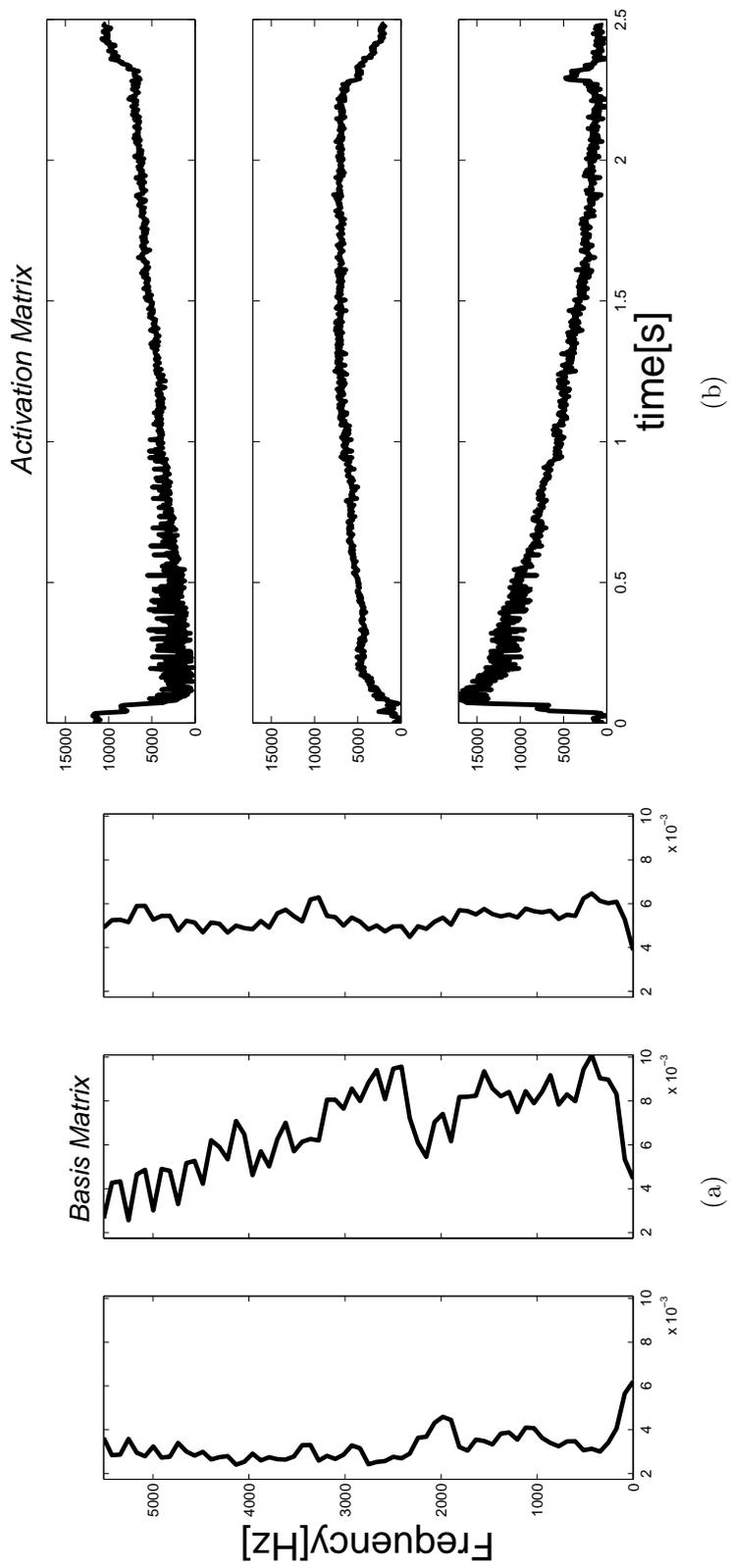


図 4.1: 分析窓幅 512 点での分析結果

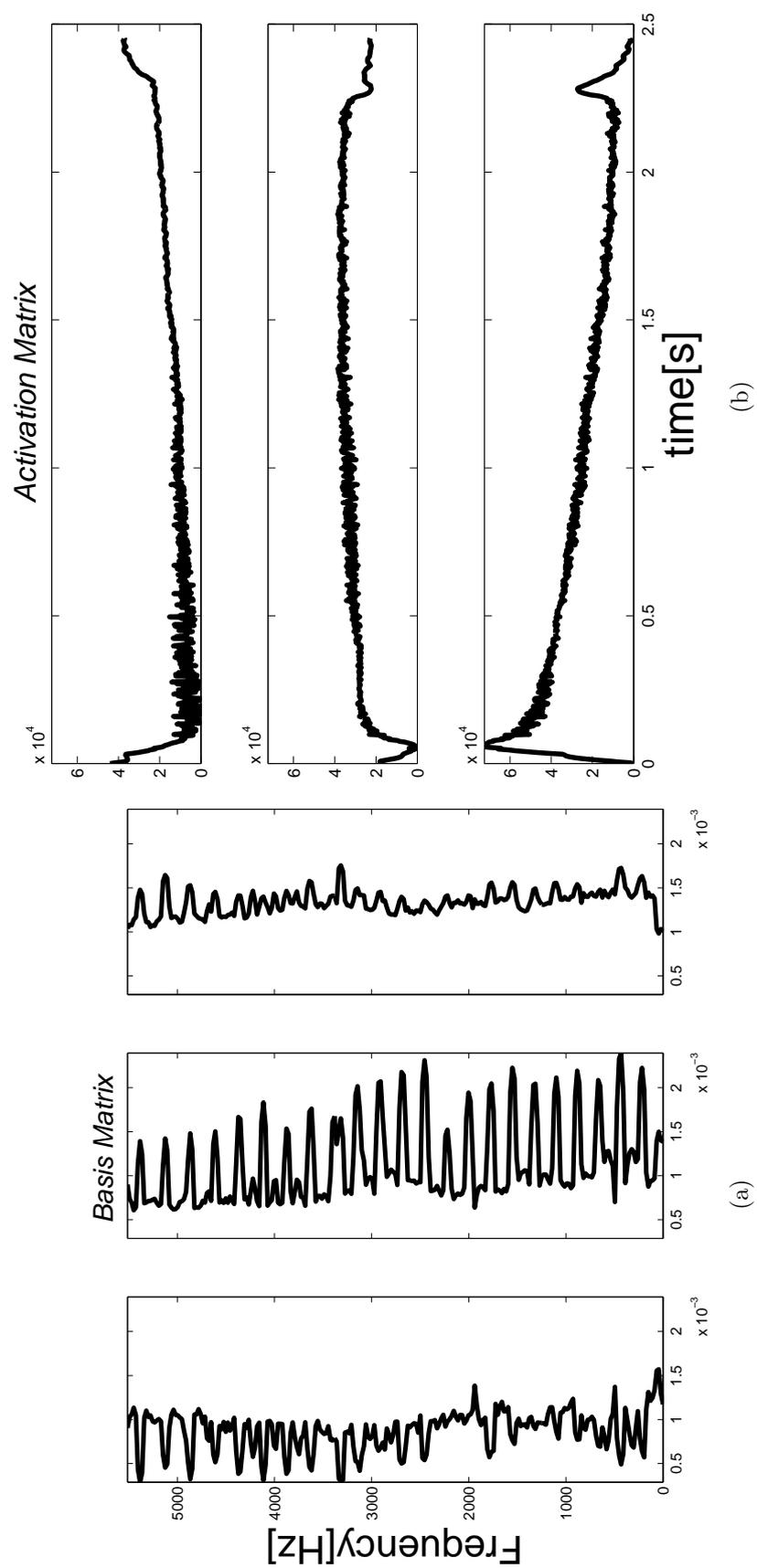


図 4.2: 分析窓幅 2048 点での分析結果

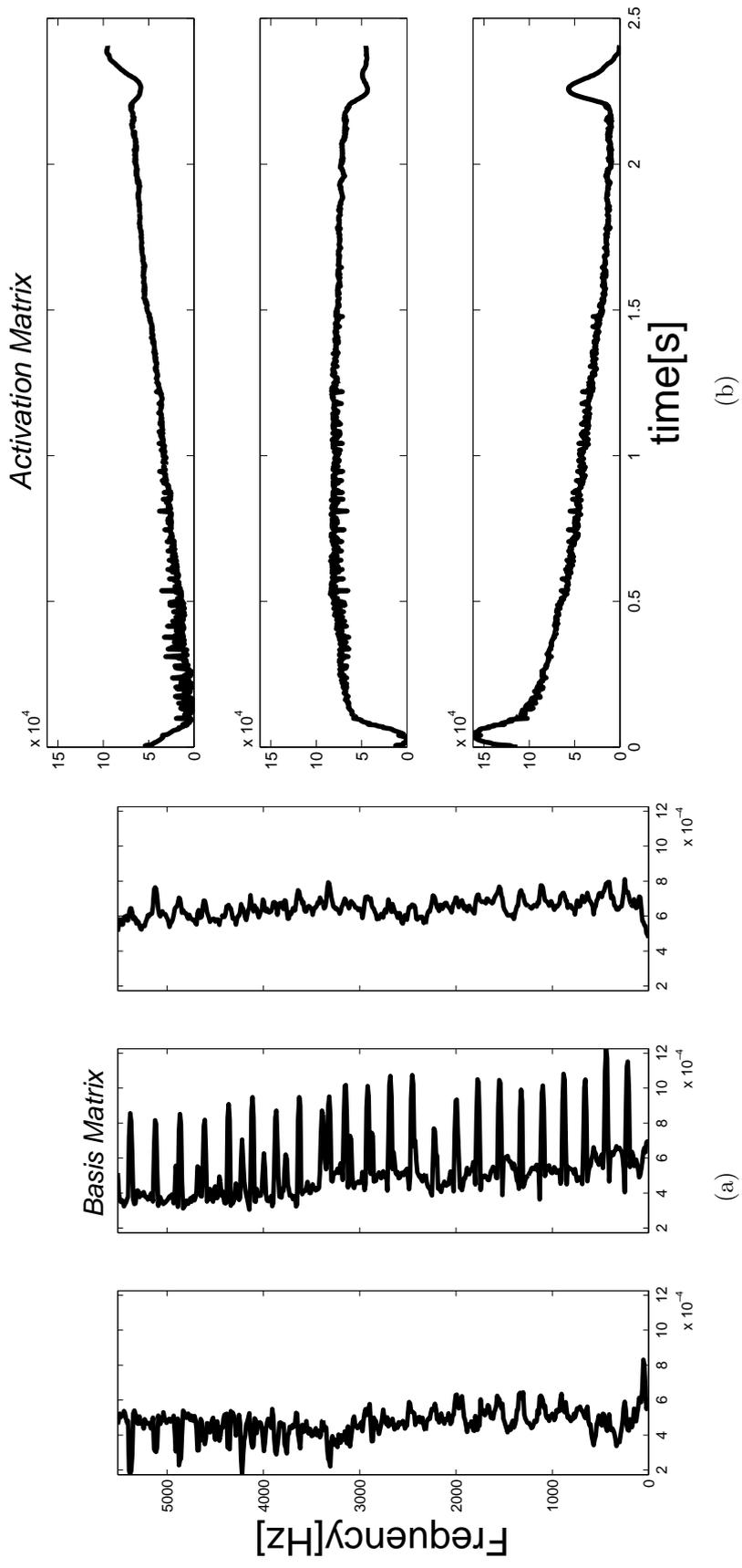


図 4.3: 分析窓幅 4096 点での分析結果

表 4.1: 基底数 $K = 3$ と $K = 4$ の各基底ベクトル間での相関係数

	1st (K=4)	2nd (K=4)	3rd (K=4)	4th (K=4)
1st (K=3)	0.89	0.62	-0.66	0.12
2nd (K=3)	-0.52	0.50	0.72	-0.33
3rd (K=3)	0.53	-0.29	-0.51	0.98

4.1.3 分析結果と考察

各基底数における分析結果の一例を図 4.5～4.9 に示す (他の結果は付録参照)。基底数 $K = 1$ の分析結果 (図 4.5) では、基底行列に調波構造が、アクティベーションに波形の概形が表れていることを確認出来た。

基底数 $K = 2$ の分析結果 (図 4.6) では、全体周波数領域を表す第 1 ベクトルと 4 kHz 以下の成分を表す第 2 ベクトルに分かれていた。第 2 アクティベーションベクトルが、徐々に値が増えている事、また、第 1 ベクトルが減衰していることから、高周波数成分が早く減衰するが、低周波数領域は長く残るとい特徴がこの 2 つのベクトルの動きから考えられる。 $K = 3$ の分析結果 (図 4.7) に着目すると、アクティベーションベクトルが定常部、音の立ち上がり部、減衰部の 3 つに分かれていた。特に、調波構造を強く示しているのが第 1 ベクトルである定常部であり、第 3 ベクトルである減衰部は $K = 1, 2$ ほど強く示してはいなかった。ピアノの発音機構は (1) 打鍵、(2) ハンマーによる打弦、(3) 弦の振動、(4) 響板という形で伝わり、発音する。このシステムと対応させると、第 1 ベクトルは調波構造を表しかつ、定常的なアクティベーションなので、(4) の響板による影響があてはまると思われる。第 2 ベクトルは最初に強い値を持つが、すぐにアクティベーションが下がるところ、また、基底ベクトルはほぼフラットな概形だが、基本周波数 (220Hz) 以下の低域成分が強く表れている事から、打鍵時の下部雑音並びに、打弦時の衝突音が混合して表れていると考えられる。

$K = 4$ の結果 (図 4.8) を $K = 3$ の結果と対比させると、定常部が $K = 4$ のときの結果の第 1 ベクトルと第 2 ベクトルに分離されていると考えられる。実際に各ベクトル間での相関を計算したところ、図 4.1 に示すように、 $K = 4$ の第 1 ベクトル、第 2 ベクトルは共に $K = 3$ の第 1 ベクトルと最も相関が高かった。

また、基底数 $K = 5$ の結果 (図 4.9) にも注目すると、第 1、第 2 ベクトルは 2 つで定常部を表現しようとしていて、第 3、第 4 ベクトルが立ち上がり部を表現しているとアクティベーションから推測される。先ほど同様、相関係数を計算すると、5 つのベクトルのうち、第 1 と第 2 ベクトルは、基底数 $K = 3$ の第 1 ベクトルと相関が高く、第 3 と第 4 ベクトルは基底数 $K = 3$ の第 2 ベクトルと相関が高く、第 5 ベクトルは基底数 $K = 3$ の第 3 ベクトルと相関が高い (図 4.2 参照)。

これらの傾向から、基底数を 4 以上にすると、新たなベクトルは定常部、立ち上がり部、減衰部のどれか軸に集約されると考えられる。したがって、ピアノ音を表現するにあたり

表 4.2: 基底数 $K = 3$ と $K = 5$ の各基底ベクトル間での相関係数

	1st(K=5)	2nd (K=5)	3rd (K=5)	4th (K=5)	5th (K=5)
1st (K=3)	0.78	0.51	0.38	-0.77	0.35
2nd (K=3)	-0.42	0.08	0.57	0.45	-0.42
3rd (K=3)	0.62	-0.34	-0.49	0.15	0.96

基底数が3つあれば基本的な性質は表現出来ると思われる。

定性的な話は出来たが、定量的なことについては評価できていない。そこで基底数 K の値を決定するためにも、分離されて出来る基底行列 U とアクティベーション行列 V の積で再合成される対数スペクトログラム \hat{Y} と元の対数スペクトログラム Y との歪み率 (Signal to distortion ratio; SDR) を計算することにより定量的に評価する。計算式は式 (4.2) に示すものである。

$$SDR = 10 \log 10 \frac{S}{S - \hat{S}} dB \quad (4.1)$$

(S は原信号, \hat{S} は目的信号を表す)

この式を用いて全音源に対して SDR を計算し、それぞれの基底数ごとの平均値をプロットしたものが図 4.4 である。SDR は値が高いほど、歪みが少ない事を表す。この結果より、基底数 $K = 3$ までは SDR の値が直線的に上昇しているが、 $K = 4$ 以降では SDR の上昇幅が小さくなり、約 11 dB でほとんど横ばいになっていた。

これらのことを踏まえると、 $K \leq 3$ では、ピアノの発音機構の要因を十分に分離出来なく、歪みも大きいという傾向があり、 $4 \leq K$ では冗長な分離結果や、発音機構と対応していない分離結果などが見られた。また、基底数を増やしても SDR の値の上昇幅は上がらず、むしろ SDR の値そのものは横ばいという傾向が見られた。つまり、ピアノ音を表現するにあたり、基底数 K の値は 3 が適していると考えられる。

Relationship between number of bases and SDR

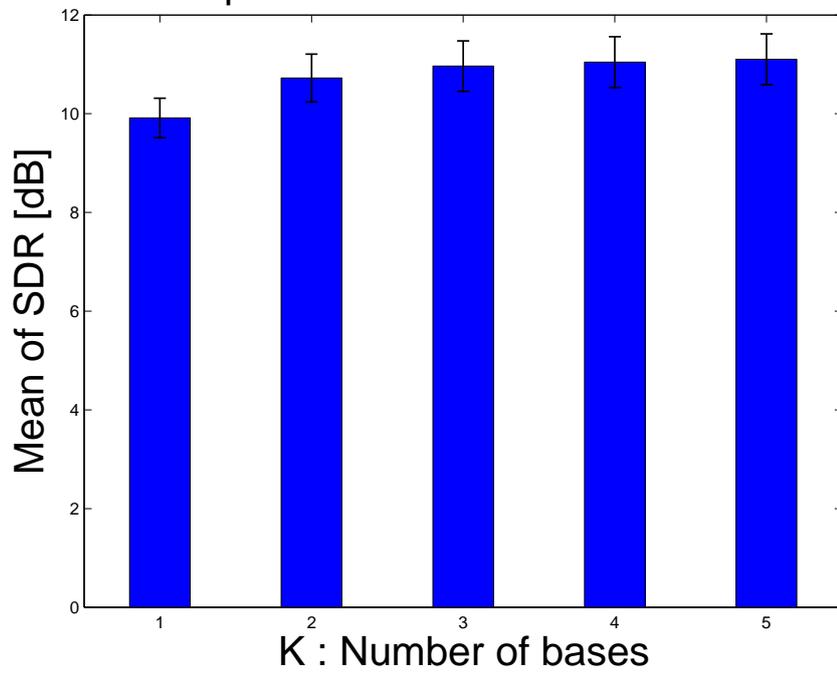


図 4.4: 基底数と SDR の関係

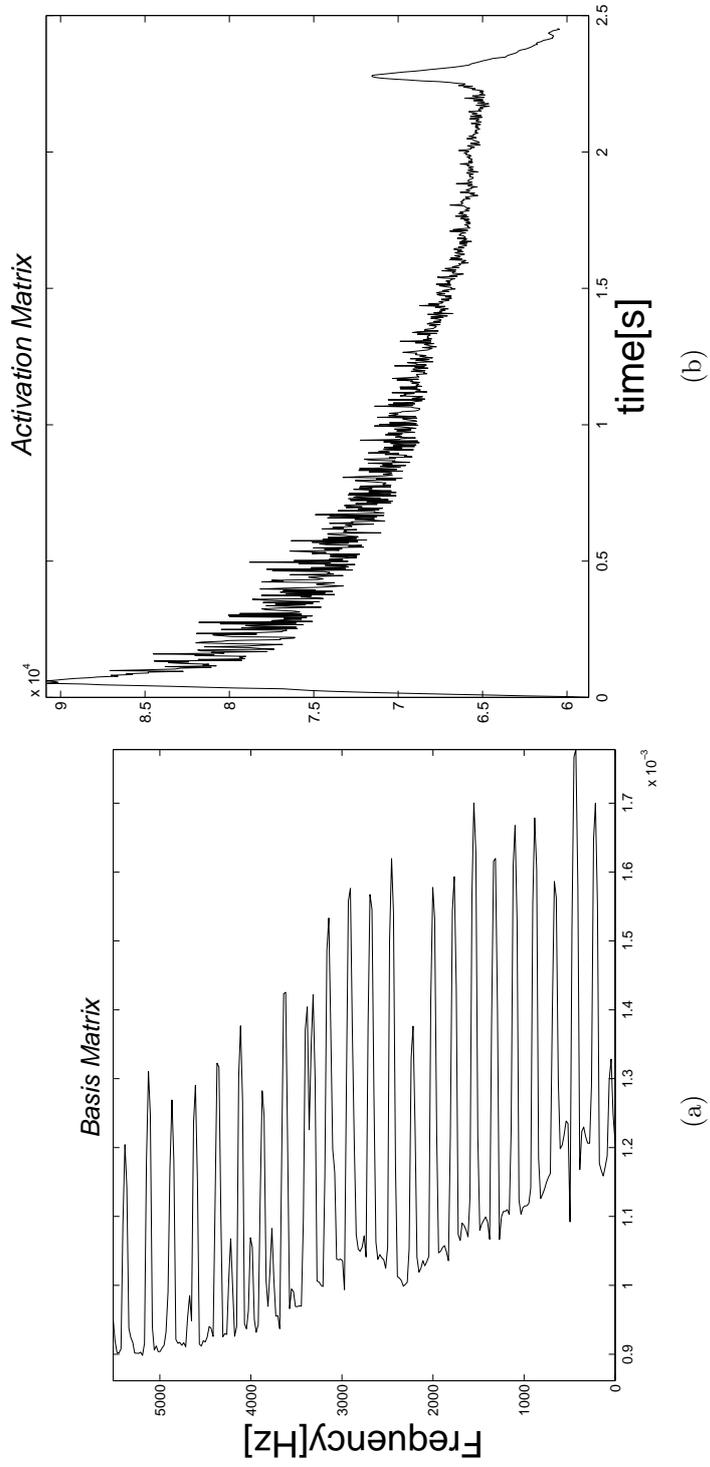


図 4.5: $K = 1$ の場合の分離結果 (RWCI)

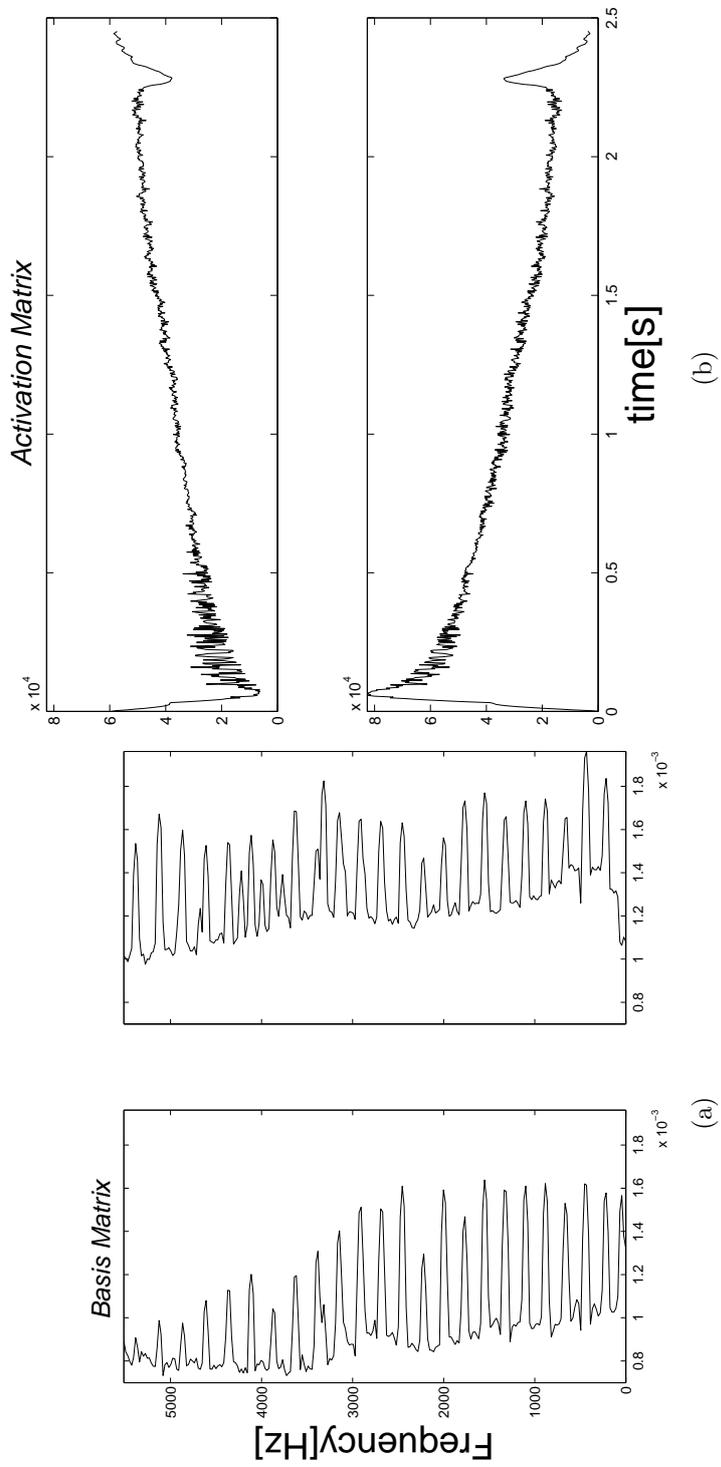


図 4.6: $K = 2$ の場合の分離結果 (RWCI)

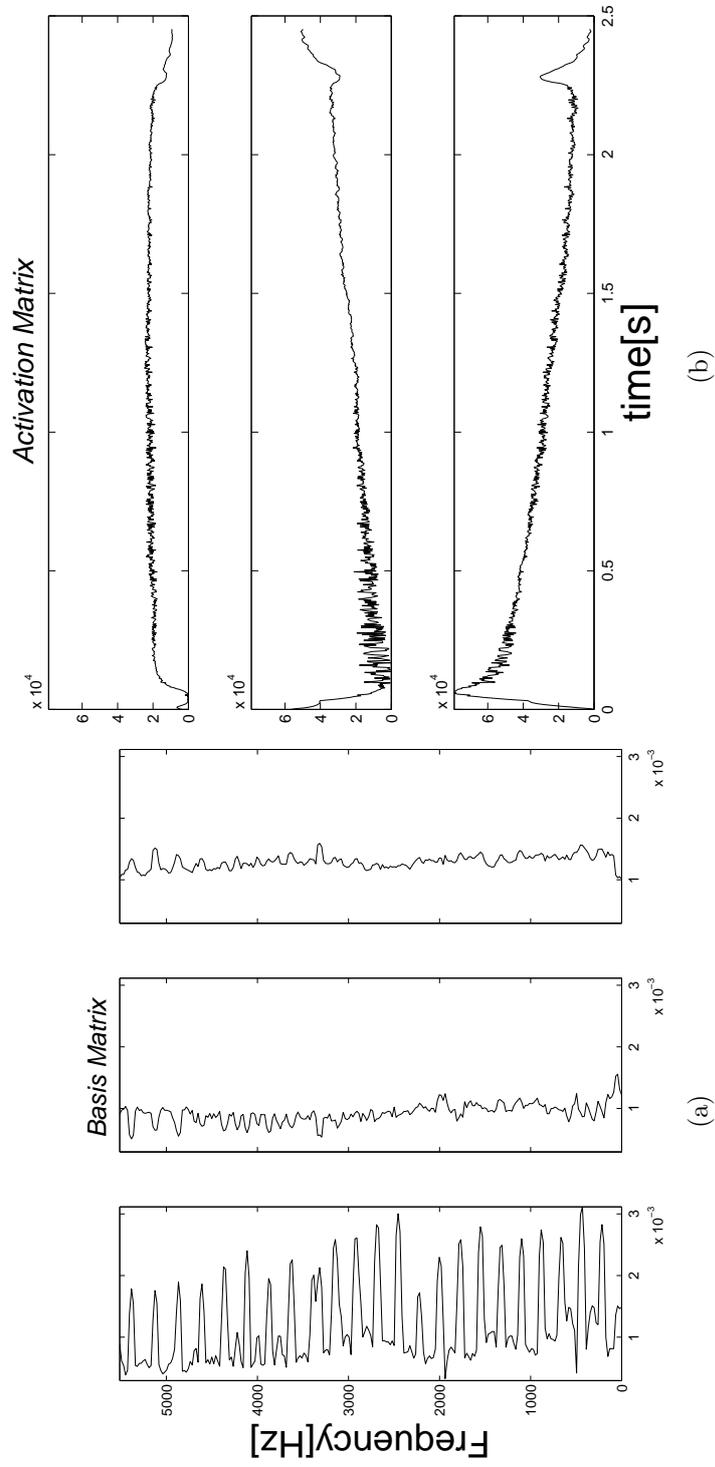


図 4.7: $K = 3$ の場合の分離結果 (RWCI)

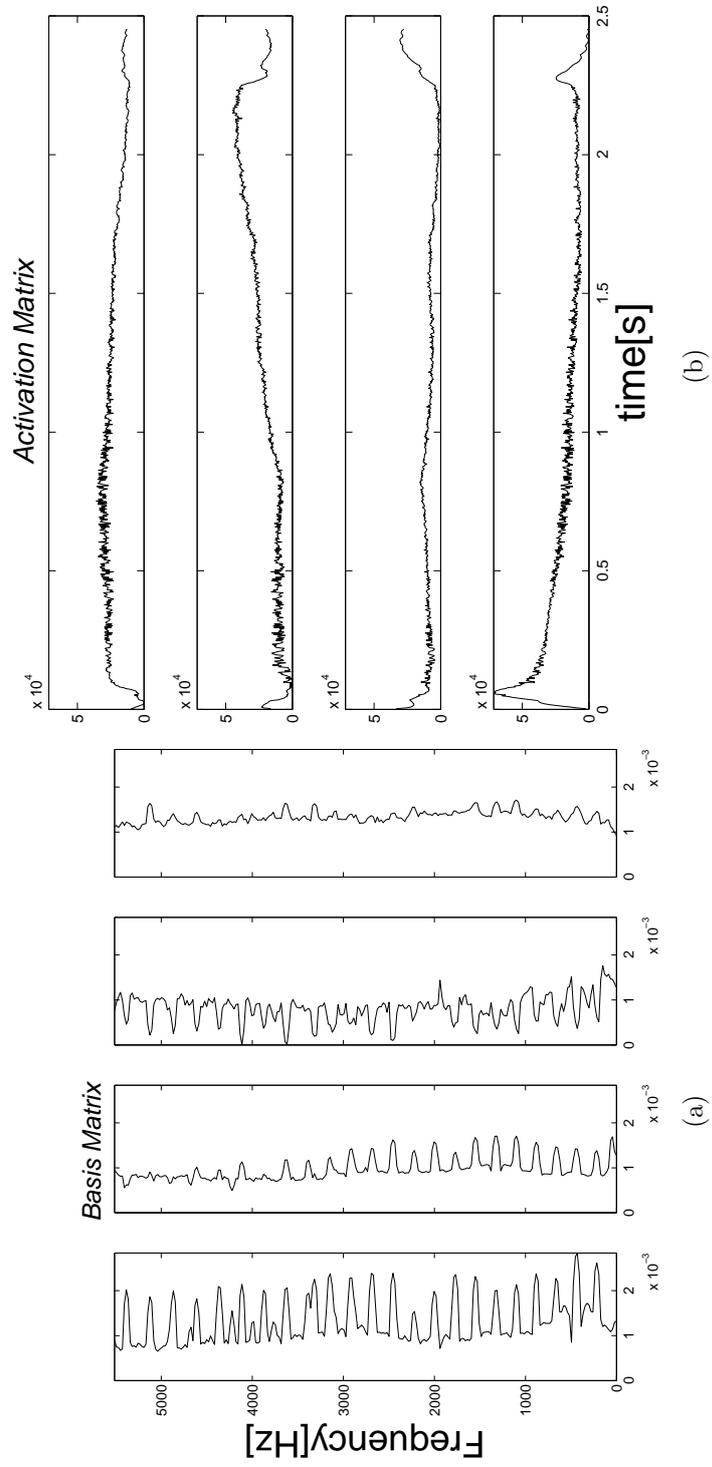


図 4.8: $K = 4$ の場合の分離結果 (RWCI)

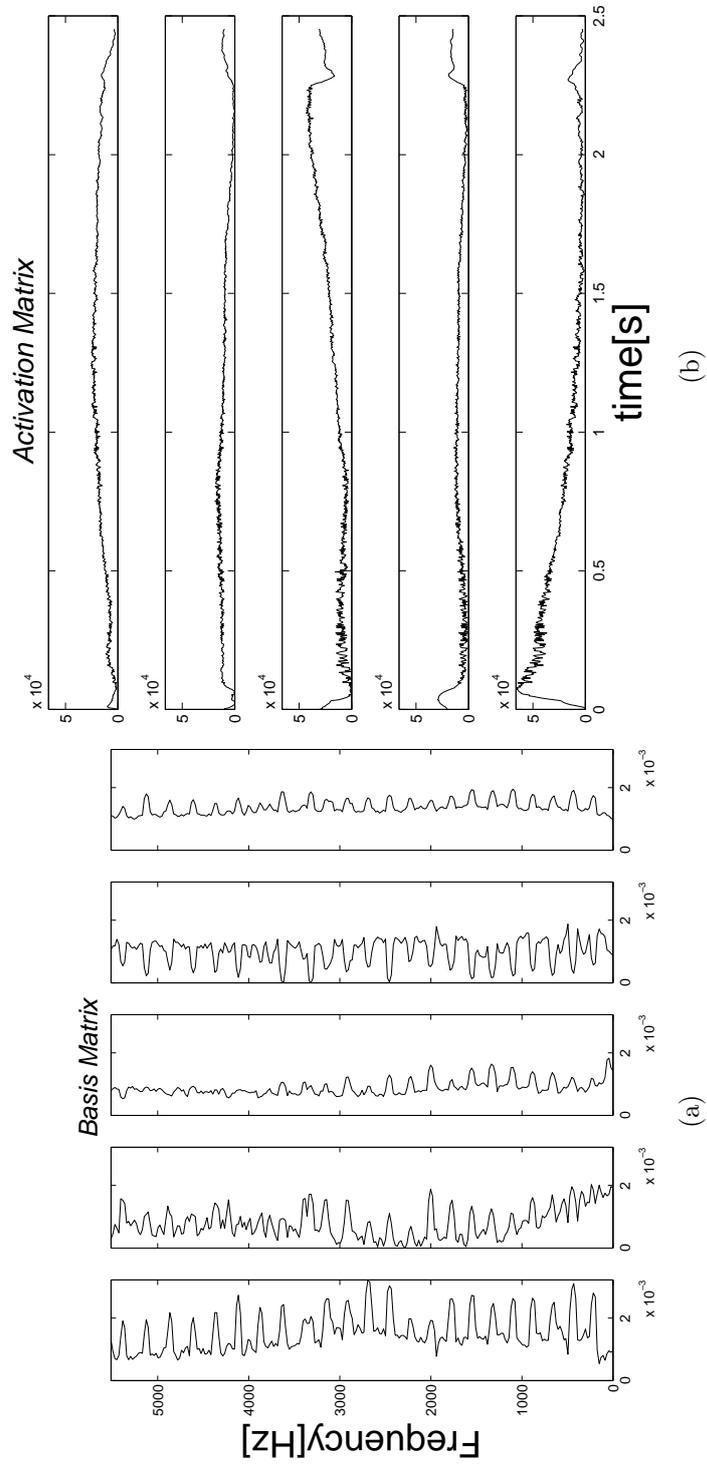


図 4.9: $K = 5$ の場合の分離結果 (RWCI)

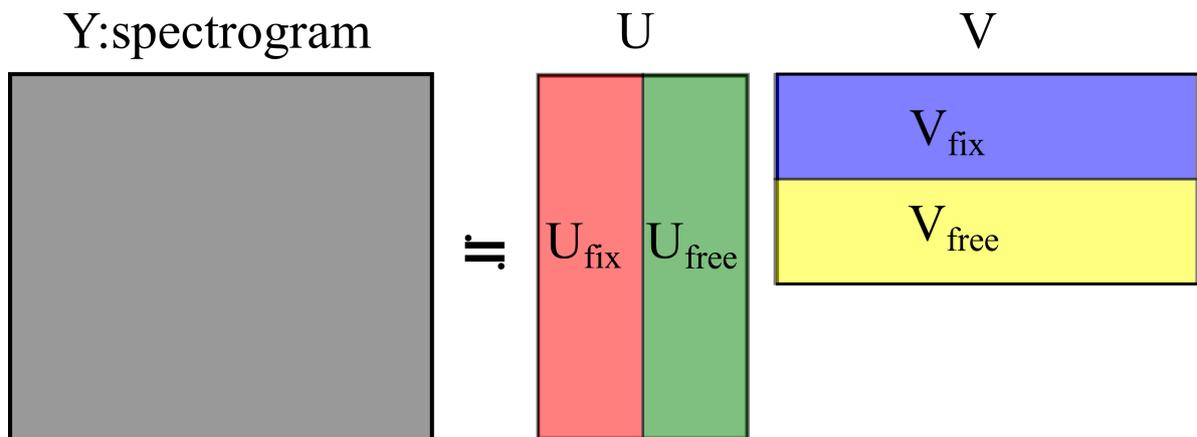


図 4.10: 基底固定を用いる NMF の概念図

4.2 基底行列の部分固定を用いた非負値行列因子分解によるピアノ音の個体差分析

4.2.1 実験目的

先ほどの分析で、基底数を3にすることにより、定常部、立ち上がり部、減衰部の3つに基底ベクトルが分類された。以降、定常部を表す第1ベクトルを定常部ベクトル、音の立ち上りを示す第2ベクトルをアタックベクトル、減衰を示す第3ベクトルを減衰部ベクトルを名付ける。

この基底数 $K = 3$ の結果では、ピアノの発音機構と関連のある結果が得られたが、そこから個体差が具体的にどのようなものなのかはわからなかった。そこで、個体差をより明確に抽出するために、基底行列の一部を固定して、再度 NMF で分析を行う。

NMF 入力非負値行列を基底行列とアクティベーション行列に近似分解する。このとき、得られる基底行列の一部(図 4.10 中の U_{fix})を予め別の成分で固定することにより、その固定した成分に対応するアクティベーションベクトル(図 4.10 中の V_{fix})に、固定した成分の時間変動が得られる。また、非固定部(図 4.10 中の U_{free})は NMF の更新により、得られ、固定した成分(U_{fix})では表現出来ない成分が表れるはずである。また、この非固定部に対応した時間変動(V_{free})も同様に NMF の更新で得られる。つまり、固定する成分として、ピアノ音の共通成分を用いれば、固定部に対応したアクティベーションベクトルに、共通成分の時間変動が、非固定部の基底行列には、その入力音特有の成分、つまり、入力音の個体差が表れてくるはずである。

4.2.2 実験条件

先ほどの結果より，3つの基底ベクトル，すなわち定常ベクトル，アタックベクトル，減衰ベクトルの3種類でピアノ音を表現したので，それぞれにおける共通成分と個体差があると仮定した。

共通成分であれば，一般的には平均値を用いる事が多い。しかし，NMFは加法による表現でかつ非負値のため，減算が一切行えない。つまり，共通成分に個体差を表すために不必要な成分が混入していた場合，それを取り除く事が出来ない。平均値を共通成分として用いると，このような事態が起こる可能性がある。そこで共通成分として，グランドピアノ音ならそれぞれのグランドピアノ音の基底ベクトルの最小値，アップライトピアノ音ならそれぞれのアップライトピアノ音の基底ベクトルの最小値を用いた。この実験では先ほどの仮定より基底数 K は6とした。

4.2.3 分析結果と考察

図 4.11 にグランドピアノ音の共通成分として用いた基底行列，図 4.12 にアップライトピアノ音の共通成分として用いた基底行列を示す。図 4.11 の第1ベクトルである定常部ベクトルは調波構造がよく見られるが，図 4.12 の定常部ベクトルは2 kHz 以上から調波構造のようなものが散見される。しかし，2 kHz 以下では局所的なピークはあるものの，いわゆる調波構造は観測されなかった。

第2ベクトルであるアタックベクトルに着目すると，図 4.11 では，細かい変動こそあるが，ほとんどフラットな形状をしている。一方図 4.12 では，定常部ベクトルほど激しく変動はしてないが，かなり定常部ベクトルと似通った特徴が表れていた。

減衰部ベクトルでは，図 4.11 では，全体的にフラットな形状をしているが，1~2 kHz, 4 kHz 以上の部分でピークが散見された。ところが図 4.12 では，2~4 kHz あたりでは調波構造を反転させたような谷の構造が観測された。また，500 Hz 以下の低周波数領域のパワーが強いのも特徴である。

そして，図 4.13, 4.14 にグランドピアノ音の分析結果の一部を，図 4.15, 4.16, 4.17 にアップライトピアノ音の分析結果の一部を示す。各図における (a),(b) はそれぞれ基底行列，アクティベーション行列である。また，図 4.13(a), 4.14 図 (a) の第1ベクトルから第3ベクトルが，図 4.11 に示した通り，共通成分として固定した基底行列である。同様に，図 4.15(a), 4.16(a), 4.17(a) の第1ベクトルから第3ベクトルが図 4.12 に示した，共通成分として固定した基底行列である。

分析結果の非固定部 (第4~第6ベクトル) のほとんどが，定常部ベクトル，アタックベクトル，減衰部ベクトルの3種類に分かれた。固定部 (第1~第3ベクトル) が共通成分であるので，非固定部には各ベクトルの個体差が表れていると考えられる。特に，第4ベクトルである定常部ベクトルでは，ごく一部にしかピークを持たないものや，中域 (2~3 kHz) に大きなピークを持つもの，そして，低域 (1 kHz 以下) と中域両方にピークを持つものなど，音源によって大きく異なっていた。

また，減衰部ベクトルでは，共通成分の第3ベクトルと，個体差の第6ベクトルでピアノ音の2段階減衰が表現されるが，最初の急峻な減衰が第3ベクトル，後半の緩やかな減衰が第6ベクトルと，2つに分かれている．このことから，緩やかな減衰が個体差として大きく影響している事が考えられる．

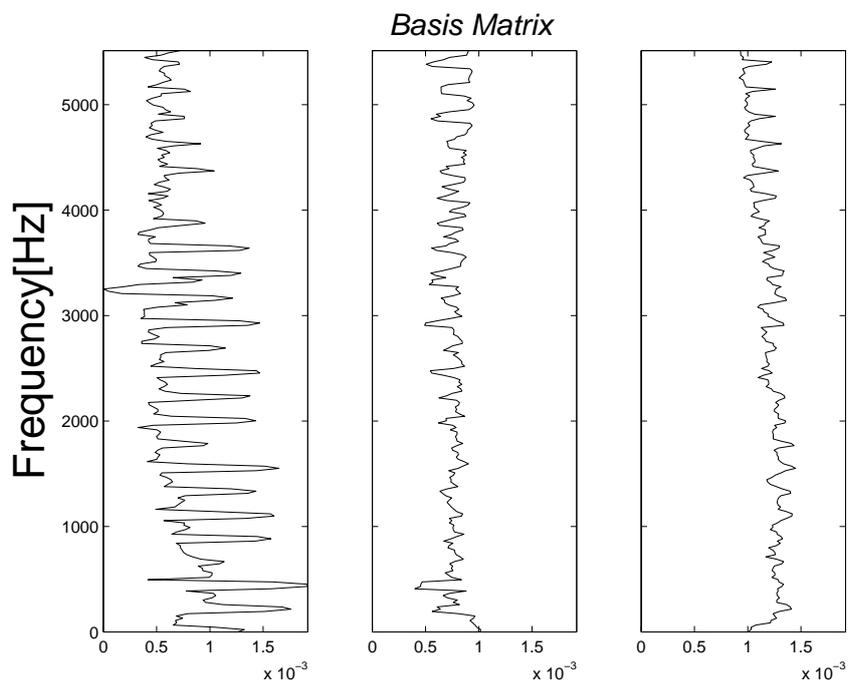


図 4.11: グランドピアノ音の共通成分

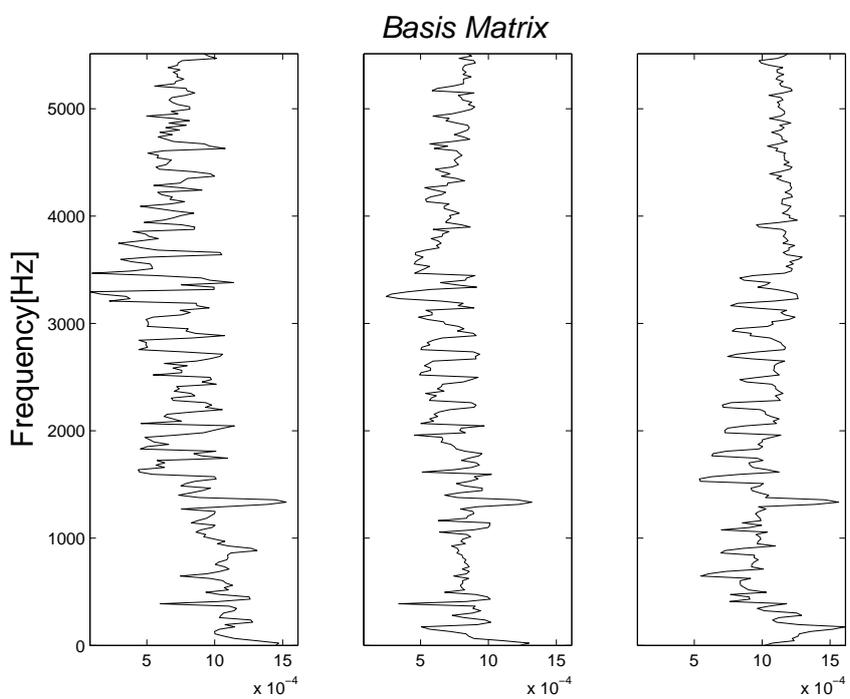


図 4.12: アップライトピアノ音の共通成分

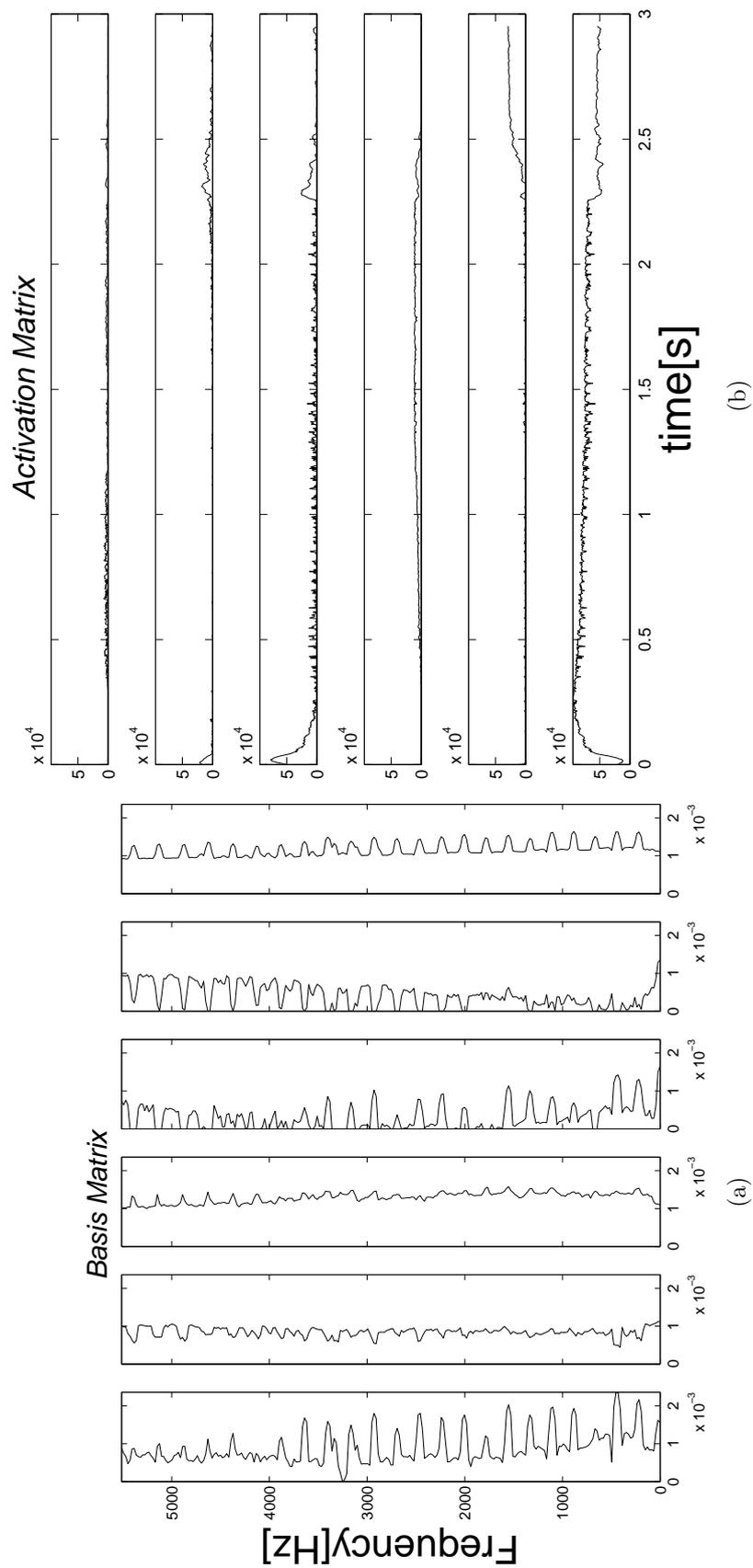


図 4.13: グランドピアノノ音の分析結果 (GP2)

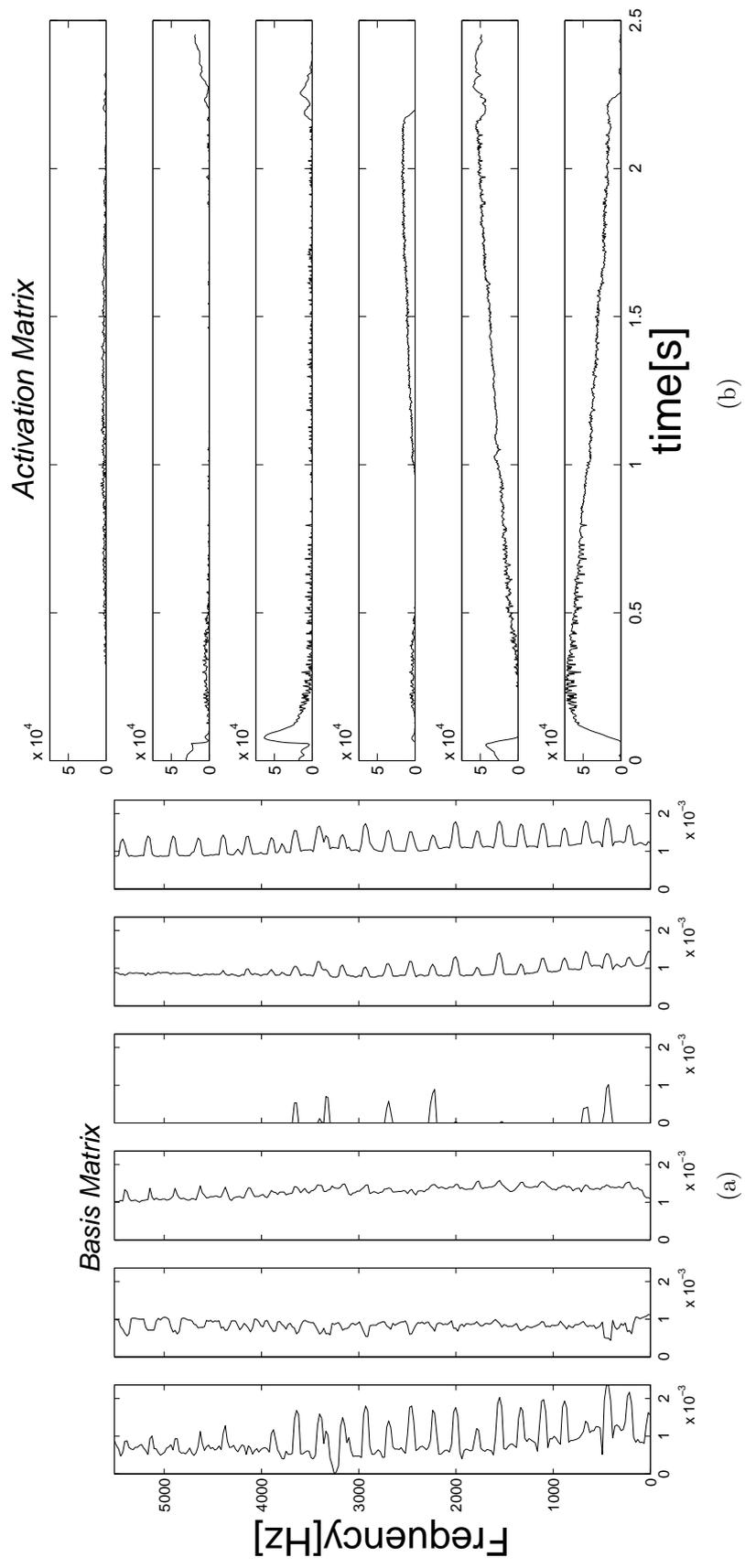


図 4.14: グランドピアノ音の分析結果 (RWC3)

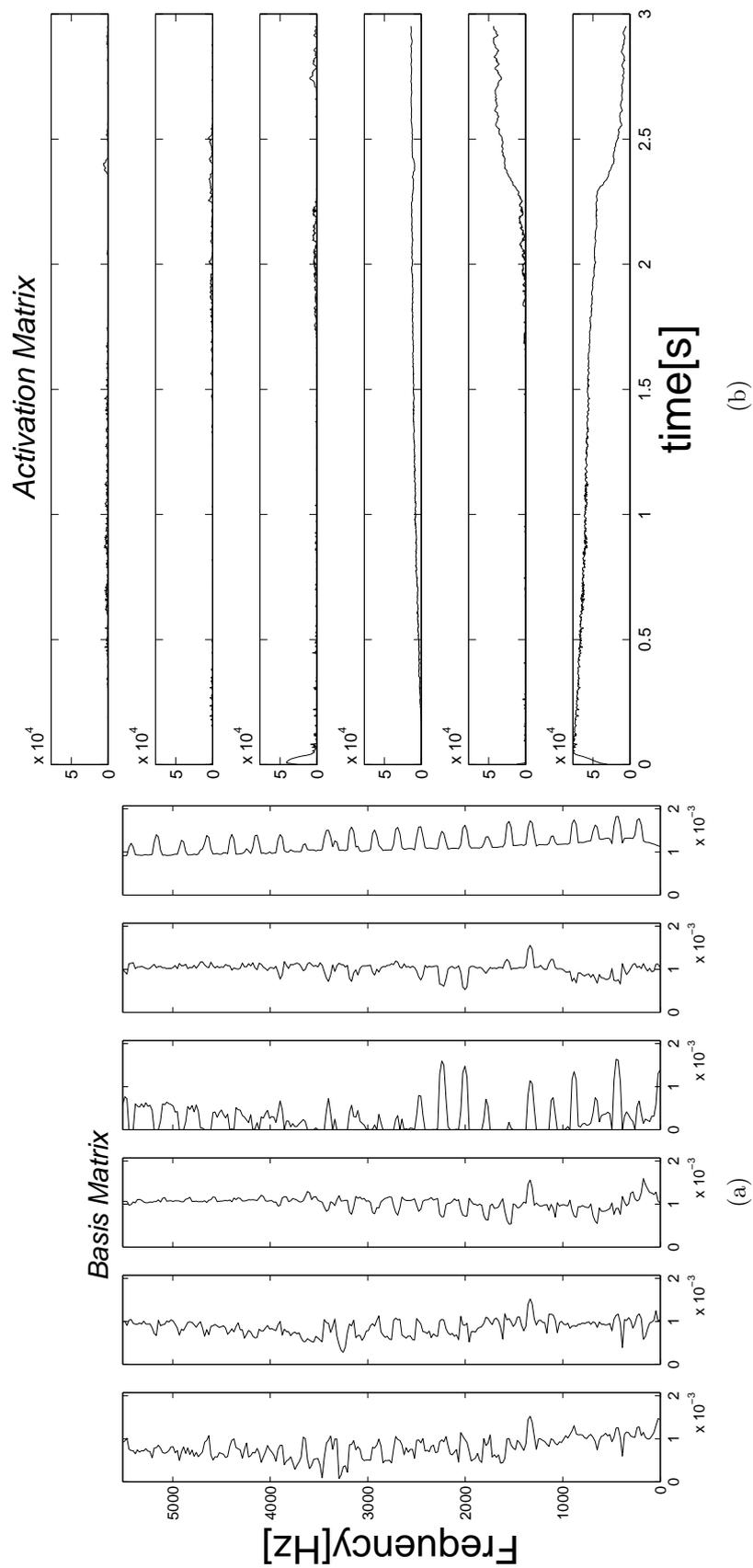


図 4.15: 基底行列の一部を固定して NMF を適応した結果 (UP1)

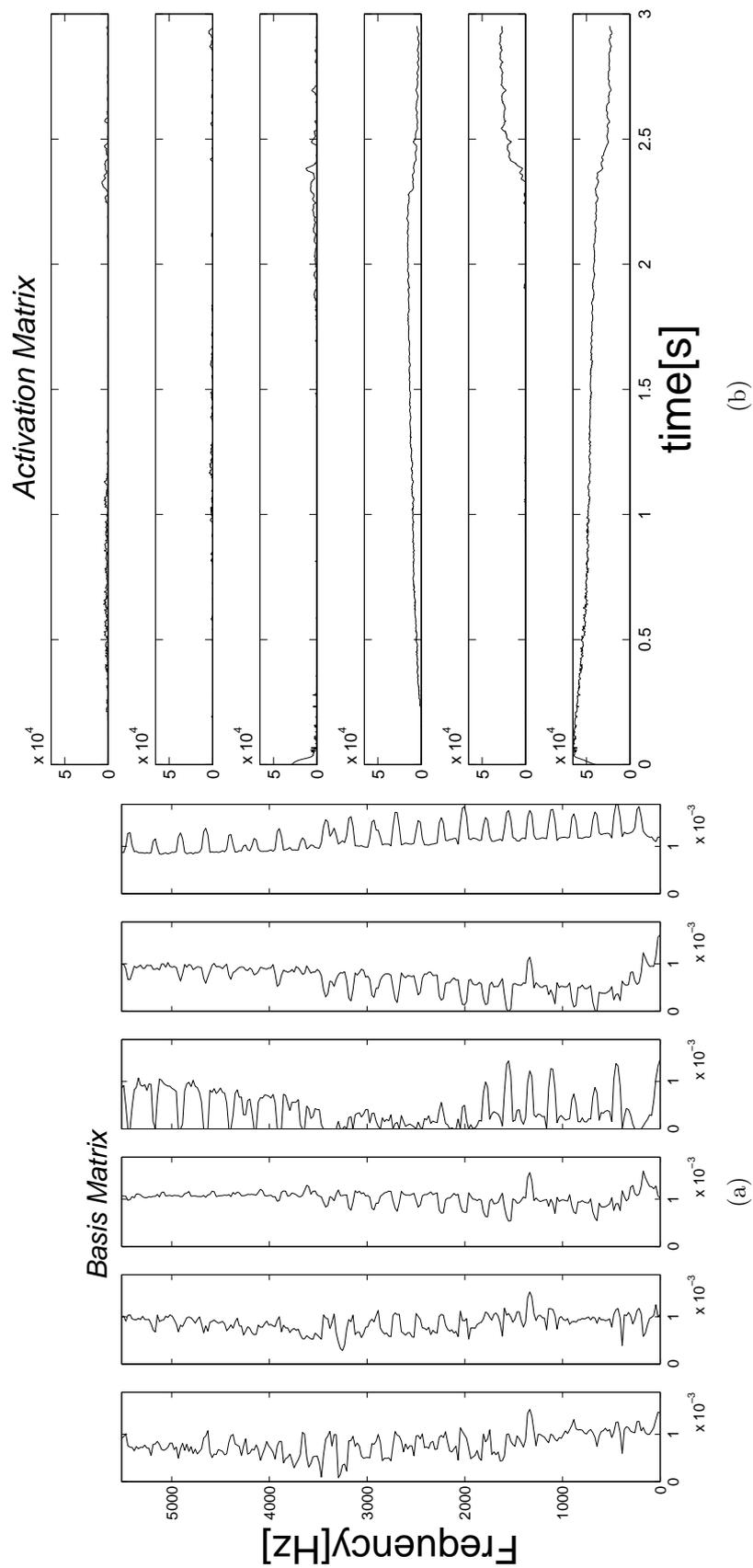


図 4.16: 基底行列の一部を固定して NMF を適応した結果 (UP2)

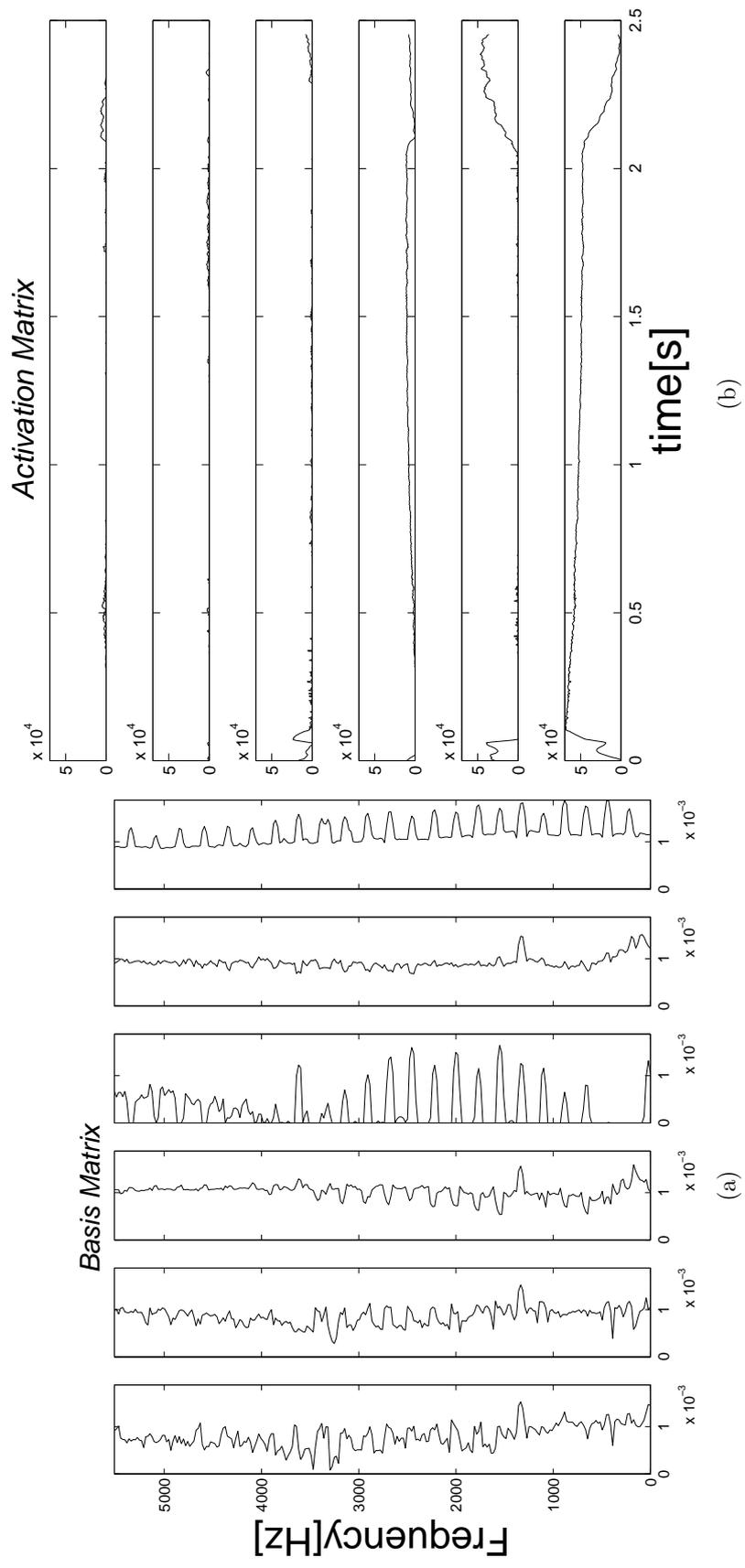


図 4.17: 基底行列の一部を固定して NMF を適応した結果 (MAPS)

スペクトルを変形させて行う音色変換では、楽器の調波成分や非調波成分を表す関数パラメータを調整して音色を作っている。今回の分析では、調波構造としては定常部ベクトルに、音源固有のピークが存在した。つまり、ある音源をより正確に表現するために、定常部ベクトルのピークを表現するような関数パラメータを学習すれば、より近い音色の音源が生成出来ると考えられる。また、減衰部ベクトルのアクティベーションで、2段階減衰のそれぞれの段階が共通成分と個体差成分とで分かれた。これより、物理モデルので減衰部を表す関数のパラメータにおいて、それぞれの減衰を表すようにパラメータを学習すれば、その音源を表す波形が生成出来ると考えられる。ただし、あくまでも今回得られたのは、対数パワースペクトログラムのある調波構造における時間変動であるため、直接時間波形に持つて行くことはあまりにも乱暴すぎる。そのため、一つの指標として今回得られた時間変動波形を利用出来るのではないかと思われる。

今回基底数6で分析を行った際、定常部のスペクトルピークと、減衰部のアクティベーションに個体差が強く反映されていることがわかった。しかし、他の部分、例えばアタックベクトルなどにはうまいこと個体差が見えてこなかった。また、一部音源に関しては、アタックベクトルが個体差に表れず、そのかわり減衰部のアクティベーションが2つ表れるなど、共通成分と個体差の和という仮定通りにいかないものもあった。ほとんどの音源では、先ほどのように個体差が表れた結果が得られたが、分離結果自体に一部例外が見られたので、基底数をいくつにすべきかというのは今後も議論が必要と考えられる。

4.3 まとめ

この章では分析結果に基づいてNMFの基底数を決定した。また、具体的に個体差を示すために、基底の一部をグランドピアノ音同士、アップライトピアノ音同士、それぞれの共通成分で固定して、再度NMFで分析を行った。その結果、定常部ベクトルに特有のピークを示し、減衰ベクトルのうち、緩やかに減衰する成分が個体差として強く表れていたことがわかった。

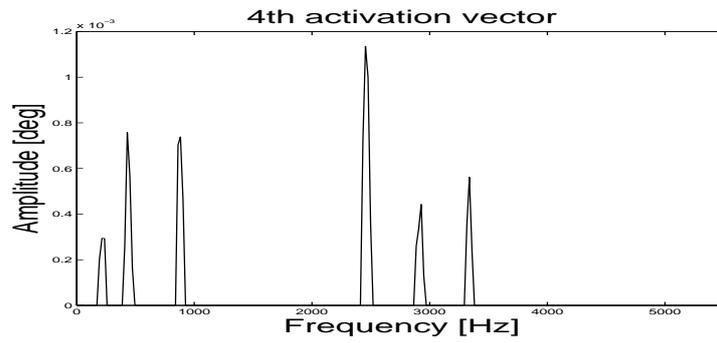
第5章 総合検討

5.1 グランドピアノ音間での個体差の違い

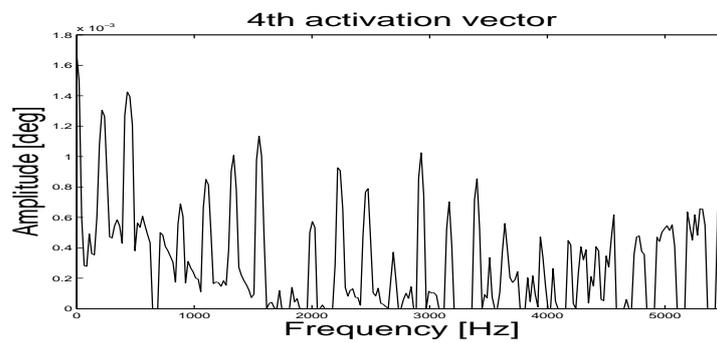
図 4.11 に示す通りグランドピアノ音間の共通成分で定常部ベクトルは調波構造が綺麗に表れている。特に 4 kHz 付近までの構造がよく保存されているため、図に示す通り、各グランドピアノ音の基底行列の第 4 ベクトルは GP1 (図 5.1(a)) と RWC3 (図 5.1(d)) では本当に必要な成分のピークしか表れていない。これらに対し、GP2 (図 5.1(b)) では第 1, 2, 4, 5, 6, 7 倍音及び、2 ~ 4 kHz の倍音成分がピークとして表れ、4 kHz 以上の高域成分ではピーク間の谷の部分が表れている。また、RWC1 (図 5.1(c)) では、第 1, 2 倍音及び、3 kHz 付近を中心としたピークが大きく表れている。これらの成分は共通成分だけでは表現出来ない成分であり、音源毎に全く異なった特徴を示しているため、個体差として大きく表れていると考えられる。特に各音源の第 4 アクティベーションベクトルの動きに注目すると、図 4.13(b) のように音の立ち上がりから約 0.5~1.0 秒ほど経過してから上がり始め、弧を描きながら、ダンパーがおりた部分でパワーが落ちている。この動きをピアノの発音機構と対応させて考えると、ふたによる反響と響板の影響が考えられる。元々基底数 $K = 3$ で分析したとき、定常部、アタック部、減衰部の 3 つに分けられた。このことを考慮すると、第 4 ベクトルは定常部の個体差と考えられるので、響板の影響が強く反映されていると推測される。

また、各分析結果の第 6 アクティベーションベクトルに注目すると減衰傾向がある。したがって、共通成分の第 3 ベクトルである減衰部ベクトルでは表すことが出来ない成分が第 6 ベクトルに表れていると考えられる。グランドピアノ音の固定部、非固定部それぞれの減衰部ベクトルに対応するアクティベーションを図 5.2 に一例を記載した。図 5.2(a) が共通成分のアクティベーションで、(b) が非固定部のアクティベーションである。この図から、ピアノ音の特徴である 2 段階減衰のうち、音の立ち上がり直後の急峻な減衰は、(a) の共通成分のアクティベーションに表れ、後半の緩やかな減衰は (b) である非固定部のアクティベーションに表れていた。音源固有の個体差を考えると、非固定部である第 6 ベクトルが表す、緩やかな減衰が要素として重要であると考えられる。そこで、各音源の第 6 ベクトルの減衰の傾きを調べたところ表 5.1 のようになった。これらの結果より、弦の減衰の特性という視点からも、グランドピアノ音の個体差が表れていると考えられる。

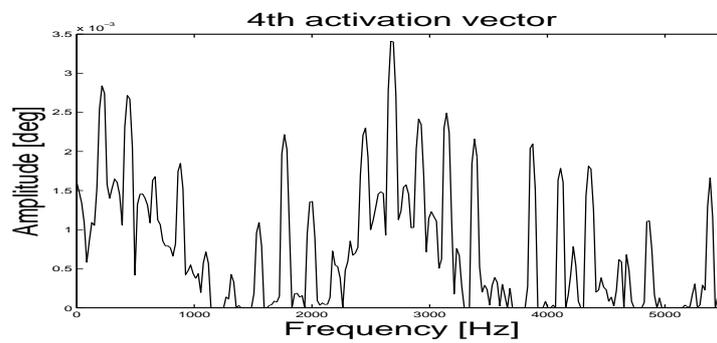
音の立ち上がり部に関しては、RWC3 (図 4.14) のように立ち上がりが見えるものもあれば、GP2 (図 4.13) のようにダンパーがおりた後のノイズ成分を捉えているものなど、比較が困難であった。



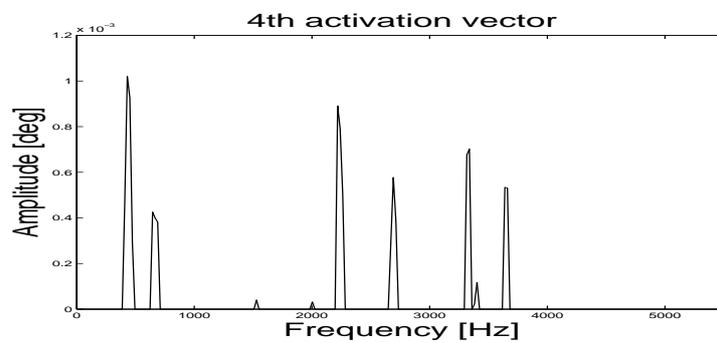
(a)



(b)

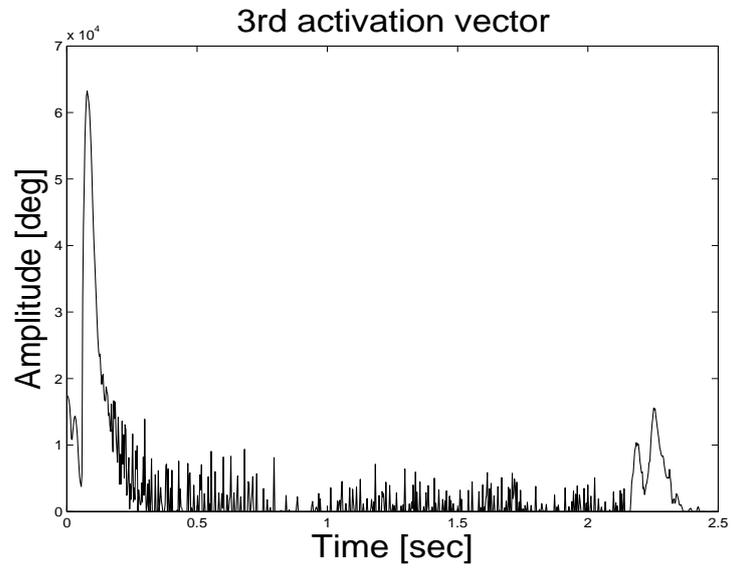


(c)

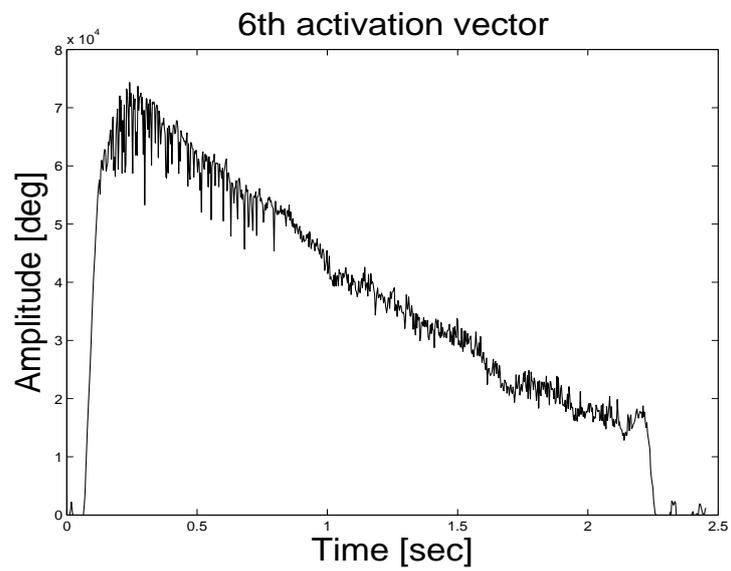


(d)

図 5.1: アップライトピアノ音の非固定部の定常部ベクトル



(a)



(b)

図 5.2: グランドピアノ音の減衰部ベクトルのアクティベーション

表 5.1: グランドピアノ音の減衰の傾き

音源	GP1	GP2	RWC1	RWC3
傾き (deg/sec)	-2.53×10^4	-1.21×10^4	-1.97×10^4	-3.26×10^4

表 5.2: アップライトピアノ音の減衰の傾き

音源	UP1	UP2	UP3	UP4	MAPS
傾き (deg/sec)	-1.51×10^4	-1.29×10^4	-1.19×10^4	-1.31×10^4	-1.53×10^4

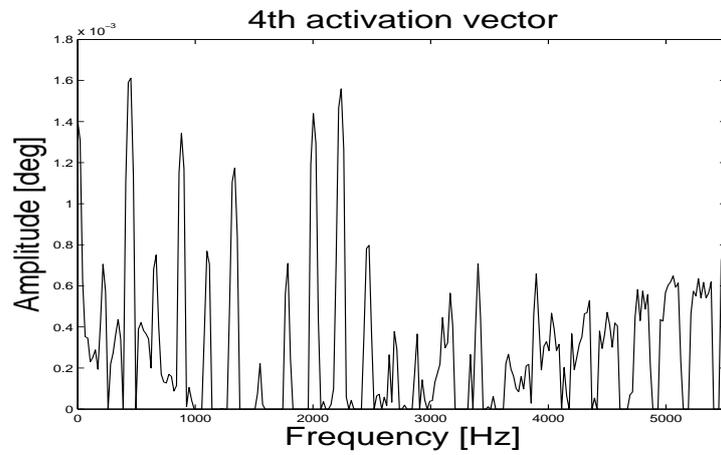
5.2 アップライトピアノ音間での個体差の違い

図 5.3 に非固定部である，第 4 基底ベクトルを記載した．図 5.3(a) は UP1 の結果，5.3(b) は UP2，そして 5.3(c) は MAPS の結果を表している．各音源の特徴として，UP1 では偶数次倍音のピークが大きく，UP2 では 0.5~2 kHz に大きなピークがあり，MAPS は 2 kHz を中心に，前後それぞれおよそ 1.2 kHz にわたって大きなピークを持っていた．この第 4 ベクトルのアクティベーションに着目すると，約 0.5 から 2.2 sec までゆるやかな弧を描いていた．したがって，これらも響板の等の反響による影響と推測される．

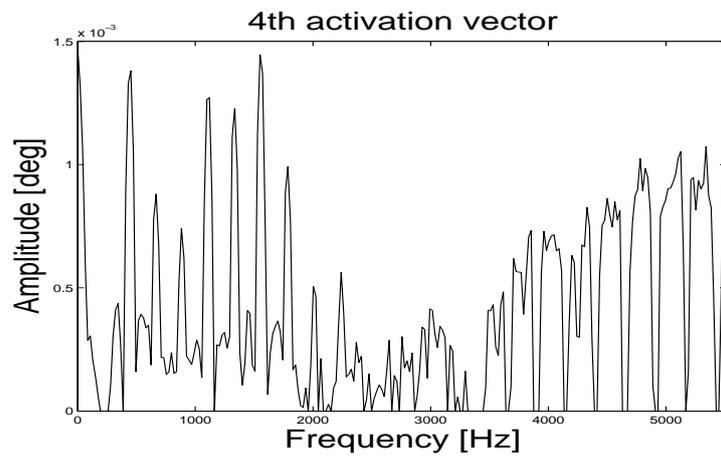
グランドピアノ音同様，アップライトピアノ音でも減衰が第 3 ベクトルと第 6 ベクトルに分かれていた．第 6 ベクトルの緩やかな減衰の傾きを表 5.2 にまとめた．グランドピアノ音同様，音源によって傾きが異なっていたが，グランドピアノ音ほど傾きにバラつきはなかった．

図 5.4 にアップライトピアノ音の基底ベクトル，アクティベーションベクトルそれぞれの第 5 ベクトル (非固定部) を記載した．図 5.4(a),(c),(e) は，それぞれ UP1, UP2, MAPS の非固定部のアタックベクトル，図 5.4(b)(d)(f) がそれぞれ UP1, UP2, MAPS アクティベーションベクトルである．MAPS ではアクティベーションにアタックらしきピークがあるが，それ以外は，ほとんどがダンパーがおりた後の成分を捉えていた．それぞれの基底ベクトルを見ると，全て第 6 倍音に大きなピークを持っている事から，ダンパが下りる影響が表れている事がここからもわかる．つまり，第 5 ベクトルは主にダンパが下りるときからの影響を表していると考えると，UP1 (図 5.4(a)) では，第 2~4 倍音成分は小さいが，ほとんど全周波数でフラットな形状をしている．UP2(図 5.4(c)) では，周波数が高くなるにつれ，値が少しずつ上がっている事から，残響として高周波数成分が残りやすい性質があると推測される．MAPS(図 5.4(e)) ではおよそ 500 Hz 以下の成分が多く，第 2 倍音以下の成分と低周波数領域の雑音が多く含まれた音が残響として響く性質であると考えられる．このようにダンパが下りた後の影響は音源によって異なっていることが，この分析によってわかった．

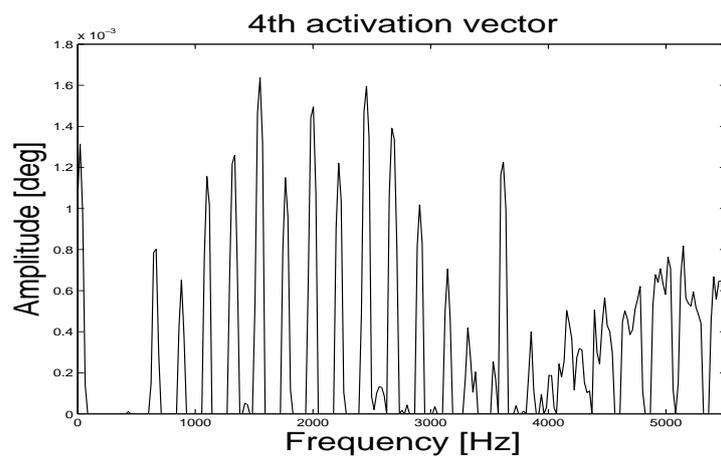
以上の事を踏まえると，アップライトピアノ音の個体差で重要な要素としては減衰の傾き，反響部分の周波数のピーク，そして，ダンパーが降りたあとの残っている成分の違い



(a)

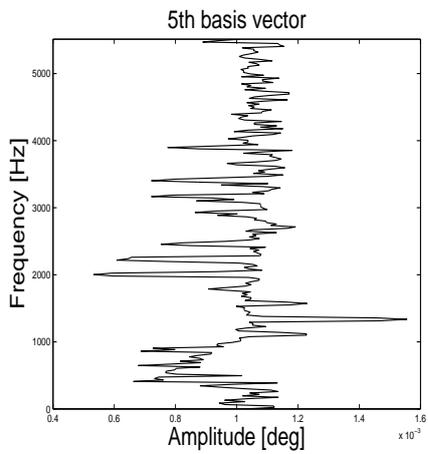


(b)

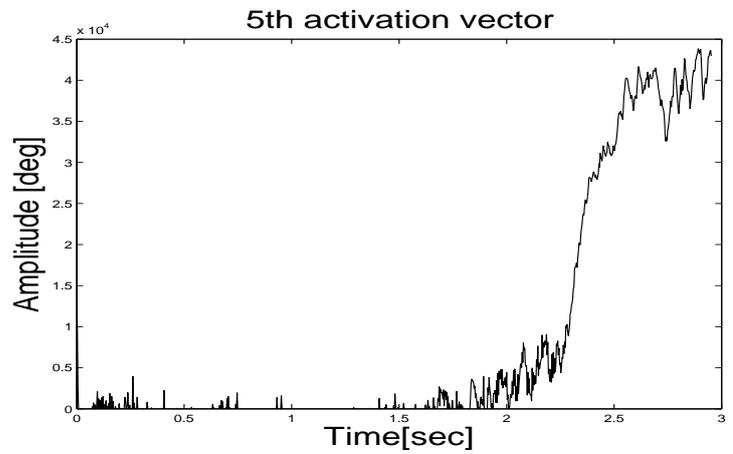


(c)

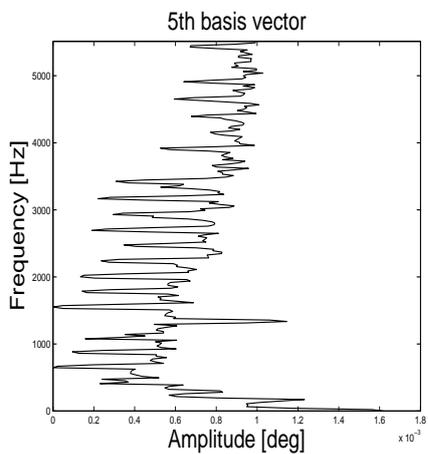
図 5.3: アップライトピアノ音の非固定部の定常部ベクトル



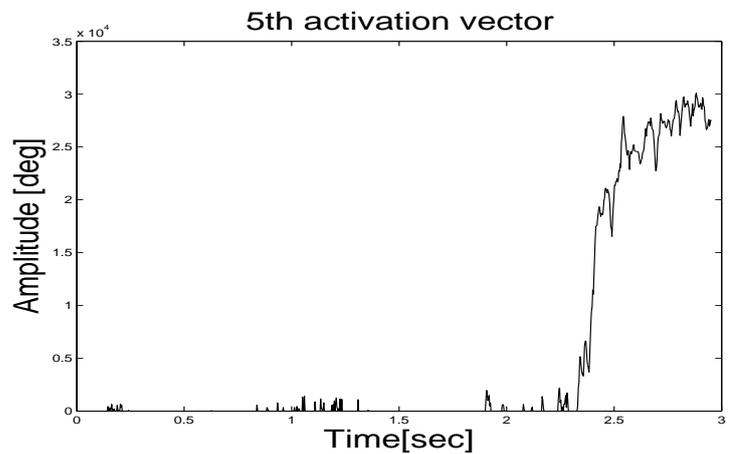
(a)



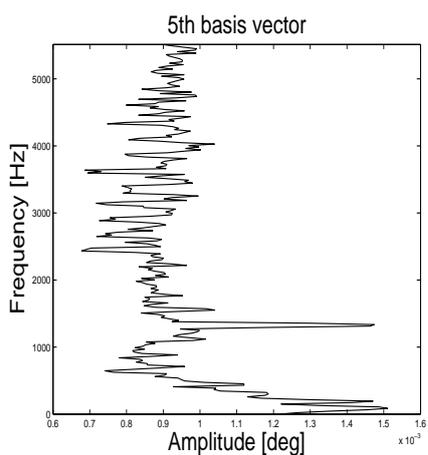
(b)



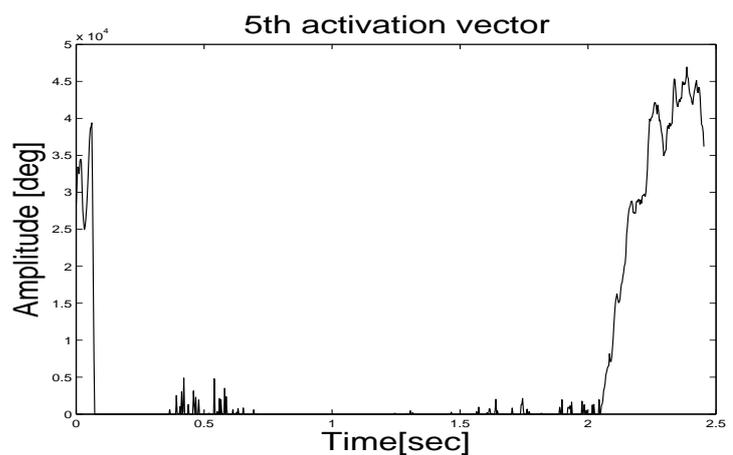
(c)



(d)



(e)



(f)

図 5.4: アップライトピアノ音の非固定部のアタックベクトル

が重要であると考えられる。

5.3 各ピアノ間での個体差の違い

グランドピアノ音、アップライトピアノ音それぞれのクラス内での個体差について触れたので、この節では、グランドピアノ音、アップライトピアノ音、MIDI音源それぞれの違いについて考察を行う。

5.3.1 アクティベーションベクトルの差異

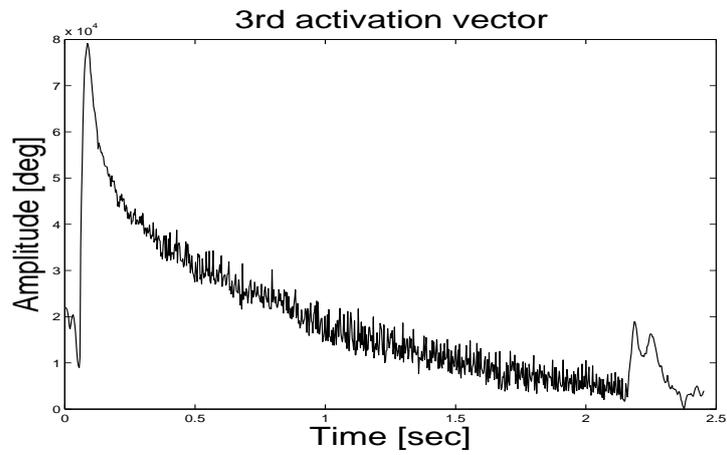
アクティベーション行列において特徴的なものから比較を行う。図 5.5(a) にグランドピアノ音、同図 (b) にアップライトピアノ音、同図 (c) MIDI音源それぞれの減衰部ベクトルのアクティベーションを記載した。

図 5.5(c) の MIDI音源の減衰部ベクトルの時間変動は、緩やかなカーブを描いて、単純な指数関数のような概形を示していた。一方、実ピアノ音では、グランドピアノ音の結果、アップライトピアノ音の結果共に立ち上がりのピークの後、急激に落ち、その後緩やかに減衰する 2 段階減衰が観測された。また、グランドピアノ音、アップライトピアノ音共に 2.25 sec 付近でダンパが下りたため、その影響が 2 つ目のピークとして表れているが、MIDI音源ではピークが表れることもなく、約 2.8 sec あたりから急激にアクティベーションの値が 0 に近づき音が無くなった。この 2 つの特徴は実音源と MIDI音源で大きく異なっている点であった。

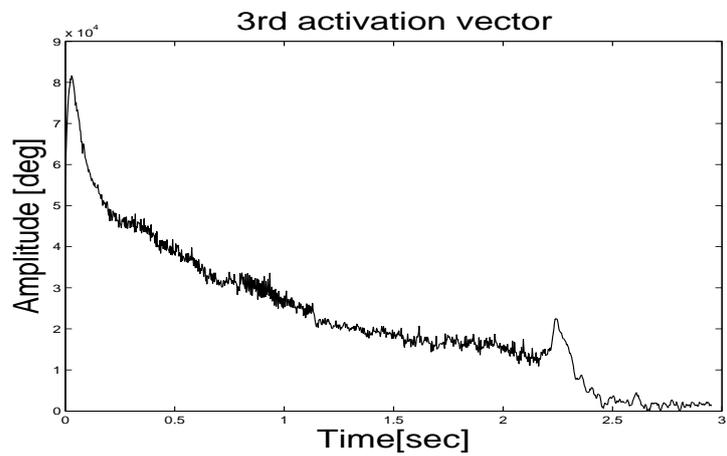
実音源同士で、特にグランドピアノ音とアップライトピアノ音では、2 段階減衰の最初の減衰に違いが表れていた。図 5.6(a) にグランドピアノ音、(b) にアップライトピアノ音の通成分の減衰部ベクトルのアクティベーションを示す。グランドピアノ音では、ピーク位置から、パワーが 0 に落ちるまで、0.22 sec かかっていた。しかしグランドピアノ音ではピーク位置からパワーが 0 になるまで、0.06 sec と凄く急峻に落ちている事が分かる。これはアップライトピアノの構造上、ハンマーと弦が触れいている時間が長いために、振動を抑制してしまうので、このように急激に減衰していると考えられる。

5.3.2 定常部ベクトルの差異

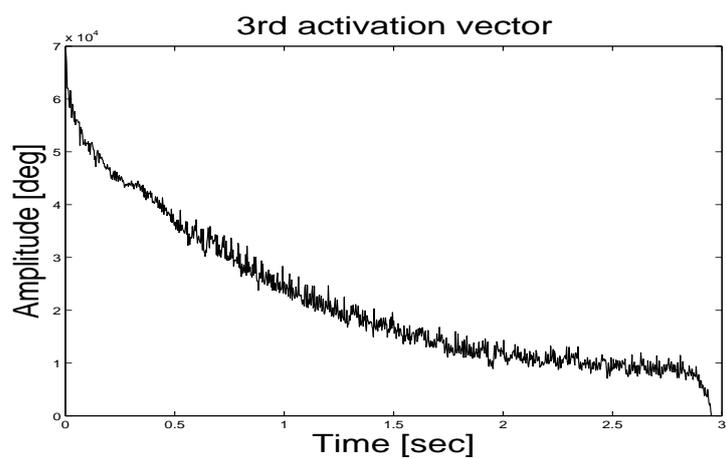
続いて、基底行列のベクトルを比較を行う。図 5.7(a) にグランドピアノ音の、(b) にアップライトピアノ音の、(c) に MIDI音源の定常部ベクトルを記載した。まず始めに注目すべき違いとしては、基本周波数とその高調波の相対的な強さの違いだろう。グランドピアノ音は基本周波数と第 2 倍音を比べたとき、第 2 倍音の方がおよそ 1.4 倍の大きさのピークを持っていた。また、アップライトピアノ音ではおよそ 1.6 倍の大きさのピークを持っていた。しかし、MIDI音源では、基本周波数のパワーの方が第 2 倍音よりおよそ 1.3



(a)

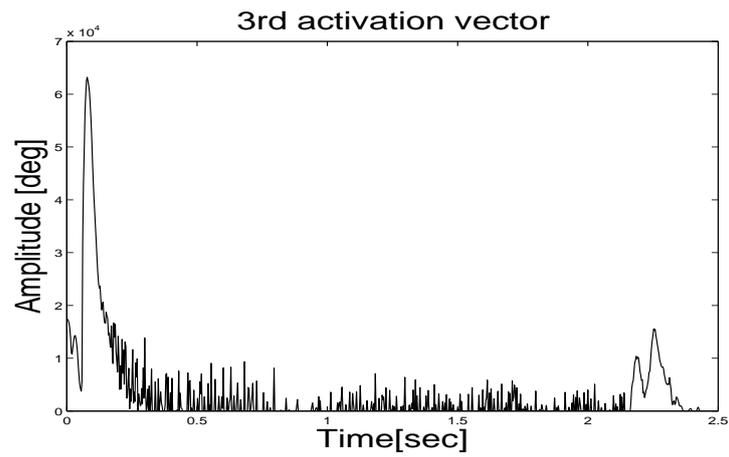


(b)

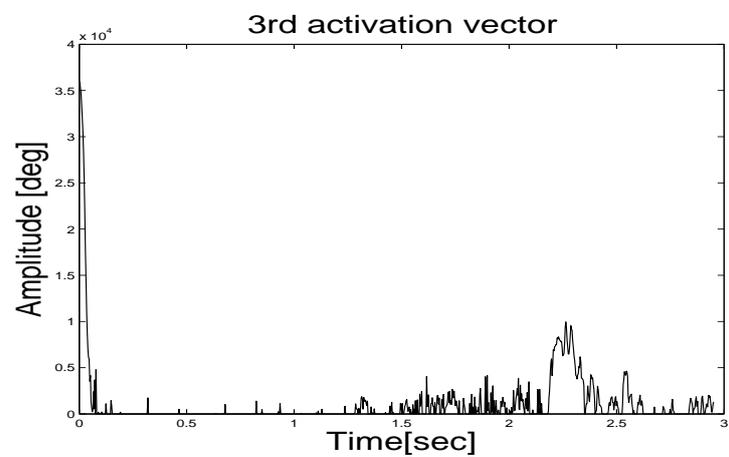


(c)

図 5.5: 各ピアノ音の減衰部ベクトルのアクティベーション

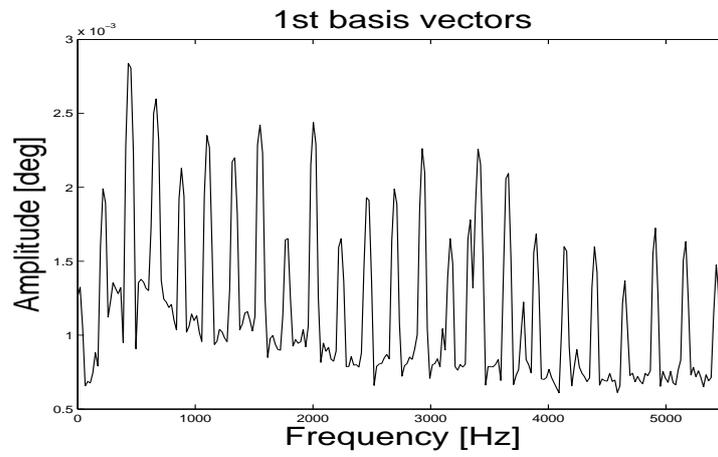


(a)

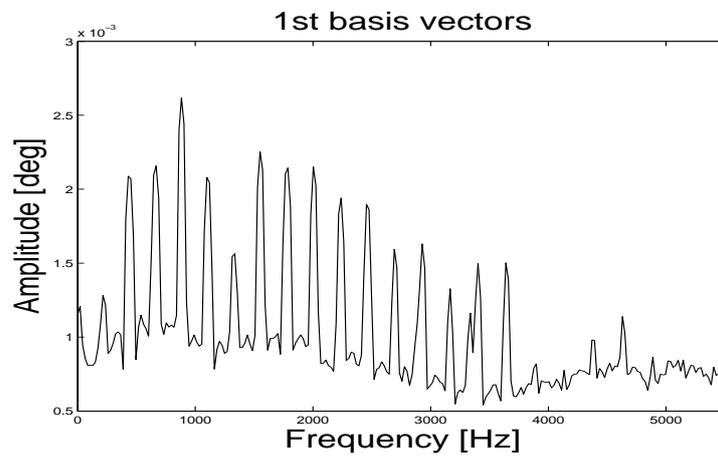


(b)

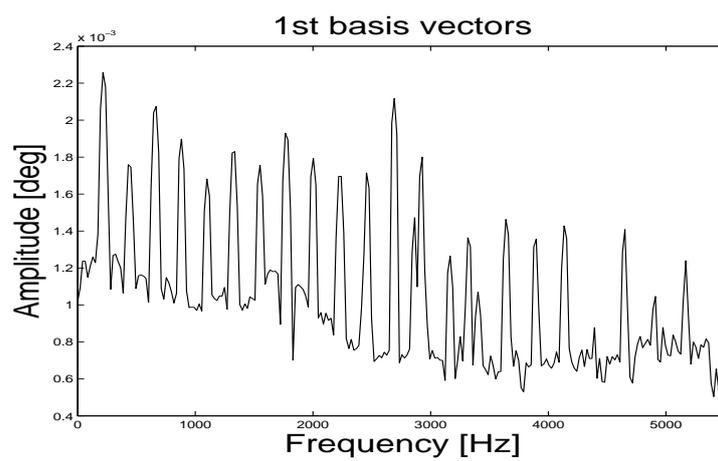
図 5.6: 各ピアノ音の共通成分の減衰部ベクトルのアクティベーション



(a)



(b)



(c)

図 5.7: 各ピアノ音の定常部ベクトル

表 5.3: 各ピアノ間の相関係数の平均値

	GRAND	UPRIGHT	MIDI
GRAND	0.77	0.63	0.73
UPRIGHT		0.78	0.67
MIDI			1.00

倍ほど大きかった。また、調波構造の違いとして、ハンマーの打弦位置の影響も各音源で異なる。グランドピアノのハンマーは一般的に弦の7分の1か8分の1のところを叩くと言われている [2]。つまり、第7, 第8倍音のピークが打弦によって小さくなるはずである。そこで各図を見てみると、グランドピアノ音は第8倍音で急激にピークの値が下がっていることから、弦の8分の1の点を打弦していることがわかる。また、アップライトピアノ音は第6倍音のピークが極端に下がっている。アップライトピアノはグランドピアノに比べ、サイズが小さいため、打弦位置が異なる。そのため、弦の6分の1のところを叩いているので、このような影響が出ていると考えられる。一方MIDI音源は第2, 第5, 第7倍音のところでピーク位置が谷になってはいるものの、上記2つの実音源のように明らかに他とは極端にピークが低いということではなかった。この打弦位置の影響も、実音源はしっかりと反映されているが、MIDI音源だと十分に反映されていないことがわかった。

この定常部ベクトルに関して、それぞれのピアノ同士で相関係数を計算した。

表 5.3 より、アップライトピアノ音同士、グランドピアノ音同士の定常部ベクトルの相関係数はそれぞれ0.77以上と高い値であった。しかし、アップライトピアノ音とグランドピアノ音間では相関係数が0.63まで低下する事から、両者の間は大きく違うということが言えた。また、MIDI音源はアップライトピアノ音、グランドピアノ音それぞれとの相関係数は低い傾向を示した。ただし、MIDI音源-アップライトピアノ音間の相関係数よりも、MIDI音源-グランドピアノ音間の相関係数の値の方が高かった。これは、利用したMIDI音源がグランドピアノを想定して作られたからと考えられる。

5.3.3 高周波数領域の影響

先ほど定常部ベクトルについて比較を行ったが、このとき4kHz以上の成分に着目すると、グランドピアノ音は調波構造のピークが観測されるのに対し、アップライトピアノ音ではごく一部でしかピークが見られなかった。そこで、アップライトピアノ音の4kHz以上の成分がどこにあるか確認したところ、図 5.8 に示す通り、減衰部ベクトルに成分がある事が分かった。

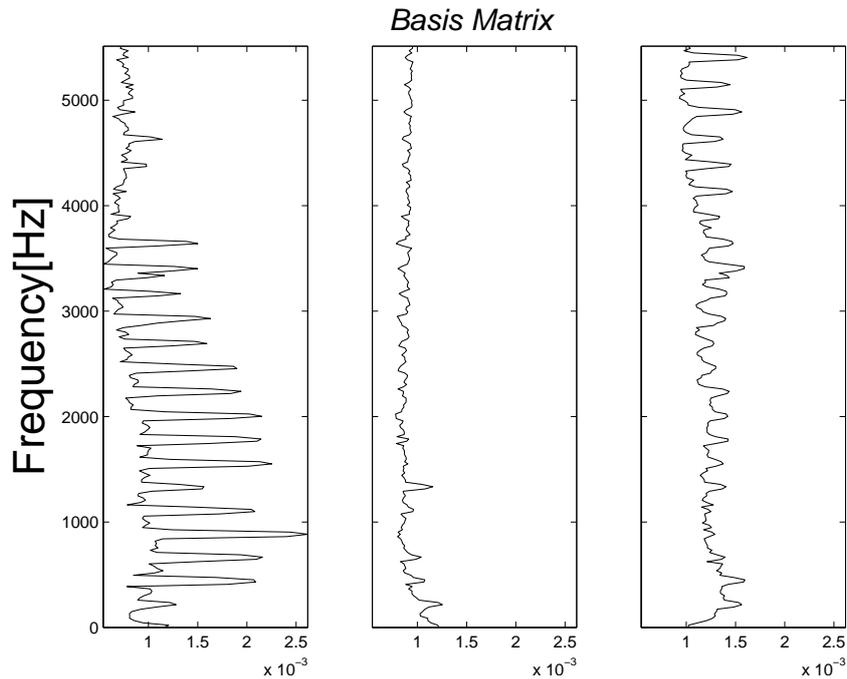


図 5.8: アップライトピアノ音の基底行列 (UP3)

つまり、アップライトピアノ音の4kHz以上の高周波数領域成分のほとんどが、すぐに減衰してしまうことがわかる。一方グランドピアノ音は定常部ベクトルに高周波数領域の成分があるため、アップライトピアノ音のようにすぐには減衰しないと言える。この違いはピアノの構造が影響しているのではないかと考えられる。グランドピアノ音では定常部に表れている事から、高域成分が響板の響きなどで強調されていると推測される。しかしアップライトピアノでは、減衰部で高域成分が観測された。これは、アップライトピアノは響板が壁に近いこと、また、打弦機構の手前にカバーとして板がついているため、グランドピアノのような響板等の強調がそこまで行えないため、弦の減衰の仕方が直接反映されたと考えられる。

5.4 まとめ

グランドピアノ音とアップライトピアノ音では、それぞれのピアノの構造であるハンマーの打源位置がしっかりと反映されていた。また、4kHz以上の高域成分の時間変動が大きく異なっていた。ピアノ音の個体差として、定常部ベクトル、特に反響による増幅のところでグランドピアノ音、アップライトピアノ音それぞれ音源毎に異なるピークを持つことが確認された。また、2段階減衰の2段階目である緩やかな減衰が個体差成分として表れている事から、この緩やかな減衰の仕方が非常に重要であると結論づけられた。

MIDI音源と実音源では、減衰の仕方が大きく異なり、ピアノ特有の2段階減衰がない

ことや、ダンパが下りたときの影響が未考慮である2点が大きく異なる点であった。また、調波構造では、ハンマーの打弦点の影響が表れていないこと、基本周波数とその倍音のパワー比の違い、定常部ベクトルの相関係数が低いことなどが挙げられる。したがって、MIDI音源を実音源に近づけるためには、調波構造では打弦点による一部ピークのパワーが小さくなっていること、基本周波数と倍音のパワー比を調整することが重要と考えられる。特に、グランドピアノ音に近づけるならば、4kHz以上の高域成分の減衰を抑える方向、アップライトピアノ音に近づけるならば、弦の振動と同じような減衰を行わせるように設計する必要がある。また、時間変動では、弦の2段階減衰をモデルとして組み込むことは当然だが、ダンパの影響は非ペダルトーンではほとんど考慮されていないので、この影響を減衰波形に組み込む必要がある。

第6章 結論

6.1 本研究のまとめ

本研究ではピアノ音の個体差をNMFを用いて分析した。その結果、NMFでピアノ音を表現するには、基底数は3つにすると、ピアノの発音機構の打鍵+打弦時の雑音、弦の振動、響板による共鳴・反響のそれぞれの成分に分けて表現することができた。そして、それぞれの成分において共通成分として最小値を算出し、固定して再度NMFで分析する事により、各音源固有の個体差が反響部分の調波構造と、弦の減衰波形などに強く表れた。

MIDI音源と実ピアノ音では減衰部の時間変動が大きく異なり、2段階減衰やダンパの影響がMIDI音源では表現が不十分であることがわかった。また、アップライトピアノ音とグランドピアノ音では4 kHz以上の高域成分の時間変動パターン、そして、ハンマーが弦に接触している時間から起因すると思われる、2段階減衰の最初の減衰の仕方が違いとして表れた。

6.2 今後の展望

今回ピアノ音の個体差分析にあたり、ピアノの発音機構を考慮したため、対数パワースペクトログラムを用いた。これは要因が並列的ではなく、直列的なことに起因する。そのため、ピアノ音に限らず他の楽器でも、個体差に関わる要因が直列的であれば、同じ分析手法にて個体差を表す事が可能であると考えられる。例えば、金管楽器のホルンであれば、演奏者の息に、唇やリードの振動が加わり、楽器に伝わる。そして最後にベル(楽器の先端の音が放出される部分)に右手を入れて音色をコントロールするので、これらの複雑な要因の解析に用いる事が出来ると考えられる。

今回は、音量、音高、演奏者を出来る限り同じになるように統一したが、演奏者を変えたものを分析すれば、演奏者の癖も、個体差の一部として表れると考えられる。つまり、この要因を解析出来れば、演奏表現の解析や、再現などに有効と考えられる。このように多楽器への応用や、拘束条件を変える事により、様々な効果が期待出来ると思われる。

NMFで対数パワースペクトログラムを分離して、そのとき基底行列、アクティベーション行列それぞれに個体差が表れた。この得られた個体差成分を別の音源のものと入れ替え、対数パワースペクトログラムを再合成、そしてそこから、時間波形に戻せば、音色変換が行えることが考えられる。今回MIDI音源では弦の2段階減衰とダンパの影響が表現出来ていなかった。そこで、NMFで分析してそのときに得られた減衰部を表すアクティ

バージョンを実音源のアクティベーションに入れ替えて、スペクトログラム、そして時間波形を再合成すれば、より実音源に近い MIDI 音源に変換出来ると思われる。

謝辞

赤木 正人 教授には、本研究を進めるにあたり、数多くの御指導ならびに御鞭撻を賜りましたこと、深く感謝致します。また、研究室会議をはじめ、様々な機会でご指導、御討論、御助言等をして頂いた鶴木 祐史 准教授、宮内 良太 助教に心より感謝致します。さらに、本研究を進めるにあたり、御指導、御討論だけではなく、ピアノ音の収録にまで携わって頂いた森川 大輔 助教には多大なる感謝を致します。本研究を進めるにあたり、使用技術について深くまで御助言を賜りました党 建武 教授、田中 宏和 准教授、末光 厚夫 助教、川本 真一 助教に誠に感謝致します。また、分析用のピアノ音の収録にあたり、快く許可して頂いた、石川県立音楽堂の関係者の方々にこの場を借りて深く御礼申し上げます。

本研究室のメンバーの皆様にも感謝します。特に、濱田 康弘氏、長田 和也氏の両名にはピアノ音の収録の協力の際、快く引き受けて頂き、深く感謝致します。また、論文の英語添削を引き受けて頂いた、研究員の Mr. Elbarougy, 博士後期課程の Mr.Chau, Mr.Ngo 各氏に改め感謝の意を示します。日頃から熱心に議論につき合って頂き、さらに公私にわたりお世話になった西野 恭生氏、安藤 将氏に深く感謝致します。それに加え、インターネットを通じて議論や問題解決につきあって頂いた、電気通信大学 情報理工学研究科の宮地 紘司氏、有北 知弘氏の両名には本当に感謝しております。

最後に、本学の研究生活において携わって頂いた全ての方々に改めて感謝致します。

参考文献

- [1] 剣持秀紀. 歌声合成技術の動向：「初音ミク」を支える技術 (小特集「音声合成に関する研究の動向」). 日本音響学会誌, Vol. 67, No. 1, pp. 46–50, 2010.
- [2] 中村勲, 鈴木英男. ピアノの発音機構と音色. 日本音響学会誌, Vol. 49, No. 3, pp. 178–183, 1993.
- [3] 高澤嘉光. ヴァーチャルピアノ. 電気通信大学紀要, Vol. 20, No. 1-2, pp. 15–21, 2007.
- [4] 橋本智樹, 梅谷征雄. ピアノ音の物理シミュレーション：弦・駒・響板の連成解析をめざして. 電子情報通信学会技術研究報告. SP, 音声, Vol. 99, No. 626, pp. 1–6, 2000.
- [5] 中村勲, 長沼大介. 三本弦によるピアノ音の減衰特性の考察. 情報処理学会研究報告. [音楽情報科学], Vol. 2002, No. 14, pp. 7–12, 2002.
- [6] 西口磯春, 佐々木真孝, 山田倫也. ピアノの2段減衰のモデル化について. 日本音響学会研究発表会講演論文集, 2013.
- [7] N.H. Fletcher and T.D. Rossing. The Physics of Musical Instruments second edition, chapter 12, pp. 352–398. springer, 1998.
- [8] 柏野邦夫, 村瀬洋. 適応型混合テンプレートを用いた音源同定：音楽演奏への適用. 電子情報通信学会論文誌. D-II, 情報・システム, II-情報処理, Vol. 81, No. 7, pp. 1510–1517, 1998.
- [9] 安良岡直希, 糸山克寿, 高橋徹, 駒谷和範, 尾形哲也, 奥乃博. 調波非調波 gmm に基づく midi 演奏音響信号に対する音色・演奏表情操作 (音楽信号処理, 学生セッション, 人工知能と認知科学, 情報処理学会創立 50 周年記念). 全国大会講演論文集, Vol. 72, No. 2, pp. 183–184, 2010.
- [10] 山家清彦, 小川信治, 山本有美子, 徳弘一路. スタインウェイおよびベーゼンドルファー社製ピアノの音色分析とプレイヤーの演奏印象の関係. 日本音響学会研究発表会講演論文集, Vol. 2003, No. 1, pp. 823–824, mar 2003.
- [11] C.T.Lee, Y.H.Yang, and H.H.Chen. Multipitch estimation of piano music by exemplar-based sparse representation. Multimedia ,IEEE transactions on, Vol. 14, No. 3, pp. 608–618, 2012.

- [12] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. Nature, Vol. 401, pp. 788–791, 1999.
- [13] P.Smaragdis and J.C.Brown. Non-negative matrix factorization for polyphonic music. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, pp. 19–22, 2003.
- [14] 糸山克寿, 高橋徹, 駒谷和範, 尾形哲也, 奥乃博. 楽譜情報を用いた nmf による音楽音響信号の音源分離. 全国大会講演論文集, No. 2, pp. 163–164, 2007.
- [15] F.Rigaud, A.Falaize, B.David, and L.Daudet. Does inharmonicity improve an nmf-based piano transcription model? Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 11–15, 2013.
- [16] 安良岡直希, 吉岡拓也, 糸山克寿, 高橋徹, 駒谷和範, 尾形哲也, 奥乃博. フレーズ置換のための調波非調波 gmm・nmf に基づく音源分離・演奏合成. 情報処理学会論文誌, Vol. 52, No. 12, pp. 3839–3852, 2011.
- [17] 平賀瑠美, 平田圭二, 片寄晴弘. 蓮根:めぎせ世界一のピアニスト. 情報処理, Vol. 43, No. 2, 2002.
- [18] 奥満久, 畝見達夫. 感性情報に基づいた音色作成支援システムの設計 (ポスター (生活支援・アート・可視化), 映像表現・芸術科学フォーラム 2013). 映像情報メディア学会技術報告, Vol. 37, No. 17, pp. 265–268, 2013.
- [19] ヤマハ楽器. グランドピアノとアップライトピアノの違い. <http://jp.yamaha.com/products/musical-instruments/keyboards/about/gp/#upgp>.
- [20] W.Goebel, R.Bresin, and A.Galemba. Once again: The perception of piano touch and tone: Can touch audibly change piano sound independently of intensity? Proceedings of the International Symposium on Musical Acoustics, pp. 332–335, 2004.
- [21] F.Rigaud, B.David, and L.Daudet. A parametric model of piano tuning. Proc. of the 14th International Conference on Digital Audio Effects, pp. 393–399, 2011.
- [22] F.Rigaud, A.Falaize, B.David, and L.Daudet. Does inharmonicity improve an nmf-based piano transcription model? ICASSP, 2013.
- [23] A.Lefèvre, F. Bach, and C.Févotte. Online algorithms for nonnegative matrix factorization with the itakura-saito divergence. In Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), Mohonk, NY, Oct. 2011.

- [24] C.Févotte. Majorization-minimization algorithm for smooth itakura-saito nonnegative matrix factorization. In Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, May 2011.
- [25] 宏澤田. 非負値行列因子分解 nmf の基礎とデータ/信号解析への応用. 電子情報通信学会誌, Vol. 95, No. 9, pp. 829–833, 2012.
- [26] 亀岡弘和. チュートリアル:非負値行列因子分解. 情報処理学会研究報告. [音楽情報科学], Vol. 2011, No. 4, p. 1, 2011.
- [27] 中野允裕, ルルージョナトン, 亀岡弘和, 中村友彦, 小野順貴, 嵯峨山茂樹. スペクトログラムのベイジアンノンパラメトリックモデリングに基づく音楽信号の解析. 情報処理学会研究報告. [音楽情報科学], Vol. 2011, No. 6, pp. 1–8, jul 2011.
- [28] 吉井和佳, 後藤真孝. 多重音基本周波数解析のための無限複合自己回帰モデル. 研究報告音楽情報科学 (MUS) , Vol. 2012-MUS-96, No. 8, pp. 1–8, 2012.
- [29] 後藤真孝, 橋口博樹, 西村拓一, 岡隆一. Rwc 研究用音楽データベース : 音楽ジャンルデータベースと楽器音データベース. 日本音響学会研究発表会講演論文集, Vol. 2003, No. 1, pp. 843–844, mar 2003.
- [30] V.Emiya, R.Badeau, and B.David. Multipitch estimation of piano sounds using a new probabilistic spectral smoothness principle. IEEE Transactions on Audio, Speech and Language Processing, No. 18, pp. 1643–1654, 2010.

付録

ピアノ音収録

今回分析に必要なピアノ音の収録を行った。収録場所は、石川県立音楽堂の練習室1～5,7を利用した。そこで、収録にあたり、収録機材や、方法について記載する。

収録機材

用いた収録機材を表1に記載する。マイクは合計で5本用いた。基本的にはマイク3本で、直接音を収録し、1本はピアノの中心位置より約1.5mほど離れた位置で、そして残りの1本は参考用として、PCなしオーディオインターフェース以下(オーディオI/F)の近くで収録するのに用いた。

収録方法

ピアノ音収録にあたり、時間との兼ね合いもあったため、収録した音の設定は表の通りである。途中にある二重線より下の部分に関しては、時間的に余裕があるものについてのみ収録した。基本的には二重線の手前までの6種類の収録を行った。

グランドピアノ音収録

グランドピアノ音の収録にあたり、収録環境及び収録方法について記載する。マイク等のセッティングは、グランドピアノ音のデータベースである、RWC-DBのものに極力近くなるよう調整した。

GP1

本文中でGP1と呼称していたピアノはYAMAHA GRAND C3(図1)を用いた。寸法は幅144cm、奥行き180cmである。ふたは全開にした。収録方法だが、図2にも示しているが、マイク位置を詳しく記すと表3のようになる。



図 1: YAMAHA GRAND C3



(a)



(b)



(c)



(d)

図 2: GP1 のマイク設置

表 1: 収録機材一覧

PC	DELL precision M6500
OS	windows 7 (32-bit)
CPU	intel core i7
ソフトウェア	MATLAB2013a
オーディオインターフェース	Roland OCTA-CAPTURE
マイク	Audio-Technica AT845Ra (1,2,4,5 ch)
	RAMSA WM-C75 (3 ch)
ヘッドホン	SENNHEISER HDA200

グランドピアノ GP2

GP2 と呼称していたピアノは YAMAHA GRAND S6A を用いた。寸法は幅 151cm, 奥行き 210cm である。ふたは全開にした。収録方法だが、図 3 にも示しているが、マイク位置を詳しく記すと表 4 のようになる。

アップライトピアノ音収録

アップライトピアノ音の収録にあたり、収録環境及び収録方法について記載する。マイク等のセッティングは、アップライトピアノ音のデータベースである、MAPS-DB のものに極力近くなるよう調整した。

アップライトピアノ UP1,UP2,UP4

UP1,2,4 は YAMAHA アップライトピアノ YU3SZ を用いて収録した。寸法は幅 150 cm, 高さ 129cm, 奥行き 65 cm である。マイクの設置に関しては図 4 でも記載しているが、詳しい内容を表 5 にも記した。

アップライトピアノ UP3

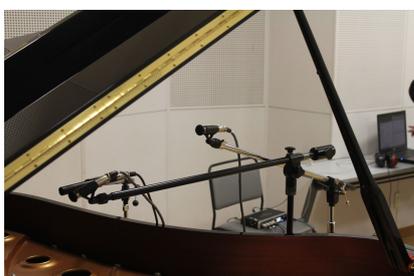
UP3 は KAWAI アップライトピアノ K80 を用いた。寸法は YAMAHA のものと同様なので、マイク位置についての説明は割愛する。



(a)



(b)



(c)

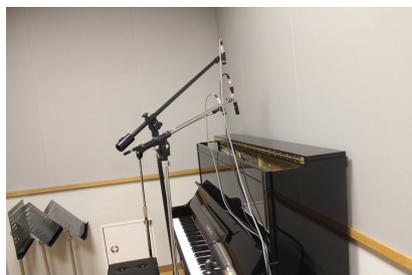


(d)

図 3: GP2 のマイク設置



(a)



(b)



(c)



(d)

図 4: アップライトピアノのマイク設置

表 2: 収録音の設定

収録番号	音高	音量	時間 [sec]	演奏者
01	A3(220 Hz)	<i>m</i>	3	P1
02	A3	<i>f</i>	3	P1
03	A3	<i>p</i>	3	P1
04	A3	<i>m</i>	30	P1
05	A4(442 Hz)	<i>m</i>	3	P1
06	A3	<i>m</i>	3	P2
07	A4	<i>f</i>	3	P1
08	A4	<i>p</i>	3	P1
09	A3	<i>m</i>	30	P2
10	A3	<i>m</i>	30	P2
11	A3	<i>m</i>	3	P3
12	A3	<i>m</i>	30	P3
12	A3	<i>p</i>	30	P2
12	A3	<i>p</i>	30	P3

表 3: GP1 収録時のマイク位置

Ch	幅	奥行き	高さ [cm]	備考
1	72 cm	90 cm	弦から鉛直上に 28	
2	108 cm	90 cm	弦から鉛直上に 42	
3	72 cm	100 cm	弦から鉛直上に 28	
4				中心位置から直線距離 140 cm
5				参照用

分析結果

各基底数における分析結果

各基底数における、全ての音源の分析結果を示す。

MIDI 音源は図 5~9, UP1 は図 10~14, UP2 は図 15~19, UP3 は図 20~24, UP4 は図 25~29, MAPS は図 30~34, GP1 は図 35~39, GP1 は図 40~44, RWC1 は図 45~49, RWC3 は図 50~54, にそれぞれ基底数 $K = 1$ から 5 までの結果を示す。また、各図において、(a) は基底行列, (b) はアクティベーション行列を表す。

表 4: GP2 収録時のマイク位置

Ch	幅	奥行き	高さ	備考
1	75.5 cm	105 cm	弦から鉛直上に 28 cm	
2	113 cm	105 cm	弦から鉛直上に 42 cm	
3	75.5 cm	110 cm	弦から鉛直上に 28 cm	
4				中心位置から直線距離 140 cm
5				参照用

表 5: アップライトピアノ音収録時のマイク位置

Ch	幅	奥行き	高さ	備考
1	70 cm	最奥より 21 cm	149 cm	
2	75 cm	最奥より 21 cm	179 cm	
3	80 cm	最奥より 21 cm	149 cm	
4			120cm	中心位置から直線距離 140 cm
5				参照用

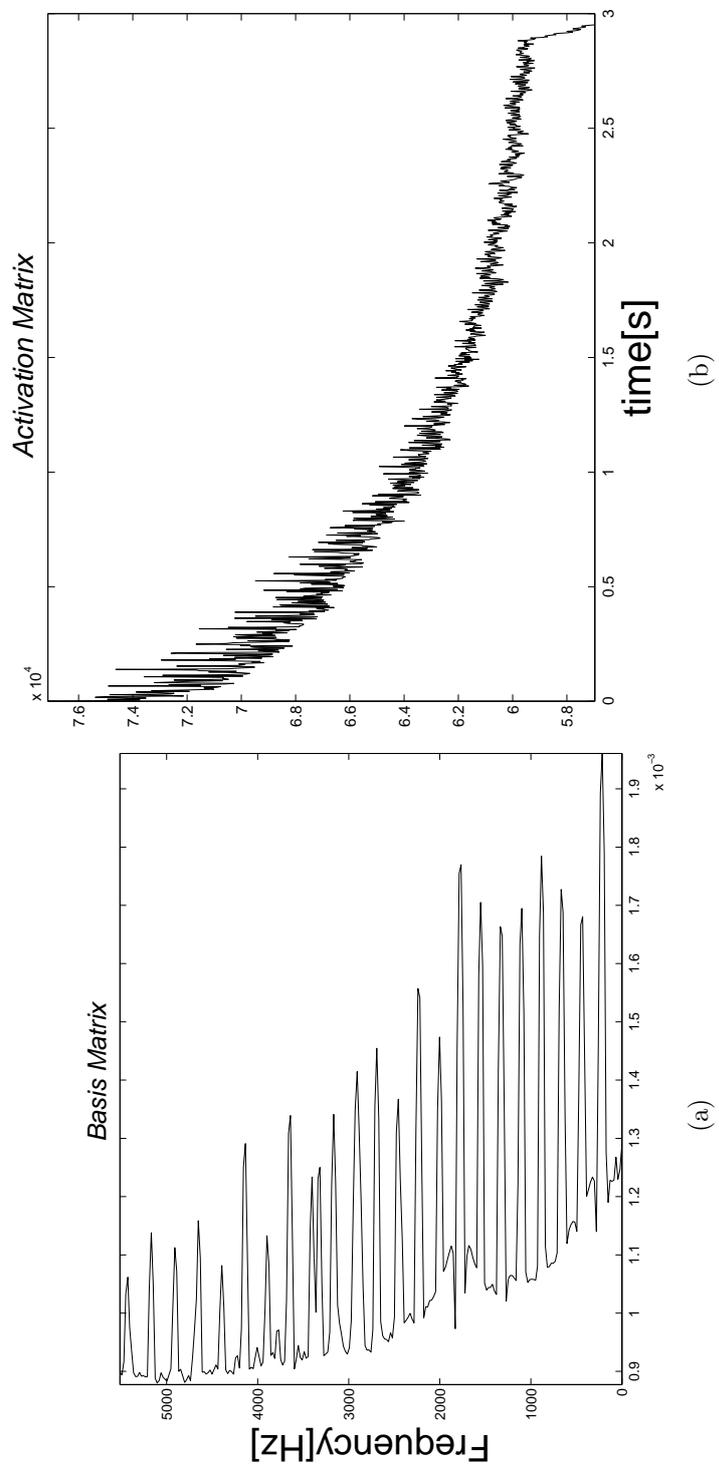


図 5: $K = 1$ の場合の分離結果 (MIDI)

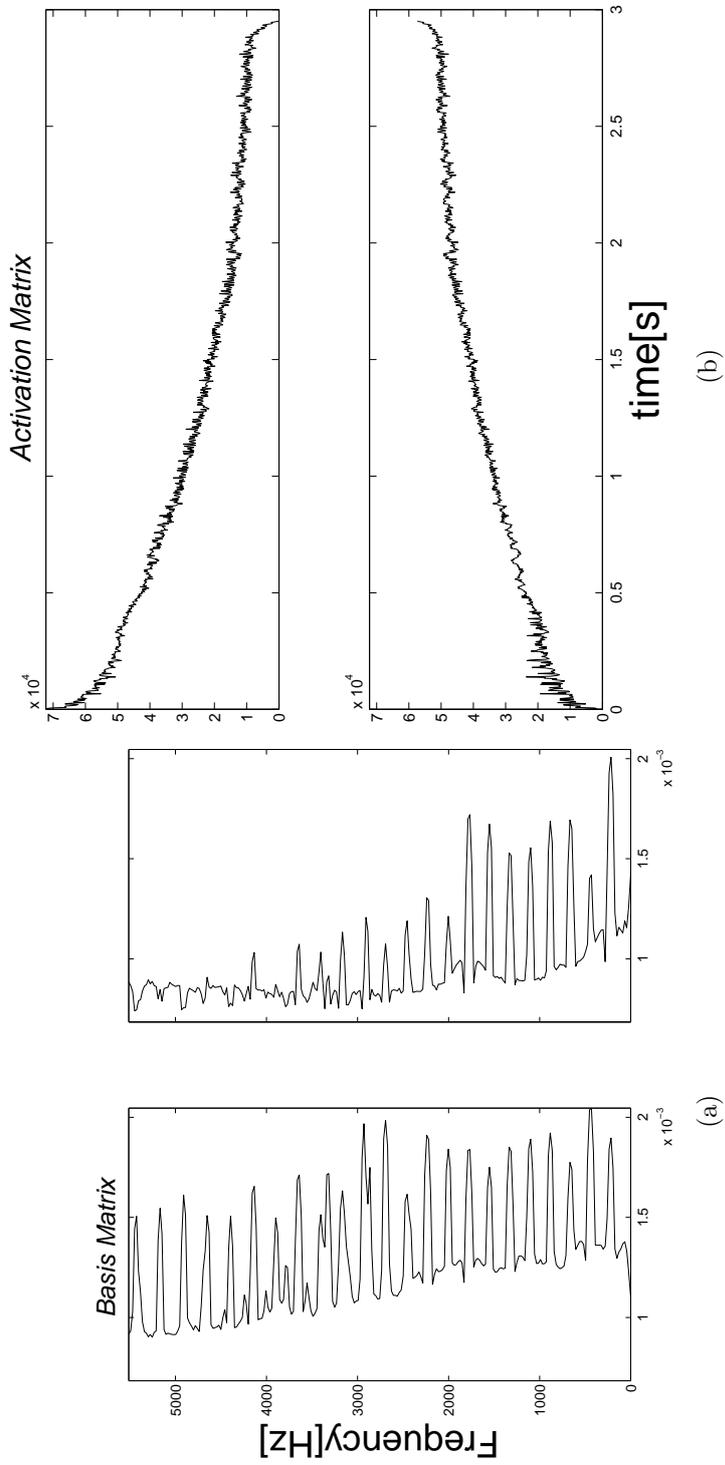


図 6: $K = 2$ の場合の分離結果 (MIDI)

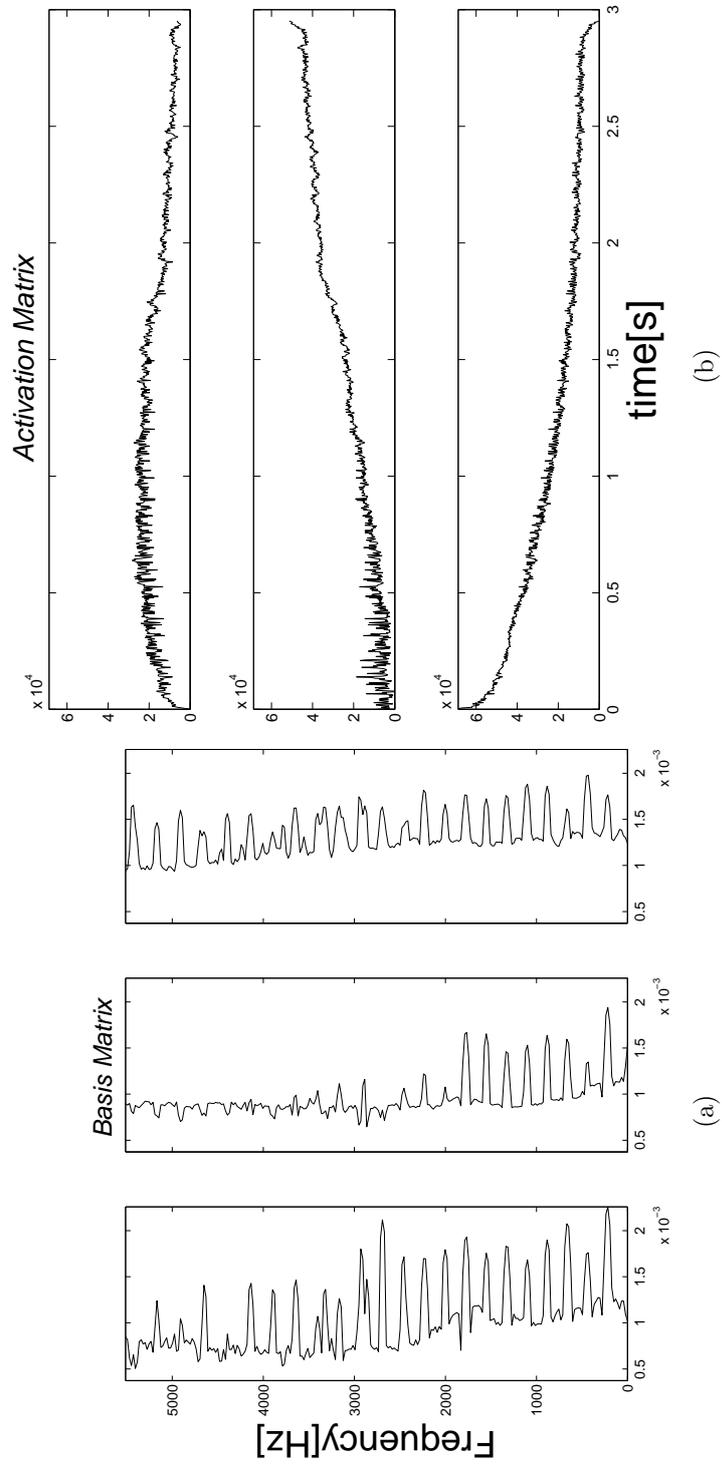


図 7: $K = 3$ の場合の分離結果 (MIDI)

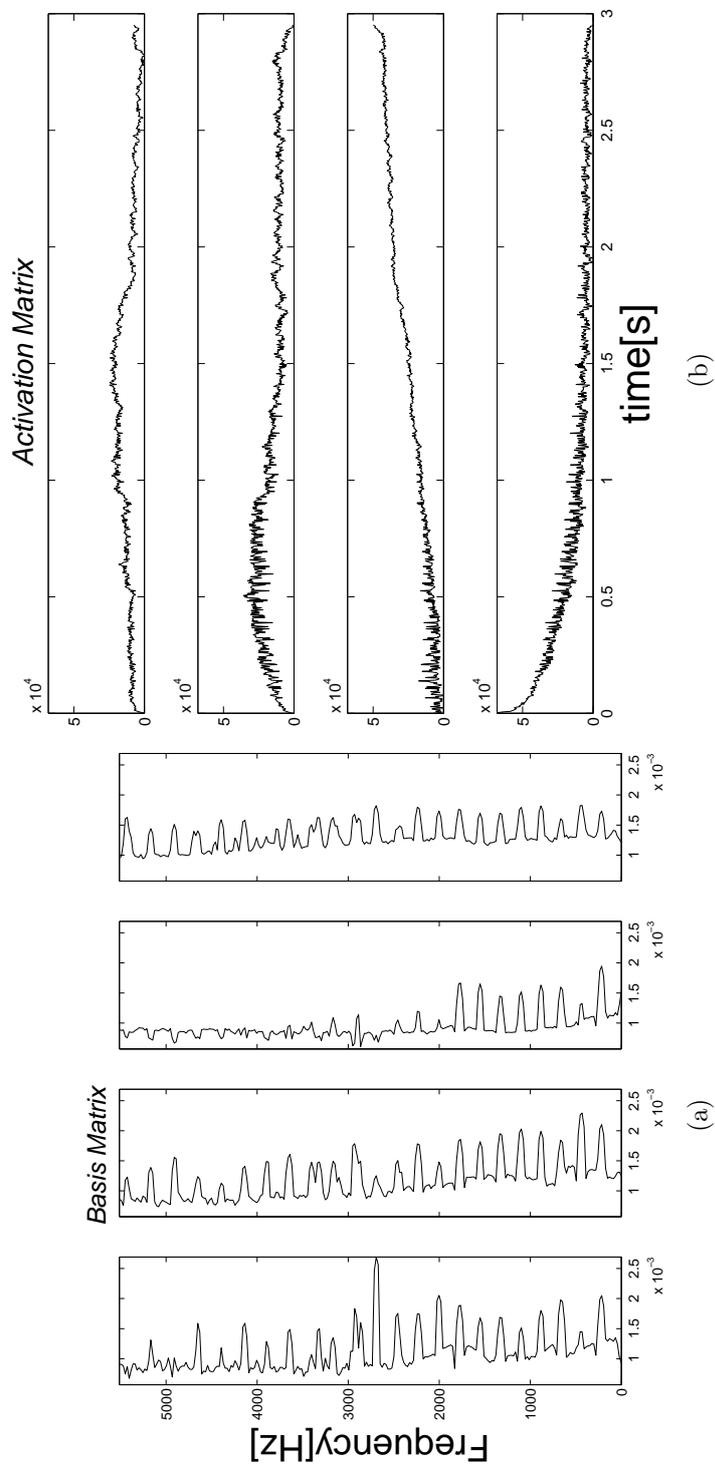


図 8: $K = 4$ の場合の分離結果 (MIDI)

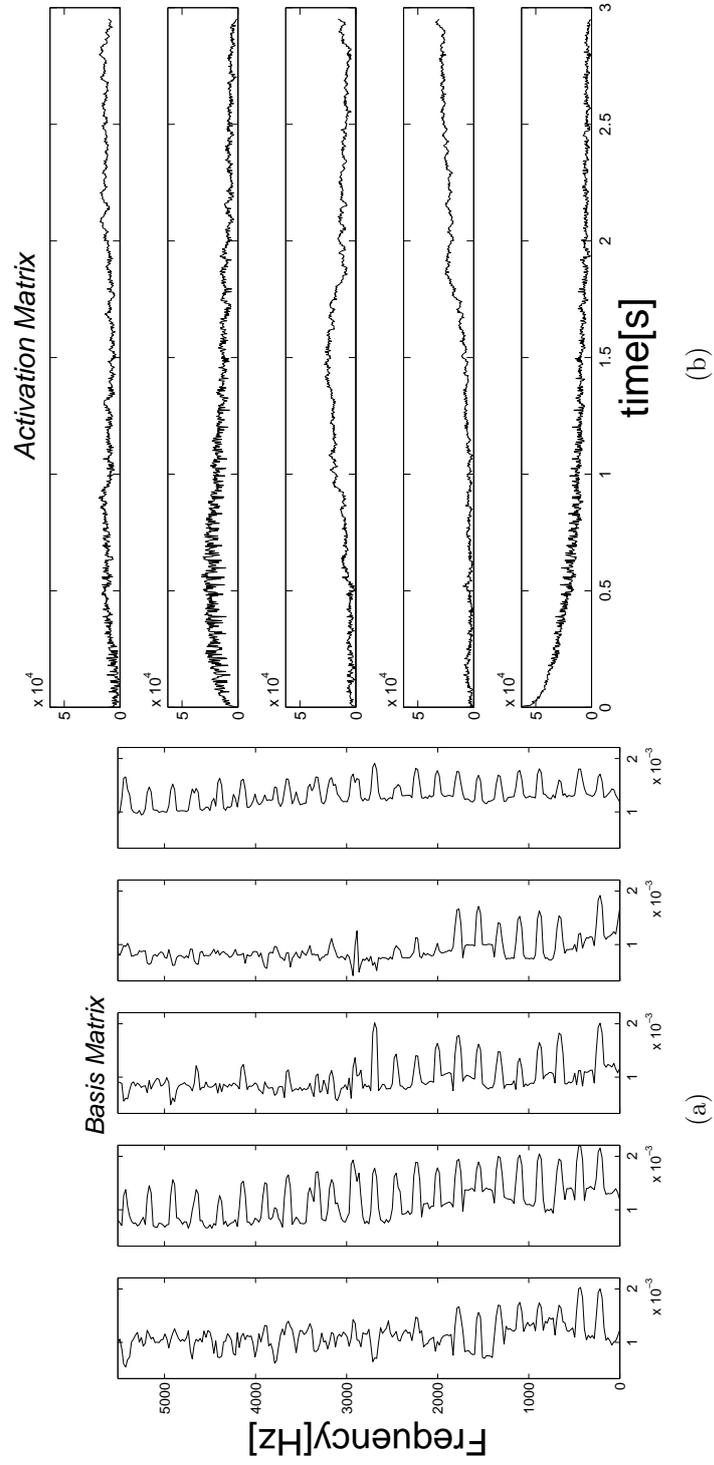


図 9: $K = 5$ の場合の分離結果 (MIDI)

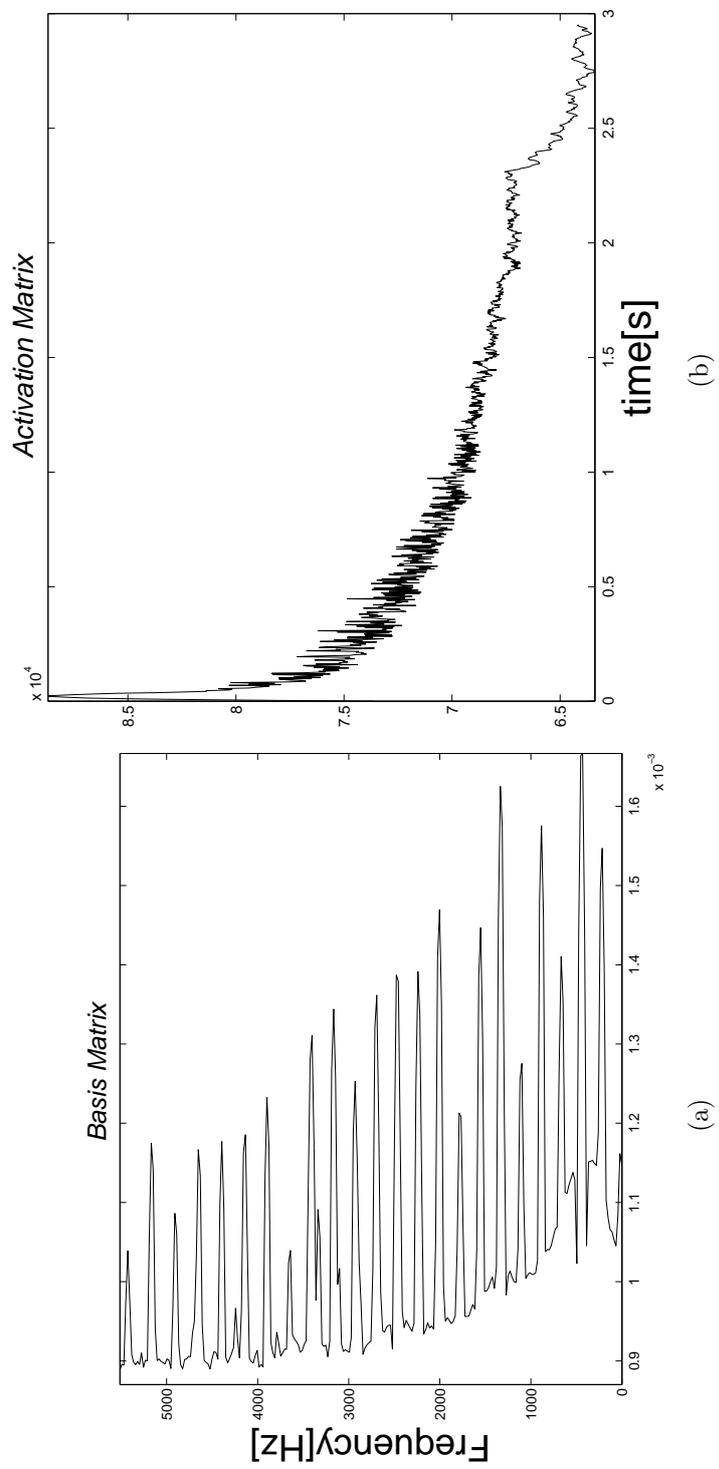


図 10: $K = 1$ の場合の分離結果 (UP1)

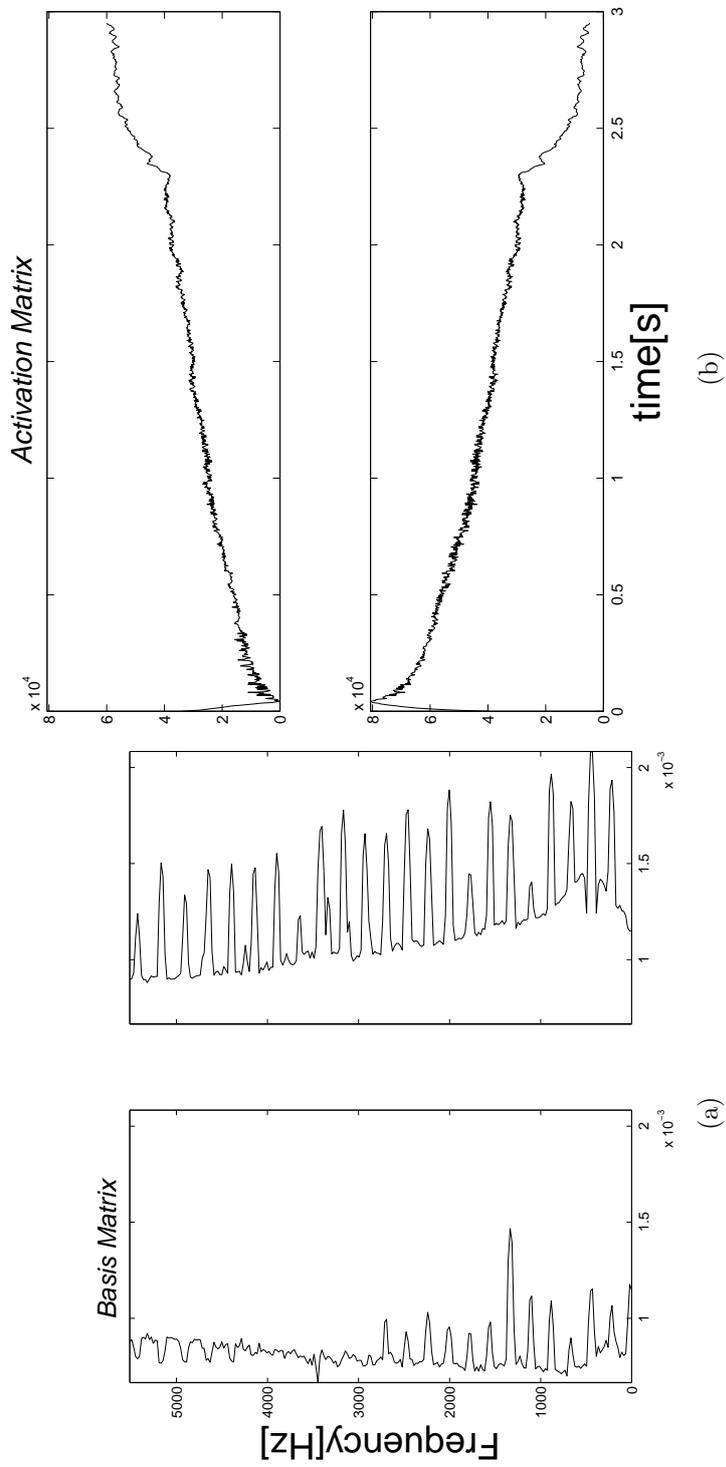


図 11: $K = 2$ の場合の分離結果 (UP1)

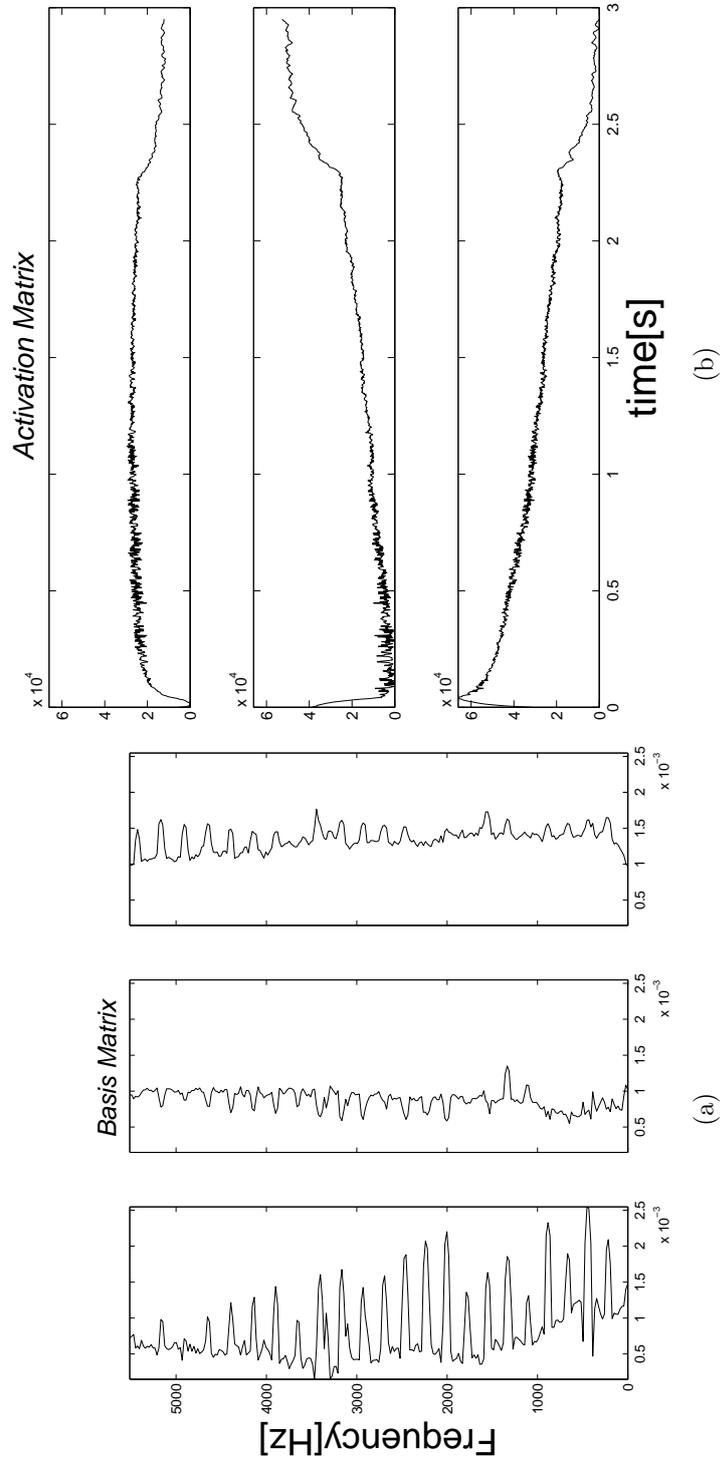


図 12: $K = 3$ の場合の分離結果 (UPI)

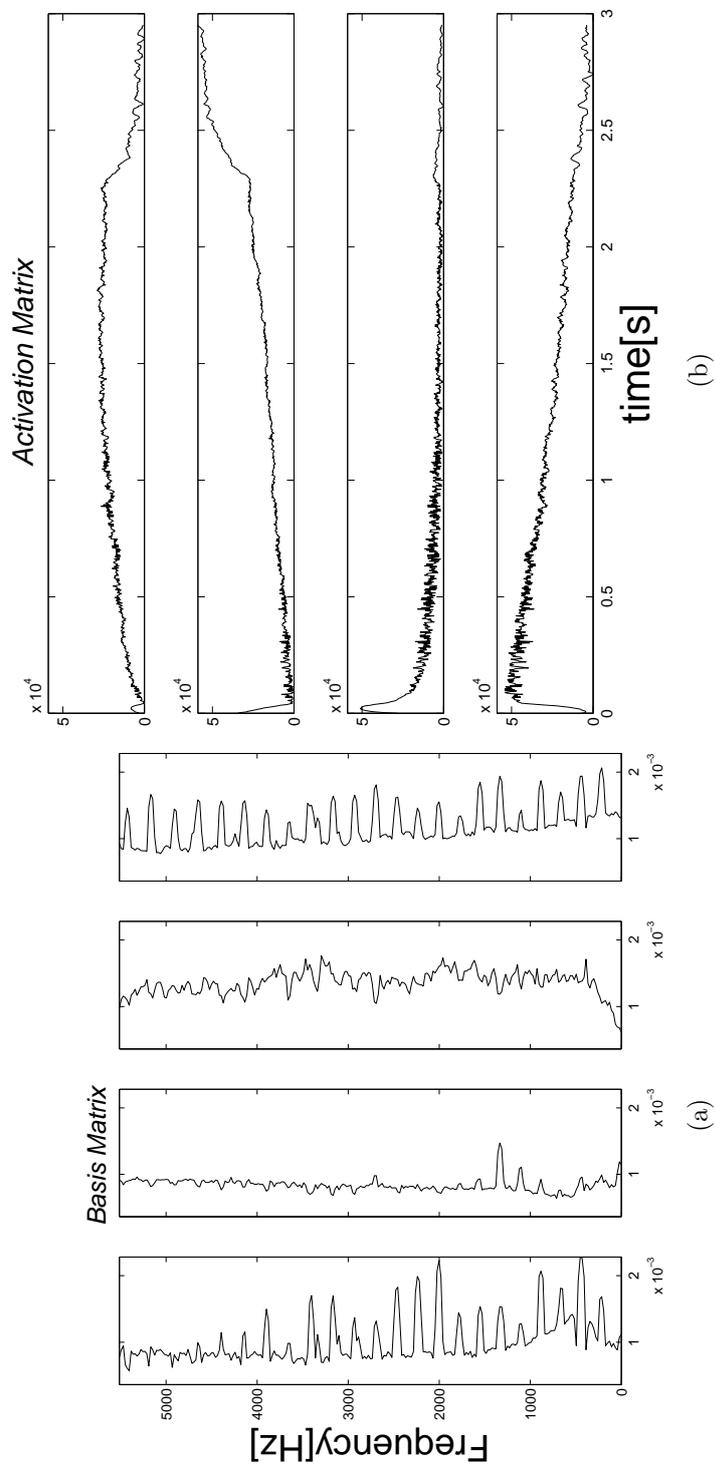


図 13: $K = 4$ の場合の分離結果 (UP1)

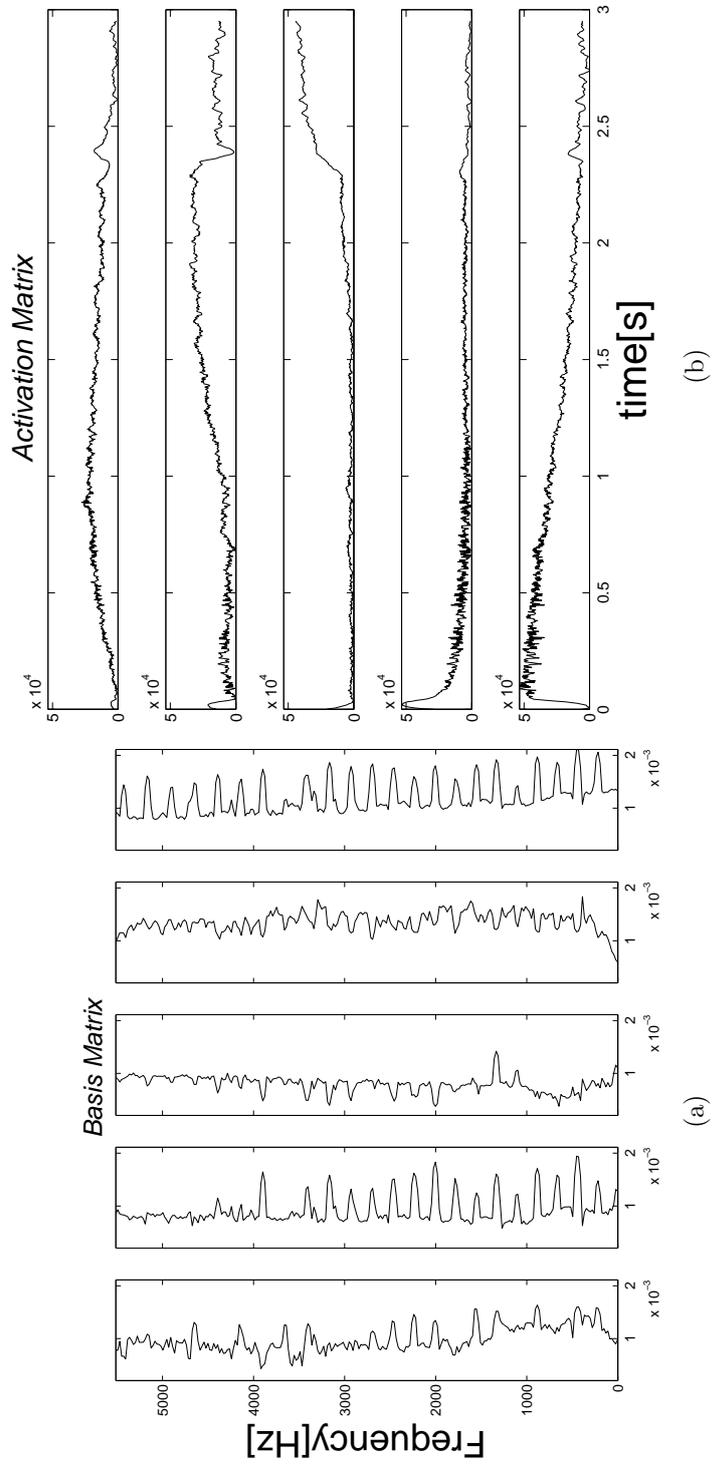


図 14: $K = 5$ の場合の分離結果 (UP1)

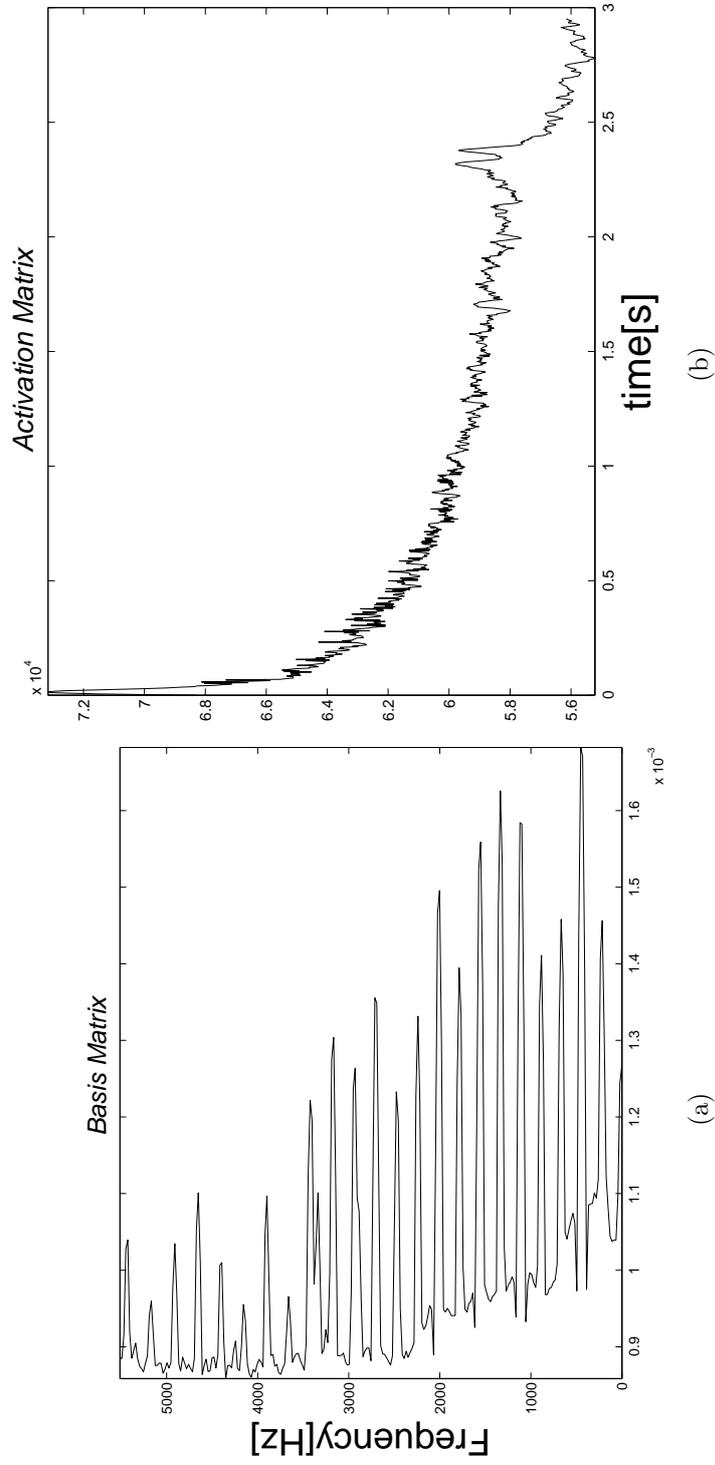


図 15: $K = 1$ の場合の分離結果 (UP2)

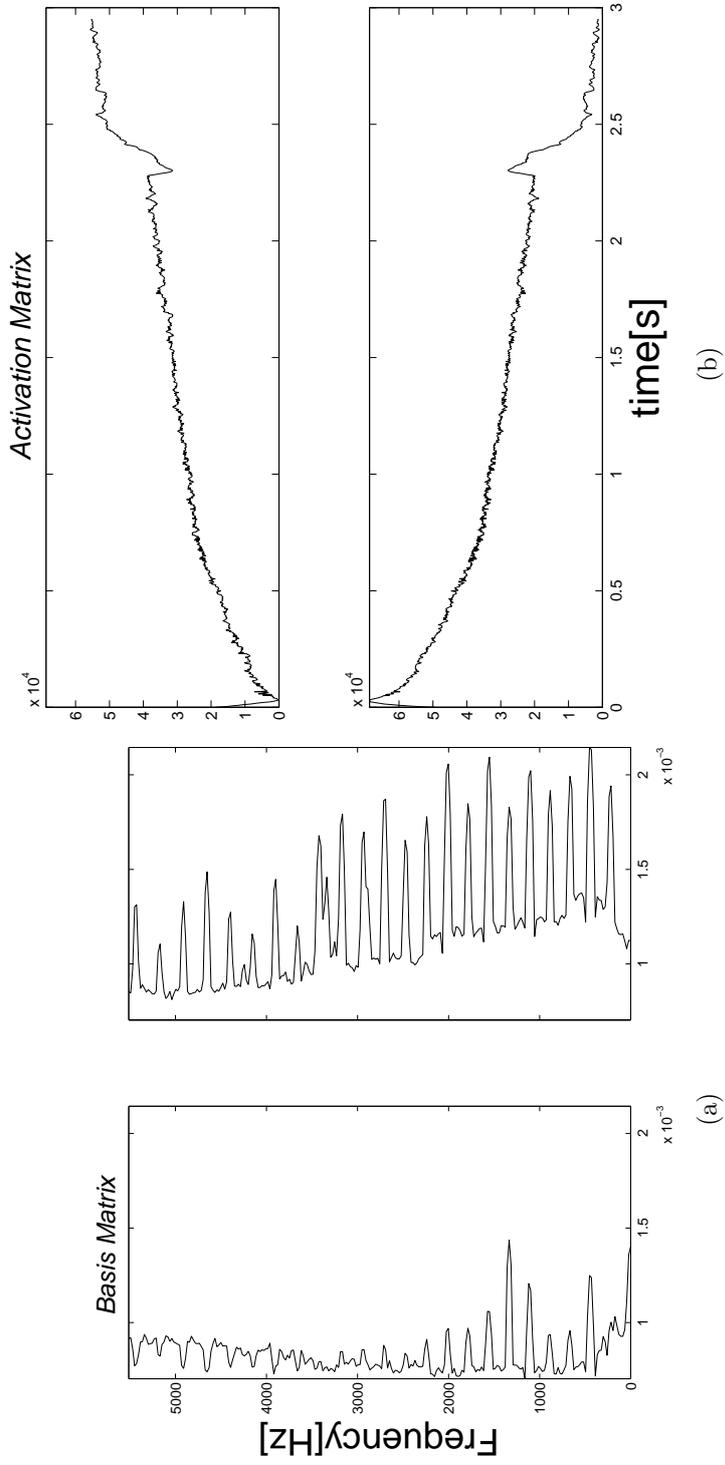


図 16: $K = 2$ の場合の分離結果 (UP2)

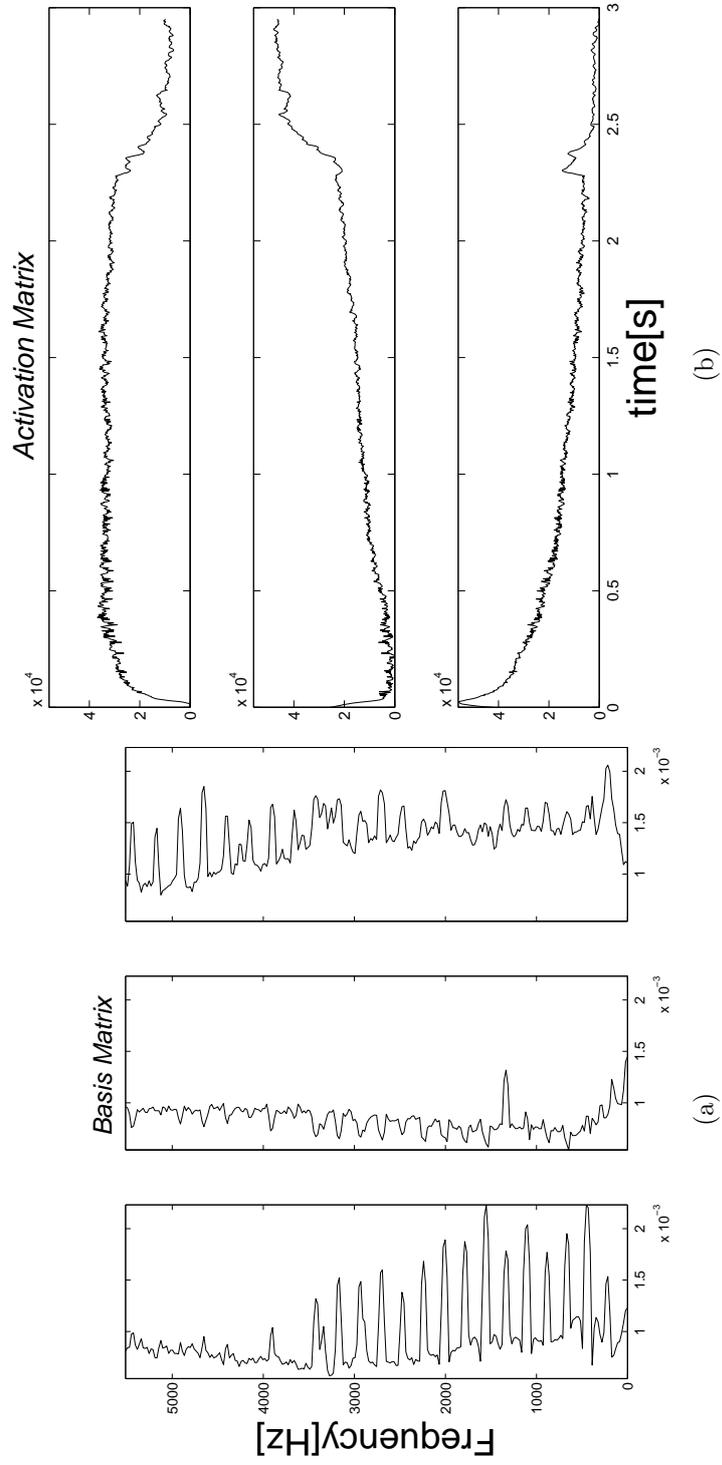


図 17: $K=3$ の場合の分離結果 (UP2)

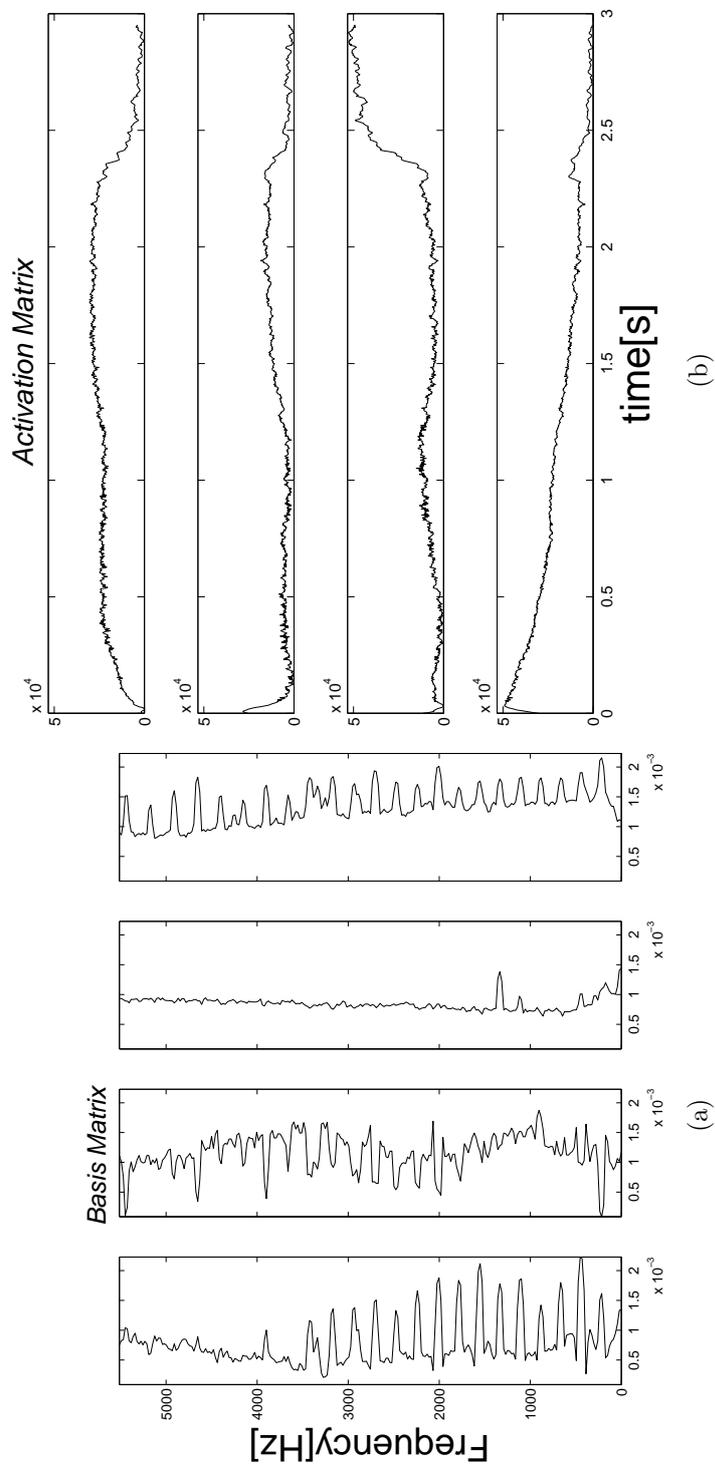


図 18: $K = 4$ の場合の分離結果 (UP2)

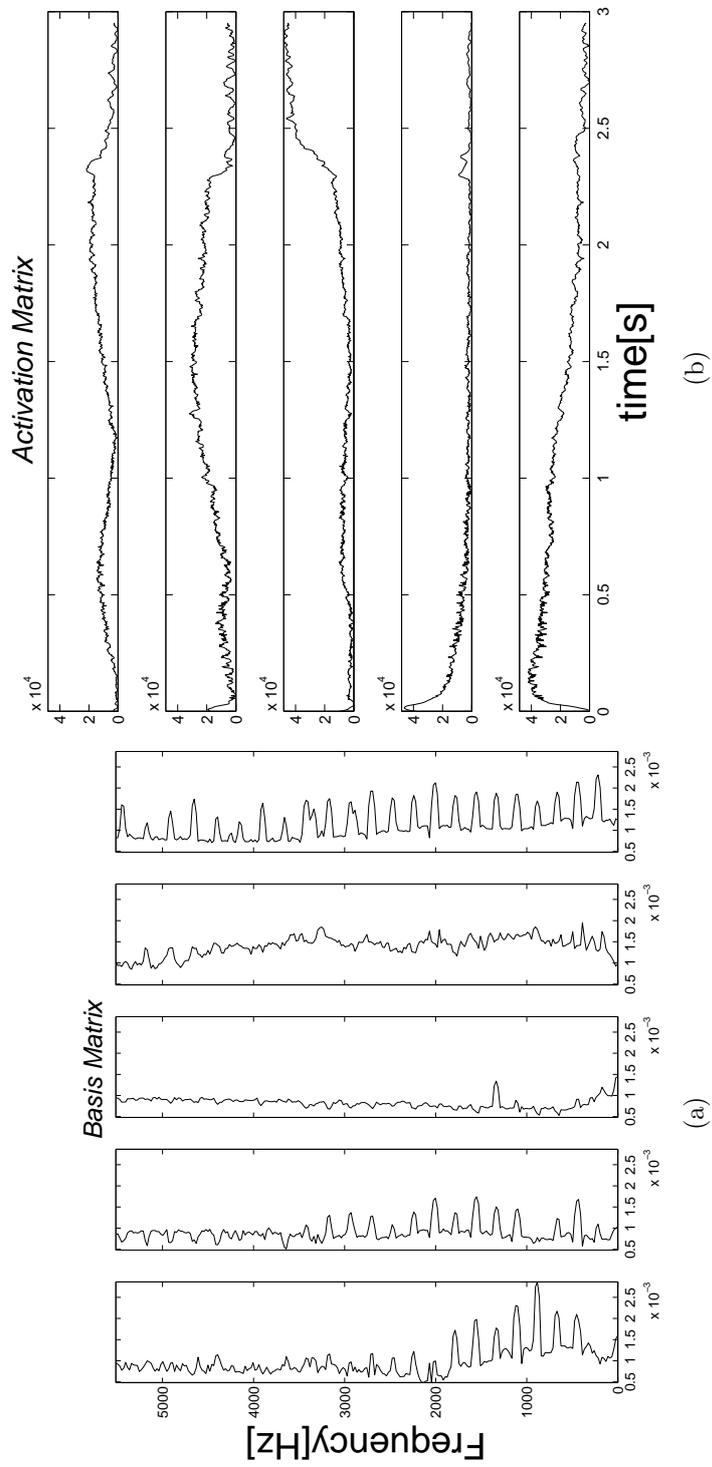


図 19: $K = 5$ の場合の分離結果 (UP2)

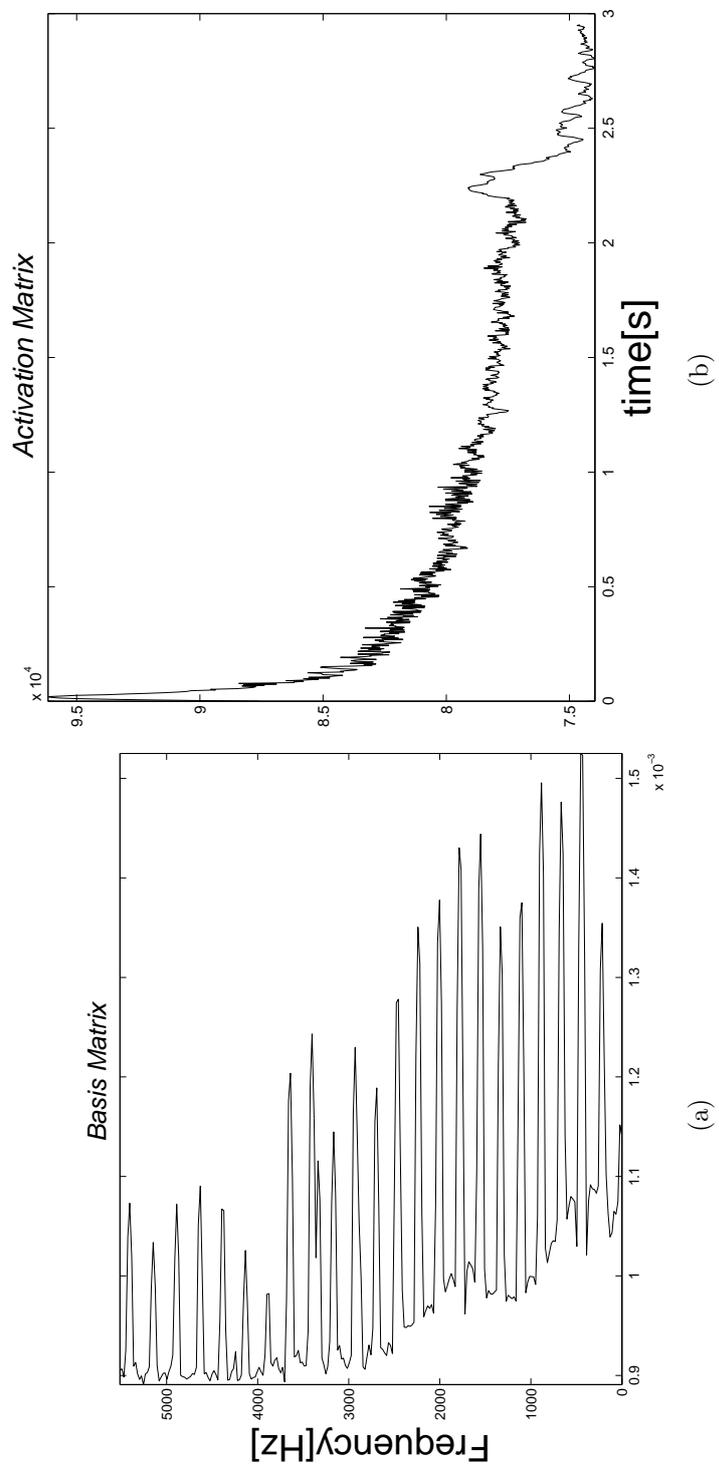


図 20: $K = 1$ の場合の分離結果 (UP3)

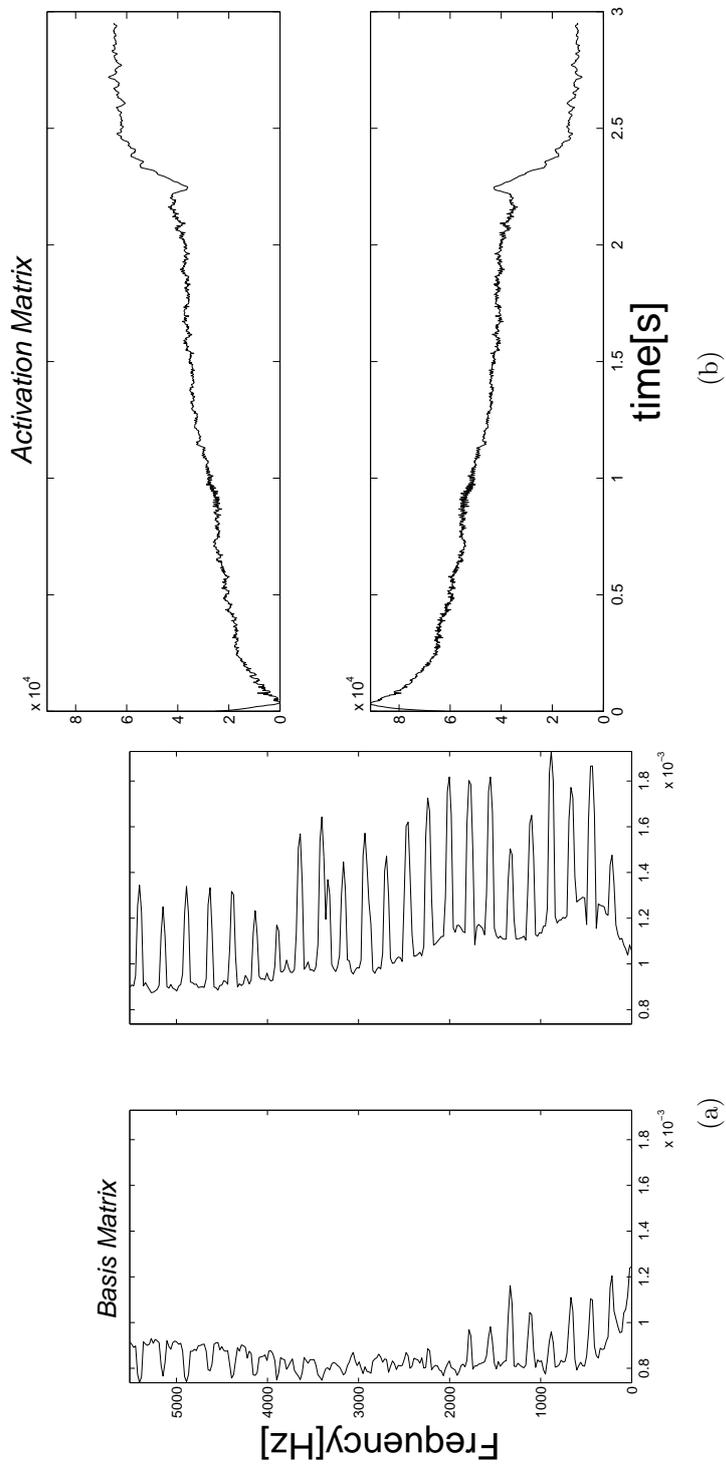


図 21: $K = 2$ の場合の分離結果 (UP3)

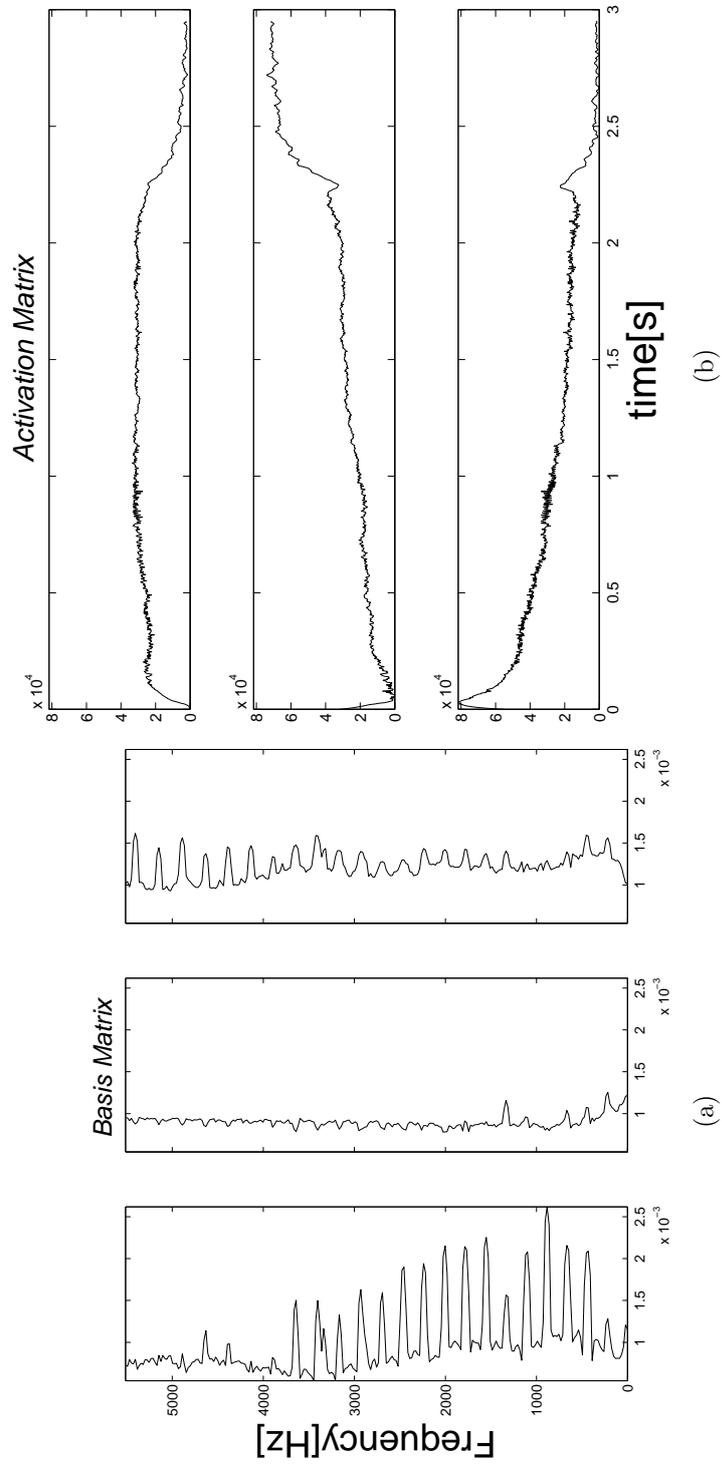


図 22: $K=3$ の場合の分離結果 (UP3)

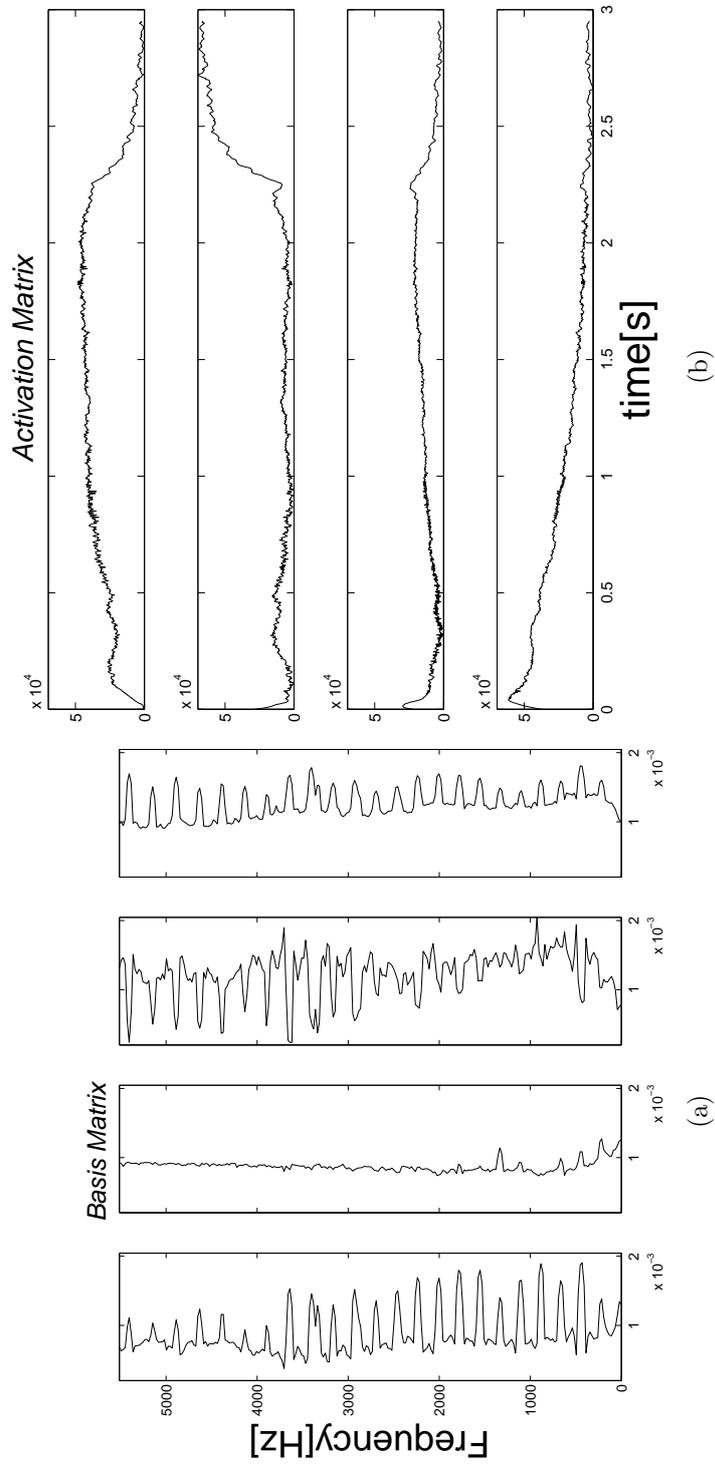


図 23: $K = 4$ の場合の分離結果 (UP3)

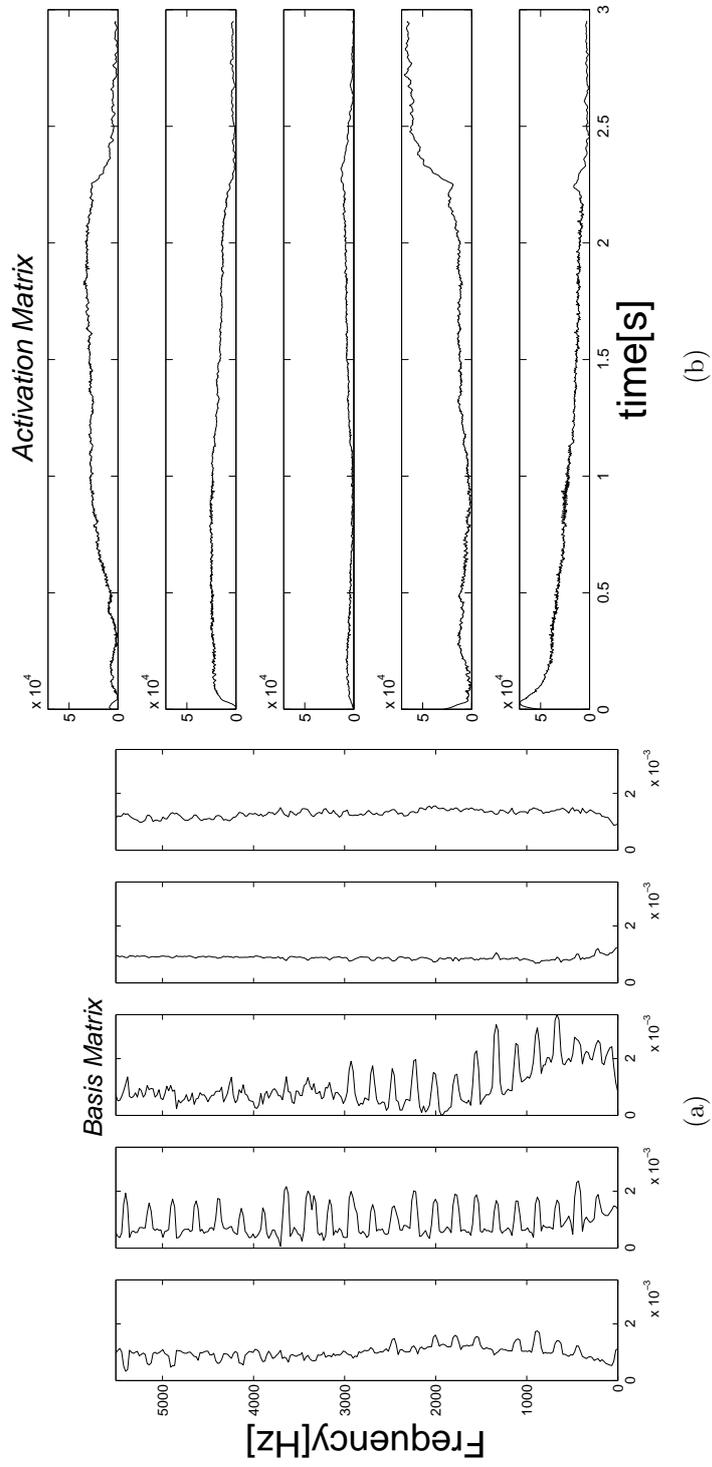


図 24: $K = 5$ の場合の分離結果 (UP3)

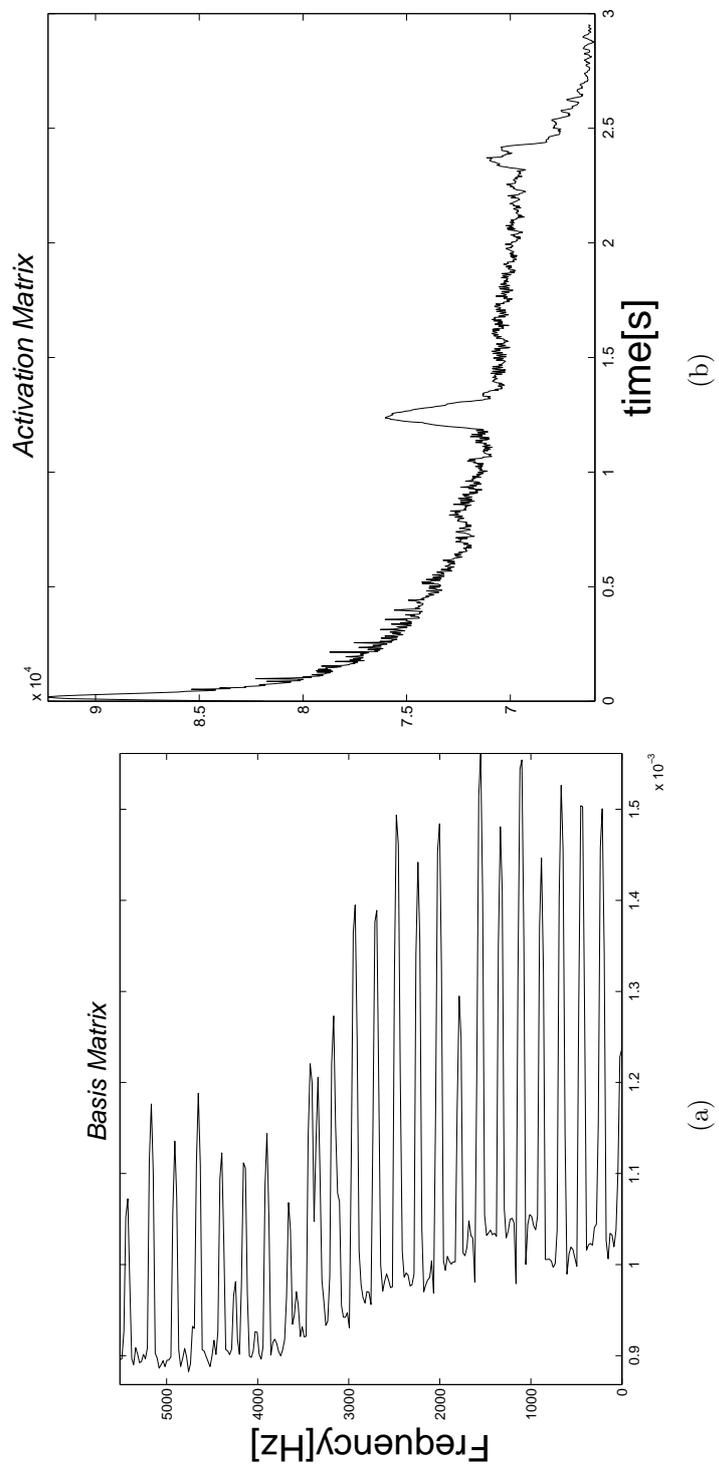


図 25: $K = 1$ の場合の分離結果 (UP4)

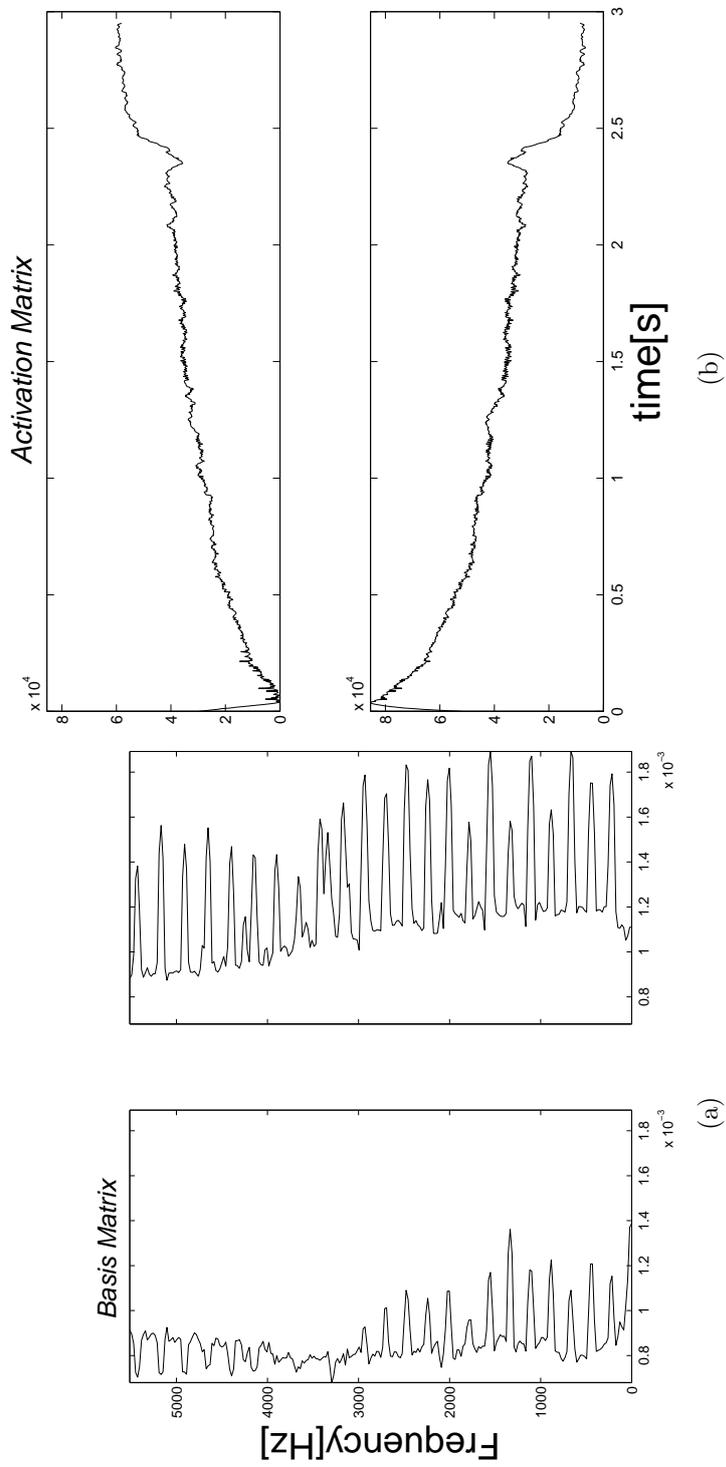


図 26: $K = 2$ の場合の分離結果 (UP4)

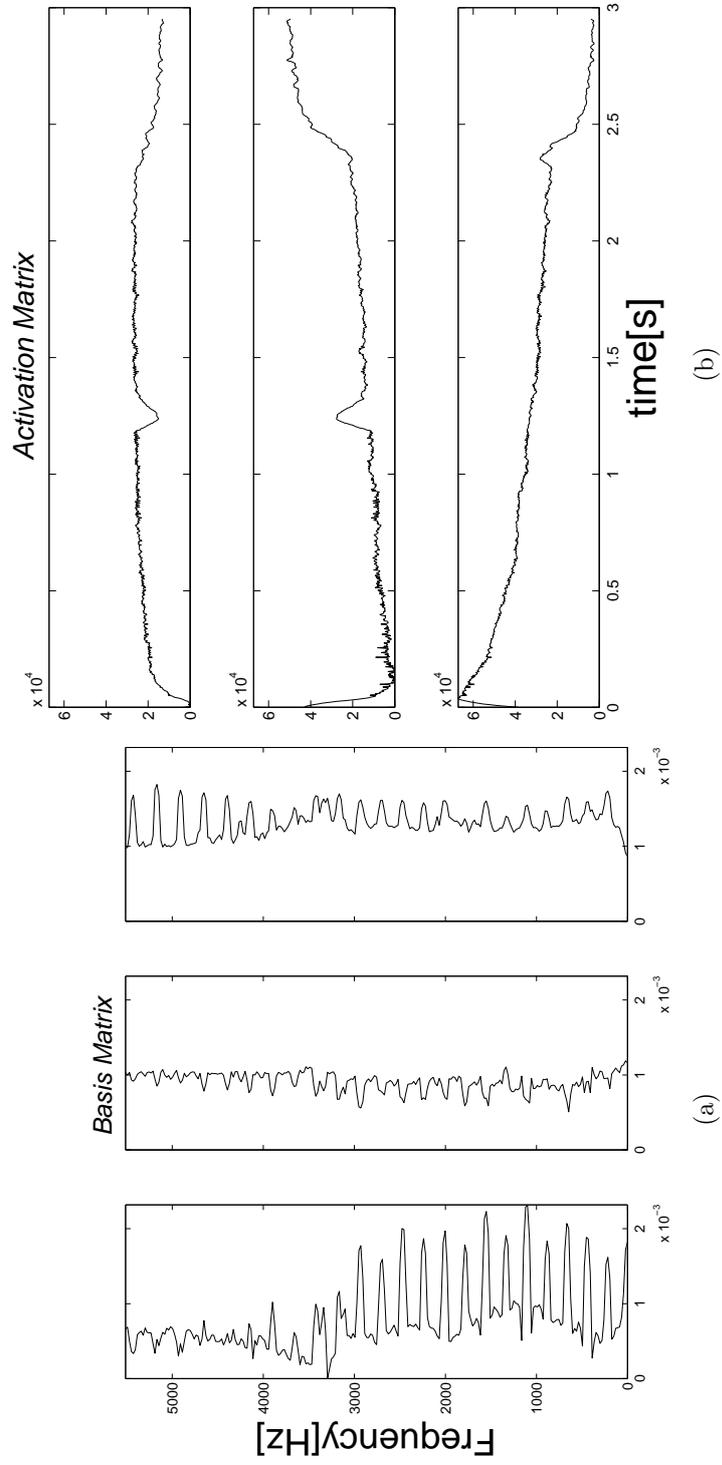


図 27: $K = 3$ の場合の分離結果 (UP4)

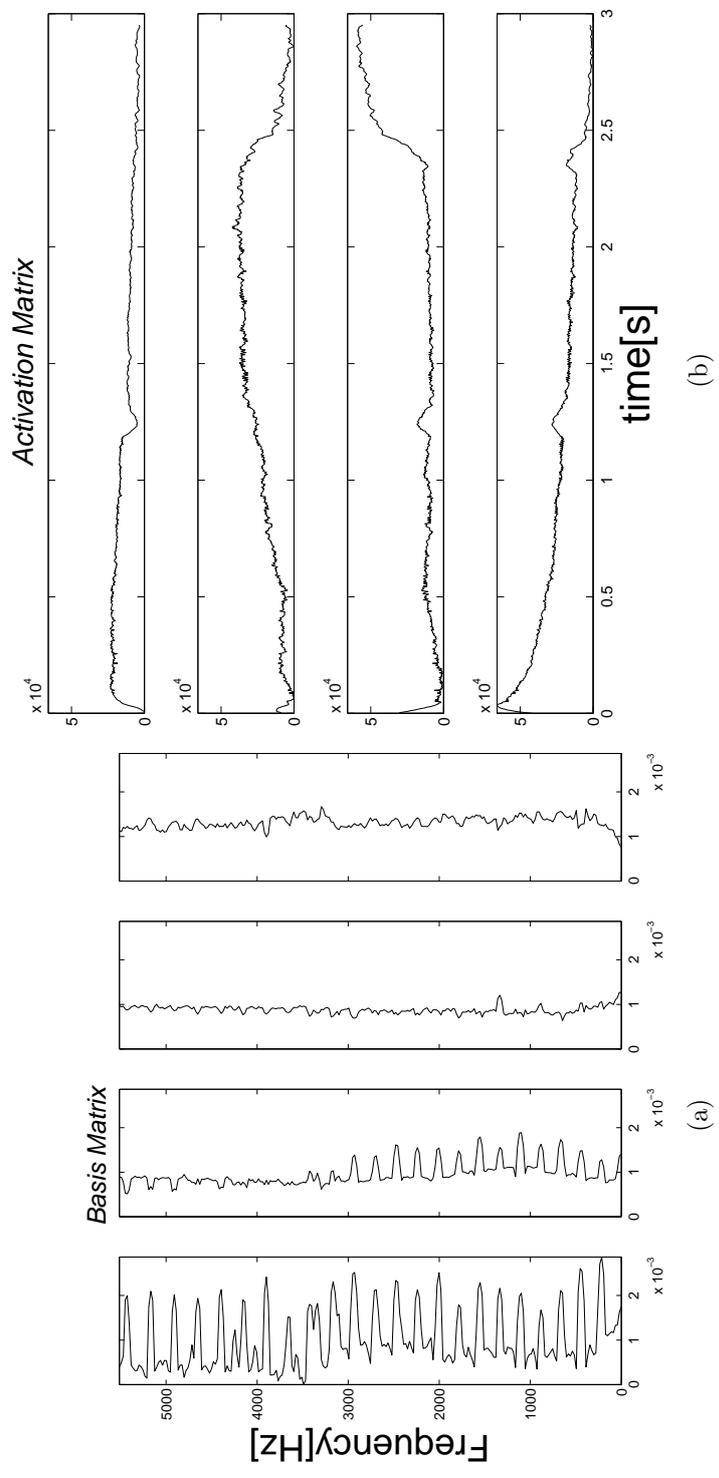


図 28: $K = 4$ の場合の分離結果 (UP4)

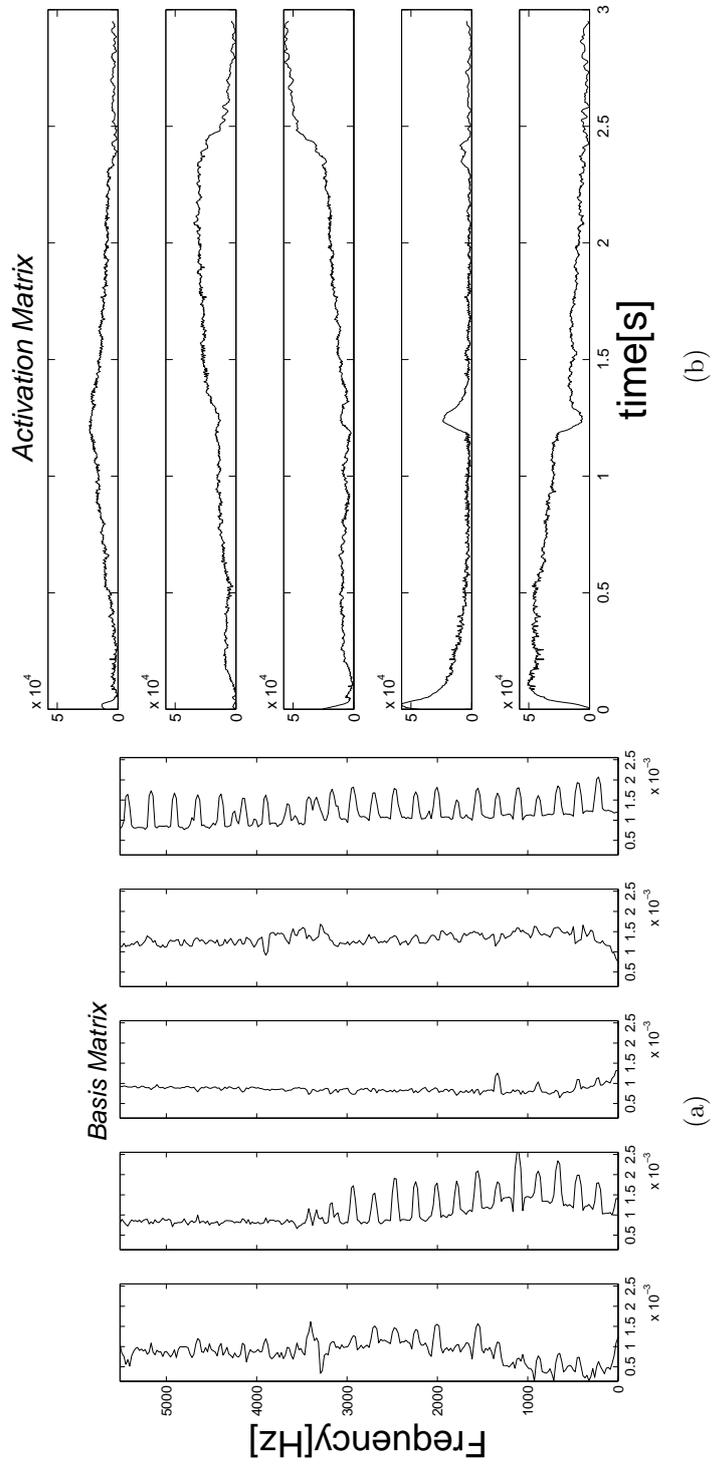


図 29: $K = 5$ の場合の分離結果 (UP4)

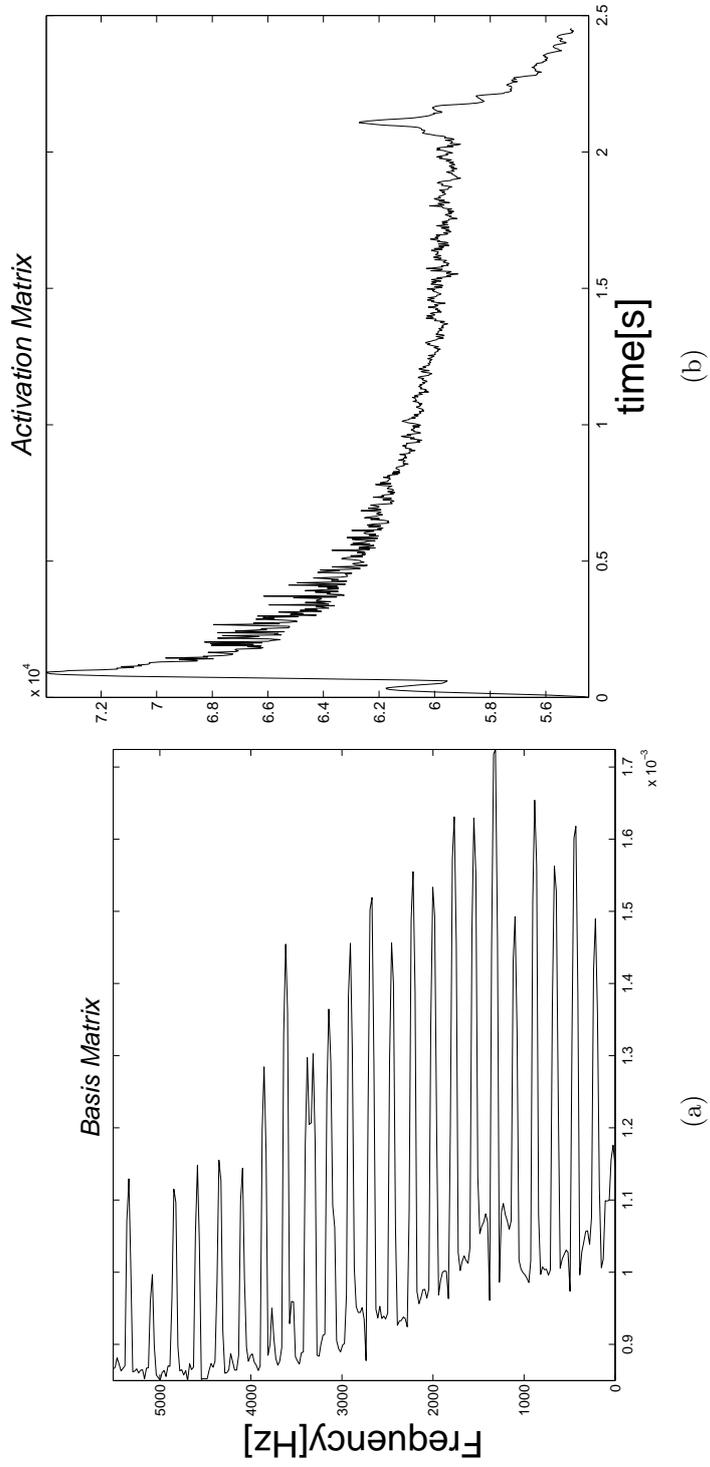


図 30: $K = 1$ の場合の分離結果 (MAPS)

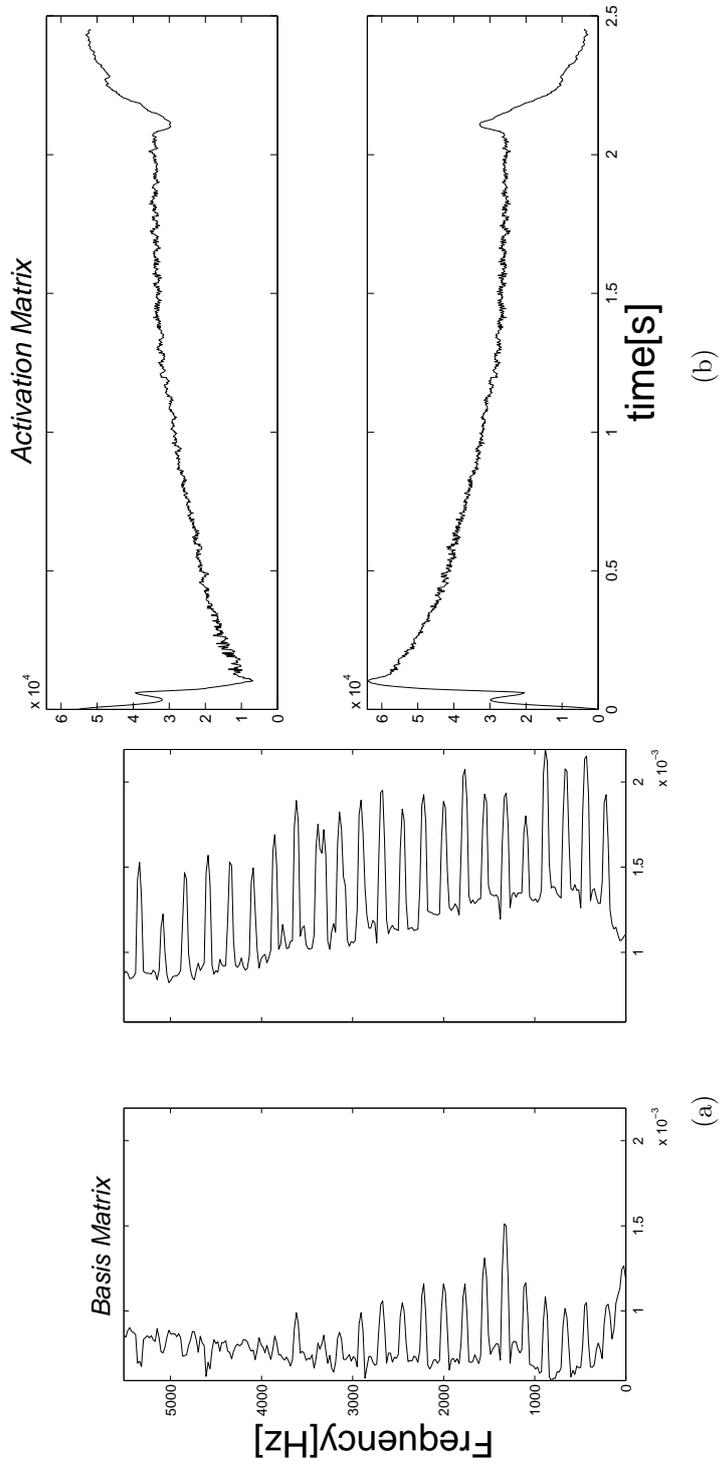


図 31: $K = 2$ の場合の分離結果 (MAPS)

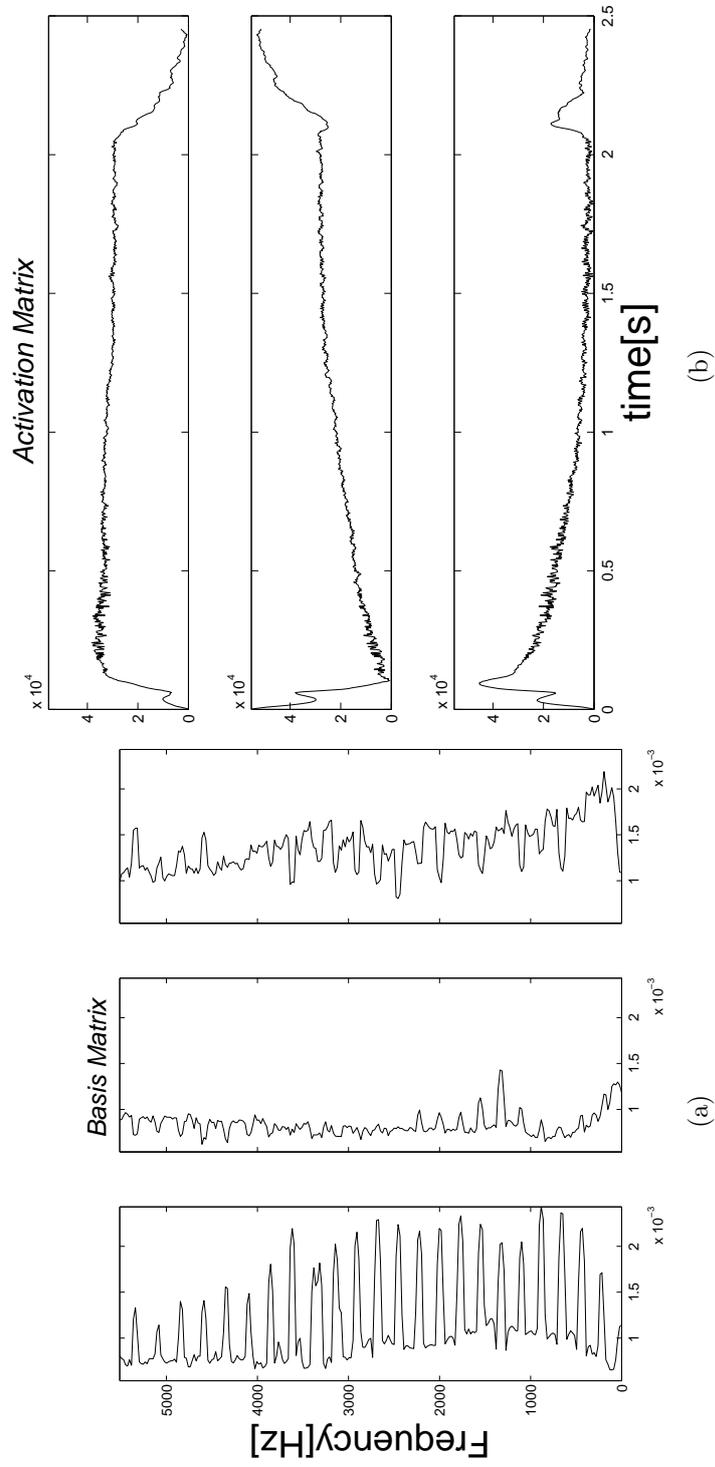


図 32: $K = 3$ の場合の分離結果 (MAPS)

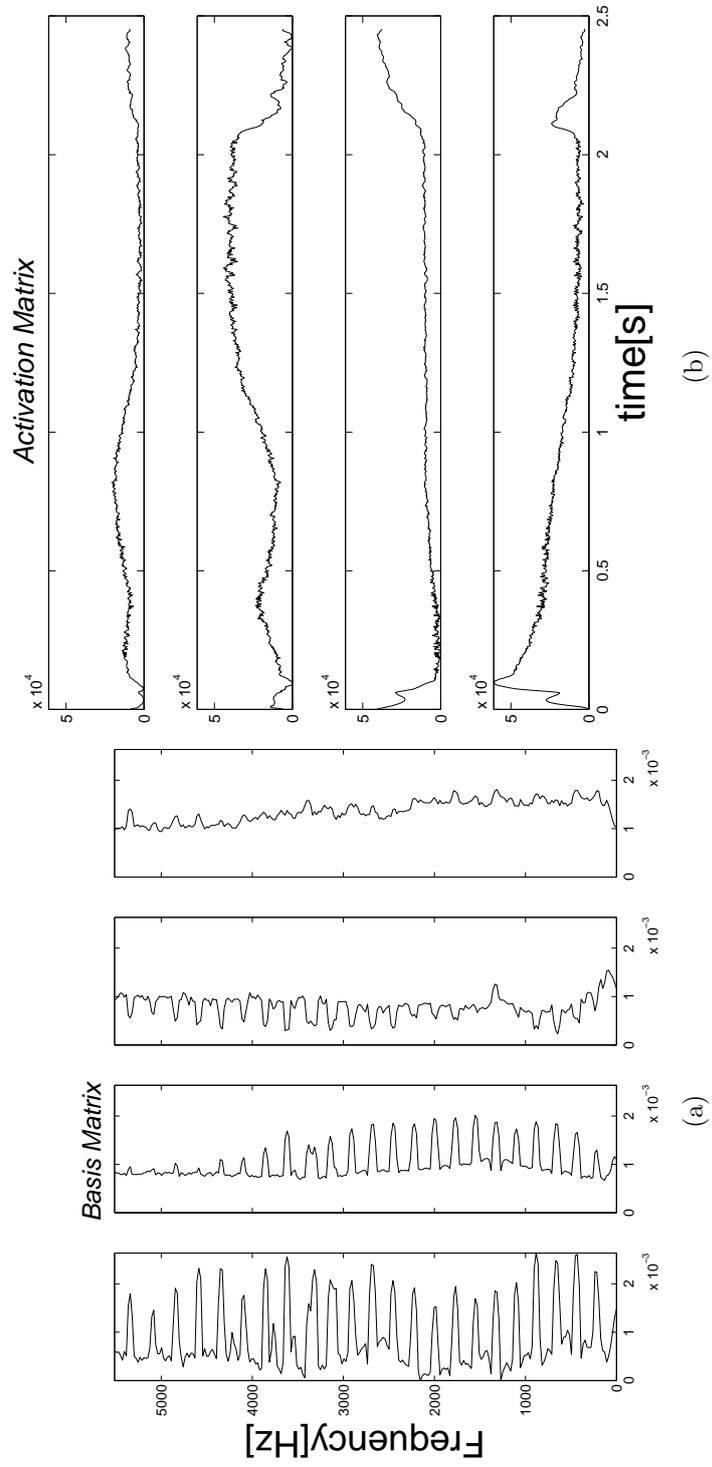


図 33: $K = 4$ の場合の分離結果 (MAPS)

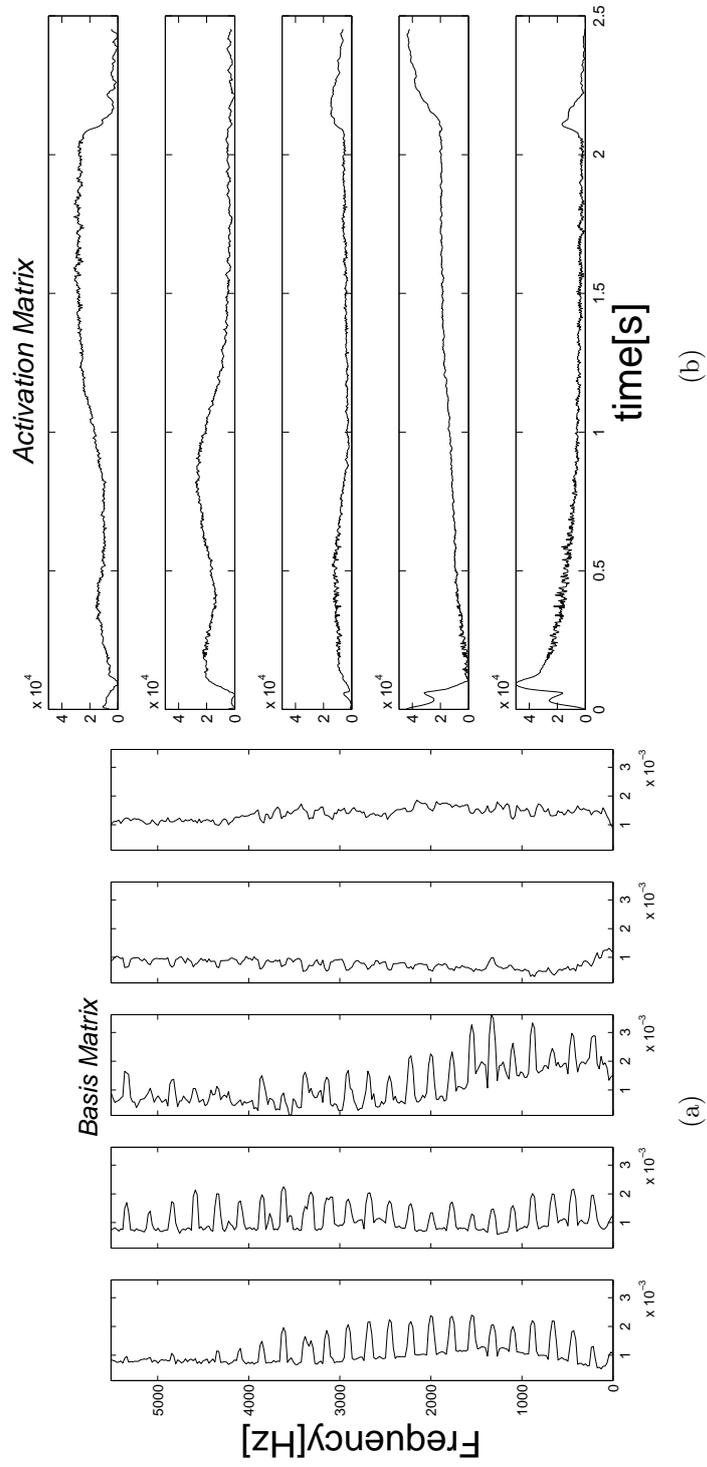


図 34: $K = 5$ の場合の分離結果 (MAPS)

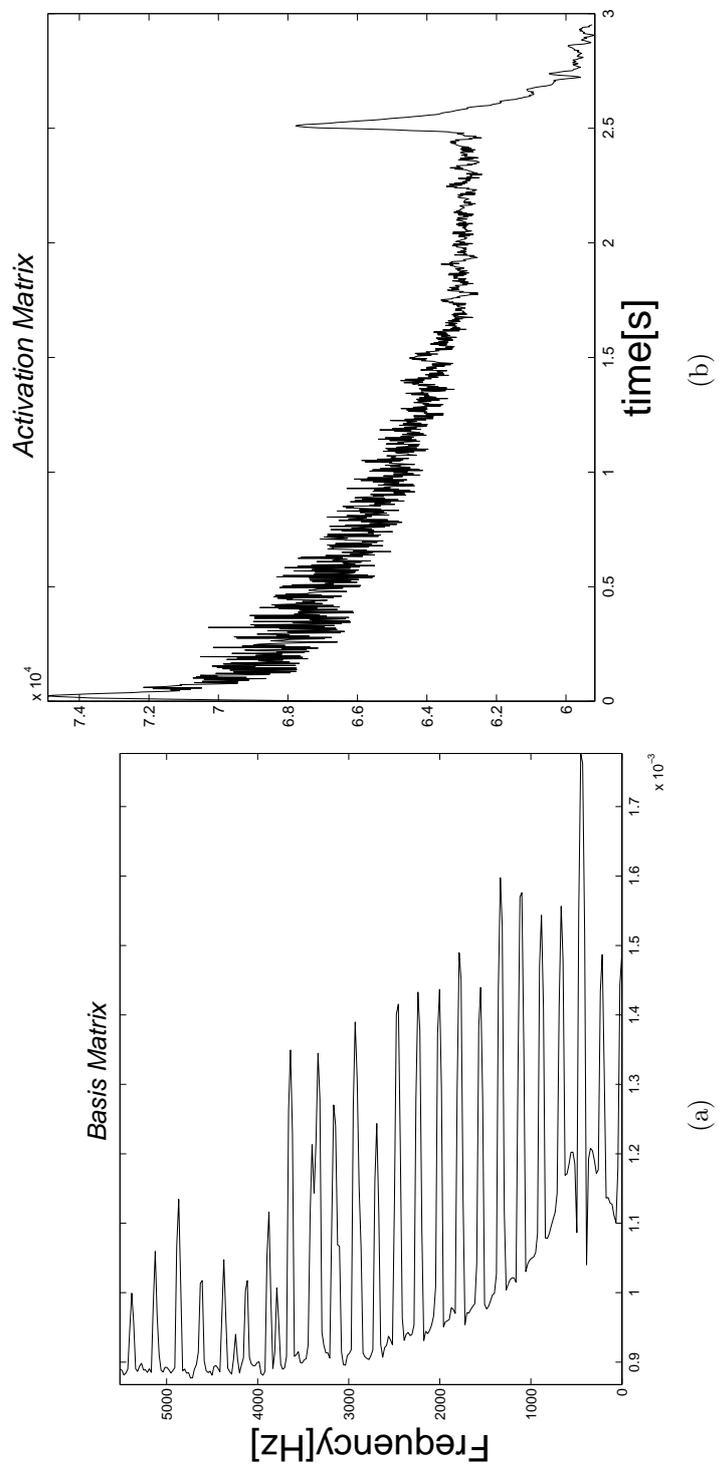


図 35: $K = 1$ の場合の分離結果 (GP1)

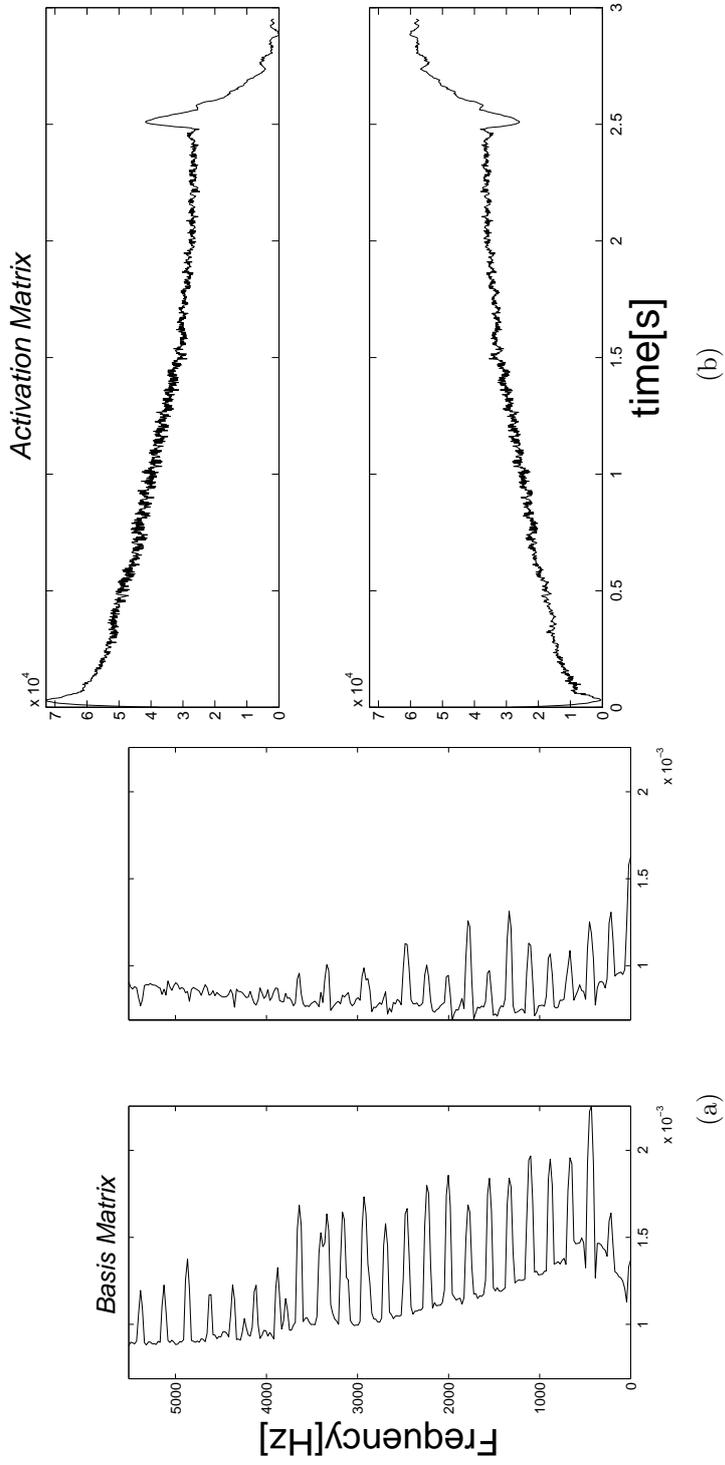


図 36: $K = 2$ の場合の分離結果 (GPI)

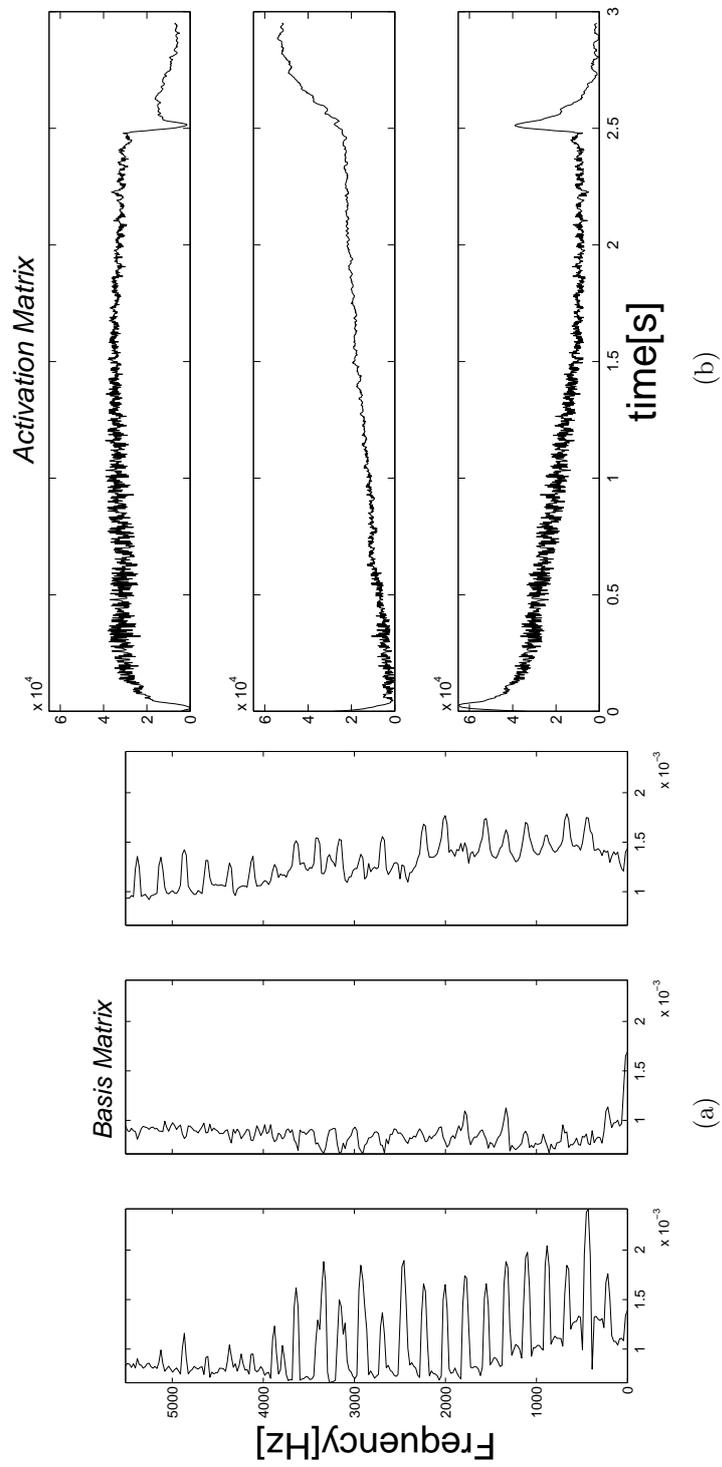


図 37: $K=3$ の場合の分離結果 (GPI)

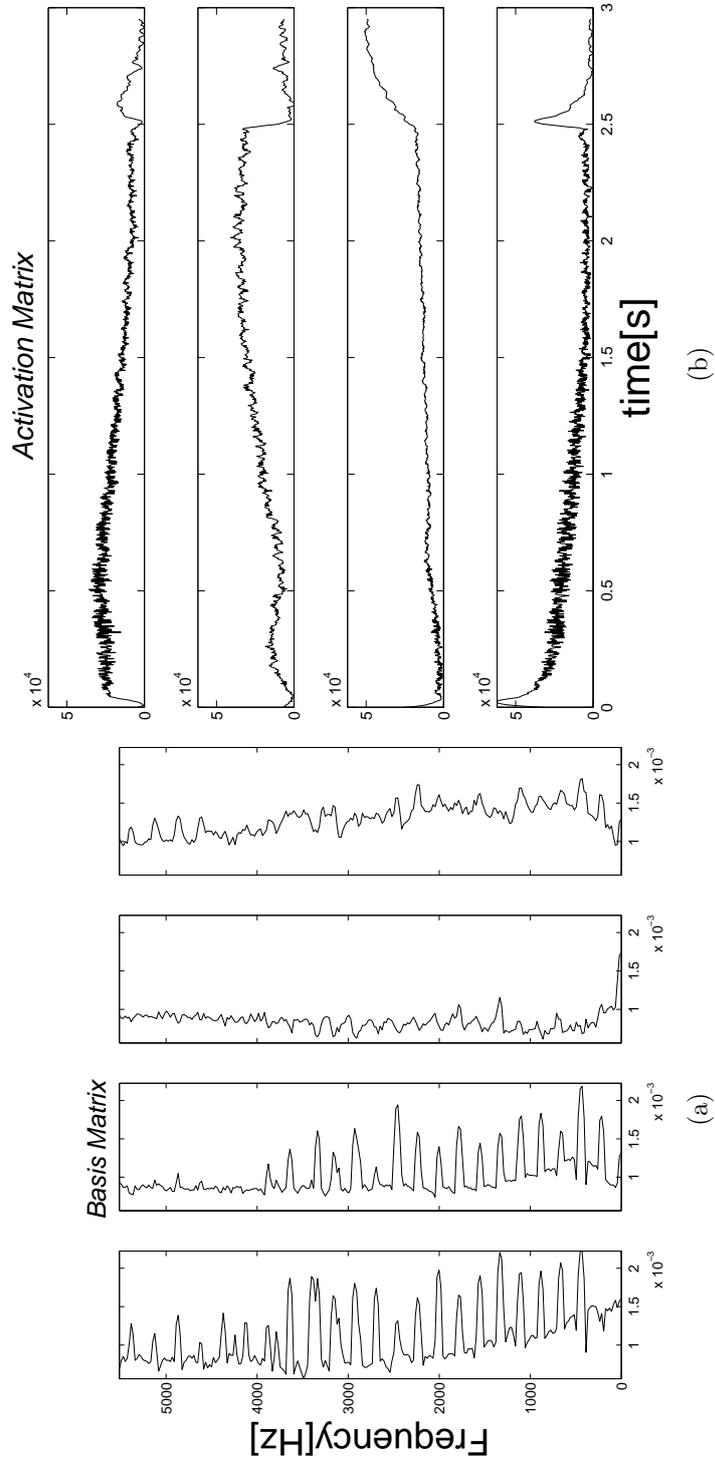


図 38: $K = 4$ の場合の分離結果 (GPI)

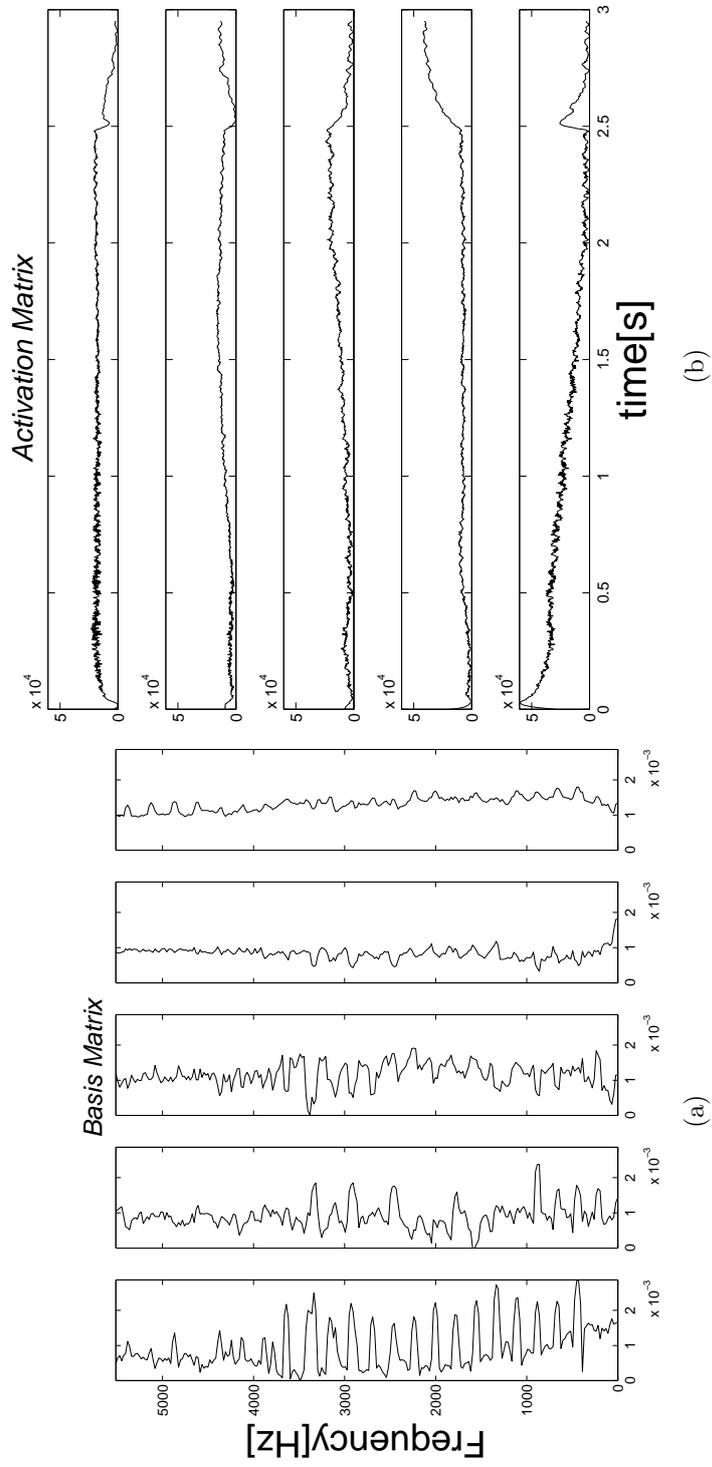


図 39: $K = 5$ の場合の分離結果 (GP1)

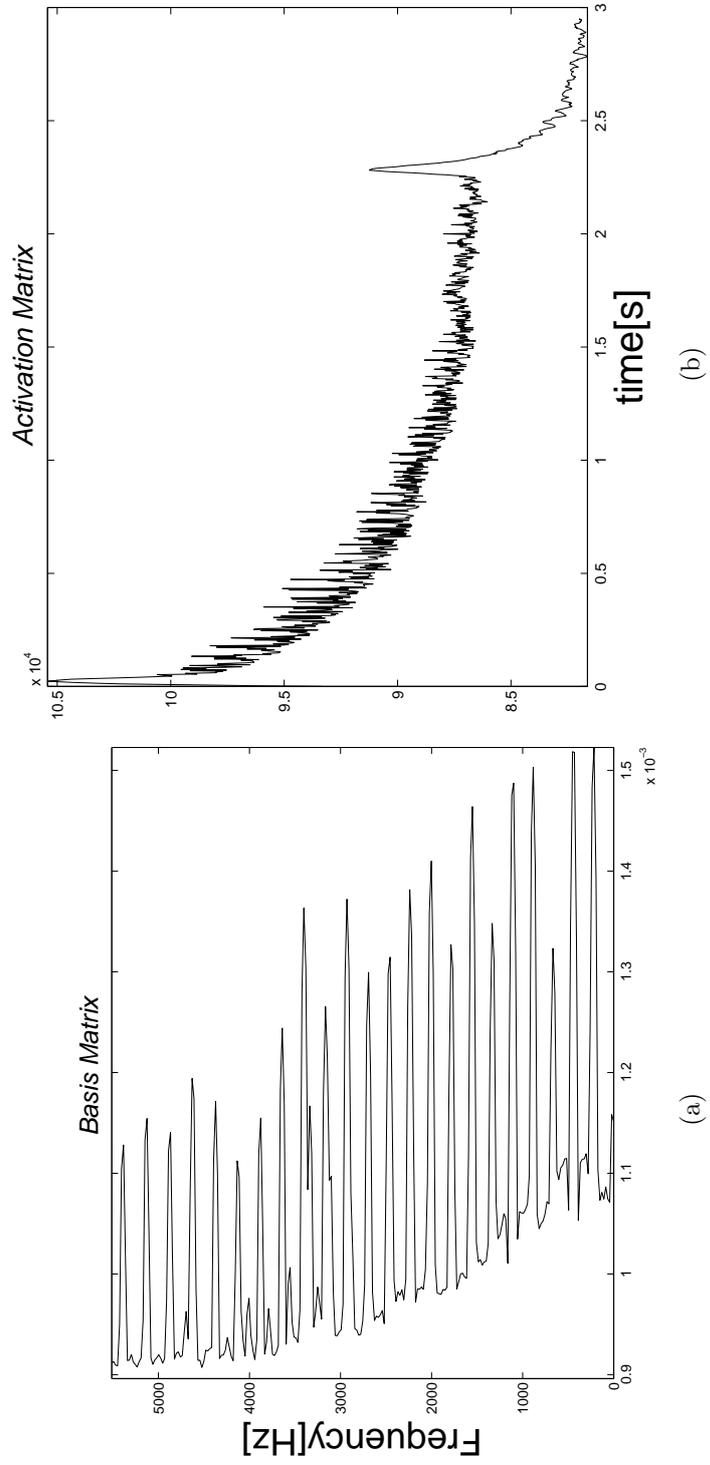


図 40: $K = 1$ の場合の分離結果 (GP2)

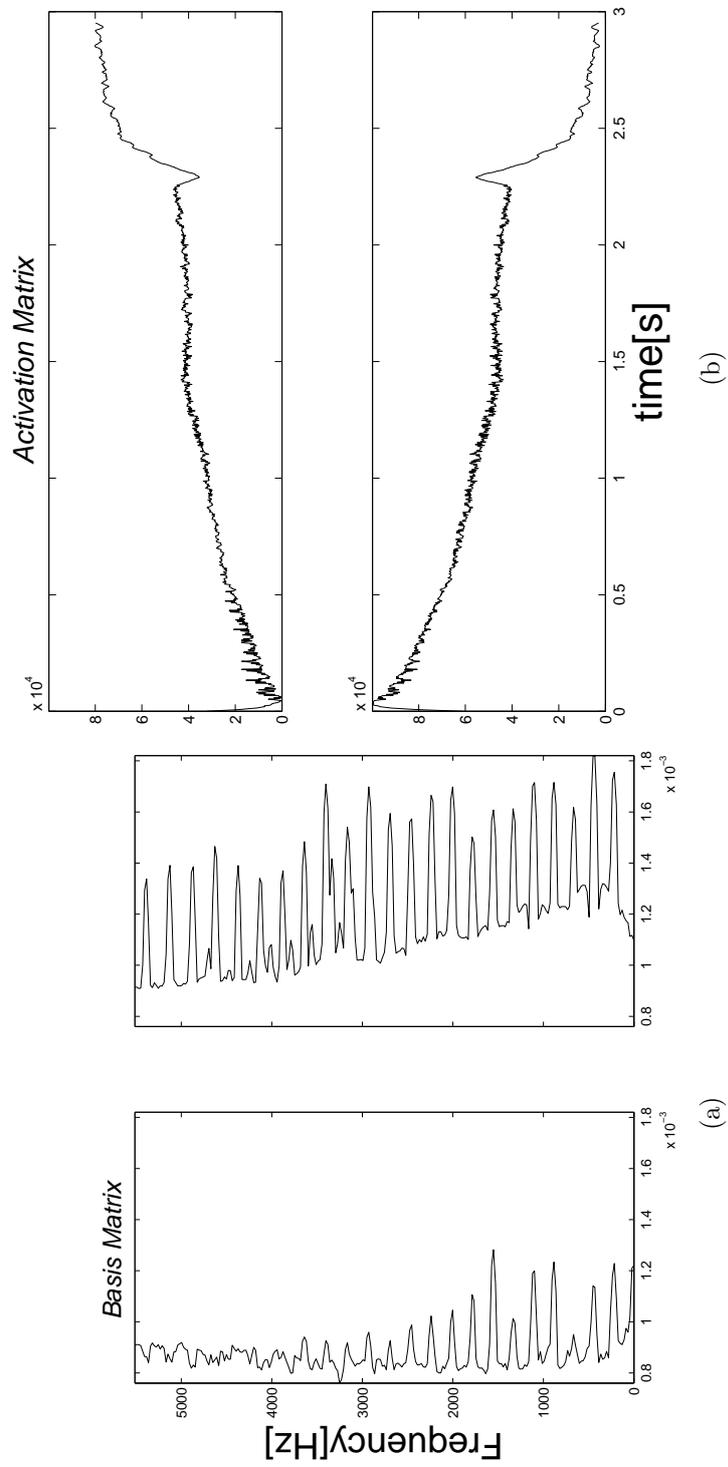


図 41: $K = 2$ の場合の分離結果 (GP2)

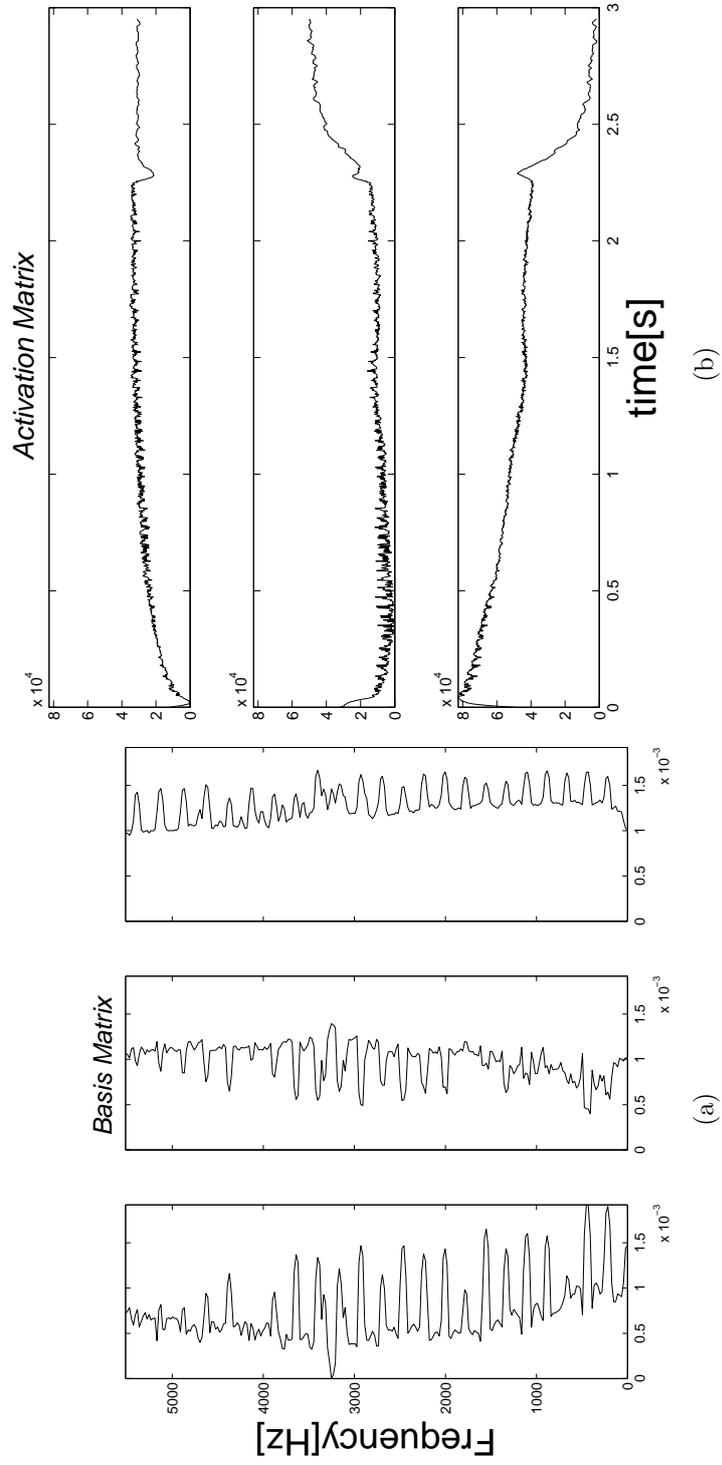


図 42: $K=3$ の場合の分離結果 (GP2)

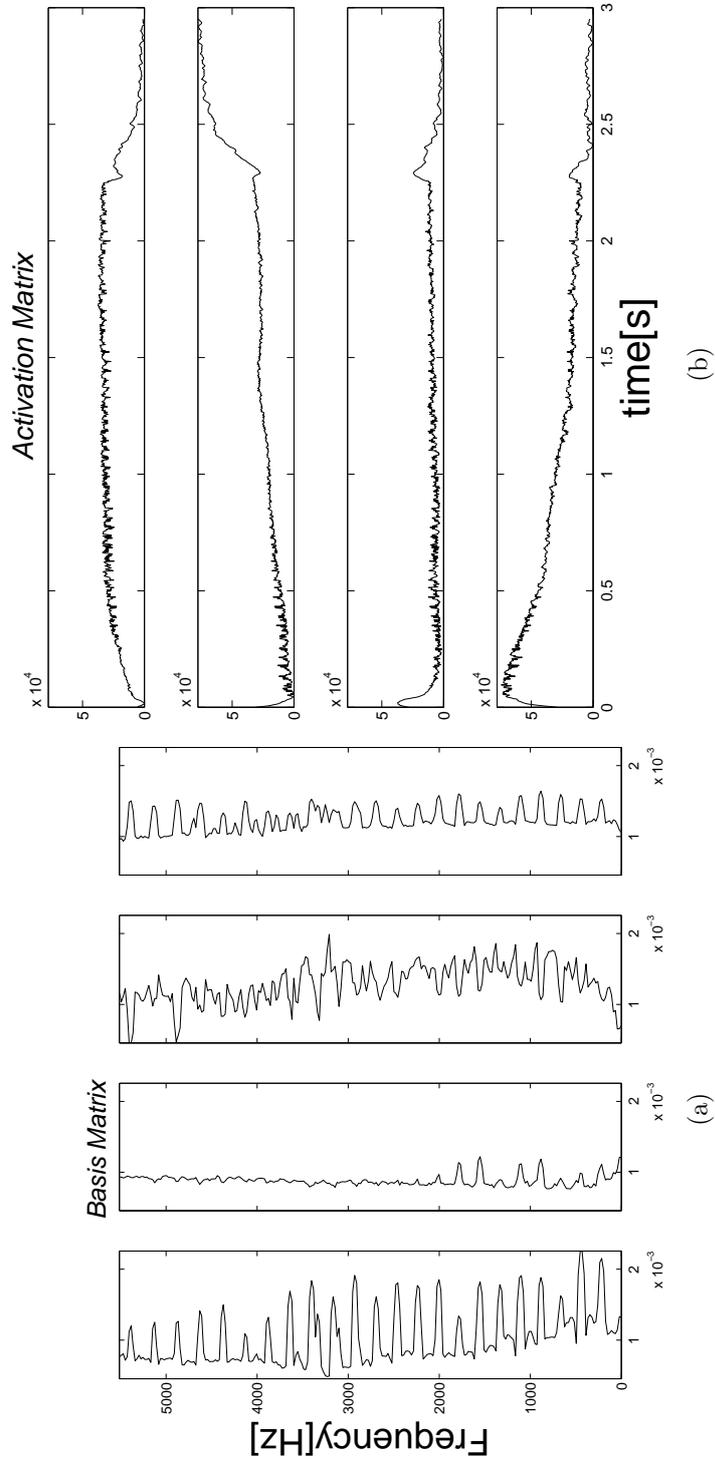


図 43: $K = 4$ の場合の分離結果 (GP2)

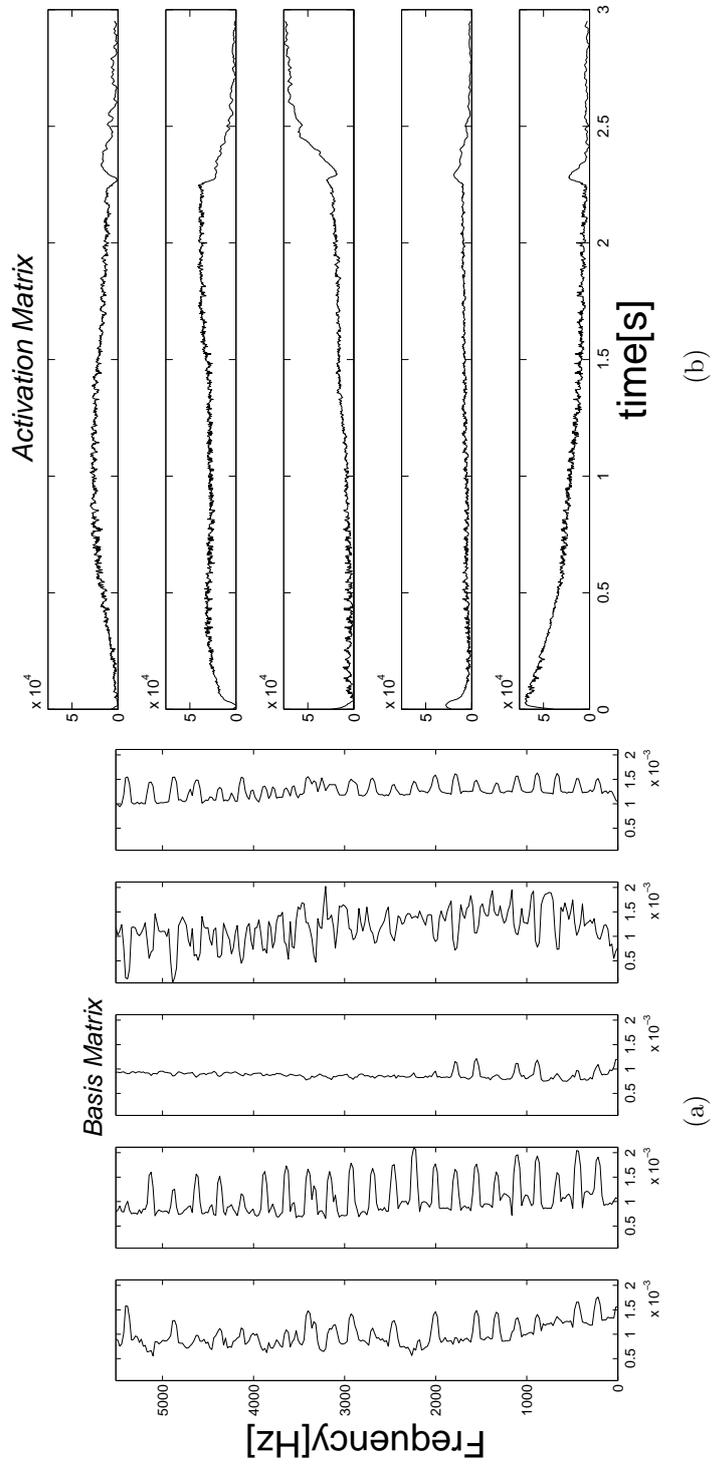


図 44: $K = 5$ の場合の分離結果 (GP2)

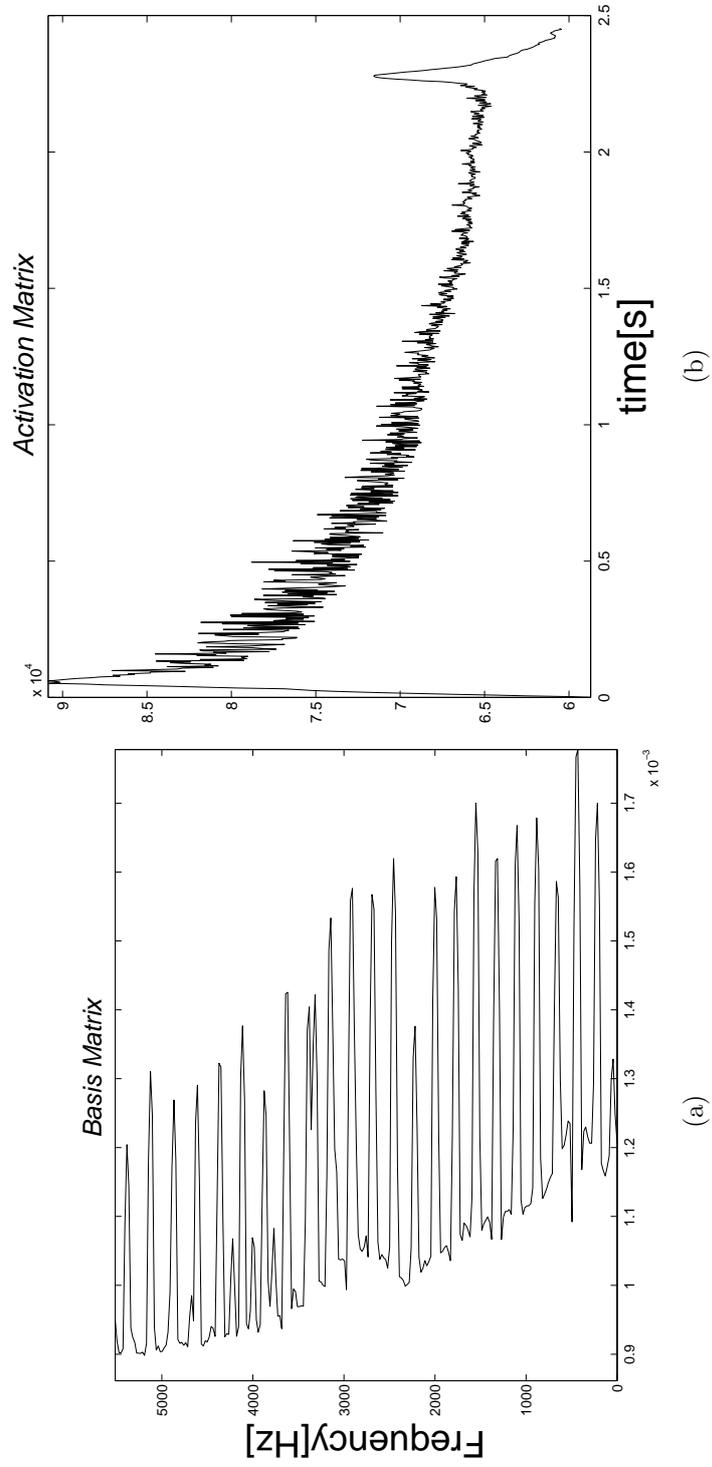


図 45: $K = 1$ の場合の分離結果 (RWC1)

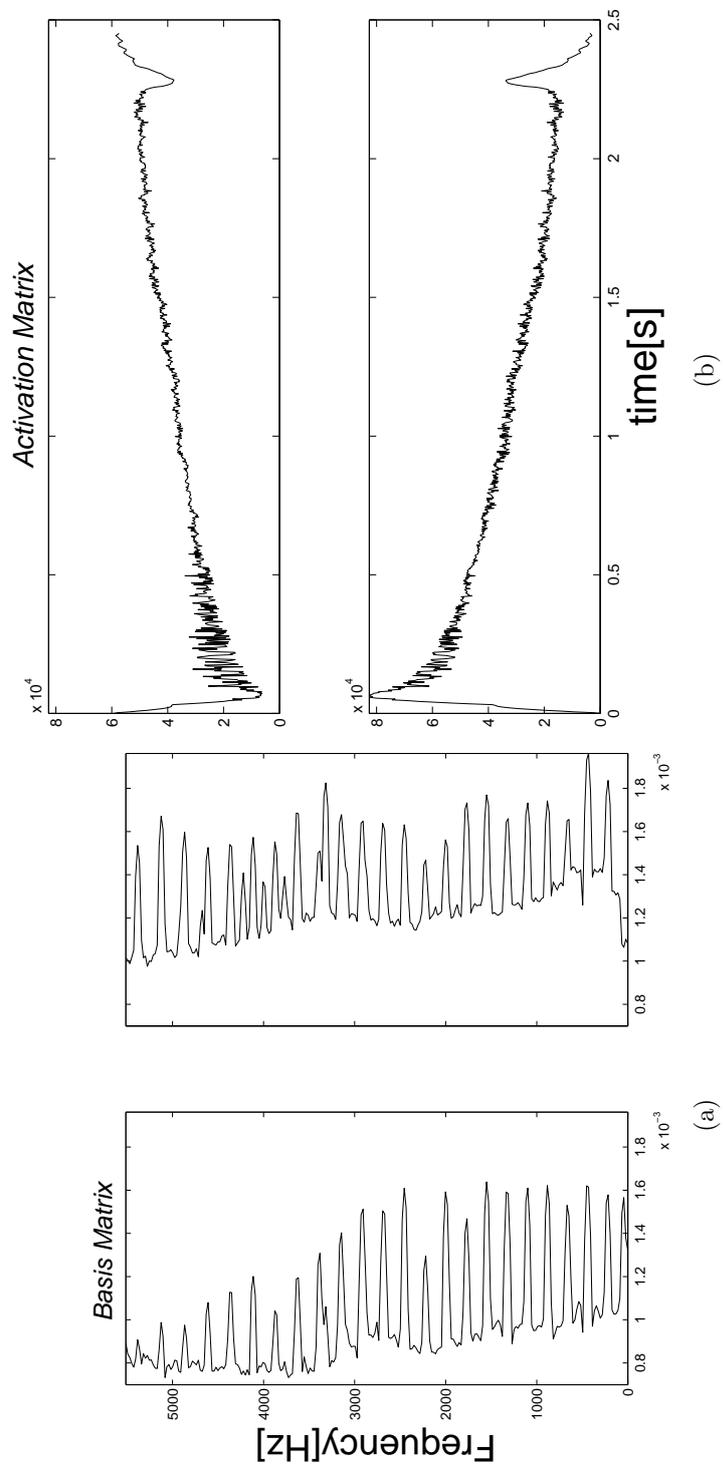


図 46: $K = 2$ の場合の分離結果 (RWCI1)

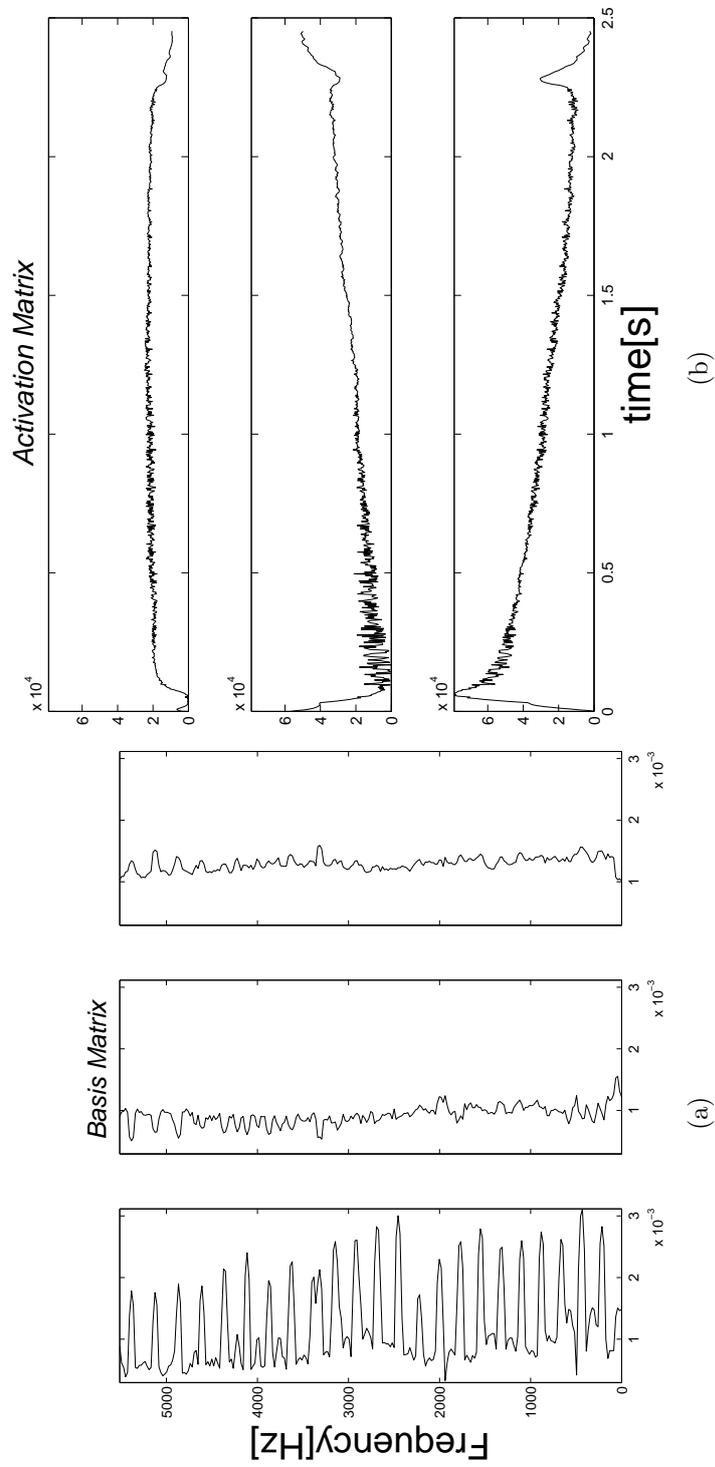


図 47: $K = 3$ の場合の分離結果 (RWC1)

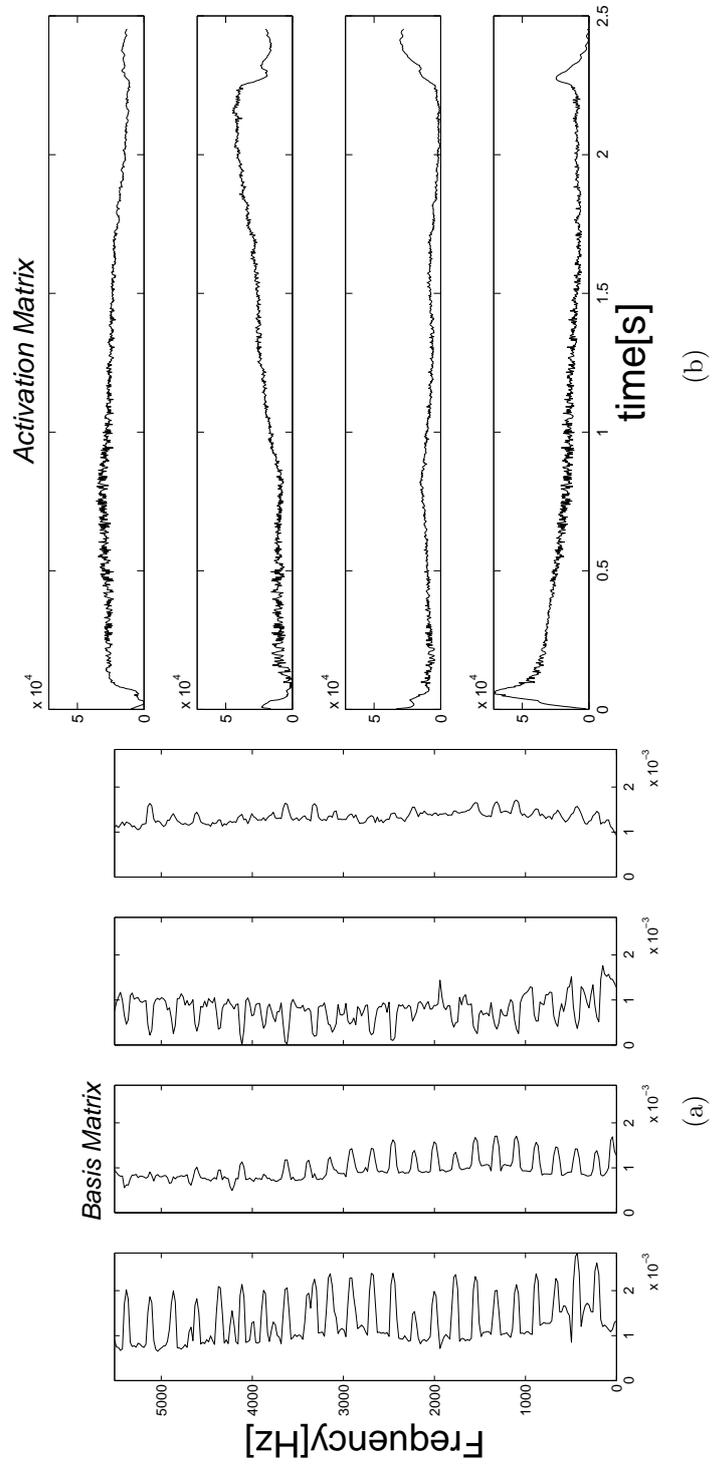


図 48: $K = 4$ の場合の分離結果 (RWC1)

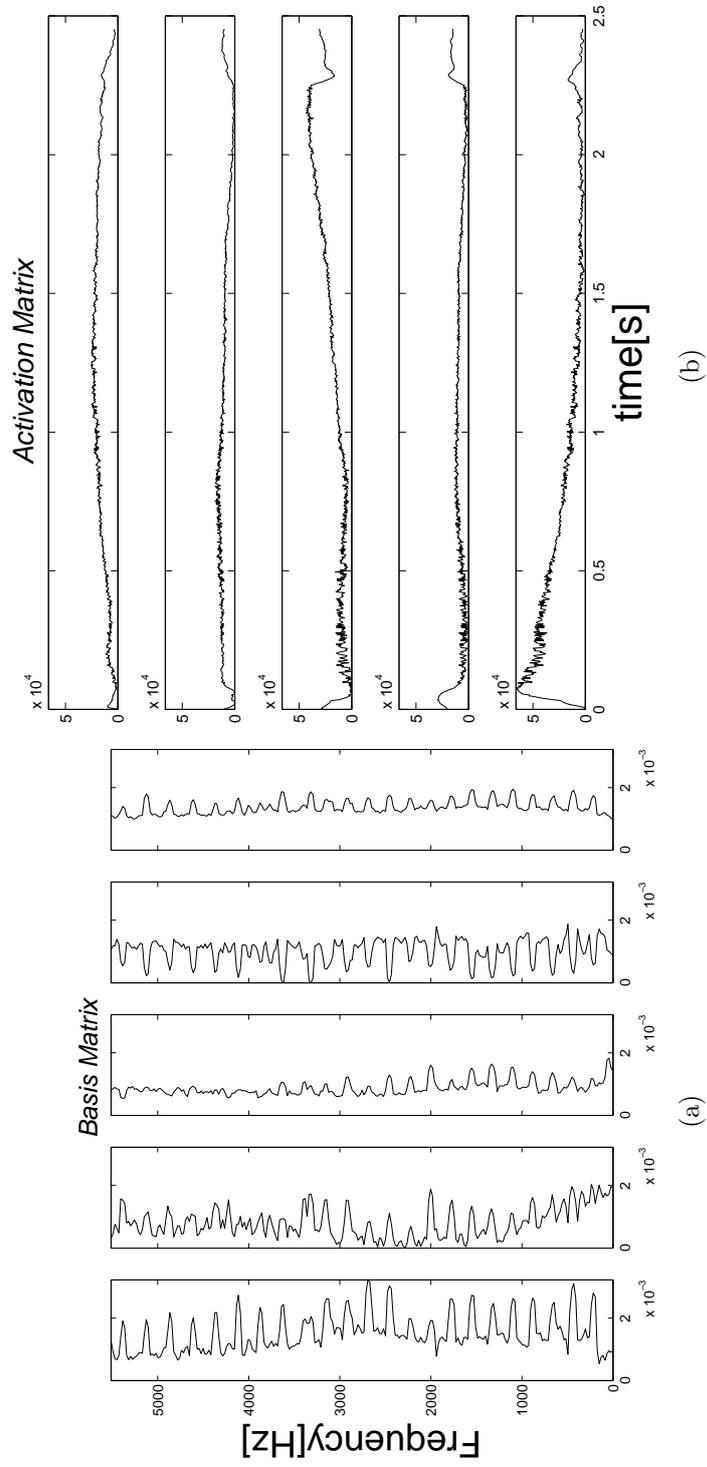


図 49: $K = 5$ の場合の分離結果 (RWC1)

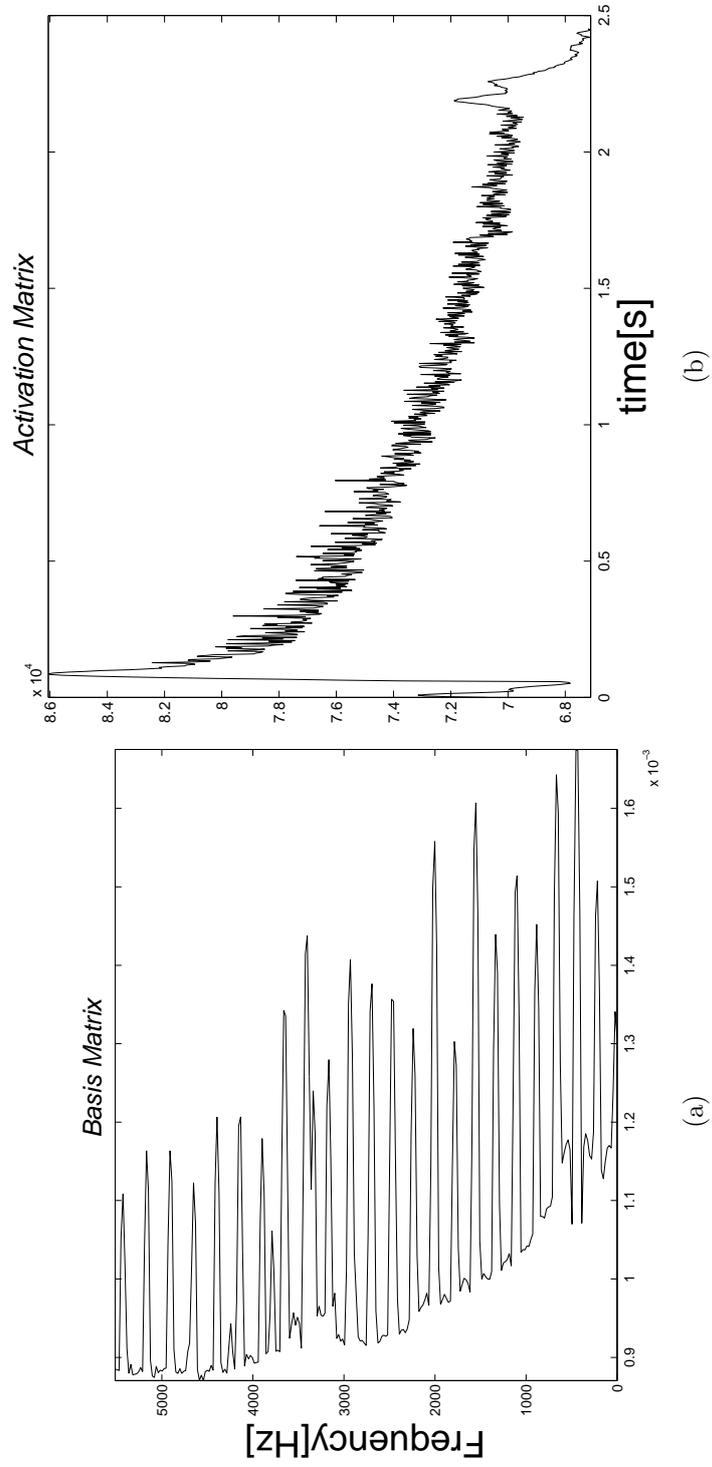


図 50: $K = 1$ の場合の分離結果 (RWC3)

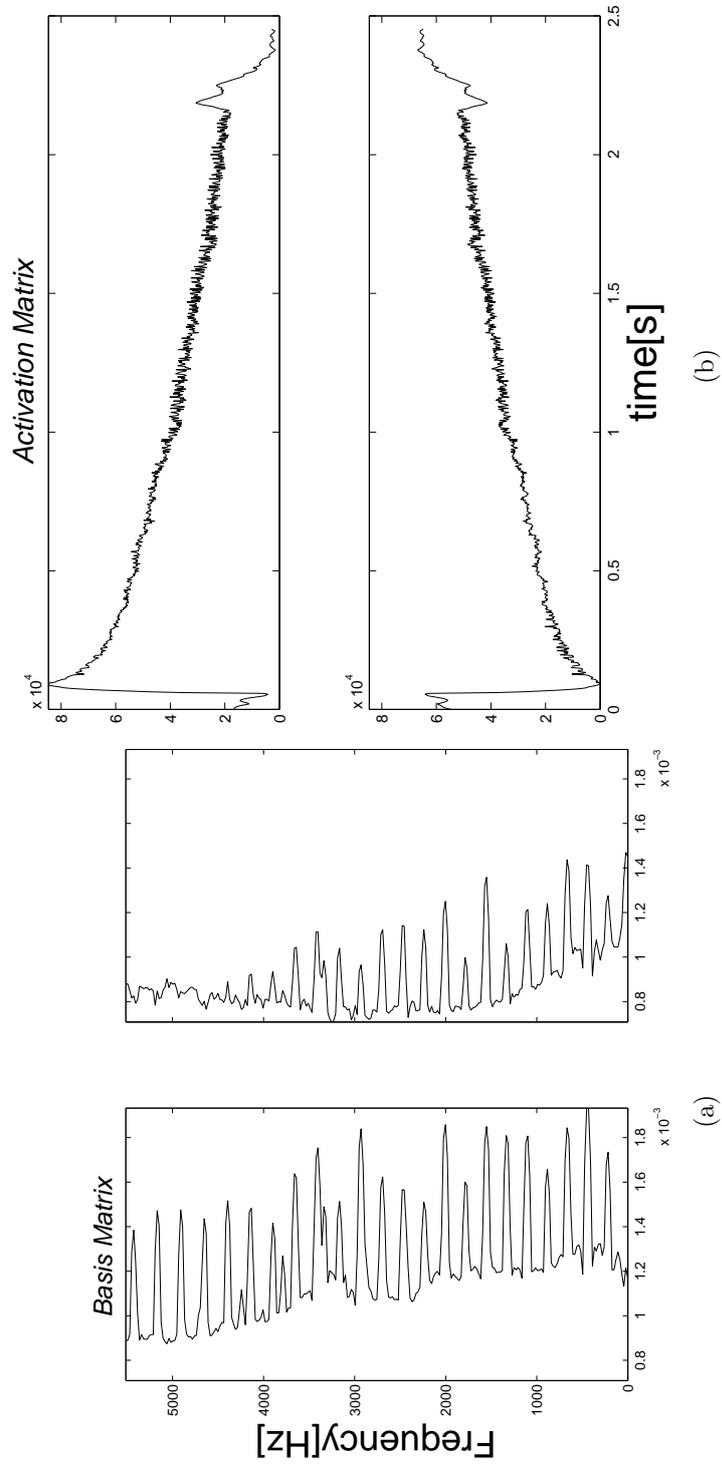


図 51: $K = 2$ の場合の分離結果 (RWC3)

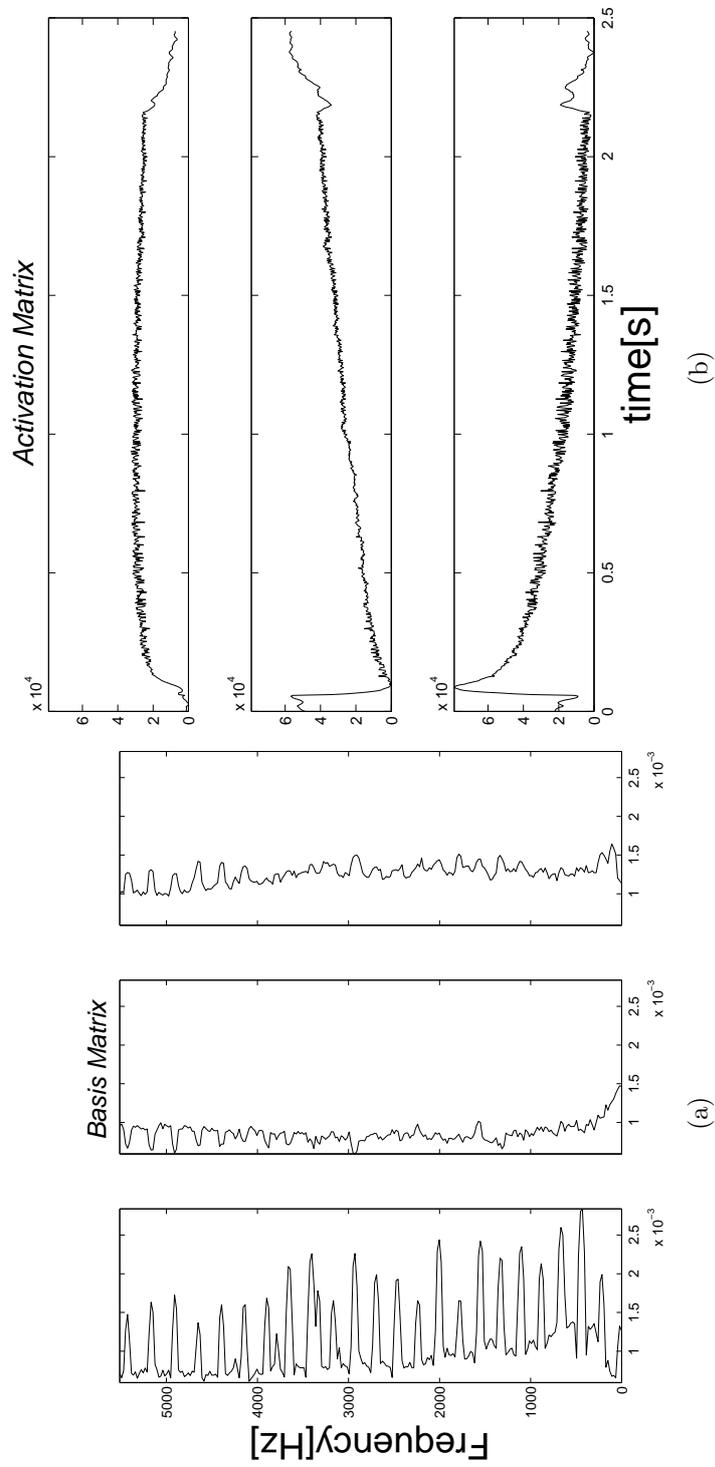


図 52: $K = 3$ の場合の分離結果 (RWC3)

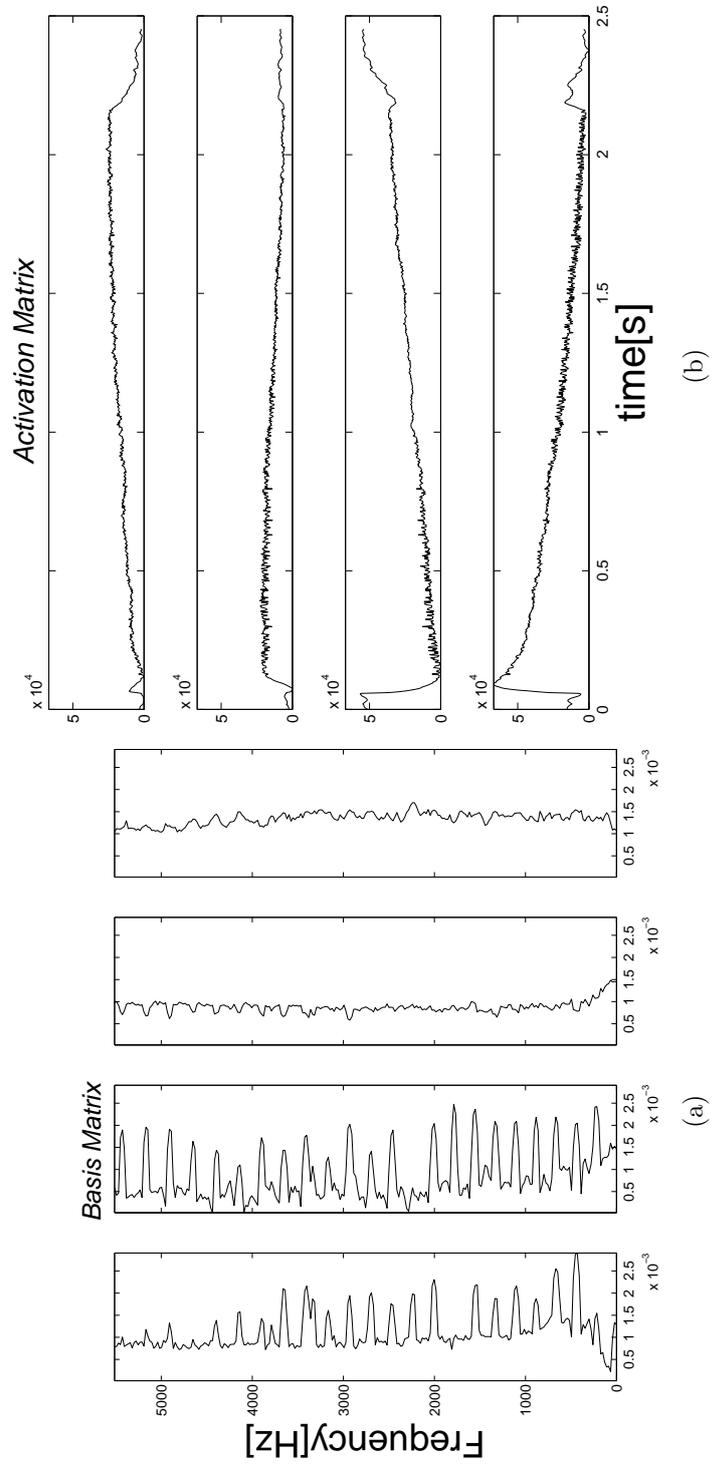


図 53: $K = 4$ の場合の分離結果 (RWC3)

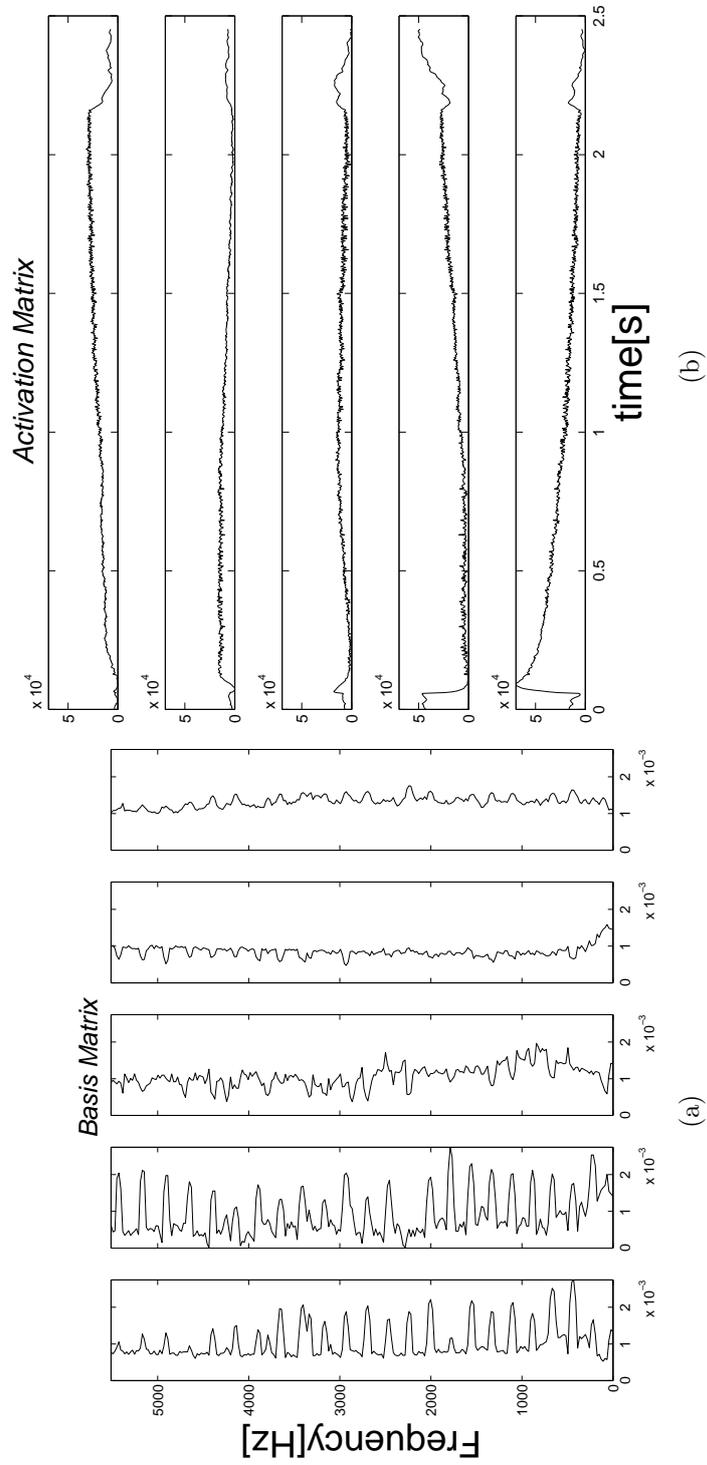


図 54: $K = 5$ の場合の分離結果 (RWC3)

基底の一部固定における分析結果

各基底数における、全ての音源の分析結果を示す。UP1は図 55, UP2は図 56, UP3は図 57, UP4は図 58, MAPSは図 59, GP1は図 60, GP2は図 61, RWC1は図 62, RWC3は図 63, にそれぞれの結果を示す。また、各図において、(a)は基底行列, (b)はアクティベーション行列を表す。

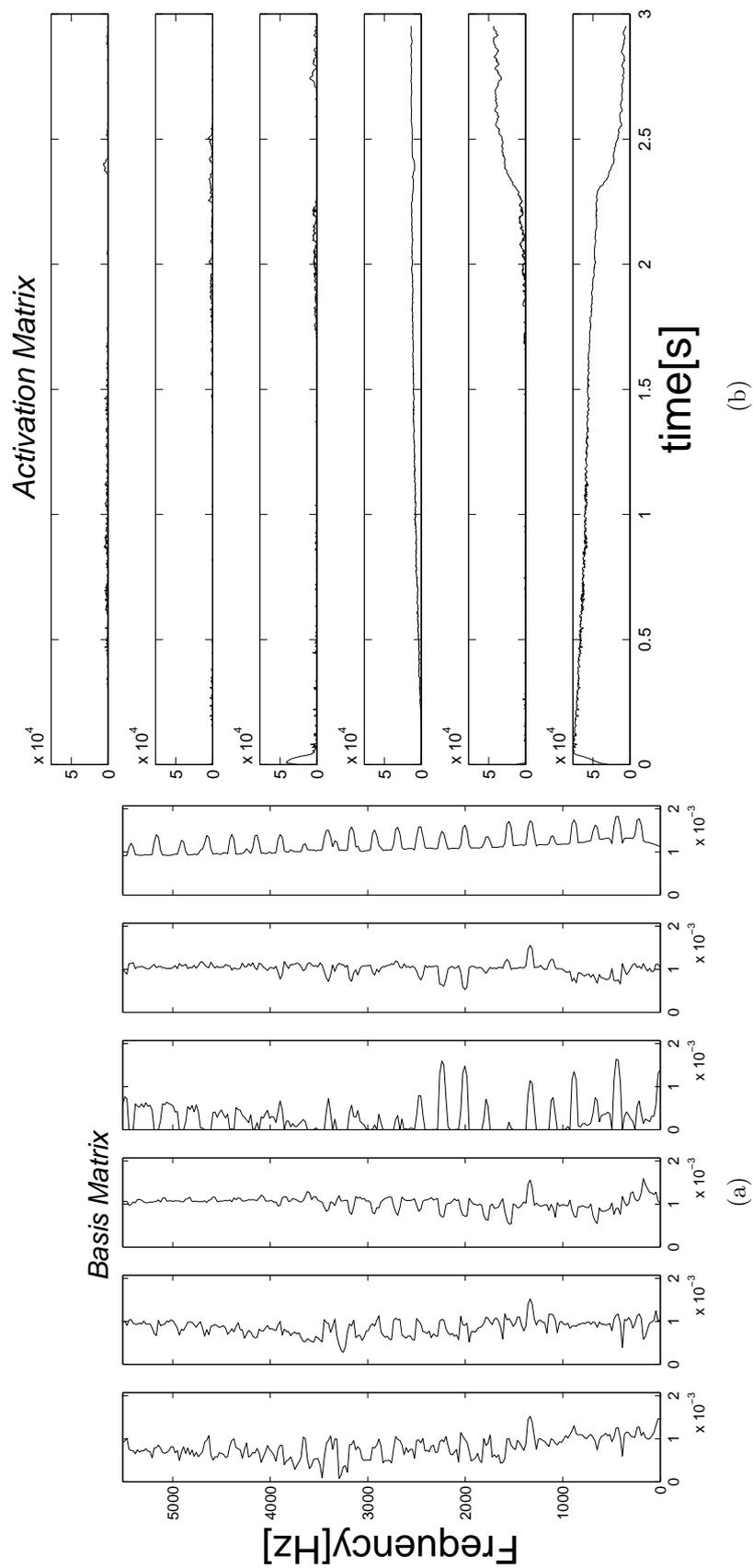


図 55: 基底行列の一部を固定してNMFを適応した結果 (UPI)

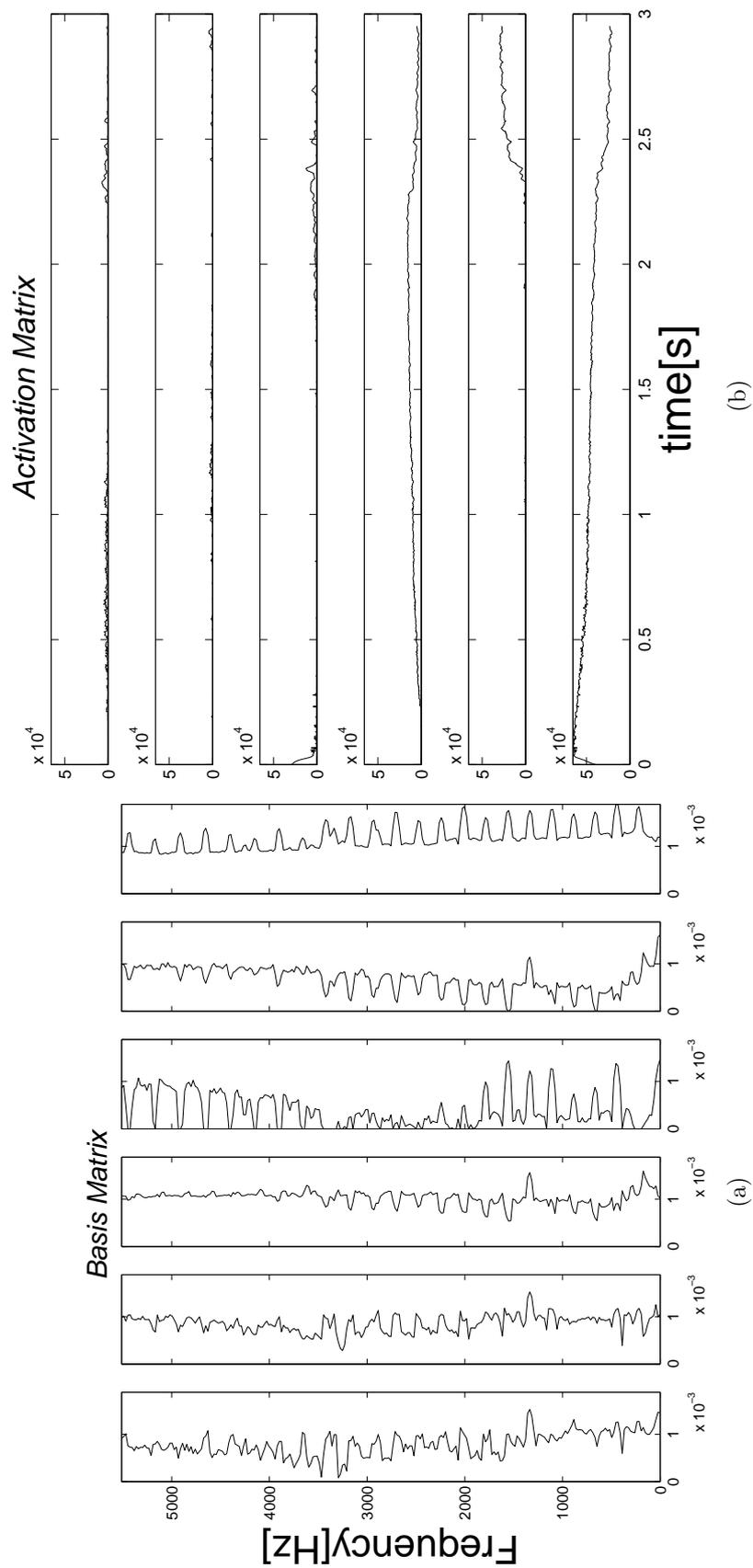


図 56: 基底行列の一部を固定してNMFを適応した結果 (UP2)

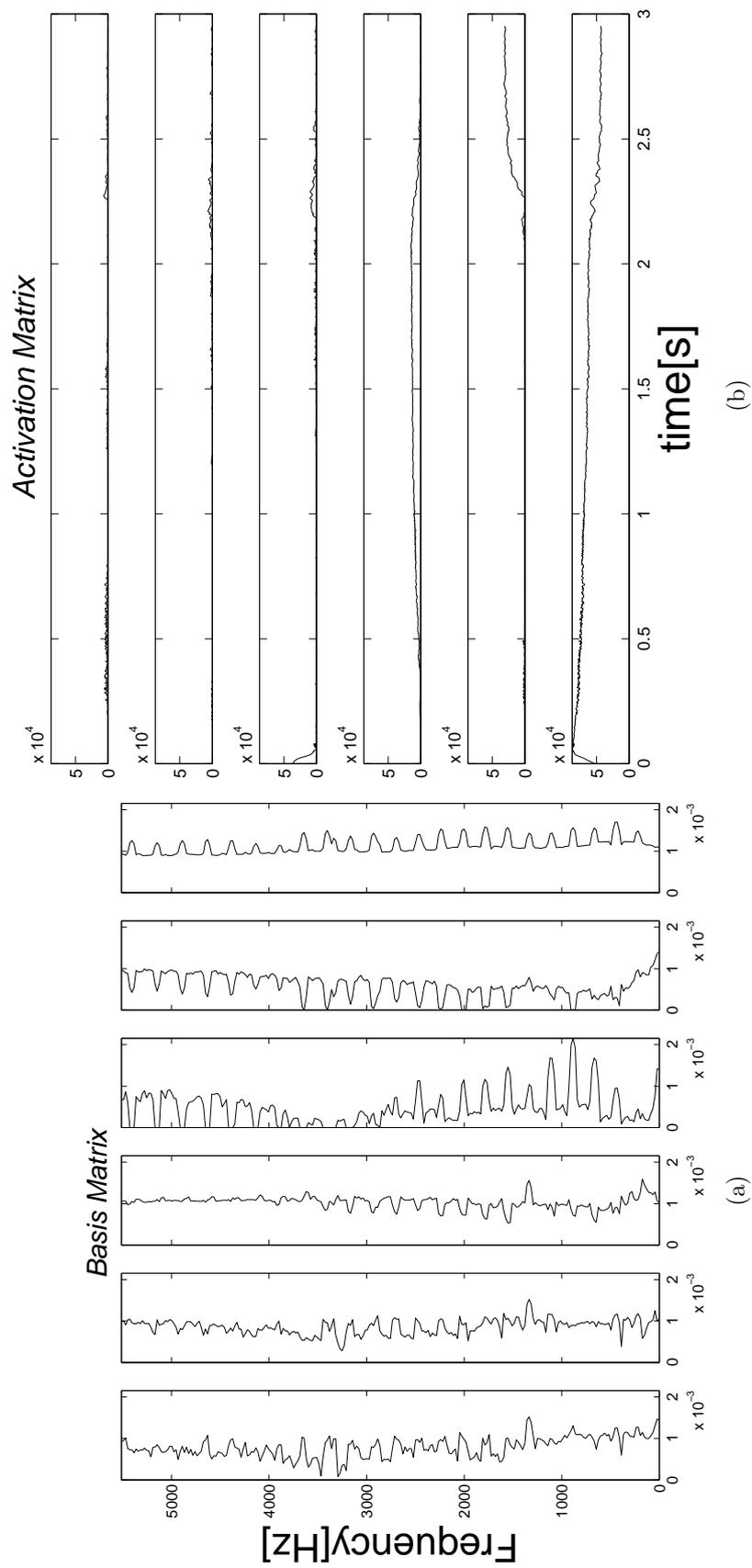


図 57: 基底行列の一部を固定してNMFを適応した結果 (UP3)

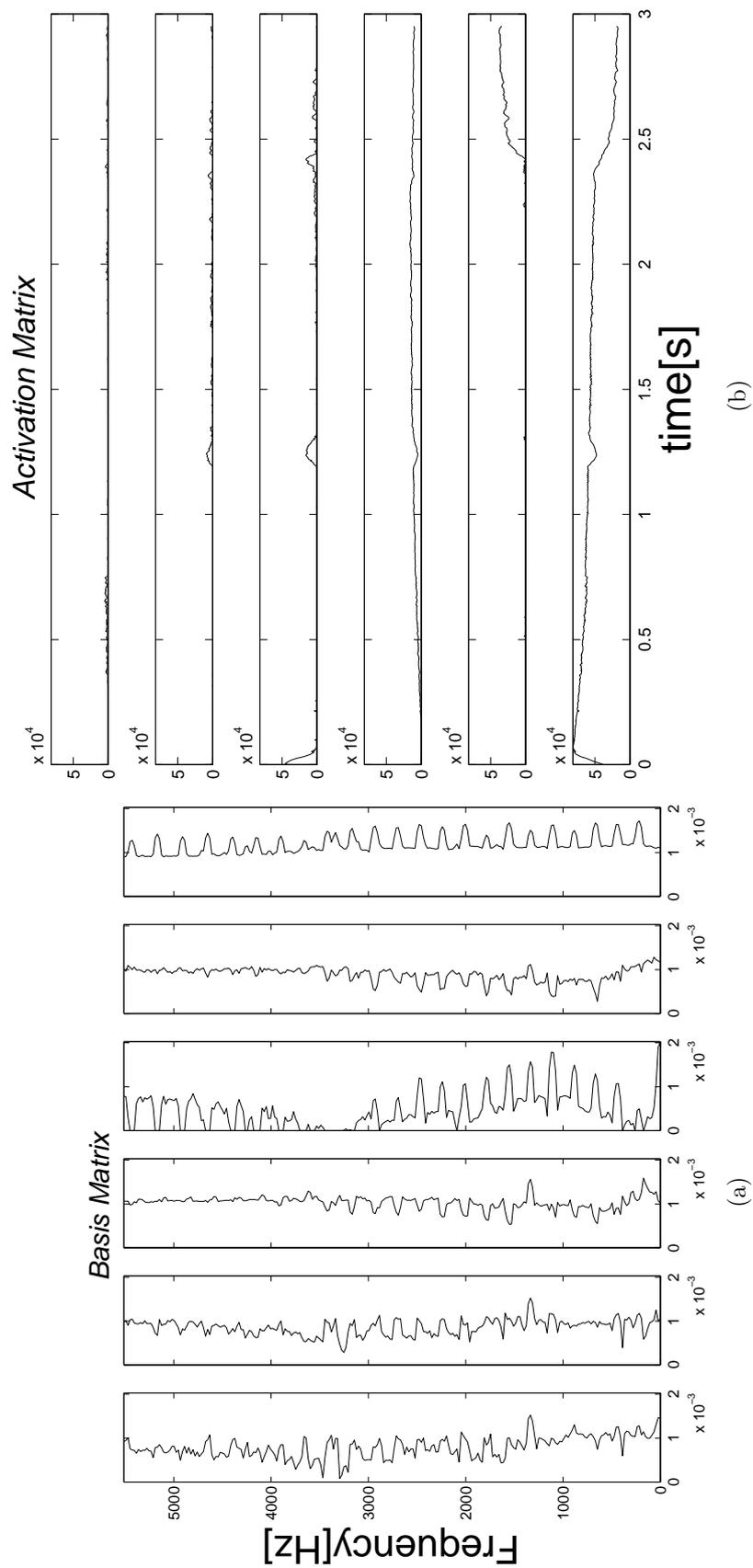


図 58: 基底行列の一部を固定してNMFを適応した結果 (UP4)

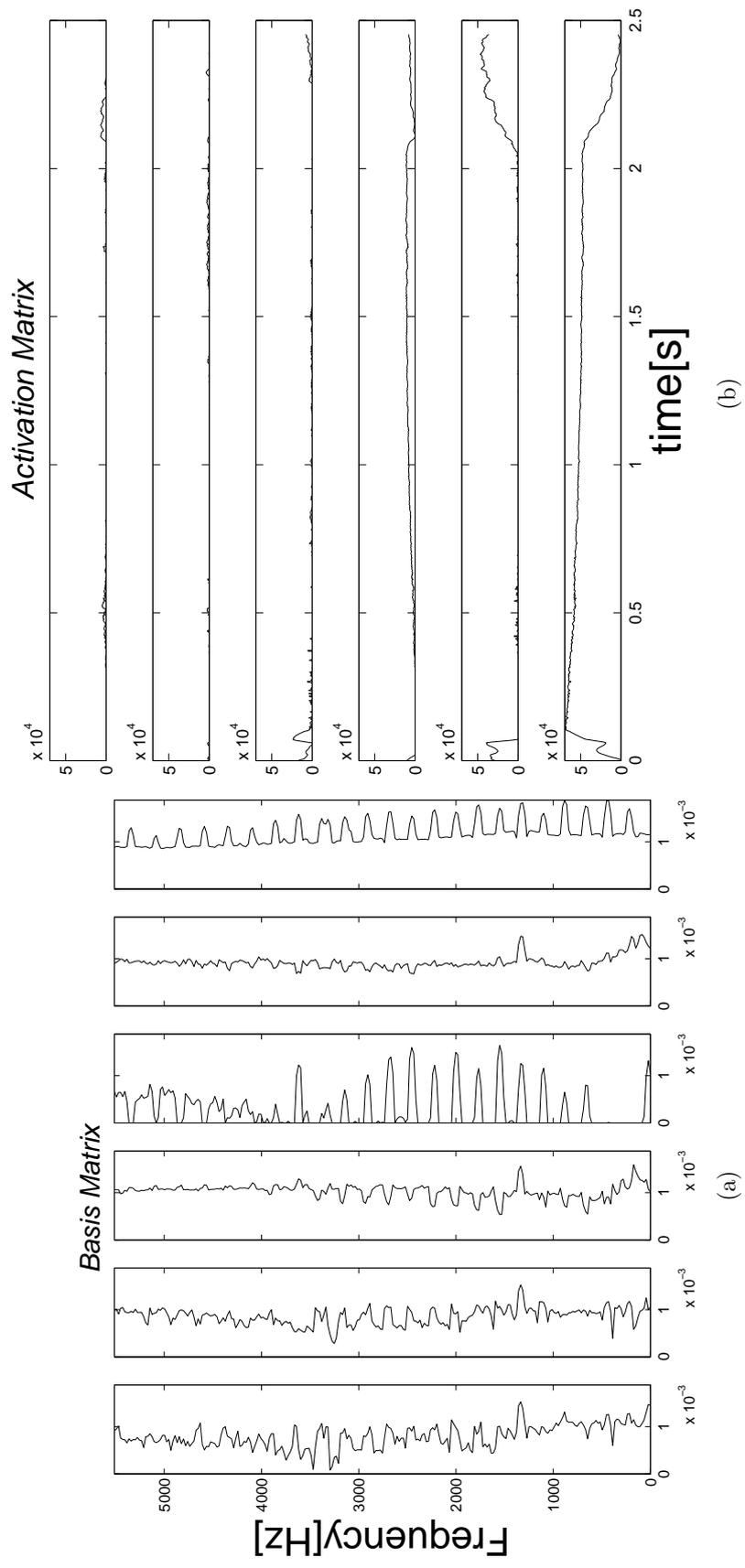


図 59: 基底行列の一部を固定して NMF を適応した結果 (MAPS)

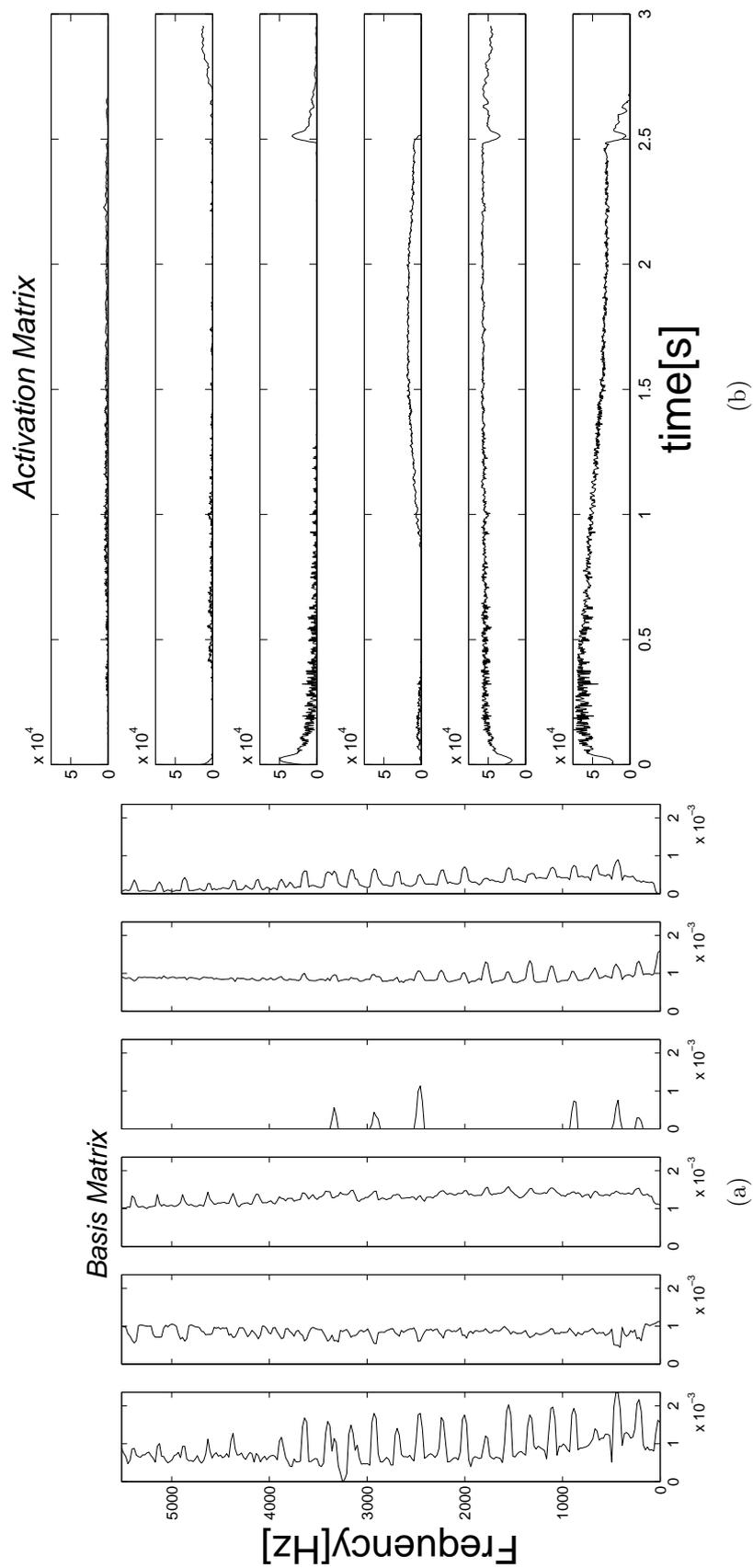


図 60: 基底行列の一部を固定してNMFを適応した結果 (GPI)

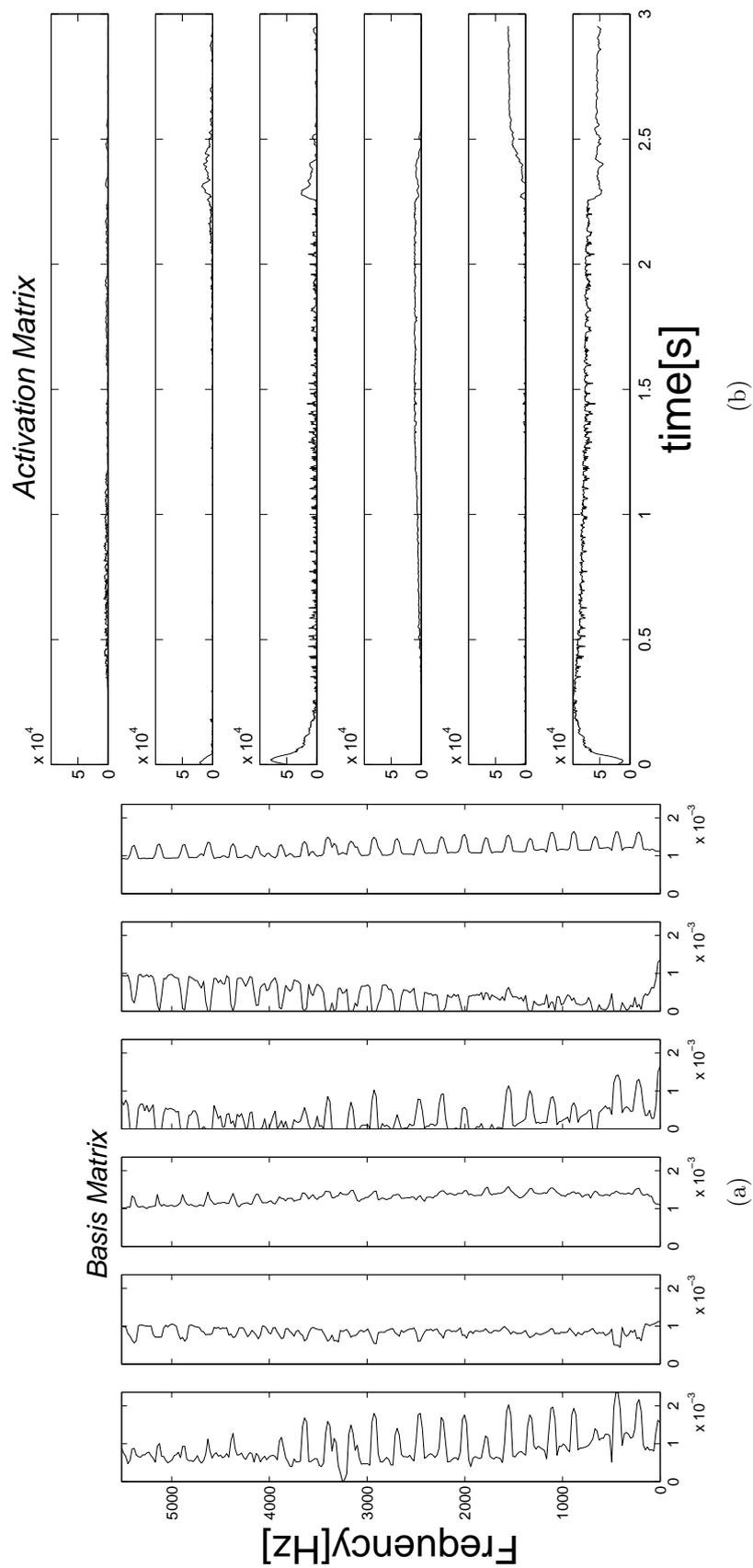


図 61: 基底行列の一部を固定してNMFを適応した結果 (GP2)

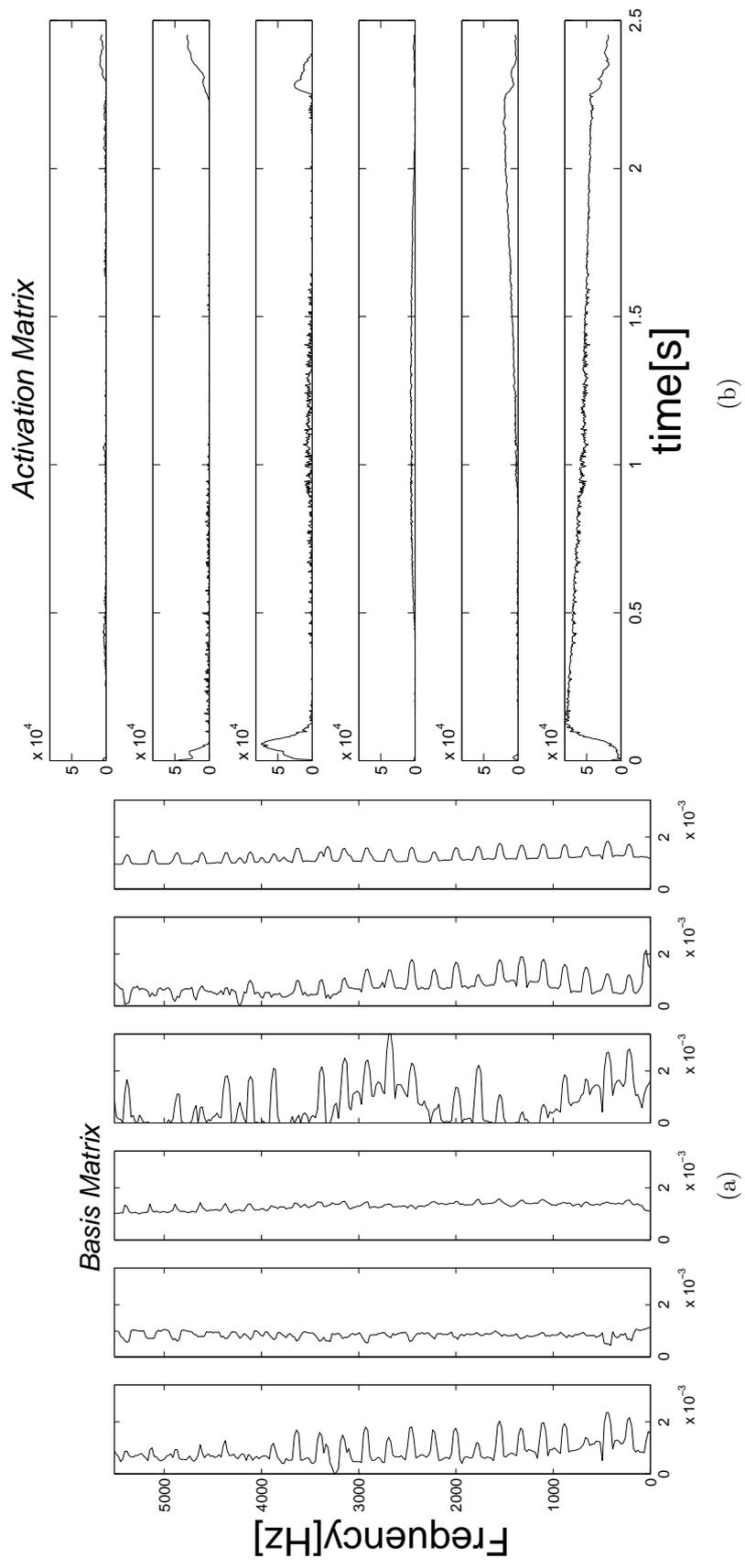


図 62: 基底行列の一部を固定して NMF を適応した結果 (RWCI1)

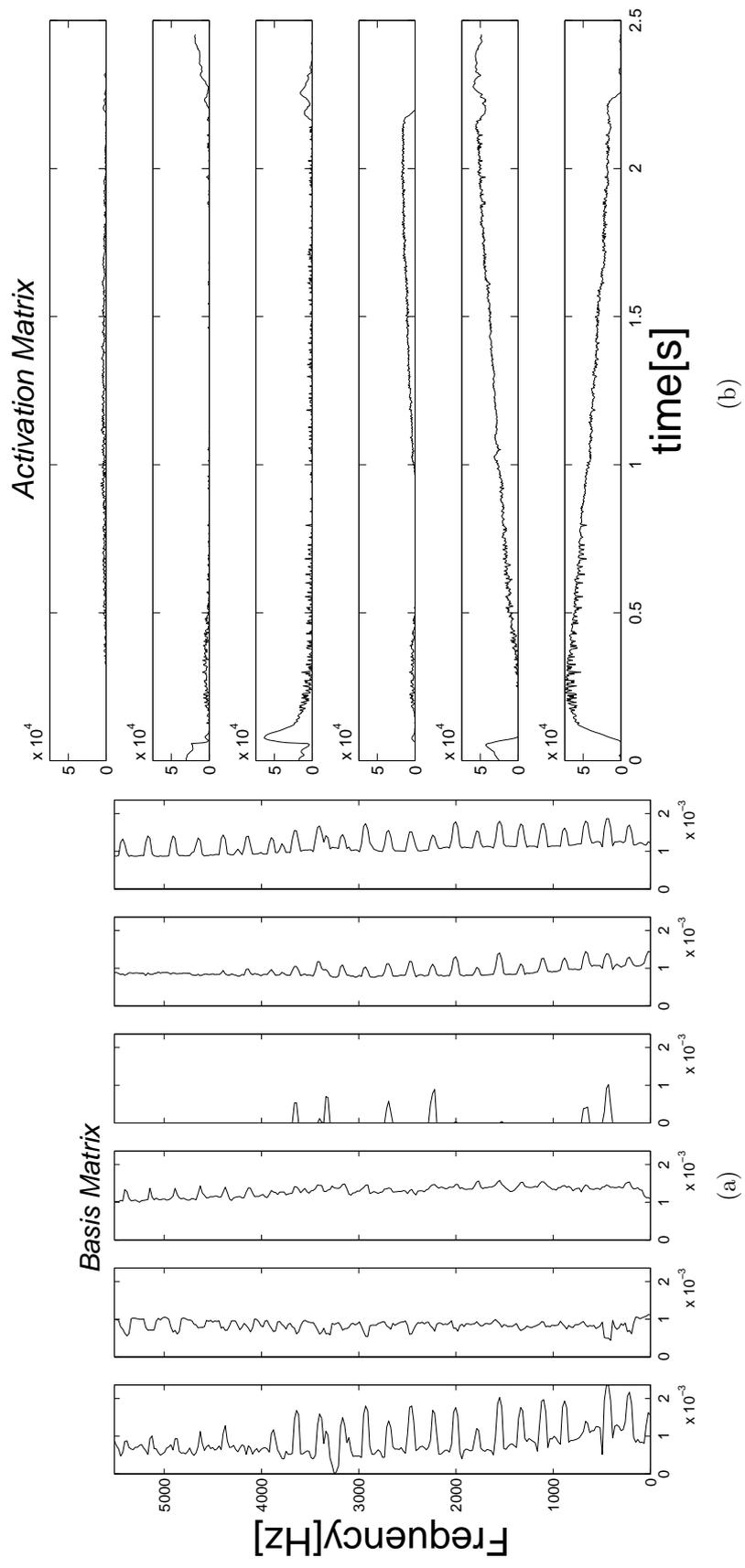


図 63: 基底行列の一部を固定して NMF を適応した結果 (RWC3)

研究業績

国際会議

- Kobayashi,K.,Morikawa,D.,Akagi,M.,”Study on Analyzing Individuality of Piano Sounds Using Non-negative Matrix Factorization,” The 6th seminar of A3 foresight program, February 2014.
- Kobayashi,K.,Morikawa,D.,Akagi,M.,”Study on Analyzing Individuality of Instrument Sounds Using Non-negative Matrix Factorization,” Proc. 2014 RISP International Workshop on Nonliner Circuits, Communications and Signal Processing, 33–36, March 2014.

研究会

- 小林慶祐, 森川大輔, 赤木正人, ”非負値行列因子分解を用いたピアノ音の個体差要因の分析,”第4回音響セミナー in 金沢, December2013.
- 小林慶祐, 森川大輔, 赤木正人, ”非負値行列因子分解を用いたピアノ音の個体差要因の分析,”日本音響学会 2014 年春季研究発表会, March 2014.