

統計的信号処理を用いた 合奏音からの楽器パート音の分離

深山幸穂[†] 田中大介^{††} 高本健生[†] 佐々木崇人[†]

本報では統計的最小二乗規範に基づき、合奏音のモノラル信号を指定された楽器パートごとの演奏音に分離する信号処理法を検討する。前提として、対象となる楽器パートの音域内の音高の異なる各音について、自己相関行列と相互相関行列を既知とした。この信号処理法では合奏音を前述の各音に音量係数を乗じた和と仮定し、まず、入力した合奏音の自己相関行列から各音の音量係数を同定し、次に、合奏音と各音との相互相関行列を算出している。そして、これらの自己相関行列と相互相関行列から、楽器パートごとの音を推定するフィルタの係数を求める。バイオリン音、トランペット音、フルート音による実施例を付す。

Separation of each Instrumental part in music sounds with a statistical signal processing

Yukio Fukayama[†], Daisuke Tanaka[†],
Kensei Takamoto[†] and Takato Sasaki[†]

An optimal filter that separates a tone of particular pitch highness played by particular instrument applying the statistical least square criterion is discussed. The filter is based on a priori autocorrelations of and cross correlations among tones. It firstly identifies current amplitude factors of tones; then, it evaluates cross correlation between current input signal and each tone. Finally it obtains the estimate of seeking tone from the input signal which is mixture of sounds played by instruments and corrupted by white noise. Some numerical examples with violin, trumpet and flute tones are also introduced.

1. はじめに

本報では統計的最小二乗規範に基づき、合奏音のモノラル信号を指定された楽器パートごとの演奏音に分離する信号処理法を検討する。よく知られているように、MIDI 標準 [2] が規定するノート番号 i の音高の名目上の周波数は、国際標準ピッチ(1939)に従う

12 音平均律[4]により $f_i = 440 \times 2^{\frac{i-69}{12}}$ [Hz] となり、 $\dots, f_{69}, f_{70}, f_{71} \dots$ は大譜表における中央の C 音の直ぐ上の $\dots, A, A\#, B, \dots$ 音に対応している。このとき、楽器パートの音域は、例えば、標準的なピアノの全鍵盤を用いれば $21 \leq i \leq 108$ となる。本研究では、バイオリン、トランペット等、第 p 番目の楽器パートの演奏音を分離するため、その楽器の音域に属するノート番号の個数だけ当該音を推定するフィルタを用意し、これらを通過した信号を加算する。

上述したフィルタの設計にあたり、合奏の各楽器パートで用いるノート番号の異なる音を「要素となる音」と称することとし、これらにユニークな音番号 $l=1, 2, \dots, L$ を付して取扱う。すなわち、合奏音を要素となる音に音量係数を乗じた和としてモデル化する。このとき、要素となる音それぞれは互いに広義の定常な確率過程によるランダムな波形であり、統計量は既知と仮定している。さらに、音量係数は対応する音の波形に比して緩慢に時間変化し、短い時間では定数と考える。以下、次節より統計的信号処理の考え方[3]に基づいて具体的に議論する。

2. 入力信号のモデル化と音量係数の同定

本節では、第 p 番目の楽器パートで演奏され、ノート番号 i の音高である、第 $l=l(i, p)$ 番目の要素となる音 $\{S_l(k)\}$, ($l=1, 2, \dots, L$) について、その音量係数 σ_l の求め方を検討する。ここに、時点 k はサンプリング間隔 Δt , 開始時刻 t_0 の下に、時刻 $t = t_0 + k\Delta t$ における値を示す。

[†] 新居浜工業高等専門学校電子制御工学科
Electronic Control Department, Niihama National College of Technology
^{††} 新居浜工業高等専門学校専攻科
Advanced Study Courses, Niihama National College of Technology

まず、入力信号 $\{Y(k)\}$ は確率過程であり、次式にモデル化する：

$$Y(k) = \sum_{l=1}^L \sigma_l S_l(k) + W(k), \quad (1)$$

ここに、要素となる音と音量係数との積和を合奏音とし、白色雑音 $\{W(k)\}$ が混入して入力信号となった。また、音量係数 σ_l は時点 k と共に緩慢に変化し、後述する $2M$ 近傍では定数とみなせる。

このとき $\{S_l(k)\}$ と $\{W(k)\}$ は互いに広義の定常(w.s.s.)な確率過程である。すなわち、統計量として、平均は零、自己相関と相互相関 ($l, j=1, 2, \dots, L$) は時点 k によらず、時点の差 m の関数と仮定する：

$$E\{S_l(k)\} = E\{W(k)\} = 0, \quad (2)$$

$$E\{S_l(k+m)S_j(k)\} = \begin{cases} r_{lj}(m), & (|m| \leq 2M) \\ 0, & (|m| \geq 2M+1) \end{cases}, \quad (3)$$

$$E\{W(k+m)W(k)\} = \begin{cases} w^2, & (m=0) \\ 0, & (m \neq 0) \end{cases}, \quad (4)$$

$$E\{S_l(k+m)W(k)\} = 0, \text{ (for all } m), \quad (5)$$

ここに、その外側において $\{S_l(k)\}$ の自己相関や相互相関が無視できるように、時点 k の $2M$ 近傍を設定した。

また、信号のエネルギー e^2 は (1) 式より次式に評価できる：

$$\begin{aligned} e^2 &= E\{Y(k)^2\} \\ &= \sigma^T \begin{pmatrix} E\{S_1^T(k)S_1(k)\} & \cdots & E\{S_1^T(k)S_L(k)\} \\ \vdots & \ddots & \vdots \\ E\{S_L^T(k)S_1(k)\} & \cdots & E\{S_L^T(k)S_L(k)\} \end{pmatrix} \sigma + E\{W(k)^2\} \\ &= \sigma^T \begin{pmatrix} r_{11}(0) & \cdots & r_{1L}(0) \\ \vdots & \ddots & \vdots \\ r_{L1}(0) & \cdots & r_{LL}(0) \end{pmatrix} \sigma + w^2, \end{aligned} \quad (6)$$

ここに、音量係数ベクトル $\sigma \in \mathbf{R}^L$ を定義した：

$$\sigma = (\sigma_1 \ \cdots \ \sigma_L)^T. \quad (7)$$

次に、入力信号を $2M$ 近傍においてベクトルで表記する：

$$\begin{aligned} Y(k) &= (\sigma^T \otimes \mathbf{I}_{2M+1}) \begin{pmatrix} S_1(k) \\ \vdots \\ S_L(k) \end{pmatrix} + W(k) \\ &= (\sigma_1 \mathbf{I}_{2M+1} \ \cdots \ \sigma_L \mathbf{I}_{2M+1}) \begin{pmatrix} S_1(k) \\ \vdots \\ S_L(k) \end{pmatrix} + W(k), \end{aligned} \quad (8)$$

ここに、 \otimes はクロネッカ積であり、次のとおり、信号ベクトル $Y(k) \in \mathbf{R}^{2M+1}$ 、第 l 番目の要素となる音のベクトル $S_l(k) \in \mathbf{R}^{2M+1}$ 、($l=1, 2, \dots, L$)、および、白色雑音ベクトル $W(k) \in \mathbf{R}^{2M+1}$ を定義した：

$$Y(k) = (Y(k+M) \ \cdots \ Y(k-M))^T, \quad (9)$$

$$S_l(k) = (S_l(k+M) \ \cdots \ S_l(k-M))^T, \quad (10)$$

$$W(k) = (W(k+M) \ \cdots \ W(k-M))^T. \quad (11)$$

このとき、自己相関行列 $R_{YY} \in \mathbf{R}^{(2M+1) \times (2M+1)}$ は $\{Y(k)\}$ から直接に評価でき、さらに、緩慢に時間変化する σ を用いて次式にモデル化できる：

$$\begin{aligned} R_{YY} &= E\{Y(k)Y^T(k)\} \\ &= (\sigma^T \otimes \mathbf{I}_{2M+1}) \begin{pmatrix} R_{11} & \cdots & R_{1L} \\ \vdots & \ddots & \vdots \\ R_{L1} & \cdots & R_{LL} \end{pmatrix} (\sigma \otimes \mathbf{I}_{2M+1}) + w^2 \mathbf{I}_{2M+1}, \end{aligned} \quad (12)$$

ここに、 $R_{lj} \in \mathbf{R}^{(2M+1) \times (2M+1)}$ は、要素となる $S_l(k)$ および $S_j(k)$ 、($l, j=1, 2, \dots, L$) について、自己相関行列 ($l=j$) または相互相関行列 ($l \neq j$) であり、時点 k によらないため、予め評価しておく：

$$\mathbf{R}_{ij} = E\{\mathbf{S}_i(k)\mathbf{S}_j^T(k)\} = \begin{pmatrix} r_{ij}(0) & \cdots & r_{ij}(2M) \\ \vdots & \ddots & \vdots \\ r_{ij}(-2M) & \cdots & r_{ij}(0) \end{pmatrix}. \quad (13)$$

このとき、 \mathbf{R}_{YY} は正定対称行列であるから次のとおり規格化直交系をなす固有ベクトル $\mathbf{v}_m \in \mathbf{R}^{2M+1}$ を選ぶことができる：

$$\mathbf{v}_m^T \mathbf{v}_n = \delta_{mn}, \quad (m, n = 1, 2, \dots, 2M+1), \quad (14)$$

ここに、対応する固有値 λ_m は以下のように置ける：

$$\mathbf{R}_{YY} \mathbf{v}_m = \lambda_m \mathbf{v}_m, \quad (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{2M+1} \geq 0). \quad (15)$$

いま、固有ベクトル \mathbf{v}_m^T および \mathbf{v}_m , ($m = 1, 2, \dots, 2M+1$) をそれぞれ左右から (12) 式に乗じると次式を得る：

$$\lambda_m = \boldsymbol{\sigma}^T \begin{pmatrix} \mathbf{v}_m^T \mathbf{R}_{11} \mathbf{v}_m & \cdots & \mathbf{v}_m^T \mathbf{R}_{1L} \mathbf{v}_m \\ \vdots & \ddots & \vdots \\ \mathbf{v}_m^T \mathbf{R}_{L1} \mathbf{v}_m & \cdots & \mathbf{v}_m^T \mathbf{R}_{LL} \mathbf{v}_m \end{pmatrix} \boldsymbol{\sigma} + w^2, \quad (16)$$

さらに、(6) 式を用いると、 w^2 を消去できる：

$$\gamma_m = \boldsymbol{\sigma}^T \mathbf{V}_m \boldsymbol{\sigma}, \quad (17)$$

ここに、 \mathbf{V}_m および γ_m を定義した：

$$\mathbf{V}_m = \begin{pmatrix} r_{11}(0) & \cdots & r_{1L}(0) \\ \vdots & \ddots & \vdots \\ r_{L1}(0) & \cdots & r_{LL}(0) \end{pmatrix} - \begin{pmatrix} \mathbf{v}_m^T \mathbf{R}_{11} \mathbf{v}_m & \cdots & \mathbf{v}_m^T \mathbf{R}_{1L} \mathbf{v}_m \\ \vdots & \ddots & \vdots \\ \mathbf{v}_m^T \mathbf{R}_{L1} \mathbf{v}_m & \cdots & \mathbf{v}_m^T \mathbf{R}_{LL} \mathbf{v}_m \end{pmatrix}, \quad (18)$$

$$\gamma_m = e^2 - \lambda_m. \quad (19)$$

従って、通常は $L \ll 2M+1$ であるから、(17) 式を満たす $\boldsymbol{\sigma} (\geq \mathbf{0}_L)$ を最小二乗解として同定できる：

$$\boldsymbol{\sigma} = \arg \min \{J(\boldsymbol{\sigma})\}, \quad (20)$$

ここに、ロバスト性を得るため、次式のあてはめ誤差 $J(\boldsymbol{\sigma})$ において重み係数

$b_m (\geq 0)$, ($m = 1, 2, \dots, 2M+1$) を固有値 λ_m に応じて与えればよい：

$$J(\boldsymbol{\sigma}) = \sum_{m=1}^{2M+1} b_m (\boldsymbol{\sigma}^T \mathbf{V}_m \boldsymbol{\sigma} - \gamma_m)^2. \quad (21)$$

このとき、勾配ベクトル $\mathbf{f}(\boldsymbol{\sigma}) \in \mathbf{R}^L$ とヘッセ行列 $\mathbf{F}(\boldsymbol{\sigma}) \in \mathbf{R}^{L \times L}$ は以下となる：

$$\mathbf{f}(\boldsymbol{\sigma}) = \frac{\partial J(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} = \sum_{m=1}^{2M+1} 2b_m (\mathbf{V}_m + \mathbf{V}_m^T) \boldsymbol{\sigma} (\boldsymbol{\sigma}^T \mathbf{V}_m \boldsymbol{\sigma} - \gamma_m), \quad (22)$$

$$\begin{aligned} \mathbf{F}(\boldsymbol{\sigma}) &= \frac{\partial^2 J(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma} \partial \boldsymbol{\sigma}^T} = \frac{\partial \mathbf{f}(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}^T} \\ &= \sum_{m=1}^{2M+1} 2b_m (\mathbf{V}_m + \mathbf{V}_m^T) [\boldsymbol{\sigma} \boldsymbol{\sigma}^T (\mathbf{V}_m + \mathbf{V}_m^T) + (\boldsymbol{\sigma}^T \mathbf{V}_m \boldsymbol{\sigma} - \gamma_m) \mathbf{I}_L]. \end{aligned} \quad (23)$$

よって、適当な初期条件 $\boldsymbol{\sigma}^{(0)}$ の下、次式の収束計算 ($i = 1, 2, \dots$) により $\boldsymbol{\sigma}$ を決定できる：

$$\boldsymbol{\sigma}^{(i)} = \boldsymbol{\sigma}^{(i-1)} - \mathbf{F}^{-1}(\boldsymbol{\sigma}^{(i-1)}) \mathbf{f}(\boldsymbol{\sigma}^{(i-1)}), \quad (24)$$

ここに、 $\boldsymbol{\sigma} = \boldsymbol{\sigma}^{(i)}$ は次式の収束判定条件 ε をもって与える：

$$\|\boldsymbol{\sigma}^{(i)} - \boldsymbol{\sigma}^{(i-1)}\| < \varepsilon. \quad (25)$$

3. 楽器パート音分離のためのフィルタ

本節では、入力信号 $\{Y(k)\}$ に含まれている、第 p 番目の楽器パートで演奏され、ノート番号 i の音高である、第 $l = l(i, p)$ 番目の音 $\{U_l(k)\}$, ($l = 1, 2, \dots, L$) について、その推定値 $\{\hat{U}_l(k)\}$ を得るフィルタを導出する。

まず、合奏音中の第 l 番目の音は次式にモデル化できる：

$$U_l(k) = \sigma_l S_l(k), \quad (26)$$

ここに、音量係数 σ_l , および、要素となる音 $\{S_l(k)\}$, ($l = 1, 2, \dots, L$) に係わる事項は前節と同様である。

いま、推定値 $\{\hat{U}_l(k)\}$ を得るため、インパルス応答 $\{h_l(m)\}$, $(-M \leq m \leq M)$ の非因果型フィルタを考える:

$$\hat{U}_l(k) = \sum_{m=-M}^M h_l(m)Y(k-m) = \mathbf{h}_l^T \mathbf{Y}(k), \quad (27)$$

ここに、インパルス応答ベクトルを定義した:

$$\mathbf{h}_l = (h_l(-M) \quad h_l(-M+1) \quad \cdots \quad h_l(M))^T. \quad (28)$$

このとき、統計的最小二乗規範により、第 l 番目の音の推定に最適なインパルス応答ベクトル \mathbf{h}_l , $(l=1,2,\dots,L)$ は次式の評価関数を最小にする:

$$\begin{aligned} C(\mathbf{h}_l) &= E\{(U_l(k) - \hat{U}_l(k))^2\} \\ &= E\{(\sigma_l S_l(k) - \mathbf{h}_l^T \mathbf{Y}(k))(\sigma_l S_l(k) - \mathbf{h}_l^T \mathbf{Y}(k))^T\} \\ &= \sigma_l^2 r_{ll}(0) - 2\sigma_l \mathbf{h}_l^T \mathbf{r}_{yl} + \mathbf{h}_l^T \mathbf{R}_{YY} \mathbf{h}_l \\ &= (\mathbf{h}_l - \sigma_l \mathbf{R}_{YY}^{-1} \mathbf{r}_{yl})^T \mathbf{R}_{YY} (\mathbf{h}_l - \sigma_l \mathbf{R}_{YY}^{-1} \mathbf{r}_{yl}) + \sigma_l^2 [r_{ll}(0) - \mathbf{r}_{yl}^T \mathbf{R}_{YY}^{-1} \mathbf{r}_{yl}] \\ &\geq \sigma_l^2 [r_{ll}(0) - \mathbf{r}_{yl}^T \mathbf{R}_{YY}^{-1} \mathbf{r}_{yl}], \end{aligned} \quad (29)$$

以上の展開において、前節に述べた互いに広義の定常な性質を適用した。

従って、(29) 式の等号成立より最適なインパルス応答ベクトルが得られる:

$$\mathbf{h}_l = \sigma_l \mathbf{R}_{YY}^{-1} \mathbf{r}_{yl}, \quad (30)$$

ここに、自己相関行列 \mathbf{R}_{YY} は正定対称で逆行列を有する。また、入力信号のと次式の相互相関ベクトル $\mathbf{r}_{yl} \in \mathbf{R}^{2M+1}$ を用いた:

$$\mathbf{r}_{yl} = (r_{yl}(M) \quad \cdots \quad r_{yl}(-M))^T. \quad (31)$$

このとき、 \mathbf{r}_{yl} は次式に示すとおり、入力信号と第 l 番目の要素となる音との相互相関行列 \mathbf{R}_{Yl} , $(l=1,2,\dots,L)$ の中央の列として求められる:

$$\mathbf{R}_{Yl} = E\{\mathbf{Y}(k)S_l^T(k)\} = \begin{pmatrix} r_{yl}(0) & \cdots & r_{yl}(2M) \\ \vdots & \ddots & \vdots \\ r_{yl}(-2M) & \cdots & r_{yl}(0) \end{pmatrix}, \quad (32)$$

ここに、(8) 式より、 \mathbf{R}_{Yl} は音量係数ベクトル $\boldsymbol{\sigma}$, および、予め評価が可能な要素と

なる音どうしの相互相関行列 \mathbf{R}_{lj} , $(l,j=1,2,\dots,L)$ を用いて次式に求められる:

$$\mathbf{R}_{Yl} = (\boldsymbol{\sigma}^T \otimes \mathbf{I}_{2M+1}) \begin{pmatrix} E\{S_1(k)S_l^T(k)\} \\ \vdots \\ E\{S_L(k)S_l^T(k)\} \end{pmatrix} = \sum_{j=1}^L \sigma_j \mathbf{R}_{jl}. \quad (33)$$

4. 統計量の評価

前節までに述べた $\boldsymbol{\sigma}$ の同定や $\{\hat{U}_l(k)\}$ の推定には、 \mathbf{R}_{lj} および \mathbf{R}_{YY} が必要であり、本節では、二次の定常確率過程におけるエルゴード性を仮定して、これらの期待値を見本過程の時間平均から評価する。

まず、第 l 番目と第 j 番目の要素となる音 $\{S_l(k)\}$ と $\{S_j(k)\}$ は確率過程であり、自己相関行列または相互相関行列 \mathbf{R}_{lj} , $(l,j=1,2,\dots,L)$ は次の構成である:

$$\mathbf{R}_{lj} = \begin{pmatrix} r_{lj}(0) & \cdots & r_{lj}(2M) \\ \vdots & \ddots & \vdots \\ r_{lj}(-2M) & \cdots & r_{lj}(0) \end{pmatrix}, \quad (34)$$

これは時点 k に依存しないため、事前にそれぞれの音を単独に演奏して見本過程 $\{s_l(\kappa)\}$ と $\{s_j(\kappa)\}$ に用いれば、上式中の $r_{lj}(m)$, $(m=0,\pm 1,\dots,\pm 2M)$ を N 個のデータの時間平均として以下のとおり評価できる:

$$r_{lj}(m) = E\{S_l(k+m)S_j(k)\} = \frac{1}{N} \sum_{\kappa=1}^N s_l(\kappa+m)s_j(\kappa), \quad (35)$$

ここに、時点 κ は事前の演奏の開始時刻 τ_0 にて、時刻 $\tau = \tau_0 + \kappa \Delta t$ の値を示す。

一方、確率過程 $\{Y(k)\}$ について、自己相関行列 \mathbf{R}_{YY} は時点 k により緩慢に変化し、見本過程として実測された入力信号 $\{y(k)\}$ を成分とするベクトル $\mathbf{y}(k) \in \mathbf{R}^{2M+1}$ の時間平均として評価できる。さらに、忘却係数 α , $(0 < \alpha < 1)$ のべき乗を重みとした時間平均で近似すれば、以下のとおり漸化式に帰着させて計算量を

低減できる:

$$\begin{aligned}
 \mathbf{R}_{YY}(k) &= \begin{pmatrix} r_{YY}(0) & \cdots & r_{YY}(2M) \\ \vdots & \ddots & \vdots \\ r_{YY}(2M) & \cdots & r_{YY}(0) \end{pmatrix} = E\{\mathbf{Y}(k)\mathbf{Y}^T(k)\} \\
 &= \frac{1}{N} \sum_{k=1}^N \mathbf{y}(k)\mathbf{y}^T(k) \approx \left(\sum_{j=0}^{\infty} \alpha^j \right)^{-1} \left(\sum_{j=0}^{\infty} \alpha^j \mathbf{y}(k-j)\mathbf{y}^T(k-j) \right) \\
 &= (1-\alpha) \sum_{j=0}^{\infty} \alpha^j \mathbf{y}(k-j)\mathbf{y}^T(k-j) \\
 &= \alpha(1-\alpha) \sum_{j=0}^{\infty} \alpha^j \mathbf{y}(k-1-j)\mathbf{y}^T(k-1-j) + (1-\alpha)\mathbf{y}(k)\mathbf{y}^T(k) \\
 &= \alpha \mathbf{R}_{YY}(k-1) + (1-\alpha)\mathbf{y}(k)\mathbf{y}^T(k), \tag{36}
 \end{aligned}$$

ここに、 $\mathbf{y}(k)$ は次式の構成であり、入力信号のエネルギー e^2 は $\mathbf{R}_{YY}(k)$ の対角要素として評価できる:

$$\mathbf{y}(k) = \begin{pmatrix} y(k+M) & \cdots & y(k-M) \end{pmatrix}^T, \tag{37}$$

$$e^2 = E\{(Y(k))^2\} = r_{YY}(0), \tag{38}$$

さらに、 \mathbf{R}_{YY}^{-1} の算出は逆行列の補題[6]により漸化式に帰着できる:

$$\begin{aligned}
 \mathbf{R}_{YY}^{-1}(k) &= (\alpha \mathbf{R}_{YY}(k-1) + (1-\alpha)\mathbf{y}(k)\mathbf{y}^T(k))^{-1} \\
 &= \frac{1}{\alpha} \mathbf{R}_{YY}^{-1}(k-1) - \frac{(1-\alpha)\mathbf{R}_{YY}^{-1}(k-1)\mathbf{y}(k)\mathbf{y}^T(k)\mathbf{R}_{YY}^{-1}(k-1)}{\alpha[\alpha + (1-\alpha)\mathbf{y}^T(k)\mathbf{R}_{YY}^{-1}(k-1)\mathbf{y}(k)]}. \tag{39}
 \end{aligned}$$

5. 実施例

図 1 から図 3 には、順に、バイオリン (Vn) 音、トランペット (Tp) 音、および、フルート (Fl) 音により、大譜表の中央の C 音の直ぐ上の A 音、すなわち、名目周波数 $f_{69} = 440$ [Hz]、を MIDI システム (YAMAHA 社 XGworks) によりサンブ

リング周波数 $f_s = 44.1$ [kHz] で発生した際の自己相関を、横軸として時点の差によって示している。これらは同一音高であるが音色の差異をもたらす自己相関は明瞭に異なる。なお、実際の楽器を用いても、自己相関には再現性があり、音色の差異に依存することを確認している。

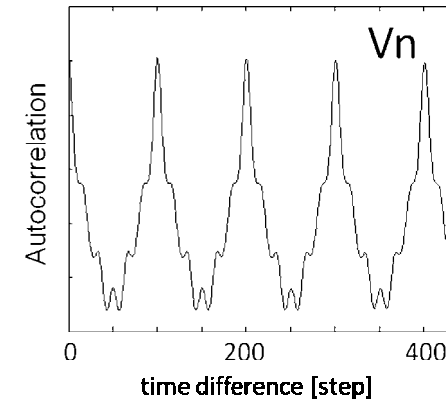


Fig.1 Autocorrelation of violin tone (440Hz)

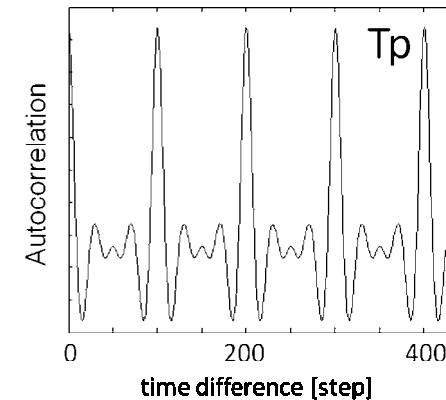


Fig.2 Autocorrelation of trumpet tone (440Hz)

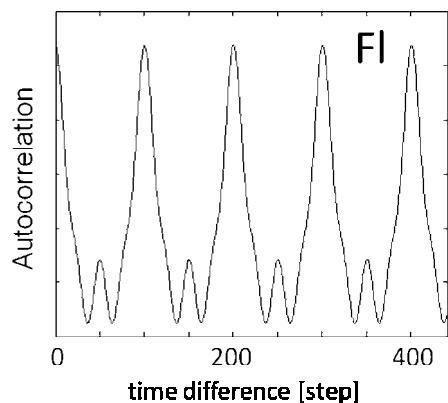


Fig.3 Autocorrelation of flute tone (440Hz)

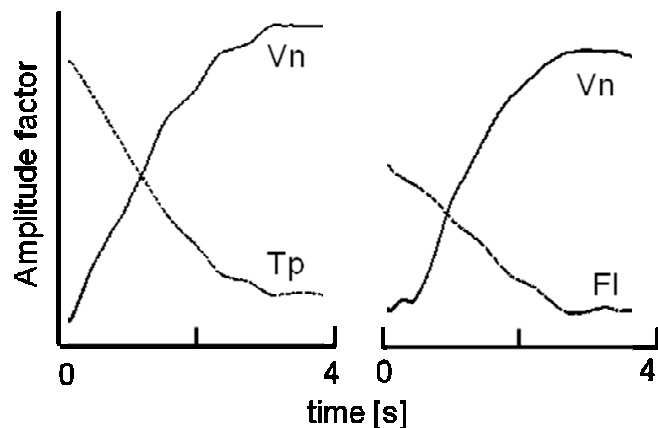


Fig.4 Identification of amplitude factor (440Hz)

図 4 には 2 節に述べた手法の適用例として、同一音高の 2 音の混合より音量係数を同定した結果を示す。これらは 4 節に従い、 $f_s = 44.1$ [kHz], $f_{69} = 440$ [Hz] の音高で MIDI システム (YAMAHA 社 XGworks) を用いて Vn 音, Tp 音, Fl 音の統計量を事前に評価した後、別の MIDI システム (ROLAND 社 VSC) にて、すべて 2 秒

間をかけて、図 4 の左右とも Vn 音の音量を増加させ、同時に左側のケースでは Tp 音を、右側のケースでは Fl 音の音量を減少した場合である。両ケースを通じ、2 節に述べた手法が妥当に機能したことを例証できる。これらの同定結果を、さらに 3 節に述べた推定手法に適用することにより、Vn 音, Tp 音, Fl 音の混合について、互いに音高が異なる場合には極めて明瞭に、同一音高の場合でも、例えば、楽器パートごとの採譜に有効な程度に、それぞれの楽器音に分離できることを確認している。

6. 結論

合奏音のモノラル信号を指定された楽器パートごとの演奏音に分離する際に適用できる下記の特徴をもつ信号処理法を提案した：

- 1) 楽器パートごとの各音高を要素となる音に選び、事前に評価したこれらの自己相関、相互相関に基づき、統計的最小二乗規範を適用
- 2) 合奏音を各要素となる音に音量係数を乗じた和と仮定し、これに白色雑音を加えて入力信号としてモデル化
- 3) まず、各要素となる音の音量係数を同定した後、入力信号と各要素となる音と相互相関を評価して、楽器パートごとの演奏音を推定
- 4) バイオリン音, トランペット音, フルート音を混合し、同一音高でも分離が可能な信号処理性能を確認

参考文献

- [1] Y. Fukayama et al.: Identification of Instruments and Keys for Music on the Time-frequency plane; Proceedings of 41st international symposium on stochastic systems theory and its applications, pp.19-24, ISCIE, Kobe JP (2009)
- [2] MIDI manufacturers association; <http://www.midi.org/aboutmidi/tutorials.php>
- [3] R. A. Roberts and C. T. Mullis: Digital Signal Processing; Addison Wesley, Reading MA (1987)
- [4] M. Kennedy: Oxford dictionary of music; Oxford, New York NY (1985)
- [5] A. Klapuri and M. Davy, editors: Signal processing methods for music transcription, Springer, New York NY (2006)
- [6] G. Golub and C. Van Loan: Matrix computations; Johns Hopkins, Baltimore MA (1983)