

# チャープレット変換と再割り当て法を用いた基本周波数推定\*

☆ 中谷優太, 赤石夏輝, 山田宏樹, 矢田部浩平 (農工大)

## 1 まえがき

チャープレット変換 (CT) は, 周波数変動を考慮した周波数解析である。CT によって, 時間・周波数・チャープレート (T-F-CR) による信号の 3次元表現が得られる。本稿では, チャープレートの情報を用いて調波構造を強調した T-F 表現を得る手法を提案する。また, T-F 表現を利用する基本周波数推定の手法に, 提案手法で得た T-F 表現を利用すると, 短時間フーリエ変換 (STFT) で得た従来の T-F 表現を用いたときと比べて推定精度が向上することを示す。

## 2 チャープレット変換

位相の時間 1 階偏微分は瞬時周波数と呼ばれ, ある時刻における周波数を表す。位相の時間 2 階偏微分はチャープレートと呼ばれ, 周波数の時間変化量を表す。チャープレット変換 (CT) [1] はチャープレートパラメータ  $\beta$  を用いて STFT を一般化したものである。信号  $f$  の CT は実数の窓関数  $g$  を用いて次のように定義される。

$$C_f^g(t, \omega, \beta) = \int_{-\infty}^{+\infty} f(\mu+t)g(\mu)e^{-\frac{j\beta\mu^2}{2}}e^{-j\omega\mu}d\mu \quad (1)$$

ただし,  $j = \sqrt{-1}$ ,  $t$  は時間,  $\omega$  は角周波数を表す。CT は時間信号を T-F-CR 領域に射影する。チャープレットは, 各成分の T-F 平面での傾きを意味する。図-1 に  $\beta = -200, 0, 200$  のときの CT で得られた T-F 平面を示す。 $\beta = 0$  のときは STFT と一致する。 $\beta = -200$  のときは右下への傾き,  $\beta = 200$  のときは右上への傾きを持つ成分をより細く表現する。

## 3 F-CR 平面での再割り当て法

CT で得られる T-F-CR 空間の F-CR 平面について考える。F-CR 平面では, 図-2 (b) のように, 信号の各成分が周波数方向とチャープレート方向の両方に広がる。これを,  $|C_f^g(t, \omega, \beta)|^2$  を  $|C_f^g(t, \hat{\omega}, \hat{\beta})|^2$  に移動させる再割り当て法によって, 図-2 (c) のように本来の周波数チャープレートビンに集めることを考える。周波数方向の再割り当て位置は,

$$\hat{\omega}(t, \omega, \beta) = -\Im\left\{\frac{X_f^{Dg}(t, \omega)}{X_f^g(t, \omega)}\right\} + \omega \quad (2)$$

によって計算できる [2]。ここで,  $X_f^g$  と  $X_f^{Dg}$  はそれぞれ窓関数  $g$  と時間微分した窓関数  $Dg$  を用いて

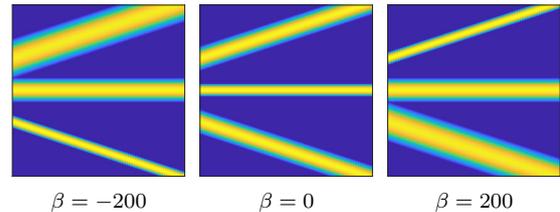


図-1 チャープレット変換の例

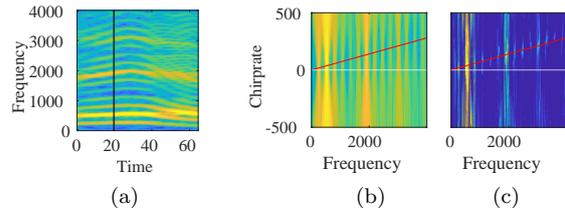


図-2 (a) は STFT で得たスペクトログラム, 黒線は時間インデックス 20 を表す。(b) と (c) は再割り当て前と後の F-CR 平面 (時間インデックス 20), 縦軸がチャープレートで横軸が周波数である。中央の白線は  $\beta = 0$  を表す。視認性のために (c) で振幅が大きい点が並ぶ直線  $\beta = \alpha\omega$  を (b) と (c) に赤線で引いた ( $\alpha = 7/100$ )。

$f$  を STFT した結果を表す。また,  $\Im\{\cdot\}$  は虚部を表す。チャープレート方向の再割り当て位置は,

$$\hat{\beta}(t, \omega, \beta) = \Im\left\{\frac{C_f^g(t, \omega, \beta)^2}{C_f^{tg}(t, \omega, \beta)^2 - C_f^{t^2g}(t, \omega, \beta)C_f^g(t, \omega, \beta)}\right\} + \beta \quad (3)$$

によって計算できる [1]。ここで,  $C_f^{tg}$  と  $C_f^{t^2g}$  はそれぞれ時間で重みづけをした窓関数  $tg$  と時間の 2 乗で重みづけをした窓関数  $t^2g$  を用いて  $f$  を CT した結果を表す。

## 4 提案手法

### 4.1 F-CR 平面における調波構造の特徴

図-2 (c) のように, 調波音は F-CR 平面で一直線に並ぶ性質がある。これは, 調波が周波数変動するとき, 変動の大きさは各調波成分の周波数に比例するためである。チャープレートは周波数の時間変化量であるから, 調波成分は F-CR 平面で  $\beta = \alpha\omega$  の直線上に並ぶ ( $\alpha$  は傾き)。この特徴を用いて調波構造を強調した T-F 表現を作成する。

$\beta = 0$  のとき CT は STFT と一致するので, 図-2 (b) の白線 ( $\beta = 0$ ) が重なっているビンは, STFT で得たスペクトログラムのその時間フレームにおける

\*Pitch estimation using chirplet transform and reassignment method. By Yuta NAKAYA, Natsuki AKAISHI, Koki YAMADA and Kohei YATABE (Tokyo University of Agriculture and Technology)

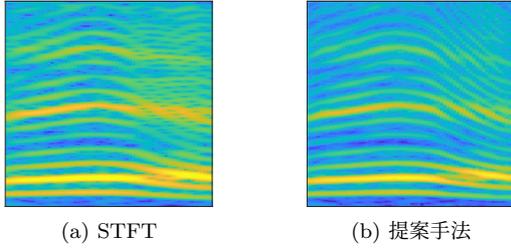


図-3 (a) STFT と (b) 提案手法のそれぞれで得られたスペクトログラム

周波数ビンと一致する。ただし、図-2 (b) では横軸が周波数になっていることに注意されたい。図-2 (b) で、赤線と白線を比べると、赤線が通っている部分の方が調波構造が鮮明になっている。そこで、赤線上のビンとその時間フレームの周波数ビンとすることで、調波を強調したスペクトログラムを構成する。

#### 4.2 調波構造を強調したスペクトログラムの作成

図-2 (c) の赤線のような、振幅が大きい点が並ぶ直線  $\beta = \alpha\omega$  の傾き  $\alpha$  を推定する手法を説明する。 $\beta = 0$  (白線) と  $\beta = \alpha\omega$  (赤線) がなす角を  $\theta (= \tan^{-1} \alpha)$  とおく。 $\theta$  は極座標系の偏角として扱い、 $\beta < 0$  のとき、 $\theta$  の符号は負とする。まず、ある角度  $\theta$  に対応する直線の周辺にある成分だけを抜き出すために、F-CR 平面で  $\theta \pm \epsilon$  の範囲を 1、それ以外は 0 とするマスク  $M(\omega, \beta, \theta)$  を作る。 $\epsilon$  は任意の角度とする。

$$M(\omega, \beta, \theta) = \begin{cases} 1 & (\theta - \epsilon < \tan^{-1} \frac{\beta}{\omega} \leq \theta + \epsilon) \\ 0 & (\text{otherwise}) \end{cases} \quad (4)$$

次に、各時刻における再割り当て後の結果  $\hat{C}_{f,t}^g(\omega, \beta)$  にマスク  $M(\omega, \beta, \theta)$  をかけた上でチャープレート方向で平均をとったものを  $A_t(\omega, \theta)$  とする。

$$A_t(\omega, \theta) = \frac{1}{n(\omega, \theta)} \sum_{\beta} \hat{C}_{f,t}^g(\omega, \beta) M(\omega, \beta, \theta) \quad (5)$$

$$n(\omega, \theta) = \sum_{\beta} M(\omega, \beta, \theta) \quad (6)$$

最後に、 $A_t(\omega, \theta)$  の全ての  $\omega$  での合計値が最も大きくなるときの  $\hat{\theta}_t$  を求める。

$$\hat{\theta}_t = \arg \max_{\theta} \sum_{\omega} A_t(\omega, \theta) \quad (7)$$

こうして求めた  $\hat{\theta}_t$  から、その時刻における推定値を  $\hat{\alpha}_t = \tan \hat{\theta}_t$  とする。この操作を各時刻で行う。

以上の手順で得たスペクトログラムを図-3 に示す。提案手法で得られたスペクトログラムの方が、STFT で得られたものよりも調波構造が強調されている。

## 5 実験

STFT で得られたスペクトログラムと、提案手法による調波を強調したスペクトログラムを用いてそ

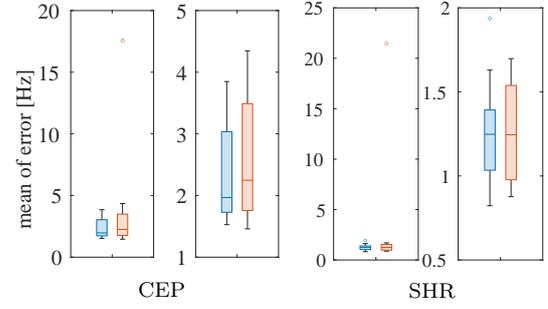


図-4 40 サンプルの音声に対して各手法で推定した結果とリファレンスとの誤差の時間平均の分布。青が提案手法で赤が従来手法である。値が小さいほど精度が高いことを示す。それぞれの手法で、左が外れ値を含む全体図で、右が主要部分の拡大図。

れぞれ基本周波数推定を行い、その精度を比較した。推定手法には、ケプストラム法 (CEP)、調波成分の和と調波間の中央の成分の和との差を利用する手法 (SHR) [3] の 2 つを用いた。

音声データは APLAWD database を用いた [4]。男女 10 人ずつで 1 人 2 サンプル、計 40 サンプルの音声で推定を行った。音声と同時に記録された電気声門図データの基本周波数を MATLAB の pitch 関数で推定した結果をリファレンスとした。

各手法による推定結果とリファレンスとの誤差の時間平均で精度を比較する。その分布を図-4 に示す。各手法の縮図を見ると、両手法ともに従来手法では精度が非常に低いサンプルが 1 つあるが、提案手法ではそれがなくなっている。1 サンプルのみの評価のため断定はできないが、提案手法によって、精度の著しい低下を防ぐことができている可能性がある。CEP の拡大図を見ると、中央値と第三四分位数、最大値が改善していることがわかる。SHR の拡大図についても、最小値、第三四分位数が改善していることから、提案手法の有効性が確認できる。

## 6 むすび

チャープレートの情報を用いて調波構造を強調したスペクトログラムを作成する手法を提案し、それを用いると基本周波数推定の精度が向上することを確認した。今後はチャープレートの情報をさらに活用した基本周波数推定に取り組む。

### 参考文献

- [1] X. Zhu, H. Yang, Z. Zhang, J. Gao and N. Liu, "Frequency-chirpate reassignment," *Digit. Signal Process.*, **104**, 102783, (2020).
- [2] 矢田部浩平, "第六回: 時間周波数領域のスパース表現," 音響学会誌, **77**, 609-616 (2021).
- [3] X. Sun, "Pitch determination and voice quality analysis using subharmonic-to-harmonic ratio," *IEEE Int. Conf. Acoust. Speech Signal Process.*, **1**, 333-336, (2002).
- [4] G. Lindsey, A. Breen, and S. Nevard, "SPAR's archivable actual-word databases," *Univ. Coll. Lond, Tech. Rep.*, (1987).