

今井 聖 「ケプストラムを用いた音声処理技術の樹立」

IMAI, Satoshi, “Establishment of Speech Processing Technology using Cepstrum”

1. 略歴

1937 年 1 月 12 日生まれ。1959 年 3 月東京工業大学卒業，
1964 年 3 月同大学大学院博士課程修了，工学博士。1964 年 4 月
東京工業大学精密工学研究所助手，1968 年 2 月同大学精密工学
研究所助教授，1979 年 1 月同大学精密工学研究所教授。1997 年
3 月同大学停年退職，名誉教授。1997 年 4 月千葉工業大学教授。
2007 年 3 月同大学定年退職。2014 年 12 月 14 日逝去（77 歳）。



今井 聖

2. 要約

音声情報処理の分野において，音素と呼ばれる音声の言語的な最小単位の性質特徴を効率よく表現する物理量として，ケプストラムと呼ばれる特徴量の重要性に早い時期から着目し，一貫してケプストラムを用いる音声処理技術の研究・開発を行った。音声波形から音響特徴量としてのケプストラムを精度よく抽出する音声分析手法の確立，当該特徴量を用いて入力文字テキストから音声波形を生成する音声合成手法，音声波形から文字テキストに変換する音声認識手法，音声信号を効率よく伝送・蓄積する音声符号化手法等，ケプストラムに基づく新たな音声処理手法を開発してこの分野の発展に貢献した。一連の研究業績に対し，第 21 回精機学会論文賞（1965 年度），電子情報通信学会フェロー称号（2002 年），日本音響学会功績賞（2005 年）が授与された。また，雇用促進事業団や大学基準協会等の委員を歴任し，社会にも貢献した。

3. 研究業績要約

3.1 ケプストラムに基づく音声分析法に関する研究

対数スペクトルの逆フーリエ変換として定義されるケプストラムを音声処理に利用する試みは，音声処理研究の歴史の中で初期の段階からなされていたが，従来のケプストラム推定法では原理的に元の対数スペクトルに対する不偏推定量を求めることができず，必ずしも良い推定値が得られていたわけではなかった。これに対し，与えられた対数スペクトルの不偏推定量に対応するケプストラムを求める手法[1]を世界に先駆けて提案した。提案法により有声音や無声音などの音声波形の種類に依らず安定した音響特徴量の推定が可能となり，その結果得られたケプストラムパラメータを音声合成，音声認識，音声符号化等のシステムに利用した場合，システム性能が向上することを明らかにした。さらに，提案法はその後の研究により，従来異質な音声分析法として扱われてきたケプストラム法と線形予測法の二つの手法を特別な場合として包含する統一的な分析手法へと発展した。

3.2 自然性の高い音声合成システムの実現に関する研究

ケプストラムパラメータから音声波形を生成するには、非線形変換（指数関数変換）を含む音声合成フィルタを実現する必要がある。この問題に対し、ケプストラムパラメータがそのままフィルタ係数となる対数振幅近似（LMA）フィルタおよびメル対数スペクトル近似（MLSA フィルタ）と呼ぶ一連の新たなフィルタ構造を提案した[2]–[5]。そして、これを音声合成フィルタに利用した実時間規則音声合成システムを国内で最も早い時期に開発した。2000 年代以降、多様な音声合成手法として統計的パラメトリック音声合成方式が注目され、その後ニューラル音声合成に主流が移るまで、国内外の多くの機関における研究開発、企業による音声合成システム実用化、歌声合成などのエンターテインメント分野応用等において、MLSA フィルタが波形生成部分に広く利用された。

3.3 聴覚特性を考慮した音声処理に関する研究

音響特徴量として人間の聴覚特性を考慮したケプストラムを用いる新たな音声分析・合成手法を提案[5][6]し、国内のこの分野における研究の先導的役割を果たした。それまで陽に考慮されてこなかった聴覚特性を反映した音響特徴量を得るために、オールパスフィルタ（双一次変換）の位相特性が聴覚の周波数分解能に関するメル周波数目盛りをよく近似することに着目し、メル周波数軸上で表現された対数スペクトルからメルケプストラムと呼ぶパラメータを抽出する分析手法を確立した。そして、これを音声合成、音声認識、音声符号化等に利用することによりその性能が向上することを明らかにした[7][8]。

4. 参考文献

- [1] Satoshi Imai, Chieko Furuichi, “Unbiased estimator of log spectrum and its application to speech signal processing, Proc. EUSIPICO-88, pp. 203–206, Grenoble, France (1988.9)
- [2] Satoshi Imai, Tadashi Kitamura, and Hiroyuki Takeya, “A direct approximation technique of log magnitude response for digital filters,” IEEE Trans. Acoustics, Speech, and Signal Processing, Vol. ASSP-25, No. 2, pp. 127–133 (1977.4)
- [3] Satoshi Imai, “Digital filter design in the quefrency domain,” IEEE Trans. Acoustics, Speech, and Signal Processing, Vol. ASSP-26 No. 3, pp.2 26–235 (1978.6)
- [4] 今井 聖, “対数振幅近似(LMA)フィルタ,” 電子情報通信学会論文誌 A, Vol. J63-A, No. 12, pp. 886–893 (1980.12)
- [5] 今井 聖, 住田一男, 古市千枝子, “音声合成のためのメル対数スペクトル近似 (MLSA) フィルタ,” 電子情報通信学会論文誌 A, Vol. J66-A, No. 2, pp. 122–129 (1983.2)
- [6] Satoshi Imai, “Cepstral analysis synthesis on the mel frequency scale,” Proc. IEEE ICASSP ’83, pp. 93–96, Boston, USA (1983.4)
- [7] 今井 聖, 音声認識, 共立出版 (1995)
- [8] 今井 聖, 音声信号処理 音声の性質と聴覚の特性を考慮した信号処理, 森北出版 (1996)

(2023-09-11 小林隆夫-談)