

聖徳太子をつくる技術

—編集にあたって—

柏野 邦夫

NTTコミュニケーション科学基礎研究所

kunio@ca-sun1.brl.ntt.co.jp

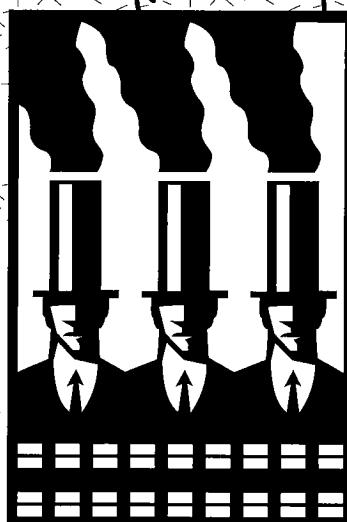


音声認識技術の進歩により、パソコンの操作なども、音声で手軽にできるようになった。

しかし、人間の耳と同じような機能をコンピュータ上に実現できたかというと、答えは否である。コンピュータの耳と人間の耳とで、機能的に最も大きく異なる点の1つは、聞き分け能力の有無である。コンピュータの音声認識では、目的とする人の声以外の音が入力信号に混ざっていると、認識精度の低下が避けられない。一方人間の場合には、パーティーや会場のようなかなり騒々しい場所でも、目的の相手と難なく会話できることが多い。

音声認識において目的以外の音に対処する方法は、これまでにも盛んに研究されてきた。複数のマイクロフォンを用いて集音の指向性を制御する方法、音の倍音構造に着目して目的の周波数成分を選択する方法、雑音を加味した認識辞書を用いる方法などはその代表的な例である。これらはいずれも成果を挙げつつ発展しているが、その一方で近年、新しいアプローチが登場した。それが本小特集で紹介する2つの技術である。

ここで紹介する2つの技術の共通点は、従来より一歩



進んだ聞き分け機能を目標として提案された手法であるという点である。従来の方法が、普通の人間並みの処理を目標としているとすれば、ここで紹介する方法は、一度に10人の訴えを聞き分けたと伝えられる聖徳太子並みの処理を目標としたものといえる。

どうしてそのようなことが可能と見込まれるのか。その具体的な考え方はそれまでまったく異なっている。1つ目は、あくまで数理的な立場から、混合された信号から独立した信号を分離する新しい手法を提案するものであり、2つ目は、あえて数理的・信号処理的な枠組みにとどまらず、知識処理を積極的に導入しようという考え方である。

本小特集は、このようなある意味できわめて対照的な手法を、それぞれの第一人者に解説していただくことを趣旨としたものである。今後本当に聖徳太子のようなコンピュータが実現できるのか、さまざまな議論が活発化する端緒となれば幸いである。末筆ながらご多忙の中で執筆くださるとともに編集にご協力いただいた著者の皆様に感謝申し上げる。

(平成12年8月1日)

