

先端技術キーワード解説

知っておきたい最新の動き

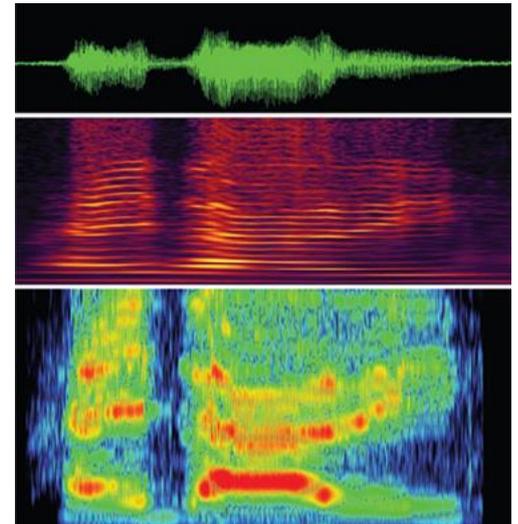
【声紋分析 (Sound Spectrogram Analysis)】

先月（2014年6月）、東京都議会でセクハラと思われるヤジがありました。直後、誰も名乗り出ないために、ヤジ発言者を特定（犯人捜し）するために、科学捜査で行う「声紋分析」をすべきという意見が出ました。都議会という決まったメンバーだけで構成する場、しかも、発言者は都民から選ばれた議員です。そこに、犯罪者を捜査するための「声紋分析」が持ち出されるとは、何とも呆れてしまいます。ところで、この「声紋分析」、どのような技術でしょうか。本当に、ヤジ発言者を特定できるのでしょうか。

始めに、人の音声は、どのように作られるかを確認しましょう。音声は、声帯の振動を声道で共鳴させることで形成されます。発生源となる声帯振動は、100~200Hzの単振動です。音叉のブーという振動とほぼ同じです。これが、声の通り道である声道を通過する時に、特定の周波数帯域が共鳴、つまり、選択、増幅されることで音声を作られます。この声道は、頬骨、上顎骨、下顎骨等で構成される口腔、さらに鼻腔、副鼻腔で構成されます。

つまり、音声の特徴を作るのは、その人の発声方法で決まる声帯振動、および、声道における骨格などの身体的な構造です。

声紋は、声を分解し、音の高低と強弱、時間的変化を模様としたものです。具体的には、スペクトログラム (Spectrogram) という形式で表されます。対象信号を、窓関数を介して周波数スペクトルを計算したものです。一般的には、横軸が時間、縦軸が周波数を表します。そして、各点の強さ（振幅）を明るさや色で示します。なお、音声を解析して声紋の形で示す装置をサウンドスペクトログラフといいます。（右図は、文献1）より引用）



それぞれのスケールには、線形と対数があります。一般に、音声信号であれば、強さを対数目盛（通常、dB）で表し、倍音の関係を示す場合は、周波数を線型目盛で表します。一方、音楽的または音色的関係を表す場合は周波数を対数目盛で表します。

声紋分析で対象とする周波数は85~8000ヘルツの範囲です。人が識別できる音の周波数は、20~20,000ヘルツ、これに対して、人が声として出せるのは80~1,000ヘルツの範囲とされています。このため、分析は、声として出せる範囲に、若干、幅を持たした周波数範囲としています。

声紋分析で対象としているのは、主に骨格などの身体的な構造による声道による個体差です。声帯振動に関連する、声の高さ・声の強さ・音色（かすれ声、鼻声など）については、分析の対象としていません。これら声帯振動に関する特徴は、変わることがあるから、あるいは、変えることが可能だからです。

声紋に表れる個人の特徴は、細かい点まで含めると百数十項目にも及ぶとも言われています。従って、すべてが一致するのは同じ指紋の人が見つかるくらいにわずかな確率だそうです。

このため、声紋鑑定は重要な証拠として、捜査や裁判に至るまで利用されているのです。

ただし、留意すべき点もあります。プライバシーに関連する問題です。一般に、個人を特定する生体情報は、犯罪の現場で、あるいは、本人の了承を得て採取します。例えば、指紋、筆跡、DNA などです。ところが、声紋は、録音さえできれば採取できる生体情報です。気づかない所で、データが採取され、ある日突然、何かの行動(?)を特定される可能性があります。

最も、議会での発言は、何であれ、特定されるべきです。議会で、誰が言ったかわからない発言が許されるようでは、民主主義の空白地帯になってしまいます。

(参考文献)

1) 科学鑑定研究所 ホームページ <http://www.e-kantei.org/>

(注)

本解説は、執筆当時の状況に基づいて解説をしております。ご覧になる時には、状況が変わっている可能性がありますので、ご注意ください。

Copyright (C) Satoru Haga 2014, All right reserved.

技術・経営の戦略研究・トータルサポーター	工学博士 中小企業診断士 社会保険労務士(登録予定)
ティー・エム研究所	代表 芳賀 知
E-Mail: info_tm-lab@mbn.nifty.com	URL: http://tm-lab@a.la9.jp/