

ヒューマンインターフェース

Report3

氏名：津波古正輝

提出日：8月11日(火曜日)

ITU-T G.711.1 広帯域音声符号化方式

内容

これからは圧縮率優先から広帯域音声再生、低遅延、低演算量を特徴とした高品質な音声符号化が求められている。NTTほか4社の韓国、フランス、中国、カナダが協力してG.711を拡張した広帯域に対応するG.711.1の標準化を行った。従来の方式では明瞭、肉声感にけるものだったが、G.711.1ではそのような損なわれている部分や、環境音、音楽、様々な広帯域のものも高品質に伝えることを可能にした。優れたパケット消失補償、少ない遅延と演算量、安価なコストでの実装を可能にし、人の声の聞き間違いが少なくなるとともに、リアルな音声を伝えることができる。

感想

今の電話でも不便は感じない。この通信技術が現在自分達が使っている技術の限界だと思っていた。NGNというキーワードを聞いた時に更に上の技術があるのは知っていたが、実用化の段階まできているのは予想外だった。広帯域というイメージができなかったが、今までより更にクリーンな音で周りの微細な音まで電話で伝えられる、というのが私の解釈だった。上等なヘッドホンは音楽がリアルで、臨場感があふれるものがある。それくらい高品質な電話ができるようになるのかなと思った。

標準化に向けての苦勞を記述した資料があったがとても多いことに驚いた。私だったら精神がどうにかしてしまう。これがこの業界で一般的な忙しさのならば私はとてもやっていけそうにないと思った。多忙なスケジュールで、異国の人との作業、仕事の期限、と胃が痛くなるようなものばかりだった。通信技術が発達している現在から推測するにこれから通信容量はどんどん増え、データの圧縮技術も更に進歩していくだろう。一体どこまで進化して、何に使われるのか。想像するのが楽しかった。

聖徳太子ロボット

内容

混合音の中から個々の音声を聞き分ける技術の開発を形にしたのが聖徳太子ロボットである。ロボットにコミュニケーション能力をつけさせ社会性を持たせたい。その為の第一歩として、多くの人の会話を別々に認識させることが必要になった。その為には複数の音の認識、音源の位置を見極める、音の分離が必要がある。HMM 音声認識 (音を個別に分ける技術)、動的計画法 (DP)、ミッシングフィーチャー理論などを利用し、複数の人の会話を聞き取れる聖徳太子ロボットを開発した。

感想

複数の音を同時に認識するというのは簡単なことだと思っていたが、想像以上に難しいということがわかった。音を認識することと音を理解することがこれほどまでにレベルが違うことに驚いた。音を認識し、理解するという人間が無意識でやっている処理がどれだけすごいことなのかを実感した講義内容だった。どうやって音を音と認識し理解しているのか、を答えようとしても私 (人間や動物) にとっては当たり前なことだったので深く考えた事はない。動物はすごい機能を備えているということを実感した。

この技術は人間の聴覚に関するものである。人間の他の感覚で味覚、触覚、視覚、嗅覚があるが、これらをコンピュータに理解させることは聴覚と同様に難しいのであろう。他の4つの感覚を研究している話も聞きたいと思った。

$$E = \sum_{n=0}^{N-1} [w(n) * (x(n) - h(n) * g_1 * a(n))]^2 \text{ の導出}$$

$$E = \sum_{n=0}^{N-1} [w(n) * x(n) - w(n) * h(n) * g_1 * a(n)]^2$$

$Xw(n) = w(n) * x(n), Sw(n) = w(n) * h(n) * a(n)$ とする。

$$E = \sum_{n=0}^{N-1} [Xw(n) - g_1 * Sw(n)]^2$$

$$E = \sum_{n=0}^{N-1} Xw(n) * Xw(n) - 2g_1 \sum_{n=0}^{N-1} Xw(n) * Sw(n) + g_1^2 \sum_{n=0}^{N-1} Sw(n) * Sw(n) \dots$$

$$\frac{\partial E}{\partial g_1} = -2 \sum_{n=0}^{N-1} Xw(n) * Sw(n) + 2g_1 \sum_{n=0}^{N-1} Sw(n) * Sw(n) = 0$$

$$g_1 = \frac{\sum_{n=0}^{N-1} Xw(n) * Sw(n)}{\sum_{n=0}^{N-1} Sw(n) * Sw(n)} \text{ より に代入}$$

$$E = \sum_{n=0}^{N-1} Xw(n) * Xw(n) - 2 \frac{\sum_{n=0}^{N-1} Xw(n) * Sw(n)}{\sum_{n=0}^{N-1} Sw(n) * Sw(n)} + \frac{\sum_{n=0}^{N-1} Xw(n) * Sw(n)}{\sum_{n=0}^{N-1} Sw(n) * Sw(n)}^2$$

$$E = \sum_{n=0}^{N-1} Xw(n) * Xw(n) - \frac{\sum_{n=0}^{N-1} Xw(n) * Sw(n)}{\sum_{n=0}^{N-1} Sw(n) * Sw(n)}^2$$