

音声信号処理技術（携帯電話）

ご利用にあたっての注意

この講座の内容は、2010年当時の情報です。予告なしに更新、あるいは掲載を終了することがあります。あらかじめご了承ください。

最終更新日 2010年10月25日

もくじ

↓ 音声信号処理技術ってなんだろう

↓ 携帯電話の音声信号処理

携帯電話「らくらくホン」の技術

- ゆっくりボイス
- はっきりボイス

↓ 小話

↓ 関連ページへのリンク



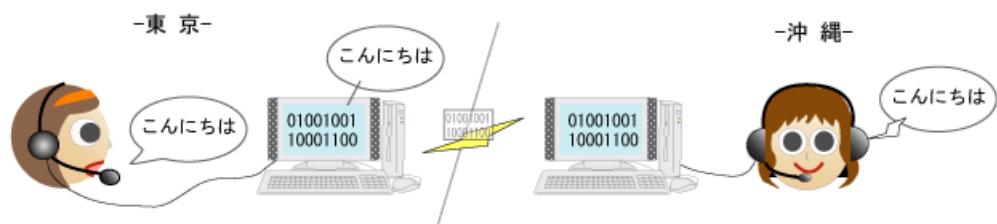
音声信号処理技術ってなんだろう

電話などで人と通話する時に、相手の声が聞き取りづらかったら誰でも困りますよね。音声信号処理とは、利用者に良い音質で通話してもらう技術です。実際に携帯電話やインターネットを使ったパソコン電話などで使われています。

音声信号処理の仕事ってどんなことかな

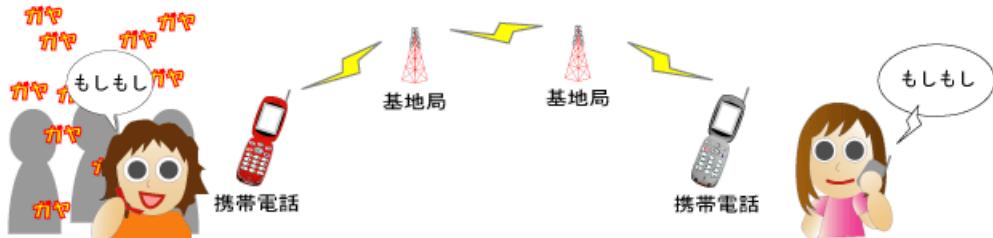
私たちが意識しなくとも、通信機器の中で自動的に音声信号処理の仕事はおこなわれています。人間の声はマイクで録るとアナログ信号となりますが、これを機械で扱いやすくするためにデジタル信号に換えます。このデジタル信号となった音声を、ノイズなどの不要な信号を取り除いたあと、相手にスムーズに送れるように小さく圧縮して送ります。受信側では圧縮された信号を復元し、聞き取りやすく加工して、最後はアナログ信号に戻して音声として聞かせます。

このように人がより快適におしゃべり(通話)できるようにしてくれる技術です。



携帯電話の音声信号処理

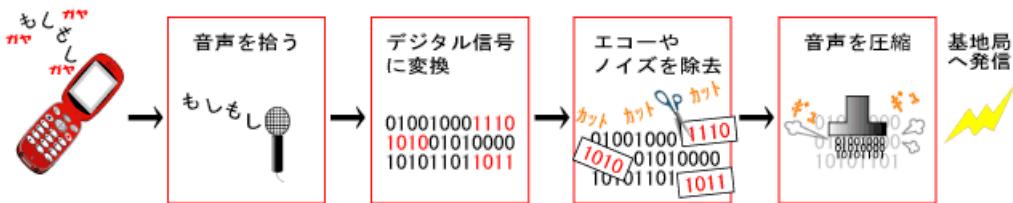
携帯電話では通話相手の声以外に周囲の雑音も同時に送られてきますが、できれば雑音は少ないほうが聞き取りやすいですよね。そこで音声処理技術を使って相手の声を聞きやすくしています。



携帯電話のしくみ

送信側

1. 人の声をアナログ信号からデジタル信号に変換します。
2. 邪魔になる音（ノイズやエコー）をカットして、音質をよくします。
3. 信号を伝送しやすいように圧縮して送ります。



受信側

1. 圧縮されて送られてきたデータを復元します。
2. 聞き取りやすいように音声をゆっくりにしたり、はっきりした声に加工します。
3. デジタル信号からアナログ信号へ変換し、人が聞き取れるようにします。

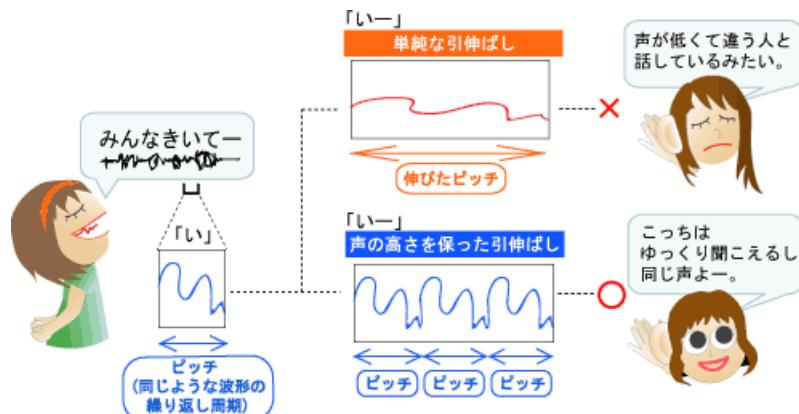


携帯電話「らくらくホン」の「ゆっくりボイス」

携帯電話の相手の声をゆっくりさせて聞きやすくする技術です。ゆっくりさせても違和感なく会話することができて、本人の声を保ったまま聞こえます。携帯電話の中では、どのような処理がおこなわれているのでしょうか。

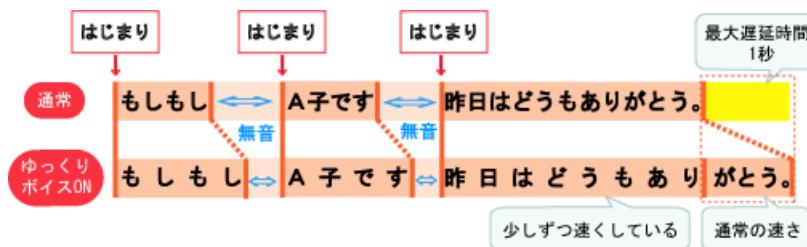
伸ばしてもアナタの声「音声伸長技術」

テープレコーダーをゆっくり再生させると本人の声より低く、別人のように感じます。それは、声を単純に引伸ばして再生しているからです。声の波形には、ピッチ(同じような波形が繰り返される周期)があります。ピッチを変えると、声の高さも変わります。この技術では、同じピッチのまま(声の高さを保ったまま)繰返すことで、本人の声のままゆっくり聞こえます。



ゆっくり再生しても会話がズレない「遅延制御技術」と音のない瞬間を見逃さない「無音検出技術」

声を引伸ばしたからといって、テレビ番組の衛星中継のようにズレてしまったら、スムーズに会話を楽しむことができません。ゆっくり再生しても楽しく会話をするために「文章中の無音(音の無い瞬間)を使って、声のはじまりをなるべくそろえる」というルールを守っています。その無音というのは、「もしもしA子です昨日はどうもありがとう」という文章の中にも、無音があります。「もしもし(無音)A子です(無音)昨日はどうもありがとう」というように隠れています。その無音の瞬間を見逃さないのが「無音検出技術」です。また、声をそろえるための無音がない場合でも、違和感を感じさせないズレの最大値を決めています。ズレが1秒以上遅延する時には、少しずつ元の速さに戻すのですが、ゆっくり感を保ちながら違和感を感じさせないように戻しかたを工夫をしています。



携帯電話「らくらくホン」の「はっきりボイス」

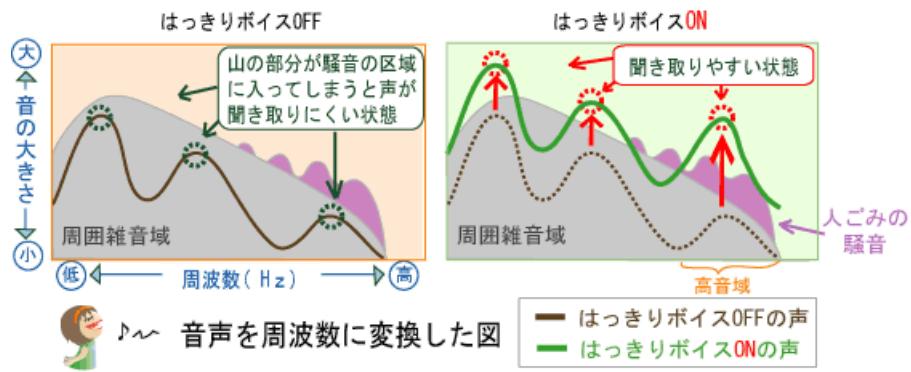
携帯電話の相手の声を聞き取りやすくする技術です。例えば駅のホームやショッピングセンターなど周囲が騒がしい場所で、携帯電話を利用した時、騒音を自動判別して、相手の声を強調してはっきり聞こえるようにします。

どうやって声をはっきりさせているのかな（概要）

音声を横軸に周波数、縦軸を音の大きさ（電力値）であらわしてみると、複数の山を持つ波のような图形になります。山の頂点が雑音域の中に埋もれてしまうと、声が聞き取りにくい状態になります。そこで、山の頂点が雑音域から飛び出した状態になるように音を大きく（電力値を高く）して、相手の声を聞き取りやすくしています。

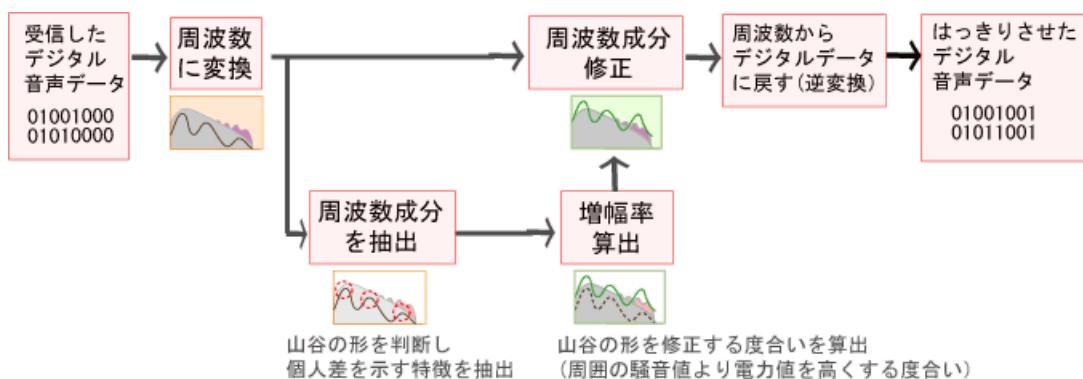
また、音声は周波数が高くなるにつれて、音が小さくなる（電力値が低くなる）傾向があります。周波数の高い高音域の音を大きくする（電力値を高める）ことで、より相手の声をはっきりさせることができます。

場内アナウンスや近くを通り過ぎる人の声など相手の声に似た騒音でも自動判別して、相手の声をはっきりさせることができます。



どうやって声をはっきりさせているのかな（詳細）

受信したデジタルデータを周波数に変換します。次に個人の特徴を示す周波数成分を抽出し、電力値が低い部分と比較して、どれくらいの修正が必要か割り出します。そして、周波数成分を修正して、デジタルデータに戻します。この後は、アナログデータに変換して、音声として聞かせます。



レーシングカーに搭載

レース中、運転するドライバーとピット（チームオーナー、マネージャー、レースエンジニアなどがレース中にいる場所）では、すごい騒音の中、音声をやりとりする必要があります。そこで、フォーミュラ・ニッポン（DoCoMo TEAM DANDELION RACING）のレーシングカーに「はっきりボイス」技術が搭載され、その効果が実証されています。



小話

「ゆっくりボイス」で必要になる「有聲音」について

私たちが普段発音する音には、有声音と無声音があります。

発音するときに、喉に手を当ててみてください。

「あー」「いー」「うー」などは、当てた手に喉の振動が伝わってくるでしょう。これが有声音です。

「すー」「つー」「ふー」などは、当てた手に何も振動が伝わってこないでしょう。これが無声音です。

その違いがわかったでしょうか。

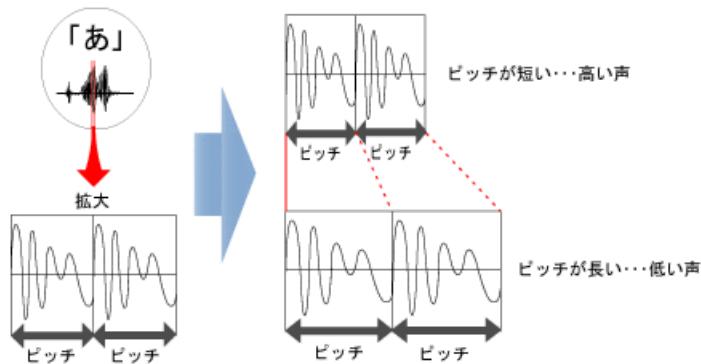


少し難しくなりますが、声の波形には、ピッチ（繰り返し周期）というものがあります。人それぞれの声の特徴は異なりますが、ピッチの違いが、声の特徴として表れます。

例えば、ピッチが短いと高い声になり、ピッチが長いと低い声になります。

そして、有声音にはピッチが存在しますが、無声音にはピッチは存在しません。

「ゆっくりボイス」では、このピッチによる本人の声の特長が変わらないように声をゆっくりさせています。



音の聞きやすさの市場調査

どういう音声が聞きやすいか、というのは人によって違うものです。

それを万人に丁度聞きやすいものにするには、どのような実験をしてデータを集めたのでしょうか。

そもそも「聞きやすさ」は人間が感じ取る感覚なので、機械で調べ、良い数値が出たからといって、必ずしも聞きやすいとは限らなかったそうです。毎日の開発では、研究員が何度も聞き比べ、聞きやすいかどうか判別していますが、製品化までには何回か一般利用者に聞いてもらい、その意見も反映させながら、より聞きやすい音を作り上げていくそうです。

そこで、研究員とは別の一般のユーザである20代から80代の人達に「聞きやすさ」を調査しました。

その結果、ゆっくり過ぎても聞き取りづらいことが分かり、「万人に聞きやすいレベル」を見つけることができました。



関連ページへのリンク

プレスリリース

- › 洗い流せるから安心「らくらくホン6」新発売（2009年7月28日）

製品紹介

- › FMWORLD（携帯電話）らくらくホン7
- › FMWORLD（携帯電話）らくらくホンプレミアム

その他

- › 富士通の研究開発