

音階および聴覚特性を考慮した高能率符号化*

中沢 誠 山崎芳男（千葉工大）

1. まえがき

音響信号を聴感上遜色なく高能率に符号化する手法はMPEG等で実用化されている。

天田らは、一般化調和解析により抽出された成分を聴覚特性の1つである周波数弁別機能を考慮し1/1200から1/24オクターブで周波数をまめ、さらにマスキング等の聴覚特性を用いて高能率符号化をこころみてきた。

本稿では、音階を考慮した1/nオクターブ分析による一般化調和解析を用いて抽出された成分に対し同時マスキングや継時マスキングといった聴覚特性を考慮した高能率符号化について報告する。

2. 一般化調和解析

一般化調和解析はN.Wienerにより提案された周波数解析の手法である。

一般化調和解析は、観測区間内で原波形からの残差が最小となる純音を抽出し、残差成分に対しても同様の処理を繰り返すという単純明快な解析手法であり、現在周波数解析に広く用いられているFFTに比べ計算量は桁違いに多い。しかし、解析区間の窓幅に関係なく正確な周波数解析が可能であり、またその解析周波数を任意に設定できる等多くの利点を有している。

3. 1/nオクターブ分析

人間の聴覚の周波数分解能は指数関数的である。よって、指数関数的な周波数間隔で抽出した成分でも聴感上原音を再現することが可能と考えられる。そこで、一般化調和解析の利点の一つである周波数を任意に設定できる点に着目し、周波数分解能を指数的に設定し、周波数解析を行う。

また、解析周波数は音階を基準とした1/36オクターブとする。これは、音階が人間の聴覚特性が顕著に表れている良い例と考える事ができるからであり、音階に相当する周波数を用いて成分を抽出すれば、高能率な符号化が期待できる。しかし通常、一般化調和解析において単体の楽器音を違和感なく再現するには基音とその10倍音程度までの周波数を考慮すればよいといわれており、1/12オクターブ、つまり12音階で抽出した場合、5倍音から倍音成分を抽出できなくなる。しかし1/36オクターブにした場合、10倍音までの抽出が可能になり、図-3に示すように聴感上原音と違いのない音に復元可能である。

4. 聴覚特性を考慮した高能率符号化

高能率符号化には情報量に着目したエントロピー符号化と人間の聴覚特性に着目した符号化がある。MPEG等の高能率符号化では両者を用いて聴感上遜色ない音質で高い符号化を実現している。両者のうち人間の聴覚特性に着目した符号化は高能率符号化を行う上で非常に有効な手段であり、例えば図-1のようにある周波数成分の影響による最小可聴値レベルが上昇し、その周波数成分が聞こえなくなるマスキング効果等が存在する。

マスキング効果にはマスクする音とマスクされる音が時間的に重なりあった時に生じる同時マスキングに加え、マスクする音とマスクされる音が時間的に差がある場合でも時間差が短ければ生じる継時マスキングという効果がある。これらを用い聴感上原音を再現するのに必要な周波数成分以外を除去することで、大幅な情報量の削減が可能である。