

Resumo

O familiar Disco Compacto permite o registo de sinais áudio que preenchem quase totalmente a região de funcionamento útil do ouvido humano. É por isso comum a sua designação de áudio de alta qualidade. O tema desta dissertação é a compressão deste tipo de sinais, sem afectar ou comprometer a sua qualidade subjectiva original. Com este objectivo, só se admitem duas premissas: os sinais áudio são representados por dois canais (estereofonia) e o receptor chama-se ouvido humano. É sobre este último elemento que se procurará fundamentar o elevado ganho de compressão.

A codificação perceptual de sinais áudio de alta qualidade é uma área de investigação que derivou da compressão de sinais de voz e que se demarcou desta e autonomizou na década de oitenta. Este percurso de desprendimento foi também o permitido pela própria evolução tecnológica e o esboçado pela sedimentação de conhecimentos acerca do sistema auditivo humano. As actividades de normalização no âmbito da ISO/MPEG, ainda em curso, estimularam também a consolidação de resultados.

A psico-acústica fornece importantes medidas de correlação entre estímulo acústico e sensação auditiva, permitindo a identificação da tolerância auditiva com a irrelevância do sinal e da acuidade auditiva com a Entropia Perceptual. Estes aspectos podem aglutinar-se numa entidade matematicamente tratável para postular um limiar de mascaramento. Neste contexto, o ganho de compressão afirma-se como resultado da extracção da redundância e da irrelevância contidas no sinal áudio. Estas componentes exibem uma grande eficiência no plano intracanal. Porém, a dimensão intercanal é também potenciadora de ganhos adicionais. Neste sentido, partindo de sugestões da psico-acústica, o limiar de mascaramento estereofónico associa-se à percepção binauricular, assim como o limiar de mascaramento monofónico retrata a percepção monauricular.

Os limiares de mascaramento são implementados num codificador perceptual e provà-se que é possível codificar transparentemente ou com muito boa qualidade, qualquer sinal estereofónico, usando uma taxa de informação total não superior a 128Kbit/s.