

MP3 VERSUS AAC

Pedro Sousa e Pedro Ferreira

Instituto Superior Técnico
Av. Rovisco Pais, 1049-001 Lisboa, Portugal
E-mail: {pemigouso;engpedroferreira}@hotmail.com

ABSTRACT

Este artigo tem como objectivo fazer um estudo comparativo à codificação utilizada em MPEG áudio. Os codificadores estudados para esta implementação são o MP3 e o AAC. O codificador MP3 foi um dos primeiros codificadores de compressão de áudio com perdas quase imperceptíveis ao ouvido humano. Por este motivo, notou-se uma grande forte expansão da utilização deste codificador, desde o seu aparecimento, permanecendo ainda nos tempos de hoje. O MPEG AAC é um codificador áudio de alta qualidade com configurações flexíveis, permitindo o uso de muitas aplicações em zonas onde a alta qualidade áudio e a limitação da largura de banda do canal de transmissão, são factores de extrema importância. Este artigo baseia-se em fazer um estudo e descrição geral aos codificadores citados, a fim de se poder fazer uma comparação rigorosa e detalhada sobre os mesmos.

Index Terms — AAC, MP3, MPEG, Audio Compression, Quantization.

1. INTRODUÇÃO

A Tecnologia Áudio tem evoluído de forma tremenda nestes últimos anos. No aparecimento dos sistemas digitais, a reprodução de som atinge o seu estado-da-arte em termos de desempenho de qualidade. No entanto, o elevado ritmo binário existente na música digital não consegue satisfazer a demanda pela limitada largura de banda no canal. Para se conseguir uma transmissão eficaz, um processo de compressão terá que ser empregue. Eficientes sistemas de codificação são aqueles que conseguem eliminar partes irrelevantes e redundantes no canal de áudio. Inicialmente, a análise psico-acústica tem a missão de reduzir as partes irrelevantes existentes. O termo de “codificador perceptivo de áudio” foi criado para explorar as propriedades da percepção auditiva humana a fim de eliminar partes que o nosso ouvido não consegue acompanhar. Entre os sistemas de codificação perceptiva de áudio disponíveis hoje, o Advanced Audio Coder (AAC) e o MP3 são referências de topo, nos mercados dos tempos actuais.

Basicamente, um codificador perceptivo consiste num modelo psico-acústico, um banco de filtros (para

transformação do tempo para a frequência), a unidade de quantização e por fim, a codificação. A figura 1 apresenta a solução da estrutura básica clássica para este tipo de codificadores perceptivos.

O MP3 e o AAC são dois formatos de áudio desenvolvidos pelo grupo Moving Pictures Expert Group (MPEG) da Organização Internacional para Padronização, com o objectivo de estabelecer normas internacionais para a codificação de áudio.

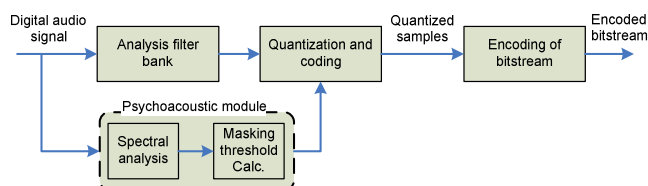


Figura 1 – Estrutura básica de um codificador áudio perceptível.

O MP3 corresponde à terceira camada da primeira norma de áudio, MPEG – 1 Áudio, existindo ainda duas camadas inferiores de menor complexidade, o MP1 e o MP2. As camadas de maior complexidade assentam-se nas camadas de menor complexidade. O AAC faz parte da segunda norma de áudio, o MPEG – 2, não compatível com o primeiro padrão, que tem como finalidade melhorar o desempenho obtido pelo primeiro.

Estas duas normas de áudio nascem da necessidade de comprimir as amostras PCM (Pulse Code Modulation) dos sinais de áudio, com o intuito de reduzir os débitos necessários para transmiti-los e armazená-los. Para se conseguir taxas de compressão elevadas (da ordem das 8 a 20 vezes em relação à codificação PCM) tanto o MP3 como o AAC utilizam codificação com perdas, que exploram a irrelevância perceptiva do sistema auditivo humano, bem como a redundância estatística dos sinais de áudio, sem que exista perda de qualidade subjectiva em relação ao sinal original. O MP3 e o AAC têm aplicações em inúmeras áreas tais como: produção de áudio, transmissão de som televisivo, armazenamento digital, produção e distribuição de programas.

Os avanços recentes na rede wireless permitiram trazer novos desafios no desenvolvimento de aplicações em dispositivos portáteis, incluindo canais de áudio digital.

Note-se que, não só é desejado um baixo ritmo binário, mas também o codificador e o decodificador devem ser capazes de executar os processamentos em baixa potência nestes dispositivos portáteis.

Na tabela 1, apresenta-se um cronograma simples da evolução da tecnologia MPEG, apresentando algumas características/ferramentas usadas.

Tabela 1 – Cronologia dos codificadores de áudio MPEG.

Designação	Ano	Características
MPEG-1 [2]	1988-1992	- Continha 3 modos de operação: Layer-1 até Layer-3; - Os 3 modos contêm diferentes níveis de complexidade e desempenho sendo o Layer-3 o mais complexo; - Qualidade alta para ritmos binários baixos: 128kbit/s para um sinal stereo).
MPEG-2 [2]	1990-1994	- Introduziu novas concepções e suportes de codificação ao nível de sinais de vídeo MPEG; - Codificação compatível com multicanais, permitindo avançar para codificadores de características multicanais tal como o conhecido <i>5.1 cinema sound</i> ; - Codificação para frequências de amostragens baixas: 16kHz, 22.05KHz e 24kHz; - Os algoritmos de codificação continuavam a ser os mesmos.
MPEG-4 [3]	1995-1998	- Surgiu para se tornar como o codificador padrão de multimédia para todo o mundo; - A ênfase mais saliente deste codificador foi na questão de criar novas funcionalidades e não na parte da eficiente compressão, tal como MPEG – 1/2; - Permite a ligação com novos serviços: Comunicações, acesso a base de dados e terminais; - Engloba uma família mais extensa de codificação áudio, entre codificação de voz, (abaixo de 2kbit/s) até codificação áudio de máxima qualidade (até 64kbit/s).
MPEG-7 [4]-[5]	1998-2001	-Description (D): é uma representação de uma característica ser definida sintáctica e semanticamente. Um único objecto passa a ser descrito por vários Descritores; - Description Schemes (DS): especifica a estrutura e a semântica da relação entre os seus componentes; - Description Definition Language (DLL): é baseado na linguagem XML e permite a alteração dos regimes de descrição e também a criação de novos descritores; - System tools: permite lidar com a binarização, a sincronização, o transporte e o armazenamento dos descritores, protegendo a propriedade intelectual.

2. TECNOLOGIAS DE CODIFICAÇÃO ÁUDIO

As tecnologias utilizadas na codificação de áudio são: codificação perceptiva e mascaramento; codificação no domínio da frequência; comutação de janelas; alocação dinâmica de bits. Para se falar de tecnologias de codificação de áudio é necessário introduzir o conceito de bandas críticas. As bandas críticas são as bandas em que as características auditivas permanecem praticamente constantes.

2.1 Mascaramento auditivo

O mascaramento simultâneo é um fenómeno no domínio da frequência em que um sinal de baixa amplitude (mascarado) se torna inaudível na presença simultânea e a frequências próximas de um sinal com uma amplitude muito maior do que o primeiro (mascarante). O mascaramento é maior na banda crítica em que o mascarador está localizado e menor nas bandas vizinhas. Define-se limiar de mascaramento como o limite em amplitude abaixo do qual todos os sinais são mascarados por um dado mascarante. O limiar de mascaramento depende da amplitude e frequência do mascarante e das características do mascarante e do mascarado. Na figura 1 está representado o mascaramento na frequência, bem como o limite de audição (*threshold in quiet*). O limiar de audição define a intensidade mínima do som em função da frequência. Como se pode observar o limiar de audição decresce com o aumento da frequência.

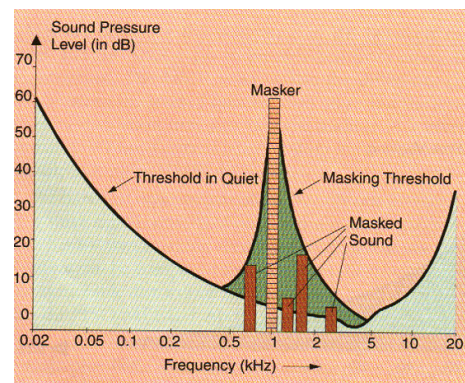


Figura 2 – Mascaramento na frequência.

A distância entre o mascarante e o mascarador é chamada *signal-ratio-mask ratio (SMR)*. Se a relação sinal ruído (SNR) é maior do que a SMR, o ruído de quantificação não é audível (1). O ruído codificado não é audível enquanto a NRM for menor do que 0.

$$NRM (\text{distorção perceptível}) = SMR - SNR(m) \text{ dB} \quad (1)$$

O sinal fonte pode ter vários mascarantes, cada um com o seu limiar de mascaramento. Um limiar de mascaramento global pode ser definido a partir destes.

Além do mascaramento na frequência existe também o mascaramento temporal que ocorre quando dois sons

aparecem num intervalo de tempo pequeno, em que o som mais forte, o mascarante, mascara os mais fracos, mascarados, mesmo que o mascarante preceda os mascarados.

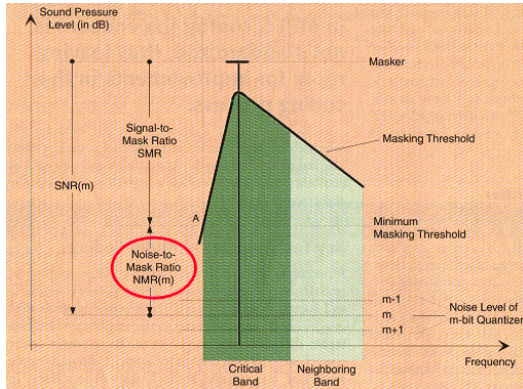


Figura 3 – Representação da SNR(m), SMR e NMR.

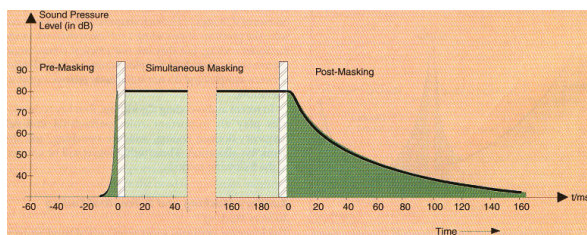


Figura 4 - Mascaramento temporal.

2.2. Codificação perceptiva

Um bom codificador deverá remover as componentes redundantes do sinal de entrada explorando as correlações entre as amostras e deve remover componentes que são perceptivamente irrelevantes ao sistema auditivo humano. A irrelevância consiste na não codificação de sinais que não são ouvidos, como os sinais que são mascarados ou que estão abaixo do limiar de audição e em quantificá-los com passos maiores em zonas em que o nosso sistema auditivo humano é menos sensível. O desenho do codificador depende então da percepção na frequência do sistema auditivo humano, acompanhado pela tolerância perceptiva a erros. Técnicas de noise-shaping permitem deslocar o ruído de quantificação para bandas onde este não é perceptível. Este deslocamento é feito de forma dinâmica e adaptado ao espectro actual do sinal de entrada e de acordo com a SMR. O processo de codificação é controlado pela curva SMR versus frequência a partir da qual a resolução da amplitude necessária em cada frequência são calculadas.

2.3. Codificação no domínio da frequência

Os codificadores no domínio da frequência exploram a redundância e a irrelevância, como objectivo de reduzir ritmo binário em relação ao PCM. Isto é conseguido

partindo o espectro do sinal original em bandas de frequência aproximadamente decorrelacionadas e quantificá-las separadamente. Existem dois algoritmos de codificação para executar esta tarefa: *subband coding* (SBC) e *transform coding* (TC).

- SBC

No SBC, o sinal de origem entra num banco de filtros passa banda contínuos na frequência, seguidamente este sofre uma decimação (diminuição da frequência de amostragem) por um factor de M, em que M é o número de filtros passa banda. O resultado da decimação é um conjunto de bandas, cujo a soma é equivalente ao sinal de origem. No receptor ocorre uma interpolação (aumento da frequência de amostragem) em cada banda para a frequência do sinal de entrada. O processo de amostragem introduz *aliasing distortion*, devido à sobreposição natural das bandas.

- TC

No TC, os blocos de amostras do sinal de origem são linearmente transformados através de uma transformada discreta, em coeficientes que são a representação destes no domínio da frequência. Estes coeficientes são quantificados e transmitidos para o decodificador, onde uma transformada inversa faz com que voltem ao domínio do tempo. Vulgarmente as transformadas utilizadas são a transformada de Fourier ou a transformada do cosseno discreta (DCT) e as suas variantes. Para evitar o efeito de bloco provocado pela utilização da DCT utiliza-se em vez desta a *modified DCT* (MDCT) *filterbank* em que existe uma sobreposição de 50% entre blocos sucessivos.

Outro método utilizado é o banco de Filtros híbrido. Os bancos de filtros híbridos são uma combinação de uma transformada discreta com um banco de filtros. Uma das vantagens de utilizar banco de filtros híbridos é que diferentes resoluções na frequência podem ser obtidas em diferentes frequências de uma maneira flexível e com baixa complexidade.

2.4. Comutação de janelas

Um dos principais problemas da codificação de sinais de áudio no domínio da frequência é o aparecimento de pré-ecos. O pré-eco é o resultado de uma quantificação inadequada de bandas que tem períodos de silêncio, seguidos de sinais de elevada amplitude. Esta quantificação inadequada conduz a erros de quantificação elevados na zona silenciosa. Exemplos de sons que provocam pré-ecos são o bater das castanholas ou em triângulos. O pré-eco pode ser evitado utilizando blocos de amostras de pequena dimensão que conduzem ao aumento do ritmo binário. Para contornar este facto utiliza-se comutação do tamanho dos blocos entre N=64 para controlar situações de pré-eco durante períodos não estacionários do sinal e N=1024 nos

outros casos. Na figura 3b para blocos com tamanho igual 1024 verifica-se que existe uma situação de pré-eco e na figura 3c com blocos de tamanho igual a 256 observa-se que o pré-eco tem uma duração temporal curta e que é mascarado temporalmente pelos sinais de maior amplitude.

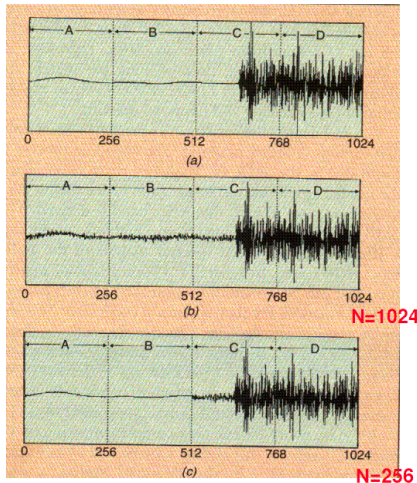


Figura 5 - Comutação de janelas: (a) sinal original, (b) sinal reconstruído com um bloco de tamanho N=1024, (c) sinal reconstruído com um bloco de tamanho N=256.

2.5. Alocação dinâmica de bits

A alocação dinâmica de bits consiste no número de bits atribuído a cada coeficiente da transformada tendo em conta o espectro a curto prazo dos blocos de áudio codificado. A alocação dinâmica de bits é feita bloco a bloco.

3. MPEG-1 ÁUDIO

3.1. Camadas I e II

As primeiras duas camadas apresentam uma estrutura muito similar, passando pelos seguintes blocos fundamentais: banco de filtro de análise, um analisador de espectro, um bloco para o cálculo de limiares de mascaramento, um quantificador e escalador, e por um bloco de alocação dinâmica de bits e factor de escala.

O sinal PCM áudio original entra no sistema sendo processado por 2 blocos em paralelo. Num bloco, constituído por um banco de filtros polifásicos é segmentado o sinal em 32 sub-bandas em frequência, equidistantes e quase incorrelacionadas. O outro bloco, constituído por um analisador de espectros calcula a FFT do sinal com 512 coeficientes, sendo o resultado posteriormente transmitido para o bloco responsável pelo cálculo dos limiares de mascaramento. A existencia destes limiares provém do facto da norma MPEG utilizar as limitações do ouvido humano já descritas, fazendo uso de modelos psico-acústicos (Sub - secção 2.2). Após o cálculo destes limiares o sinal passa para o bloco de alocação dinâmica de bits e factor de escala.

Este bloco recebe a margem de mascaramento do sinal (SMR) que é aproveitada para otimizar o número de bits alocados a cada bloco. Sabe-se que não existirá ruído perceptível se a relação sinal-ruído (SNR) do sinal quantificado for maior que a relação sinal-limiar de mascaramento (SMR). Depois de calculados os bits necessários para a quantização de cada bloco ($32 \cdot 12 = 384$ (camada I) e $32 \cdot 36 = 1152$ (camada II – três blocos da camada I equivalem a um da camada II)) e de um factor de escala, o penúltimo passo consiste na quantificação bloco a bloco, de cada banda. Finalmente é efectuada uma multiplexagem, conjugando estes dados com a informação do número de bits utilizados na quantização e do factor de escala de cada bloco.

O resultado é uma trama com um cabeçalho, informação áudio codificada e por dados secundários auxiliares. Cada trama é passível de ser decodificada, independentemente, estando toda a informação necessária para a sua decodificação no cabeçalho.

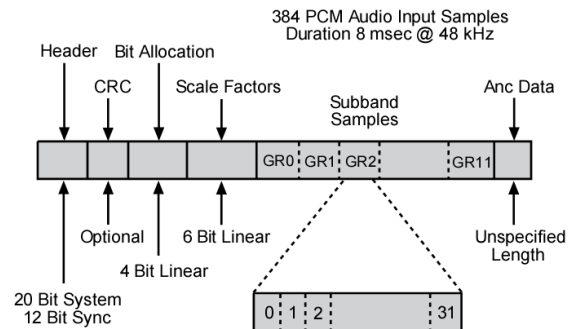


Figura 6 – Bitstream MPEG – 1 áudio.

Para a decodificação de uma trama MPEG-1 (MP1 e MP2) o decodificador necessita de desmultiplexar a informação, desquantificar os dados e o factor de escala, procedendo por fim, para a obtenção do sinal PCM de saída, à síntese de cada uma das suas sub-bandas de frequência (através de um banco de filtros de síntese) juntando estas aditivamente.

Uma das grandes diferenças entre estas duas camadas passa pela utilização de tamanhos de blocos diferentes: 12 e $3 \times 12 = 36$ amostras por sub-banda na camada I e II, respectivamente.

Este factor associado à eliminação de redundâncias oriundas dos 3 factores de escala de cada bloco da camada II, permite uma redução de 50% na compressão de dados PCM da camada II em relação à I.

3.2. MP3

O MPEG-1 Layer III, apresentado na figura 7, introduz várias novidades, sendo de realçar a utilização de um banco de filtros polifásico seguido do cálculo de uma MDCT (transformada discreta do coseno modificada, Subsecção 2.3), que permite uma maior resolução na frequência. O

MP3 utiliza 2 janelas, uma de 6 e outra de 18 pontos (comutação de janelas). Utiliza a de maior dimensão para maximizar a resolução na frequência (totalizando $32 \times 18 = 576$ amostras por bloco), e a de menor dimensão para uma melhor resolução temporal.

É também utilizada a janela de menor dimensão quando se espera a ocorrência de denominados pré-ecos. Outra das novidades desta norma passa pelo uso de codificação entrópica, que é responsável por adaptar o comprimento médio dos símbolos de saída do codificador à entropia dos símbolos saídos do quantificador.

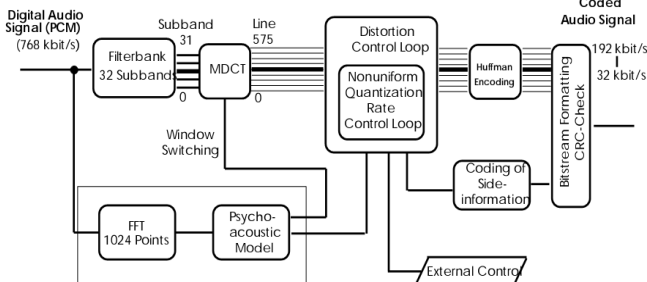


Figura 7- Arquitetura de um codificador MP3 [2].

4. MPEG-4 AAC

Advanced Audio Coder (AAC) é a mais recente norma de compressão áudio com base em Moving Picture Experts Group (MPEG). Este codificador é um dos formatos de compressão de áudio definido pela norma MPEG-2. Também costumava ser chamado por NBC (Non-Backward-Compatible), por não ser compatível com os formatos de áudio MPEG-1. Assim, para o caso deste codificador, teremos que efectuar algumas modificações na estrutura básica do codificador perceptivo, é preciso um processo espectral extra realizado antes da quantização. Este processo espectral é usado para reduzir componentes redundantes, existentes nas ferramentas principais de predição, sendo o objectivo desta, remover a correlação entre as amostras.

Tal como o codificador MP3, descrito anteriormente, o AAC também utiliza MDCT com 50% de sobreposição sobre o módulo do banco de filtros. Depois da adição deste processo de sobreposição, devido ao cancelamento do domínio do tempo, é possível obter uma reconstrução perfeita do sinal original. Contudo, isto nem sempre é verdade, devido ao erro que é introduzido durante o processo de quantização. A ideia do codificador perceptivo é esconder ao máximo este erro existente na quantização de modo que os nossos ouvidos não consigam captar. Aquelas componentes espectrais segundo o qual o ouvido humano não consegue ouvir serão eliminados no canal de codificação. A qualidade de um codificador perceptível depende, essencialmente, do módulo psico-acústico, porque é aqui, onde toda a análise psico-acústica é realizada.

O AAC usa dois tipos de dimensões de janelas, dependendo se o sinal usado é estacionário ou transiente. A decisão para trocar de uma janela para outra é também determinada pelo módulo psico-acústico, fazendo com que este módulo seja ainda mais importante para o desempenho do codificador.

O módulo de quantização opera em dois ciclos agregados. O ciclo interior quantifica o vector de entrada e aumenta o tamanho da etapa de quantização até que o vector de saída possa ser codificado com o número de bits disponíveis. De seguida, o ciclo exterior verifica a distorção de cada factor de escala de banda e, se a distorção permitida for ultrapassada, amplifica a escala do factor de banda e volta a invocar o ciclo interior, novamente. Note que, o AAC utiliza um quantizador não uniforme.

A figura 8 ilustra, como exemplo, o diagrama de blocos completo para o MPEG4-AAC [3].

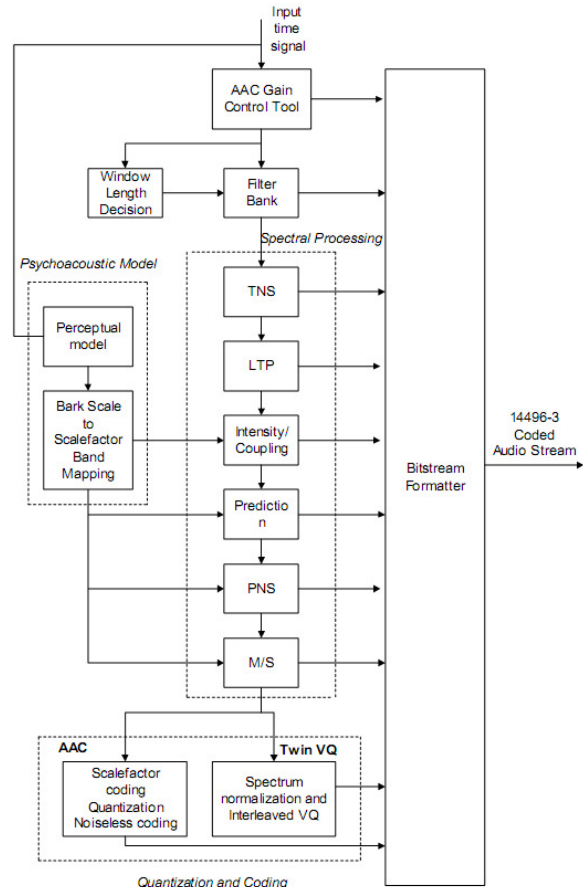


Figura 8 – Diagrama de bloco de um MPEG-4 AAC.

Existem 3 tipos de perfis definidos na norma standard:

- Perfil Principal, com todas as ferramentas activas exige uma potência de processamento considerável.

- Perfil de Complexidade Baixa (LC), com menos factor de compressão para armazenar dados de processamento e de memória RAM.
- Perfil de Factor de Amostragem escalável, com habilidade para adaptar a várias larguras de banda.

A norma AAC apenas garante que um andar AAC seja válido e correctamente descodificado por todos os descodificadores AAC. O codificador permite acomodar variações na implementação, adaptadas a diferentes regiões válidas e áreas de aplicações. O AAC-LC é o perfil que contém uma menor carga computacional, comparado com os outros perfis citados. No entanto, a eficiência global ainda depende no nível de detalhe na implementação do seu próprio codificador.

4.1. Ferramentas

O AAC pertence à classe dos codificadores de forma de onda, tradicionais orientados, perceptivamente. Isto significa que destina-se a reproduzir a forma de onda do sinal de entrada de áudio original com uma quantidade mínima de dados, tendo em conta princípios psico-acústicos para minimizar a audibilidade dos efeitos de codificação. Hoje em dia, pode-se afirmar que o AAC constitui a compressão da forma de onda mais eficiente no mercado. Assim sendo, as ferramentas mais importantes deste codificador são:

- Banco de filtros de (MDCT): transformar o sinal numa representação espectral é a chave para se poder aplicar princípios psico-acústicos e algoritmos de redução redundante para conteúdos áudio. Para tal, o AAC emprega 1024 linhas espectrais de banco de filtro MDCT, criando espectros correspondentes a 1024 amostras de entrada PCM.

- Processamento Stereo: os processamentos stereo *Intensity* e *Mid-Side* estão disponíveis a fim de aumentar a eficiência de compressão para sinas stereo. Enquanto que o primeiro é raramente usado na prática, o ultimo é usado e melhorado com base nos codificadores anteriores, tal como acontece com o MP3.

- Temporal Noise Shaping (TNS): esta ferramenta permite ao codificador moldar o ruído de quantização no domínio do tempo, através da execução de uma previsão nos dados espectrais através da frequência. Isto evita efeitos indesejáveis causados pela resolução grosseira do tempo dos filtros de bancos MDCT. Salienta-se pelo facto da ferramenta TNS ter sido desenvolvida recentemente para o AAC.

- Quantização e Codificação: as ferramentas de quantizar e codificar o espectro são similares ao que é realizado pelo MP3, com significativas melhorias na fase de codificação entrópica, resultando num melhoramento da eficiência de compressão. A sintaxe de bitstream do AAC tem sido definida de forma muito mais flexível que nos

codificadores anteriores, suportando diferentes configurações e extensões futuras.

4.2. Desempenho

Devido ao impacto significativo sobre os aspectos complexos na percepção da qualidade áudio, uma confiável avaliação para a qualidade intermédia dos sinais audio (podendo aparecer num codificador áudio de ritmo binário baixo) continua a ter uma avaliação subjectiva (avaliação pelos nossos ouvido, “ouvidómetro”), em vez de serem usadas avaliações objectivas, especialmente, quando o sinal submetido para o teste foi processado pela ferramenta de codificação paramétrica (tal como o *Spectral Band Replication* e o *Parametric Stereo*).

A figura 9 mostra, como exemplo, os recursos computacionais utilizados num codificador standard AAC-LC, implementado de uma referência MPEG [3], executado para um ritmo binário a 64 kbps e com um factor de amostragem de qualidade de um CD de 44.1 kHz. Note-se que o módulo psico-acústica ocupa 22% do processamento total devido à pesada análise espectral que efectua. Apesar disso, verifica-se que a demanda mais pesada acaba por ser no módulo do quantizador. Isto ocorre pela presença do ciclo interno de controlo citado, anteriormente. Novas abordagens para otimizar os recursos computacionais, neste tipo de codificadores, são usadas e implementadas a fim de minimizar os custos nesta questão. Em [D], é proposto um método de optimização para o codificador AAC e verifica-se que se consegue obter 50% da demanda total para recurso não utilizado.

A tecnologia AAC tem sido mostrada a fim de satisfazer as recomendações da ITU-R BS.1115 para a qualidade de codificação áudio perceptiva em aplicações broadcast a um ritmo binário de 64 Kbit/s por canal de áudio, operando a stereo e num multicanal (isto é, nenhum dos itens do teste teve uma média abaixo de -1 grau em relação ao item original não codificado). Numa comparação entre um vasto conjunto de testes stereo, o AAC a 96 kbit/s provou ter uma qualidade média em comparação ao codificador AC-3 (“Dolby Digital”), por exemplo. Para estes tipos de ritmos o AAC é definido como uma referência de topo para muitos codificadores de áudio de alta qualidade.

Se partimos para codificadores mais elaborados e optimizados dentro dos codificadores AAC, verifica-se que a eficiência de codificação pode ser aumentada cerca de 25% para um ritmo binário da ordem dos 24 kbit/s [6].

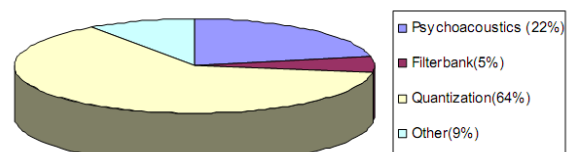


Figura 9 – Distribuição dos recursos computacionais num codificador AAC-LC.

4.3. Produtos

O codificador AAC tem sido introduzido por várias aplicações standards. Algumas dessas organizações que optaram por usar este codificador nos seus produtos estão apresentadas na tabela 2.

Tabela 2 – Produtos software e hardware com codificador AAC.

Standardization organizations, consortia and systems providers	<ul style="list-style-type: none"> • Mobile: 3GPP, 3GPP2 • Mobile TV: DVB-H, T-DMB, S-DMB, MediaFlo, ISDB 1-seg • Consumer electronics: DVD, HD-DVD, SDcard, Nero Digital • Mobile phone operations: KDDI, SK Telecom, NTT DoCoMo, Vodafone, T-Mobile, Sprint, Verizon, Telefonica, Telenor, etc. • Radio & TV providers: BBC, Deutsche Welle, RTL, Voice of Russia, Radio Netherlands, XM Satellite Radio, etc.
Hardware Products	<ul style="list-style-type: none"> • Mobile: Nokia, Motorola, Samsung, SonyEricsson, LG, Panasonic, iPhone, etc • Radios: Delphi, Pioneer, Samsung, Morphy Richards, Roberts, Audiofox, etc. • Chipsets and hardware platforms: ARM, Analog Devices, Texas Instrument, Broadcom, Sigma Design, AMD, NXP, etc • Professional Encoders: Tandberg, Harmonic, Orban, Mayah, Telos Systems, Modulus, etc.
Software Products	<ul style="list-style-type: none"> • Professional encoder software: Sorensen, Vadiator, Spacial Audio, Cube Tec • Consumer software: AOL/Winamp, RealNetworks, Nero, Yahoo, Poiksoft

6. AVALIAÇÕES FINAIS E CONCLUSÕES

Verifica-se que o codificador AAC foi projectado para melhorar o formato e, conseqüentemente ser o sucessor do MP3 (que foi especificado em e MPEG-2) pela ISSO/IEC em 11172-3 e 13818-3, por demonstrar grandes qualidades e transparência de áudio para o mesmo ritmo binário usado pelo MP3.

Em geral, o formato AAC permite uma melhor flexibilidade para projectar codificadores que o formato MP3, e corrige muitas das escolhas efectuadas, originalmente, nas especificações de áudio do MPEG-1. Este aumento de flexibilidade conduz a uma estratégia corrente na codificação e, como resultado, a uma compressão mais eficiente. No entanto, as vantagens do AAC não são totalmente decisivas, e as especificação existentes no MP3, embora já desactualizadas, revelaram-no surpreendentemente robusto, apesar das suas falhas notáveis. O AAC é melhor que o MP3 para ritmos binários baixos (tipicamente abaixo dos 128 kbit/s). Contudo, consoante o

ritmo binário aumenta, a eficiência do formato áudio torna-se o menos importante relativamente à implementação da eficiência do codificador. Assim, a vantagem que o AAC detinha em comparação com o MP3 deixa de ser dominante.

Apesar das restrições de patente referida em relação ao codificador AAC, a perpetuação do formato MP3 continua muito forte no mercado actual. As razões para tal, parecem serem os efeitos que a internet causou por:

- Familiaridade com o formato,
- A grande quantidade de música agora disponível no formato MP3,
- A grande variedade de hardware e software existente que tira vantagem do formato de arquivo,
- A ausência de restrições DRM, o que torna fácil para editar arquivos MP3, copiar e ouvir em diferentes leitores digitais portáteis (Samsung, Creative, entre outros),
- A maioria dos utilizadores domésticos desconhecem ou não se preocupam com a disputa das patentes e muitas das vezes não consideram essas questões jurídicas na escolha dos seus formatos musicais para uso pessoal.

Quando os codificadores são comparados lado a lado, o AAC destaca-se de longe do velho codificador MP3 e espera ser o novo standard codificador áudio a tomar o mercado da internet. Isto deve-se às vantagens que o codificador AAC oferece:

- A melhoria na compressão oferece resultados de alta qualidade com menores tamanhos nos arquivos,
- Suporte para multicanais de áudio, oferecendo até 48 canais de frequência na totalidade,
- Maior resolução de áudio, resultando em ritmos de amostragem até 96 kHz,
- Melhoria na eficiência da descodificação, requerendo menos potencia durante o processo de descodificação.

Após uma pesquisa exaustiva e de ser efectuada toda a explicação do funcionamento dos codificadores apresentados, e finalizando com as vantagens e desvantagens de cada um fazendo eleger o codificador predominante num futuro próximo, apresentam-se nas tabelas 3, 4 e 5, comparações interessantes relativamente ao suporte, informações gerais e técnicas e custos de licenciamento entre os 2 codificadores apresentados neste trabalho [7].

REFERÊNCIAS

- [1] Peter Noll, “MPEG Digital Audio Coding” in IEEE SignalProcessing Magazine, Vol. 14, No. 5, pp. 59–81, Setembro 1997.
- [2] K. Brandenburg, “MP3 and AAC explained”, in Proceedings of the AES 17th International Conference on High Quality Audio Coding, 1999.
- [3] Kurniawati, E.; Lau, C.T.; Premkumar, B.; Absar, J.; George, S.; “New implementation techniques of an efficient MPEG

- [4] **advanced audio coder**", Consumer Electronics, IEEE Transactions on Volume 50, pp.655-665, May 2004.
- [5] Avaro, O.; Salembier, P.; "**Overview of the MPEG-7 standard**", Circuits and Systems for Video Technology, IEEE Transactions on Volume 11, pp 668-695, June 2001.
- [6] Martinez, J.M.; Koenen, R.; Pereira, F.; "**MPEG-7: the generic multimedia content description standard**", Multimedia, IEEE Volume 9, pp.78-87, April 2002.
- [7] Herre, J.; Dietz, M.; "**MPEG-4 high-efficiency AAC coding**", Signal Processing Magazine, IEEE, Volume 25, pp. 137-142, May 2008.
- [8] http://en.wikipedia.org/wiki/Comparison_of_audio_codecs



Pedro Sousa is a MSc Student from Instituto Superior Técnico, Lisbon, Portugal and a researcher at Instituto de Telecomunicações. His research interests are in analog IC design automation and emerging evolutionary computation techniques.



Pedro Ferreira is also a MSc Student from Instituto Superior Técnico, Lisbon, and a research at Instituto de Telecomunicações. His research interests are in CAD tools and reconfigurables A/D converters.

Tabela 3 – Informação Geral Codificação Áudio.

Codificador	Algoritmo	Sample Rate	Bit rate	Bits per sample	CBR	VBR	Stereo	Multicanais
ACC	Lossy, Hybrid	8 kHz a 192 kHz	8 a 529 kbit/s (stereo)	Qualquer (tipicamente usa fp interiormente)	Sim	Sim	Yes: Dual, Mid/Side, Intensity, Parametric	Sim: Até 48 canais
MP3	Lossy	8 kHz a 48 kHz	8 a 320 kbit/s	Qualquer (tipicamente usa fp interiormente)	Sim	Sim	Yes: Dual, Mid/Side, Intensity	Não

Tabela 4 – Detalhe Técnico sobre os Codificadores Áudio.

Codificador	Criador	Data Versão Pública	Última Versão estável	Implementações prioritárias	Implementações sem custos	Patenteada	DRM
ACC	ISO/IEC MPEG Audio Committee	1997	ISO/IEC 14496-3	iTunes, Nero Digital Audio	FAAC, FAAD2, FFmpeg	Sim	FairPlay
MP3	ISO/IEC MPEG Audio Committee	1993	ISO/IEC 11172-3	FhG	LAME, FFmpeg	Sim	Sim (opcional, raramente usada)

Tabela 5 – Custos de Licenciamento.

Codificador		Aplicações Software PC	Produtos Hardware	Distribuições electrónicas de música/Streaming
ACC	Descodificador	US \$ 2,00 por unidade para profissionais US \$ 0,25 por unidade para consumidores	US \$ 2,00 por unidade	N/A
	Codificador	US \$ 20,00 por unidade	US \$ 20,00 por unidade para profissionais US \$ 0,30 por unidade para consumidores	
MP3	Descodificador	US \$ 0,75 por unidade ou US \$ 50.000 na totalidade	US \$ 0,75 por unidade	2% da receita relacionada
	Codificador	De US \$ 0,25 a US \$ 5,00 por unidade	US \$ 1,25 por unidade	