

Sistema de Marca d'Água Digital no Domínio do Tempo para Sinais de Áudio*

Karoline M. O. Nunes, Marcelo S. Pinho

Divisão de Engenharia Eletrônica
Instituto Tecnológico de Aeronáutica
Praça Marechal Eduardo Gomes, 50 – 12.228-900 São José dos Campos, SP

karolm@ita.br, mpinho@ieee.org

Abstract. *In this work, it is presented a variant of the algorithm introduced by Lie and Chang to watermark an audio signal. The original algorithm is based on amplitude modification of two frames of a group of samples (GOS) which is divided in three frames. Based on the absolute mean of samples of the three frames, it is established the rules to embed one bit in this GOS. The sample amplitudes of two frames are modified in accordance to the bit which should be embedded. This paper shows an alternative procedure, working in the three GOS frames, to modify the sample amplitudes. The new procedure reduces the distortion produced by the watermark. The two schemes are tested in a group of 5 songs and the results show an improvement around 1.5dB in the mean square error.*

Resumo. *Nesse trabalho, é apresentada uma variação do algoritmo proposto por Lie e Chang para introduzir uma marca d'água em um sinal de áudio. O algoritmo original é baseado na modificação das amplitudes de dois quadros de um grupo de amostras (GOS - Group Of Samples) que contém três quadros. Baseado na média dos valores absolutos das amostras, são estabelecidas regras par embutir um bit em um GOS. As amplitudes das amostras de dois quadros são alteradas de acordo com o bit que deve ser embutido. Esse artigo propõe um método alternativo, operando nos três quadros do GOS, para modificar as amplitudes das amostras. Esta nova proposta reduz a distorção produzida pela marca d'água. Os dois métodos são testados em um conjunto com 5 músicas e os resultados mostram uma melhora de aproximadamente 1.5dB no erro médio quadrático.*

1. Introdução

Com o crescimento do número de usuários das redes públicas de telecomunicações e com o avanço da tecnologia digital, surgiram diversos sistemas de distribuição de sinais multimídia através destas redes. Estes sistemas de distribuição estão gerando uma verdadeira revolução no mercado. Devido a facilidade de se copiar e alterar um sinal digital, junto com os novos sistemas de distribuição, surge uma série de questões envolvendo a segurança da informação. A técnica de inserir informação adicional em sinais digitais é uma ferramenta útil para combater os problemas relativos à segurança da informação. De

*Trabalho desenvolvido com o apoio da CAPES

fato, a inserção de uma assinatura digital, de uma marca d'água frágil ou de uma marca d'água robusta pode auxiliar na solução de questões sobre autenticidade, integridade e propriedade dos sinais. Por esta razão, nos últimos anos, muita atenção foi dada a esta área [Petitcolas et. al. 1999].

No caso em particular da distribuição de sinais de áudio, a questão mais discutida é a preservação dos direitos autorais. De fato, estas discussões motivaram a implementação de vários mecanismos contra a pirataria, tais como o sistema que impede a cópia de uma mídia e o sistema que impede a reprodução de um sinal em dispositivos não autorizados. Ainda por conta do problema de cópias não autorizadas de sinais de áudio, recentemente, muito esforço foi realizado em estudos de técnicas de marca d'água robusta para áudio digital [Swanson et. al. 1998, Kirovski and Malvar 2003, Lie and Chang 2006].

Em [Lie and Chang 2006] foi apresentado um esquema robusto para embutir uma marca d'água em sinais de áudio, através da alteração das amplitudes das amostras. Na proposta em questão, o sinal de áudio é dividido em grupos de amostras e cada grupo (*GOS* - *Group Of Samples*) é subdividido em três quadros. Cada bit da marca d'água é embutido em um *GOS*, através da alteração da amplitude de dois quadros.

Este artigo propõe uma variação do esquema proposto em [Lie and Chang 2006], em que os três quadros de um *GOS* são alterados de acordo com o bit da marca d'água. Esta nova variante permite atingir a mesma robustez do esquema original, distribuindo melhor a distorção provocada pela introdução da marca d'água. Este trabalho está dividido da seguinte forma. A segunda seção apresenta e analisa o algoritmo proposto em [Lie and Chang 2006]. A descrição da nova versão é apresentada na seção 3. Os resultados obtidos através da aplicação dos dois esquemas em sinais de áudio são mostrados na quarta seção. Fechando o artigo, a conclusão está na seção 5.

2. Marca d'Água Baseada na Modificação das Amplitudes das Amostras

Em [Lie and Chang 2006], foi apresentado um sistema de marca d'água digital para sinais de áudio, baseado na modificação das amplitudes das amostras. A idéia central do algoritmo de Lie e Chang (*LCA* - *Lee Chang Algorithm*) é modificar as amplitudes de um grupo de amostras (*GOS*), de acordo com um dos bits do sinal da marca d'água. Se o bit que se deseja adicionar é igual a 0, as amplitudes são ajustadas para satisfazer uma determinada condição, que representa a introdução do bit 0. Caso contrário, as amplitudes são ajustadas para satisfazer outra condição, que representa a introdução do bit 1.

Seja $x[n]$ o sinal de áudio que será marcado e seja $b[k]$ a seqüência de bits da marca d'água. O *LCA* divide a seqüência $x[n]$ em *GOS*'s com ℓ amostras cada e subdivide cada *GOS* em três quadros, com ℓ_1 , ℓ_2 e ℓ_3 amostras respectivamente. Sendo assim, o k ésimo *GOS* é composto pelas amostras $x[(k-1)\ell+1], \dots, x[k\ell]$, e os quadros podem ser representados pelos vetores

$$\mathbf{q}_1^k = (x[(k-1)\ell+1], \dots, x[(k-1)\ell+\ell_1]), \quad (1)$$

$$\mathbf{q}_2^k = (x[(k-1)\ell+\ell_1+1], \dots, x[(k-1)\ell+\ell_1+\ell_2]), \quad (2)$$

$$\mathbf{q}_3^k = (x[(k-1)\ell+\ell_1+\ell_2+1], \dots, x[k\ell]), \quad (3)$$

onde \mathbf{q}_i^k representa o quadro i do k ésimo *GOS*. Para os três quadros, as médias dos

valores absolutos são calculadas através das seguintes expressões.

$$a_1^k = \frac{1}{\ell_1} \sum_{j=(k-1)\ell+1}^{(k-1)\ell+\ell_1} |x[j]|, \quad (4)$$

$$a_2^k = \frac{1}{\ell_2} \sum_{j=(k-1)\ell+\ell_1+1}^{(k-1)\ell+\ell_1+\ell_2} |x[j]|, \quad (5)$$

$$a_3^k = \frac{1}{\ell_3} \sum_{j=(k-1)\ell+\ell_1+\ell_2+1}^{k\ell} |x[j]|. \quad (6)$$

Os valores obtidos são ordenados, dando origem às variáveis

$$a_{min}^k = \min_{i=1,2,3} \{a_i^k\}, \quad (7)$$

$$a_{med}^k = \text{med}_{i=1,2,3} \{a_i^k\}, \quad (8)$$

$$a_{max}^k = \max_{i=1,2,3} \{a_i^k\}, \quad (9)$$

onde med representa o operador mediana. A idéia básica utilizada pelo LCA para embutir o bit $b[k]$ no $k^{\text{ésimo}}$ *GOS*, é ajustar as amplitudes de tal forma que $a_{max}^k - a_{med}^k < a_{med}^k - a_{min}^k$, se $b[k] = 0$ e $a_{max}^k - a_{med}^k > a_{med}^k - a_{min}^k$, caso contrário. No entanto, para que a marca d'água seja robusta a ataques (intencionais ou não), o algoritmo utiliza um limiar de proteção, t_p . Sendo assim, os valores das amplitudes são modificados para satisfazer as seguintes condições.

$$a_{max}^k - a_{med}^k + t_p < a_{med}^k - a_{min}^k, \text{ se } b[k] = 0, \quad (10)$$

$$a_{max}^k - a_{med}^k > a_{med}^k - a_{min}^k + t_p, \text{ se } b[k] = 1, \quad (11)$$

Após modificar as amplitudes, o *LCA* verifica se a distorção introduzida pela adição da marca d'água não compromete a qualidade do áudio. Esta verificação é feita através do modelo psicoacústico utilizado no padrão MPEG [Noll 1997], sem considerar qualquer efeito de mascaramento. Sendo assim, para cada *GOS*, o *LCA* realiza dois passos: (a) alteração das amplitudes e (b) verificação da qualidade do áudio marcado. Estas duas etapas são detalhadas nas próximas subseções.

2.1. Ajuste das Amplitudes

Com o objetivo de atender as condições apresentadas nas equações (10) e (11), o *LCA* altera as amplitudes de dois quadros do *GOS*, seguindo a seguinte regra.

- $b[k] = 0$,
Se a condição em (10) já é satisfeita pelo *GOS*, nada é feito. Caso contrário, aumenta-se a_{med}^k e diminui-se a_{min}^k .
- $b[k] = 1$,
Se a condição em (11) já é satisfeita pelo *GOS*, nada é feito. Caso contrário, aumenta-se a_{max}^k e diminui-se a_{med}^k .

É interessante observar que se $b[k] = 0$ e se a condição em (10) não for satisfeita, o acréscimo da mediana, δ_{med} , e o decréscimo do mínimo, δ_{min} devem ser tais que

$$a_{max}^k - (a_{med}^k + \delta_{med}) + t_p < a_{med}^k + \delta_{med} - (a_{min}^k - \delta_{min}).$$

Portanto,

$$2\delta_{med} + \delta_{min} > t_p + a_{max}^k + a_{min}^k - 2a_{med}^k \quad (12)$$

Analogamente, no caso em que $b[k] = 1$,

$$2\delta_{med} + \delta_{max} > t_p - a_{max}^k - a_{min}^k + 2a_{med}^k \quad (13)$$

No *LCA*, os ajustes são realizados de tal forma que o acréscimo de um dos blocos é idêntico ao decréscimo aplicado no outro bloco. Sendo assim, se $b[k] = 0$, $\delta_{med} = \delta_{min}$ e se $b[k] = 1$, $\delta_{max} = \delta_{med}$. Portanto, as equações (12) e (13) se transformam em

$$\delta_0 = \frac{t_p + a_{max}^k + a_{min}^k - 2a_{med}^k}{3}, \quad (14)$$

$$\delta_1 = \frac{t_p - a_{max}^k - a_{min}^k + 2a_{med}^k}{3}, \quad (15)$$

onde δ_s representa o acréscimo (ou decréscimo) no caso em que $b[k] = s$, $s = 0, 1$. Para produzir um acréscimo (ou decréscimo) de δ_s em um quadro, o *LCA* multiplica cada uma das amostras do quadro i por um fator w_s . Sendo assim, a alteração da média dos valores absolutos é dada por

$$\begin{aligned} \delta_s &= |a_i^k - w_s a_i^k| \\ &= |1 - w_s| a_i^k, \end{aligned}$$

No caso em que a média dos valores absolutos deve ser acrescida de δ_s , $w_s > 1$ e conseqüentemente,

$$w_s = 1 + \frac{\delta_s}{a_i^k}. \quad (16)$$

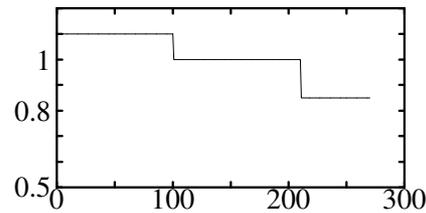
No caso inverso, $w_s < 1$ e portanto

$$w_s = 1 - \frac{\delta_s}{a_i^k}. \quad (17)$$

A multiplicação de dois quadros adjacentes por fatores diferentes pode produzir uma mudança abrupta no sinal de áudio, gerando um efeito de borda que pode ser notado pelo usuário. De fato, se este efeito de borda não for retirado, o usuário irá escutar *clicks* na transição de alguns quadros. Com o objetivo de retirar este efeito, o *LCA* utiliza uma função que altera o fator ao longo do quadro, progressivamente. A Figura 1 ilustra a diferença dos fatores utilizados para ajustar os valores das amostras ao longo de um *GOS* com 270 amostras, onde o primeiro quadro terá suas amplitudes ampliadas de 10% e o terceiro quadro terá suas amplitudes reduzidas de 15%.

Com a introdução da função que torna progressiva a alteração da amplitude de um quadro, o fator w_s deixa de produzir uma modificação de δ_s na média do valor absoluto. Para resolver este problema, o *LCA* utiliza um procedimento iterativo, aumentando (ou reduzindo) o valor de w_s gradativamente até atingir a modificação, δ_s desejada.

Ganho das Amostras sem Sistema Progressivo



Ganho das Amostras com Sistema Progressivo

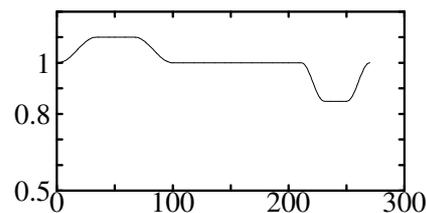


Figura 1. Sistema Progressivo do Ajuste das Amplitudes

2.2. Verificação da Qualidade

Após o ajuste das amplitudes, o algoritmo de Lie e Chang verifica se a distorção provocada por este ajuste compromete a qualidade do áudio. Esta verificação é realizada através do modelo psicoacústico utilizado no padrão MPEG, sem considerar o efeito de mascaramento. Seja $\hat{x}[(k-1)\ell+1], \dots, \hat{x}[k\ell]$ o sinal do $k^{\text{ésimo}}$ GOS após o ajuste das amplitudes. Sendo assim, o sinal erro é definido por

$$e[j] = \hat{x}[j] - x[j]. \quad (18)$$

O *LCA* considera que a distorção introduzida pelo ajuste das amplitudes é aceitável se o sinal erro possui um espectro tal que mais de 85% das magnitudes do espectro estão abaixo do modelo apresentado em [Terhardt et. al.], que representa o nível mínimo de audição, quando não existe qualquer outro som no ambiente (*threshold in quiet*). A Figura 2 ilustra a curva do limiar apresentado em [Terhardt et. al.], que pode ser obtida através da expressão

$$\text{limiar}(f) = 3.64f^{-0.8} - 6.5e^{-0.6(f-3.3)^2} + 0.001f^4 \quad (19)$$

onde f representa a frequência em *KHz* e o limiar é dado em *dB*.

Caso mais do que 15% das componentes do espectro do sinal erro estejam acima do limiar de audição estabelecido em (19), o *LCA* reduz o limiar para a proteção contra ataques, introduzido na etapa de ajuste das amplitudes. Utilizando este limiar reduzido, o *LCA* recalcula os ganhos da primeira etapa. Em seu trabalho original, Lie e Chang não informam qual a estratégia adotada quando não existe um limiar para proteção contra ataques tal que a condição imposta para a qualidade do áudio seja atendida.

3. Reduzindo a Distorção Produzida pelo LCA

O método denominado *Quantization Index Modulation* (QIM), proposto em [Chen and Wornell 2001], é um dos métodos mais populares para a inserção de

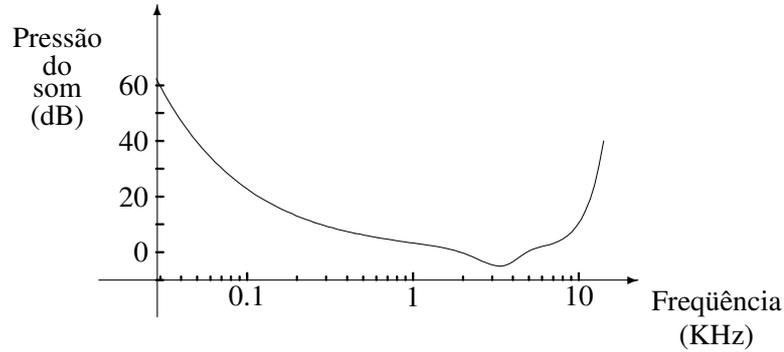


Figura 2. Limiar de Audição no Silêncio

informação adicional em sinais digitais. Sabendo que um sinal digital com ℓ amostras pertence a uma grade de pontos em \mathbb{R}^ℓ , uma solução para inserir um bit neste sinal, é a utilização de um subconjunto de pontos da grade, \mathcal{A}_0 , para indicar a inserção do bit 0 e a utilização de um outro subconjunto, \mathcal{A}_1 , para indicar a inserção do bit 1. Para que o sistema consiga recuperar o bit da marca d'água é necessário que os conjuntos sejam disjuntos, i.e., $\mathcal{A}_0 \cap \mathcal{A}_1 = \emptyset$. No algoritmo de inserção da marca d'água, o sinal original é mapeado (ou quantizado) no ponto mais próximo, pertencente ao conjunto \mathcal{A}_0 ou \mathcal{A}_1 , dependendo do bit a ser inserido. A Figura 3 ilustra a idéia do central do QIM, para um sinal com $\ell = 2$ amostras, onde cada amostra é representada com 3 bits. Neste caso, a grade possui 64 pontos, $2^3 \times 2^3$. Na Figura 3, o símbolo \circ indica os pontos do conjunto \mathcal{A}_0 , o símbolo \star indica os pontos do conjunto \mathcal{A}_1 , o símbolo \times indica que o ponto não pertence a nenhum dos dois conjuntos e o símbolo \odot representa um possível sinal em que se deseja inserir um bit da marca d'água.

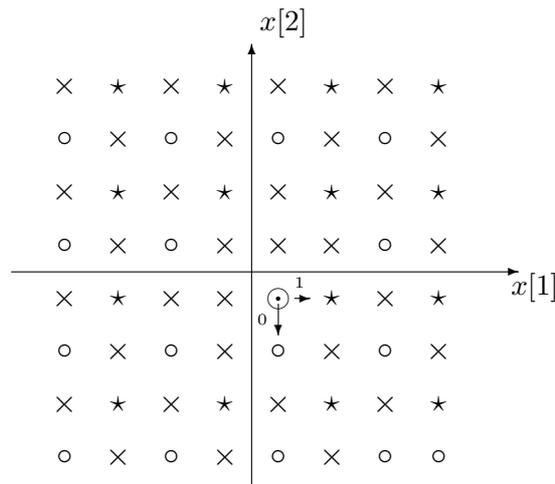


Figura 3. QIM operando em \mathbb{R}^2

De uma forma geral, é possível observar que o algoritmo de Lie e Chang também cria dois conjuntos em \mathbb{R}^ℓ que representam o bit 0 ou o bit 1. De fato, no *LCA*, o conjunto \mathcal{A}_0 é a coleção de todos os sinais que satisfazem a condição estabelecida em (10).

Da mesma forma, o conjunto \mathcal{A}_1 é composto por todos os sinais que satisfazem a desigualdade apresentada em (11). Na verdade, a diferença fundamental entre a idéia básica do QIM e o *LCA* está na forma de mapear o sinal original no sinal marcado. Devido a complexidade de se buscar o ponto mais próximo quando ℓ cresce significativamente, o mapa utilizado no QIM é inviável de ser utilizado no *LCA*. Por esta razão, este último utiliza uma solução heurística para encontrar um ponto em \mathcal{A}_0 (ou em \mathcal{A}_1) que seja próximo do sinal original. Esta solução heurística é baseada no vetor $(a_{min}^k, a_{med}^k, a_{max}^k)$. No entanto, é interessante notar que o *LCA* não encontra nem mesmo um ponto em \mathcal{A}_0 (ou em \mathcal{A}_1) que produza a mínima distância para o vetor $(a_{min}^k, a_{med}^k, a_{max}^k)$. De fato, isto é uma consequência de se optar por alterar apenas dois quadros do *GOS*.

Com o objetivo de reduzir a distorção provocada pela introdução da marca d'água, sem alterar o limiar de proteção a ataques, este artigo propõe uma variante do *LCA*. Esta variante, que será denominada, *vLCA*, utiliza uma heurística para mapear o sinal original no sinal em \mathcal{A}_s , $s = 0, 1$, tal que a distância do vetor $(a_{min}^k, a_{med}^k, a_{max}^k)$ para o vetor correspondente do sinal em \mathcal{A}_s é mínima. As equações (10) e (11) podem ser reescritas da seguinte forma

$$a_{med}^k - \frac{a_{max}^k + a_{min}^k}{2} > \frac{t_p}{2}, \text{ se } b[k] = 0, \quad (20)$$

$$\frac{a_{max}^k + a_{min}^k}{2} - a_{med}^k > \frac{t_p}{2}, \text{ se } b[k] = 1, \quad (21)$$

As equações acima mostram claramente que o bit a ser adicionado estabelece um relacionamento de ordem entre a_{med}^k e o ponto médio entre a_{min}^k e a_{max}^k . Sendo assim, a solução que minimiza a distância (considerando a norma L_2) entre o vetor $(a_{min}^k, a_{med}^k, a_{max}^k)$ e o seu correspondente no sinal marcado, é deslocar (para cima ou para baixo) cada uma das componentes dos vetores de um valor constante.

É interessante observar que em alguns casos particulares, não é possível deslocar as três componentes de um mesmo valor. Isto ocorre ou porque o a_{min}^k não pode ser negativo, ou porque a_{max}^k não pode ultrapassar um valor máximo, ou porque um deslocamento desta forma iria inverter as posições do mínimo com a mediana ou da mediana com o máximo. Em todos estes casos, a alteração de uma ou duas componentes está limitada. A solução nestes casos é produzir a máxima alteração possível nas componentes com restrições e alterar a(s) componente(s) restante(s) de forma a atingir a condição apresentada em (20) ou em (21).

Ainda é importante observar que seguindo a heurística do *vLCA*, caso seja necessário alterar as amplitudes de um *GOS*, os três quadros do *GOS* serão modificados.

4. Resultados

O *LCA* e o *vLCA* foram implementados e o desempenho foi medido através da aplicação dos dois métodos em um conjunto de sinais de áudio para testes. Este conjunto de teste foi composto por 5 trechos de 30 segundos de diferentes tipos de música, conforme indicado na Tabela 1.

O tamanho dos blocos utilizados foram $\ell_1 = 341$, $\ell_2 = 341$ e $\ell_3 = 342$, gerando assim grupos com 1024 amostras.

Tabela 1. Sinais de Áudio para Teste

	Título	Artista	Álbum	Início do Trecho
audio1	O mundo é um moinho	Cartola	Cartola	2 min.
audio2	s'Wonderful	Ray Conniff	16 Most requested songs	30 seg.
audio3	School	Supertramp	Paris	4 min.
audio4	Think of me	Rosemary Ashe Sarah Brightman Steve Barton	The phantom of the opera	2 min.
audio5	Passa em casa	Tribalistas	Tribalistas	30 seg.

O limiar utilizado para combater ataques contra a marca d'água foi calculado da mesma forma como em [Lie and Chang 2006]. Inicialmente, este limiar é projetado para que a marca d'água seja recuperada corretamente sempre que o ataque provocar um desvio no vetor $(a_{min}^k, a_{med}^k, a_{max}^k)$ menor ou igual a

- (a) $(\epsilon a_{min}^k, -\epsilon a_{med}^k, +\epsilon a_{max}^k)$, caso $b[k] = 0$,
- (b) $(-\epsilon a_{min}^k, \epsilon a_{med}^k, -\epsilon a_{max}^k)$, caso $b[k] = 1$,

onde $\epsilon > 0$. É interessante observar que os desvios apresentados no item (a) representam o pior caso para $b[k] = 0$, quando se limita a um desvio máximo de $\pm \epsilon (a_{min}^k, a_{med}^k, a_{max}^k)$. Também é possível observar que o item (b) representa o pior caso para $b[k] = 1$. Para que os desvios apresentados não provoquem erros na marca d'água, é necessário que o limiar t_p seja tal que

$$t_p = (a_{min}^k + 2a_{med}^k + a_{max}^k)\epsilon \quad (22)$$

Nos testes realizados nesse trabalho, foi adotado o valor de $\epsilon = 0.01$.

Para retirar o efeito de borda que seria introduzido caso dois quadros consecutivos fossem multiplicados por ganhos diferentes, os algoritmos utilizaram a função, $f[j]$, definida a seguir, para tornar progressivo o ganho do bloco i do k ésimo GOS.

$$f[j] = \begin{cases} 1, & \text{se } |j - \frac{\ell_i}{2}| \in \left(0, \frac{(1-\alpha)\ell_i}{2}\right); \\ \frac{1}{2} \left\{ 1 + \cos \left[\left(j - \frac{(2-\alpha)\ell_i}{2} \right) \frac{\pi}{\alpha\ell_i} \right] \right\} & \text{se } |j - \frac{\ell_i}{2}| \in \left(\frac{(1-\alpha)\ell_i}{2}, \frac{(1+\alpha)\ell_i}{2} \right); \\ 0, & \text{caso contrário,} \end{cases} \quad (23)$$

onde α é uma constante dentro do intervalo $[0, 1]$. Na verdade, a função definida em (23) é muito popular na área de telecomunicações e é conhecida como função cosseno levantado. A constante α é usualmente chamada de fator de *roll-off* e tem como objetivo suavizar a descontinuidade de um pulso retangular. Se $\alpha = 0$, $f[j]$ é um pulso retangular. Se $\alpha = 1$, $f[j]$ é um período de uma função senoidal deslocada. Nesse trabalho, foi utilizada a função cosseno levantado, com fator de *roll-off* igual a 0.1.

Nas duas variantes, foi adotado o critério de qualidade baseado no limiar de audição em silêncio. Sendo assim, quando a solução encontrada pelo ajuste das amplitudes não satisfaz o critério de qualidade, o limiar t_p é reduzido e o algoritmo recalcula os ajustes das amplitudes.

Como nas duas variantes, o limiar de proteção foi o mesmo, os dois algoritmos produzem sinais marcados com a mesma robustez. Sendo assim, a comparação entre os dois sistemas deve focar a distorção produzida por cada um deles. A Tabela 2 apresenta o erro médio quadrático entre o sinal original e o sinal marcado para o *LCA* e para o *vLCA*. Através da Tabela 2, é possível observar que na média, o *vLCA* produz um ganho de 1.5 dB ($10 \log_{10} \left(\frac{7.7e-5}{5.4e-5} \right)$) sobre o *LCA*, para o grupo de teste utilizado.

Tabela 2. Erro Médio Quadrático entre o Sinal Original e o Sinal Marcado

Sinal	<i>LCA</i>	<i>vLCA</i>
audio1	1.7e-5	1.2e-5
audio2	1.5e-5	9.2e-6
audio3	4.8e-5	3.3e-5
audio4	5.4e-6	4.4e-6
audio5	3.0e-4	2.1e-4
média	7.7e-5	5.4e-5

Um fato interessante observado nos testes, foi a existência de *GOS*'s que mesmo sem a utilização do limiar, t_p , o sinal erro gerado pelo algoritmo possuía um espectro que não satisfazia o critério de qualidade de [Lie and Chang 2006]. Este fenômeno ocorre nos dois algoritmos. Sendo assim, foi necessário estabelecer um valor mínimo para o limiar t_p . Nestes casos, quando o algoritmo atinge este valor mínimo, ele adiciona a marca d'água, mesmo que esta provoque uma distorção acima do desejado. Para avaliar o efeito deste método, foi calculado o espectro médio ao longo de todos os *GOS*'s de um sinal de áudio. Este espectro médio foi comparado com o limiar de audição em silêncio. A Tabela 3 apresenta esta comparação. Na 2ª coluna e na 4ª coluna são apresentadas as porcentagens das componentes do espectro média que estão acima do limiar de audição, para o *LCA* e para o *vLCA* respectivamente. Nas colunas 3 e 5 são apresentadas o valor máximo da diferença entre o espectro médio e o limiar de audição. Através dos resultados da Tabela 3, é possível notar que o sinal erro produzido pelos dois algoritmos está significativamente acima do limiar de audição em silêncio. No entanto, é importante ressaltar que o sinal erro está embutido no sinal marcado. Sendo assim, existe o efeito do mascaramento do sinal erro. De fato, o critério original proposto em [Lie and Chang 2006] é muito rigoroso, pois o sinal erro será sempre mascarado pelo sinal marcado. Uma extensão interessante para este trabalho seria analisar a qualidade subjetiva dos sinais marcados e qual seria o efeito de se considerar critérios menos rígidos, como por exemplo, os critérios adotados na quantização das amostras no padrão MPEG [Noll 1997].

Tabela 3. Comparação do Espectro Médio do Sinal Erro com o Limiar de Audição em Silêncio

Sinal	<i>LCA</i>		<i>vLCA</i>	
	nf	d_{max}	nf	d_{max}
audio1	18.6%	15.5 dB	18.0%	14.2 dB
audio2	20.1%	14.1 dB	19.7%	12.2 dB
audio3	27.9%	16.4 dB	24.4%	14.5 dB
audio4	13.1%	12.0 dB	9.6%	10.8 dB
audio5	41.2%	23.4 dB	39.1%	21.1 dB



5. Conclusão

Este trabalho apresentou uma variante do algoritmo proposto por Lie e Chang [Lie and Chang 2006], onde o procedimento para modificar as amostras do sinal de áudio é melhorado. O novo algoritmo para introdução de marca d'água digital, que é a principal contribuição deste trabalho, produz uma distorção menor que o algoritmo original, sem perder em robustez. Os dois esquemas foram implementados e testados em um conjunto composto por 5 trechos de músicas. Os resultados obtidos mostram um ganho de aproximadamente $1.5dB$ no erro médio quadrático.

Referências

- Petitcolas, F.A.P. and Anderson, R.J. and Kuhn, M.G. (1999). Information hiding - a survey *Proceedings of the IEEE*, 87(7):1062-1078.
- Swanson, M.D. and Zhu, B. and Tewfik, A.H. and Boney, L. (1998). Robust audio watermarking using perceptual masking *Signal Processing*, 66:337-355.
- Kirovski, D. and Malvar, H.S. (2003). Spread spectrum watermarking of audio signals *IEEE Transactions on Signal Processing*, 51(4):1020-1033.
- Lie, W. and Chang, L. (2006). Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification *IEEE Transactions on Multimedia*, 8(1):46-59.
- Noll, P. (1997). MPEG digital audio coding *IEEE Signal Processing Magazine*, 14(5):59-81.
- Terhardt, E. and Stoll, G. and Seewann, M. (1982) Algorithm for extraction of pitch and pitch salience from complex tonal signals *Journal of the Acoustical Society of America* 71:679-688.
- Chen, B. and Wornell, G.W. (2001) Quantization index modulation: a class of provably good methods for digital watermarking and information embedding *IEEE Transactions on Information Theory* 47(4): 1423-1443.