

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

Karoline M. O. Nunes e Marcelo S. Pinho

Resumo— Neste trabalho, é apresentada uma nova versão do algoritmo de Lie e Chang para inserir marca d'água em sinais de áudio. O algoritmo original opera no domínio do tempo e aplica diferentes ganhos em três seções de um grupo de amostras para inserir um bit da marca d'água. A distorção introduzida neste processo é controlada através do uso do limiar de audição em silêncio. O uso deste limiar produz uma severa limitação na robustez da marca d'água. Este trabalho propõe a utilização de um modelo psicoacústico mais realista, gerando uma nova versão do algoritmo de Lie e Chang. Resultados de simulação mostram que esta nova versão produz um ganho significativo em robustez com uma pequena degradação na fidelidade (a marca d'água é classificada como imperceptível segundo o critério PEAQ — recomendação ITU-R BS. 1387-1).

Palavras-Chave— marca-d'água em áudio, marca d'água robusta, marca d'água imperceptível.

Abstract— In this work, it is introduced a new version of the algorithm introduced by Lie and Chang to watermark an audio signal. The original algorithm works in time domain and it applies different gains in three sections of one group of samples to embed a bit of the watermark. The distortion introduced in the audio is controlled by the use of the absolute hearing threshold. The use of this control procedure sets a severe bound to the watermark robustness. This work proposes the use of a more realistic psychoacoustic model, leading to a new version of the algorithm. Simulation results show that the new version produces a significant improvement on robustness with a small degradation on fidelity (the watermark is classified as imperceptible according to PEAQ criteria — recommendation ITU-R BS. 1387-1).

Keywords— audio watermarking, robust watermarking, imperceptible watermarking

I. INTRODUÇÃO

Os avanços nas áreas de processamento digital de sinais e de comunicação digital estão provocando uma verdadeira revolução na forma de se armazenar e distribuir informação. O crescimento do número de serviços que estão sendo introduzidos nas diferentes redes de comunicação tem gerado uma série de desafios na área de transmissão da informação. Na última década, técnicas de inserção de informação têm sido utilizadas em diferentes problemas envolvendo transmissão de dados, tais como proteção de propriedade intelectual [1], estimação de canal [2] e codificação conjunta fonte-canal [3], [4]. As propriedades que um método de inserção de informação deve possuir dependem da aplicação desejada. Em sistemas que visam a proteção da propriedade intelectual, uma solução bastante comum é a utilização de métodos de inserção de marca d'água imperceptível e robusta. Neste caso,

a informação adicional (marca d'água) é inserida no sinal original (sinal hospedeiro) dando origem ao sinal marcado. A marca d'água é dita imperceptível se um usuário comum não é capaz de distinguir o sinal original do sinal marcado. Além disso, a marca é robusta se ela não é perdida quando o sinal marcado é atacado (i.e., é introduzida uma certa distorção neste sinal). Existem diferentes tipos de ataques que podem ser classificados como intencionais (que visam remover a marca) e não intencionais (produzidos por manipulações usuais, tais como um sistema de compressão ou a transmissão através de um canal ruidoso).

Com o crescimento significativo do número de sistemas de distribuição de músicas nas redes de comunicação, os métodos de marca d'água para sinais de áudio têm recebido muita atenção nos últimos anos [5], [6], [7]. Em [7], Lie e Chang introduziram um método de inserção de marca d'água imperceptível e robusta para sinais de áudio, baseado na modificação das amplitudes das amostras do sinal no domínio do tempo. No trabalho original, o método foi testado utilizando três diferentes tipos de música. A fidelidade entre o sinal marcado e o sinal original foi avaliada através de testes subjetivos e a robustez foi medida através da taxa de erro de bits. Os resultados obtidos mostraram que para sinais de áudio estereofônicos amostrados em 44,1kHz e quantizados com 16 bits por amostra, o método é capaz de inserir 43bps com alta fidelidade e com taxa de erro relativamente baixa para diferentes tipos de ataques, tais como a compressão para formato mp3 e a conversão digital-analógico-digital. Em [8], foi apresentada uma nova versão para o método de Lie e Chang capaz de melhorar a fidelidade do sinal marcado sem qualquer redução na robustez e na taxa de inserção. Em [9], a robustez do método é calculada quando o sinal marcado é atacado por ruído aditivo gaussiano branco. Neste último trabalho, através da probabilidade de erro calculada para o ataque gaussiano e através de simulações com diferentes tipos de áudio, foi possível observar que sob certas condições, a robustez dos métodos propostos em [7] e [8] não é satisfatória.

Este trabalho propõe uma alteração no critério de fidelidade utilizado no método de Lie e Chang com o objetivo de melhorar a robustez da marca d'água. A modificação proposta utiliza o modelo psicoacústico apresentado em [10], que também é a base para diferentes métodos de compressão de áudio [11]. O desempenho do método com o novo critério é avaliado através de simulações com diferentes tipos de áudio. A fidelidade entre o sinal marcado e o sinal original é avaliada através do método PEAQ (*Perceptual Evaluation of Audio Quality*) [12] e a robustez é avaliada através da probabilidade de erro quando o sinal é atacado por ruído aditivo gaussiano branco. Este

trabalho está dividido da seguinte forma. A Seção II apresenta o método proposto por Lie e Chang em [7] e sua variação proposta em [8]. A alteração proposta no critério de fidelidade é introduzida na Seção III. Os resultados são discutidos na seção IV e encerrando o trabalho, a conclusão é apresentada na última seção.

II. ALGORITMO DE LIE E CHANG

O algoritmo de Lie e Chang (LCA — *Lie and Chang Algorithm*) opera no domínio do tempo, alterando ligeiramente as amostras do sinal para inserir a marca d'água. Seja $x[i]$, $i = 1, 2, \dots$ as amostras de um canal de áudio e seja $m[j]$, $j = 1, 2, \dots$ uma seqüência de bits que neste trabalho representa a marca d'água. Para inserir $m[j]$ em $x[i]$, o LCA quebra o sinal de áudio em grupos de ℓ amostras, gerando uma seqüência de grupos de amostras $s[j] = (x[(j-1)\ell + 1], \dots, x[j\ell])$. O algoritmo subdivide cada grupo de amostra (GoS — *Group of Samples*) em três seções com ℓ_1, ℓ_2 e ℓ_3 amostras cada. Neste caso, as seções do j -ésimo GoS são dadas por

$$s_1[j] = (x[(j-1)\ell + 1], \dots, x[(j-1)\ell + \ell_1]),$$

$$s_2[j] = (x[(j-1)\ell + \ell_1 + 1], \dots, x[(j-1)\ell + \ell_1 + \ell_2]),$$

$$s_3[j] = (x[(j-1)\ell + \ell_1 + \ell_2 + 1], \dots, x[j\ell]),$$

sendo que $\ell = \ell_1 + \ell_2 + \ell_3$. A média dos valores absolutos (MVA) das amostras é calculada para cada seção, gerando $a_k[j] = \frac{1}{\ell_k} \sum_{i=1}^{\ell_k} |s_{k,i}[j]|$, onde $s_{k,i}[j]$ é a i -ésima componente de $s_k[j]$, $k = 1, 2, 3$. Após este cálculo, o algoritmo ordena as MVA's, dando origem a

$$a_{max}[j] = \max_{k=1,2,3} \{a_k[j]\},$$

$$a_{med}[j] = \text{med}_{k=1,2,3} \{a_k[j]\},$$

$$a_{min}[j] = \min_{k=1,2,3} \{a_k[j]\},$$

onde 'med' representa o operador mediana. A idéia básica do LCA é alterar ligeiramente o vetor $(a_{min}[j], a_{med}[j], a_{max}[j])$ de acordo com $m[j]$, produzindo um vetor $(a'_{min}[j], a'_{med}[j], a'_{max}[j])$ tal que

$$\begin{aligned} a'_{med}[j] - a'_{min}[j] &> a'_{max}[j] - a'_{med}[j], & \text{ se } m[j] = 0, \\ a'_{max}[j] - a'_{med}[j] &> a'_{med}[j] - a'_{min}[j], & \text{ se } m[j] = 1. \end{aligned} \quad (1)$$

Após a modificação das amplitudes, o algoritmo verifica se a distorção introduzida no sinal marcado é aceitável. Caso a qualidade esteja abaixo de um determinado limiar, o LCA diminui a alteração introduzida nas amplitudes. Este processo continua até que a qualidade desejada seja atingida ou a modificação seja a mínima que garanta um certo nível de robustez. Nas próximas subseções, são apresentados alguns detalhes importantes do LCA e da versão do LCA introduzida em [8].

A. Modificação das Amplitudes no LCA

O primeiro passo do LCA é alterar as amostras de tal forma a atender as condições apresentadas em (1). Para garantir a robustez da marca contra ataques, o algoritmo utiliza um limiar de proteção. Seja $(\hat{a}_{min}[j], \hat{a}_{med}[j], \hat{a}_{max}[j])$ o vetor após o sinal marcado ter sofrido um ataque. Seja d o desvio máximo produzido nas MVA's, i.e.,

$$d = \max_{\kappa=min,med,max} \left\{ \frac{|\hat{a}_{\kappa}[j] - a'_{\kappa}[j]|}{a'_{\kappa}[j]} \right\}.$$

Quando $m[j] = 0$, dado o desvio máximo d , o pior ataque possível seria aquele que diminui $a'_{med}[j]$ de d e aumenta $a'_{max}[j]$ e $a'_{min}[j]$ de d . Sendo assim, para garantir que o ataque não irá provocar um erro na detecção da marca, é necessário garantir que

$$(a'_{med}[j] - a'_{min}[j]) - (a'_{max}[j] - a'_{med}[j]) \geq t[j], \quad (2)$$

onde

$$t[j] = d(a'_{min}[j] + 2a'_{med}[j] + a'_{max}[j]) \quad (3)$$

é o limiar de proteção. Analogamente, quando $m[j] = 1$, é possível mostrar que um ataque que provoca um desvio máximo igual a d não irá produzir erro de detecção se

$$(a'_{max}[j] - a'_{med}[j]) - (a'_{med}[j] - a'_{min}[j]) \geq t[j]. \quad (4)$$

O procedimento para a modificação das amplitudes utilizado pelo LCA é apresentado a seguir.

- $m[j] = 0$: Se a desigualdade em (2) é satisfeita com $a'_{\kappa}[j] = a_{\kappa}[j]$, $\kappa = min, med, max$, nada é feito. Caso contrário,

$$a'_{min}[j] = a_{min}[j] - \delta_0, \quad (5a)$$

$$a'_{med}[j] = a_{med}[j] + \delta_0, \quad (5b)$$

$$a'_{max}[j] = a_{max}[j]. \quad (5c)$$

- $m[j] = 1$: Se a desigualdade em (4) é satisfeita com $a'_{\kappa}[j] = a_{\kappa}[j]$, $\kappa = min, med, max$, nada é feito. Caso contrário,

$$a'_{min}[j] = a_{min}[j], \quad (6a)$$

$$a'_{med}[j] = a_{med}[j] - \delta_1, \quad (6b)$$

$$a'_{max}[j] = a_{max}[j] + \delta_1. \quad (6c)$$

□

Para garantir que as condições apresentadas em (2) e (4) serão satisfeitas após a execução do algoritmo, os parâmetros utilizados em (5) e em (6) são dados por

$$\delta_0 = \frac{1}{3} (t[j] + a_{max}[j] + a_{min}[j] - 2a_{med}[j]), \quad (7a)$$

$$\delta_1 = \frac{1}{3} (t[j] + 2a_{med}[j] - a_{max}[j] - a_{min}[j]). \quad (7b)$$

Para modificar as MVA's, o LCA aplica um ganho em todas as amostras de cada seção que precisa ser alterada. Este ganho é calculado de forma a produzir o efeito desejado nas MVA's. Por exemplo, se $m[j] = 0$ e o algoritmo precisa ajustar $a'_{min}[j]$ de acordo com (5a), o LCA aplica um ganho $w_{0,min}[j]$ em todas as amostras. Neste caso, $a'_{min}[j] = w_{0,min}[j]a_{min}[j]$

e para que (5a) seja satisfeita, é necessário que $w_{0,min}[j] = 1 - \frac{\delta_0}{a_{min}[j]}$. Analogamente, é possível observar que os ganhos aplicados nas seções relativas às equações (5b), (6b) e (6c) são dados por $w_{0,med}[j] = 1 + \frac{\delta_0}{a_{med}[j]}$, $w_{1,med}[j] = 1 - \frac{\delta_1}{a_{med}[j]}$ e $w_{1,max}[j] = 1 + \frac{\delta_1}{a_{max}[j]}$, respectivamente.

Aplicar dois ganhos diferentes em duas seções consecutivas pode gerar uma mudança abrupta no sinal de áudio. Conforme relatado em [7], este procedimento gera efeitos que podem ser percebidos pelo usuário. Por esta razão, o LCA ajusta o ganho aplicado em uma seção utilizando uma função que suaviza esta transição. Mais detalhes sobre este processo podem ser obtidos em [7], [8].

B. Critério de fidelidade

Após a modificação das amplitudes, o LCA verifica se esta modificação não produz uma distorção perceptível. Para fazer esta verificação, o algoritmo utiliza o limiar de audição em silêncio do modelo psicoacústico apresentado em [13]. Este limiar representa o menor nível de áudio que pode ser percebido caso não exista outro som no mesmo ambiente. Seja f uma frequência em kHz. O limiar de audição em silêncio para um sinal na frequência f é dado (em dB) por

$$t_{ah}(f) = 3.64f^{-0.8} - 6.5 \exp\left\{-0.6(f - 3.3)^2\right\} + 0.001f^4, \quad (8)$$

Seja $s'[j] = (x'[(j-1)\ell + 1], \dots, x'[j\ell])$ as amostras do j -ésimo GoS após o processo de modificação das amostras. Para avaliar a fidelidade do áudio marcado, o LCA calcula a densidade espectral de potência (DEP) do erro $e[j] = s'[j] - s[j]$ e compara com o limiar em (8). Caso mais de 15% das componentes da DEP do erro estejam acima do limiar, o algoritmo declara que a qualidade é inaceitável. Neste caso, o desvio máximo d é reduzido de um valor δd e o processo de modificação das amplitudes é refeito. Este procedimento é repetido até que a qualidade seja declarada aceitável ou até que o parâmetro d atinja um valor mínimo.

É interessante observar que através do ajuste do parâmetro d , é possível alterar a robustez e a fidelidade da marca d'água. Quanto maior o valor de d , maior é a robustez (pois o limiar de proteção aumenta) e menor é a fidelidade (pois as amostras serão modificadas de forma mais acentuada). Sendo assim, o processo iterativo utilizado no LCA (que reduz o valor de d caso a qualidade seja inadequada) procura atingir a maior robustez possível dada uma qualidade mínima aceitável.

C. Variação do LCA

A partir da descrição do processo de modificação das amplitudes, é possível notar que o LCA não encontra o vetor $(a'_{min}[j], a'_{med}[j], a'_{max}[j])$ mais próximo do vetor $(a_{min}[j], a_{med}[j], a_{max}[j])$ que satisfaça as condições em (2) ou em (4). Em [8] foi proposta uma nova versão do LCA que altera este processo de modificação das amplitudes, procurando o vetor $(a'_{min}[j], a'_{med}[j], a'_{max}[j])$ mais próximo do vetor $(a_{min}[j], a_{med}[j], a_{max}[j])$ que satisfaça as condições em (2) ou em (4). Resultados de simulação relatados em [8] e em [9] mostram que esta estratégia é capaz de melhorar a

fidelidade do sinal marcado, sem reduzir a robustez da marca d'água.

Como a distância entre dois vetores depende da métrica utilizada, é possível gerar diferentes versões para o LCA. Neste trabalho, a versão denominada vLCA/ p modifica as amplitudes substituindo $(a_{min}[j], a_{med}[j], a_{max}[j])$ por um vetor $(a'_{min}[j], a'_{med}[j], a'_{max}[j])$ que satisfaça as condições em (2) ou em (4) e que

$$\left(\sum_{\kappa=min,med,max} |a'_{\kappa}[j] - a_{\kappa}[j]|^p \right)^{\frac{1}{p}}$$

seja mínimo. Obviamente que o algoritmo depende do parâmetro p . No caso da métrica quadrática, i.e., $p = 2$, é possível mostrar que a solução seria a apresentada a seguir.

- $m[j] = 0$: Se a desigualdade em (2) é satisfeita com $a'_{\kappa}[j] = a_{\kappa}[j]$, $\kappa = min, med, max$, nada é feito. Caso contrário,

$$\begin{aligned} a'_{min}[j] &= a_{min}[j] - \delta'_0, \\ a'_{med}[j] &= a_{med}[j] + 2\delta'_0, \\ a'_{max}[j] &= a_{max}[j] - \delta'_0, \end{aligned}$$

onde $\delta'_0 = \frac{1}{6}(t[j] + a_{min}[j] + a_{max}[j] - 2a_{med}[j])$.

- $m[j] = 1$: Se a desigualdade em (4) é satisfeita com $a'_{\kappa}[j] = a_{\kappa}[j]$, $\kappa = min, med, max$, nada é feito. Caso contrário,

$$\begin{aligned} a'_{min}[j] &= a_{min}[j] + \delta'_1, \\ a'_{med}[j] &= a_{med}[j] - 2\delta'_1, \\ a'_{max}[j] &= a_{max}[j] + \delta'_1, \end{aligned}$$

onde $\delta'_1 = \frac{1}{6}(t[j] + 2a_{med}[j] - a_{min}[j] - a_{max}[j])$. □

É importante observar que se todos os parâmetros d e ℓ forem mantidos iguais aos utilizados no LCA, a robustez não irá se alterar. Sendo assim, fica evidente que o vLCA/ p é uma tentativa de melhorar a fidelidade do LCA. Conforme apresentado em [8], [9], resultados de simulação mostram que de fato o vLCA/ p é capaz de melhorar a fidelidade obtida através do LCA. No entanto, é interessante observar também que a fidelidade e a robustez estão relacionadas. Sendo assim, o vLCA/ p pode ser utilizado também para melhorar a robustez da marca d'água (basta ajustar o parâmetro d de forma adequada).

III. ALTERANDO O CRITÉRIO DE FIDELIDADE

Conforme ressaltado na seção I, os resultados de simulação obtidos no trabalho original de Lie e Chang mostraram que o LCA é capaz de inserir 43 bps com alta fidelidade e com taxa de erro relativamente baixa para diferentes tipos de ataques. No entanto, as taxas de erro relatadas em alguns casos chegam a atingir aproximadamente 5% dos bits da marca d'água, o que pode inviabilizar a sua utilização em várias aplicações. Embora o vLCA/2, proposto em [8], possa ser utilizado para aumentar a robustez do algoritmo original, em geral, esta melhora é insuficiente.

Na verdade, as piores taxas de erro são obtidas quando o critério de fidelidade utilizado no algoritmo declara a qualidade inaceitável e o algoritmo reduz o parâmetro d . Através de resultados de simulação, é possível observar que devido ao critério de fidelidade utilizado, em geral, uma boa parte dos GoS's do LCA opera com um parâmetro d reduzido. Embora em menor intensidade, este efeito também é notado no vLCA/2. De fato, através do cálculo da probabilidade de erro para um ataque gaussiano e através de resultados de simulação, é possível observar sinais de áudio em que mais de 50% dos GoS's produzidos pelo LCA estariam operando com probabilidade de erro acima de 10^{-4} , mesmo quando a potência do ataque está abaixo 20dB do sinal de áudio [9]. Caso o vLCA/2 seja utilizado, este número é reduzido para aproximadamente 45%.

A fragilidade observada nos resultados de simulação é uma consequência do critério de fidelidade utilizado no LCA e no vLCA/2. É interessante observar que um critério de fidelidade baseado no limiar de audição em silêncio é extremamente exigente. De fato, a própria existência do áudio irá produzir um efeito de mascaramento da distorção introduzida (o efeito de mascaramento é amplamente utilizado nos sistemas de compressão de áudio). Sendo assim, a utilização de um critério de fidelidade menos exigente pode produzir um sistema mais robusto e que ainda atenda o requisito de imperceptibilidade. Este trabalho propõe a utilização do modelo psicoacústico apresentado em [10] como forma de avaliar a fidelidade do sinal marcado dentro do LCA (ou do vLCA/2, introduzido em [8]). Além do limiar de audição em silêncio, o efeito do mascaramento em frequência também é considerado. Nesta seção, este modelo é apresentado e os resultados de simulação são relatados na próxima Seção.

A imperceptibilidade da marca d'água só é possível devido às limitações do sistema auditivo humano. De fato, para que um sinal seja audível, é necessário que ele possua uma potência mínima. Esta potência mínima depende da frequência, do tempo e da existência ou não de outros sinais no mesmo ambiente. De fato, a existência de um sinal de áudio pode mascarar a existência de um outro sinal com potência mais baixa. Na literatura, é comum encontrar dois tipos diferentes de mascaramento. O mascaramento temporal ocorre quando um sinal impede a audição de um sinal que surge logo após o (ou imediatamente antes do) primeiro. Neste caso, os sinais ocorrem em tempos distintos. No mascaramento em frequência, um sinal de potência mais elevada impede que um segundo sinal na mesma frequência, que ocorre no mesmo tempo, seja percebido. Este efeito pode ser visto como uma elevação do limiar de audição. Através do modelo psicoacústico apresentado em [10], dado a existência de um sinal de áudio, é possível calcular o efeito que este provoca no limiar de audição. Este modelo utiliza o conceito de banda crítica, que divide o espectro audível em 25 bandas conforme apresenta a Tabela I. Para calcular o efeito do mascaramento, a potência do sinal mascarador contida em cada uma das bandas críticas é calculada. Seja $\mathbf{p} = (p_1, \dots, p_{25})$ um vetor onde cada componente representa a potência contida em cada banda crítica. Como a existência de sinal em uma banda crítica pode mascarar um segundo sinal em uma banda crítica vizinha, o

TABELA I
BANDAS CRÍTICAS

Banda Crítica	Frequência (Hz)		
	inferior	central	superior
1	0	50	100
2	100	150	200
3	200	250	300
4	300	350	400
5	400	450	510
6	510	570	630
7	630	700	770
8	770	840	920
9	920	1000	1080
10	1080	1170	1270
11	1270	1370	1480
12	1480	1600	1720
13	1720	1850	2000
14	2000	2150	2320
15	2320	2500	2700
16	2700	2900	3150
17	3150	3400	3700
18	3700	4000	4400
19	4400	4800	5300
20	5300	5800	6400
21	6400	7000	7700
22	7700	8500	9500
23	9500	10500	12000
24	12000	13500	15500
25	15500	19500	22050

modelo considera um espalhamento da potência por bandas críticas vizinhas. Este espalhamento é calculado através de

$$\mathbf{p}_e = \mathbf{p}\lambda,$$

onde

$$\lambda = \begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & \dots & \lambda_{1,25} \\ \lambda_{2,1} & \lambda_{2,2} & \dots & \lambda_{2,25} \\ \vdots & \vdots & & \vdots \\ \lambda_{25,1} & \lambda_{25,2} & \dots & \lambda_{25,25} \end{pmatrix},$$

sendo que $\lambda_{i,j}$ representa o efeito que a potência da banda crítica j produz na banda crítica i . No modelo, estes efeitos são calculados de acordo com

$$10 \log_{10} \lambda_{i,j} = 15,81 + 7,5(i-j) - 17,5 \left\{ 1 + [(i-j) + 0,474]^2 \right\}^{\frac{1}{2}}$$

O efeito de mascaramento é distinto para diferentes tipos de sinais. Neste sentido, o modelo verifica se o sinal se parece mais com um tom puro ou com ruído. Esta verificação é baseada no coeficiente de tonalidade, definido por

$$\beta = \min \left\{ \frac{SFM}{SFM_{tom}}, 1 \right\},$$

onde SFM (*Spectral Flatness Measure*) verifica a forma do espectro (se é constante, é mais próximo do ruído) e é definido pela razão entre a média geométrica e a média aritmética (em dB). Além disso, SFM_{tom} representa um valor típico de um SFM para um tom com baixo nível de ruído e assume o valor de $-60dB$.

O limiar de audição devido ao mascaramento é dado por

$$t_{ahm}[i] = \frac{10^{\log_{10} p_e[i] - \frac{\alpha_i}{10}}}{\sum_{j=1}^{25} \lambda_{j,i}}, \quad (9)$$

onde $\alpha_i = \beta(14,5 + i) + (1 - \beta)5,5$. Seguindo o modelo, um sinal será audível se seu nível estiver acima de $t_{ahm}[i]$ e de $t_{ah}(f)$.

O algoritmo que considera o efeito de mascaramento é denominado neste trabalho de *vLCA/m* (caso o algoritmo original seja utilizado) ou de *vLCA/p/m* (caso o *vLCA/p* seja utilizado). Quando se utiliza o efeito do mascaramento, além da modificação do limiar de audição, o critério de qualidade só é considerado satisfeito caso 100% das amostras estejam abaixo do limiar. Este critério foi adotado neste trabalho pois além de ser o mais usual, uma única componente acima do limiar pode comprometer a qualidade.

IV. RESULTADOS

As versões *vLCA/2* e *vLCA/2/m* foram implementadas em *software*. Os desempenhos destas duas versões foram avaliados através de testes que utilizaram cinco sinais de gêneros distintos (ópera, rock e popular) e gravados em diferentes situações (em estúdio e ao vivo). Todos os sinais de áudio utilizados são estereofônicos, tem duração de 30 segundos, foram amostrados em 44,1 kHz e quantizados com 16 bits. Os testes foram realizados inserindo bits distintos nos dois canais de áudio e que foram gerados aleatoriamente. Os tamanhos das seções foram $\ell_1 = \ell_2 = \ell_3 = 340$, que são os mesmos que foram utilizados em [7]. A função cosseno-levantado com fator de *roll-off* igual a 0,1 foi utilizada para suavizar as modificações nas amplitudes entre diferentes seções. Inicialmente, o parâmetro d foi ajustado para 0,5 e quando o critério de qualidade não era satisfeito, ele era reduzido de $\delta d = 0,01$. O valor mínimo considerado para d foi de 0,01, i.e., quando d atingia o valor de 0,01, a marca era introduzida independente do critério de qualidade ter sido satisfeito.

O primeiro resultado importante, que já havia sido observado em trabalhos anteriores [8], [9], é a existência de GoS's em que é impossível introduzir a marca d'água sem desobedecer o critério de qualidade do *vLCA/2*. De fato, mesmo tomando $d = 0$ (i.e., eliminando qualquer limiar de proteção) é impossível satisfazer o critério de qualidade em todas as GoS's. É importante citar que este efeito é mais crítico ainda no *LCA*. Quando o critério de qualidade é substituído pelo critério que considera o mascaramento, este efeito ocorre em um número muito reduzido de GoS's.

Conforme relatado em [9], para se obter uma probabilidade de erro no intervalo $[10^{-6}, 10^{-4}]$, é necessário que a relação sinal-hospedeiro-ruído, em dB, esteja tipicamente no intervalo [9, 76; 11, 89] se $d = 0,05$. Para $d = 0,04, \dots, 0,01$ este intervalo muda para [11.69, 13.82], [14.19, 16.32], [17.71, 19.84] e [23.74, 25.87], respectivamente. É possível observar através destes resultados que a robustez do algoritmo fica bastante comprometida quando o valor de d é pequeno. De fato, uma relação sinal-hospedeiro-ruído de $20dB$ é muito alta e neste caso, a marca pode ser destruída com relativa facilidade. Para avaliar o ganho na robustez quando o efeito do mascaramento é considerado, foram calculados os valores dos parâmetros d utilizados para as diferentes GoS's. A Tabela II mostra a distribuição dos GoS's de acordo com os valores utilizados para o parâmetro d , considerando as duas versões. Através

TABELA II
DISTRIBUIÇÃO DOS GoS'S DE ACORDO COM O PARÂMETRO d UTILIZADO

d		Sinal de Áudio					dist. (%)
		1	2	3	4	5	
0,01	vLCA	1040	15130	1597	215	1522	45,39
	vLCA/m	2	0	13	1	10	0,20
0,02	vLCA	355	289	252	147	192	9,52
	vLCA/m	0	0	0	0	2	0,02
0,03	vLCA	316	213	185	164	133	7,79
	vLCA/m	0	0	0	0	1	0,01
0,04	vLCA	261	160	147	185	121	6,74
	vLCA/m	1	0	0	0	2	0,02
0,05	vLCA	622	419	413	1883	626	30,56
	vLCA/m	2591	2594	2581	2593	2579	99,75

TABELA III
SIGNIFICADO DO ODG

Efeito da Distorção	ODG
Imperceptível	0
Perceptível, mas sem incomodar	-1
incomoda ligeiramente	-2
incomoda	-3
incomoda bastante	-4

da Tabela II é possível notar que a utilização do efeito do mascaramento melhora de forma significativa a robustez da marca d'água. De fato, no *vLCA/2* a maior parte dos bits são inseridos em GoS que operam com $d = 0,01$, cuja robustez é muito pequena. A substituição do critério de fidelidade faz com que 99,75% dos GoS's operem com $d = 0,05$, i.e., com uma robustez significativamente melhor.

É evidente que ao mesmo tempo em que a utilização do efeito do mascaramento aumenta a robustez do algoritmo, ela produz uma perda na qualidade do sinal marcado. De fato, se um GoS que operava com $d = 0,01$ no *vLCA/2* passa a ser ajustado com $d = 0,05$ no *vLCA/2/m*, a distorção introduzida será maior. No entanto, se o modelo psicoacústico utilizado é bom, este aumento na distorção não será tão significativo e a diferença entre o sinal marcado e o sinal original continuará imperceptível. Para avaliar a perda de fidelidade, este trabalho utiliza a medida objetiva de qualidade conhecida como *PEAQ* (*Perceptual Evaluation of Audio Quality*), que é uma recomendação da União Internacional de Telecomunicações (UIT) [12]. O *PEAQ* utiliza uma série de medidas de diferença entre os sinais de entrada e estima através de uma rede neural um grau objetivo de diferença entre os sinais de entrada (*ODG - Objective Difference Grade*). Na versão básica do *PEAQ*, o parâmetro ODG varia dentro do intervalo $[-3,98; 0,22]$. Quanto maior o valor do parâmetro, menor é a diferença entre os sinais e de forma subjetiva, o efeito da distorção introduzida no sinal marcado é dado de acordo com a Tabela III.

Para calcular o ODG entre os sinais marcados e os sinais originais, este trabalho utilizou a versão básica do *PEAQ* disponível na página da Internet do *Telecommunications & Signal Processing Laboratory* da *McGill University* (<http://www.tsp.ece.mcgill.ca>). Os resultados obtidos para as versões *vLCA/2* e *vLCA/2/m* são apresentadas na Tabela IV. Através dos resultados apresentados, é possível notar que a distorção introduzida no algoritmo que considera o efeito de mascaramento é ligeiramente superior mas sem comprometer

TABELA IV
RESULTADOS DO PEAQ (ODG)

Áudio	vLCA/2	vLCA/2/m
1	0,065	-0,091
2	0,112	-0,003
3	-0,012	-0,036
4	0,011	-0,056
5	-0,027	-0,067

a qualidade do sinal marcado. De fato, os valores de ODG estão muito próximos de 0, o que significa que a distorção é praticamente imperceptível. Os valores obtidos estão de acordo com resultados relatados para sinais de áudio comprimido através do padrão mp3, que também utiliza o efeito do mascaramento [14].

Avaliando conjuntamente os resultados apresentados nas Tabelas II e IV, é possível notar que o método vLCA/2/m, proposto neste trabalho, é capaz de melhorar consideravelmente a robustez da marca d'água sem afetar de forma significativa a qualidade do sinal marcado. No entanto, é importante ressaltar que mesmo no vLCA/2/m, existem GoS's que operam com $d < 0,05$ e sendo assim, possuem uma robustez reduzida. Na verdade, devido ao fato das versões do LCA trabalharem com um nível de fidelidade fixo, é natural que a robustez varie ao longo da seqüência de GoS's. Uma possível solução para reduzir este efeito é a utilização de códigos corretores de erro.

V. CONCLUSÕES

Este trabalho apresentou uma nova versão para o algoritmo proposto por Lie e Chang para inserir uma marca d'água robusta e imperceptível em sinais de áudio. A contribuição deste trabalho é a modificação do processo de controle de fidelidade do áudio marcado. Enquanto o algoritmo original utiliza apenas o limar de audição em silêncio para verificar a qualidade do áudio marcado, a nova versão considera o efeito do mascaramento da marca d'água pelo sinal original. Considerando que a marca d'água está inserida no sinal, o modelo psicoacústico proposto é mais próximo da realidade. Os dois modelos para controlar a fidelidade do áudio marcado foram utilizados dentro de uma variante do algoritmo de Lie e Chang proposta por Nunes em 2007. Foram utilizados cinco sinais diferentes para avaliar a melhora da nova versão proposta. Através dos resultados de simulação foi possível observar que o modelo psicoacústico que considera o efeito do mascaramento melhora de forma significativa a robustez da marca d'água sem comprometer a qualidade do sinal marcado. Nos resultados de simulação, foi observado que se o sinal marcado pelo algoritmo original sofresse um ataque aditivo gaussiano branco com relação sinal-hospedeiro-ruído de $9,76dB$, 69,44% dos GoS's teriam uma probabilidade de erro significativamente menor que 10^{-4} . Nas mesmas condições, um sinal marcado com o novo algoritmo teria 99,75% dos GoS's operando com uma probabilidade de erro de 10^{-4} . Através do critério PEAQ foi possível observar nas simulações que a degradação na qualidade é pouco relevante. De fato, o sinal de áudio com pior desempenho obteve um grau objetivo de diferença de $-0,091$ que representa uma distorção que não é perceptível.

REFERÊNCIAS

- [1] M. Barni, F. Bartolini, "Data hiding for fighting piracy," *IEEE Signal Processing Magazine*, vol. 21, pp. 28-39, 2004.
- [2] X. Wang, Y. Wu, H-C. Wu, "Iterative channel estimation and papr reduction for ofdm system with overlay watermarks," *IEEE GLOBECOM Proceedings*, 2006.
- [3] M. O. C. Altoé, *Codificação conjunta fonte-canal utilizando codificadores universais adaptativos*, Tese de mestrado, Instituto Tecnológico de Aeronáutica, 2007.
- [4] S. Lonardi, W. Szpankowski, M. D. Ward, "Error resilient lz77 data compression algorithms, analysis, and experiments," *IEEE Transactions on Information Theory*, vol. 55, pp. 1799-1813, 2007.
- [5] M. D. Swanson, B. Zhu, A. H. Tewfik, L. Boney, "Robust audio watermarking using perceptual masking," *Signal Processing*, vol. 66, pp. 337-355, 1998.
- [6] D. Kirovski, H. S. Malvar, "Spectrum watermarking of audio signals," *IEEE Transactions on Signal Processing*, vol. 51, pp. 1020-1033, 2003.
- [7] W. N. Lie, L. C. Chang, "Robust and high-quality time-domain audio watermarking based on low frequency amplitude modification," *IEEE Transactions on Multimedia*, vol. 8, pp. 46-59, 2006.
- [8] K. M. O. Nunes, M. S. Pinho, "Sistema de marca d'água digital no domínio do tempo para sinais de áudio," *Anais do Simpósio Brasileiro de Computação Musical*, pp. 61-70, 2007.
- [9] M. S. Pinho, K. M. O. Nunes, "On the performance of audio watermarking based on low-frequency amplitude modification," *submetido ao Globecom 2008*.
- [10] J. D. Johnston, "Transform coding of audio signals using perceptual noise criteria," *IEEE Journal on Selected Areas in Communications*, vol. 6, pp. 314-323, 1988.
- [11] P. Noll, "MPEG audio coding standards," *IEEE Signal Processing Magazine*, vol. 14, pp. 59-81, 1997.
- [12] ITU-R Recommendation Bs.1387, *Method for Objective Measurements of Perceived Audio Quality*, Dec. 1998.
- [13] E. Terhardt, G. Stoll, M. Seewann, "Algorithm for extraction of pitch and pitch salience from complex tonal signals," *Journal of the Acoustical Society of America*, vol. 71, pp. 679-688, 1982.
- [14] G. Vecellesi, A. Vitali, M. Zerbini, "Mp3 audio quality for single and multiple encoding," *IEEE International Conference on Multimedia and Expo*, pp. 1279-1282, 2007.