

A UTILIZAÇÃO DE REDES NEURAIIS PARA CÁLCULO DE "LOUDNESS"

Belisário Nina Huallpa⁽¹⁾; Stelamaris Rolla Bertoli⁽²⁾; José Roberto de França Arruda⁽¹⁾

(1) DMC-FEM-UNICAMP – Brasil, Caixa Postal 6122 – CEP: 13083-970, +19 3788-7906

(2) DCC-FEC-UNICAMP

beli@fem.unicamp.br, rolla@fec.unicamp.br, arruda@fem.unicamp.br

RESUMO

Recentemente, sistemas de qualidade sonora têm chamado a atenção na avaliação do som. Um dos parâmetros usados em sistemas de qualidade sonora é o "loudness", que mede a sensação ou percepção de intensidade sonora. O propósito deste trabalho é apresentar técnicas baseadas numa rede neural artificial para calcular o índice de "loudness" pelo método de Stevens. No treinamento da rede neural multicamada utiliza-se o método de retropropagação. As frequências centrais das bandas de oitava e o nível de pressão sonora por 1/1 oitava são utilizados como informações de entrada, e o índice de "loudness" de Stevens é usado como padrão de saída. Para melhorar o desempenho na fase de treinamento, processos de normalização das informações de entrada e saída são empregados. Após a fase de treinamento, a rede neural é testada usando informações de entrada próximas às utilizadas na fase de treinamento obtendo-se resultados satisfatórios. Três tipos de sons são usados para validar o método proposto. O primeiro exemplo é a utilização de uma voz dizendo "shhhh", o segundo é um sinal de ruído dentro de uma cabine de um veículo e o último é o ruído emitido por um compressor hermético medido numa sala com baixa reverberação. Os resultados deste trabalho são comparados com o método de cálculo de "loudness" de Stevens. A comparação pode ser considerada satisfatória, pois todos os casos testados apresentam erros abaixo de 5%. Além disso, este método oferece a possibilidade de correção ou deve-se adicionar em outros padrões, tornando o sistema personalizado, permitindo a avaliação de "loudness" para grupos específicos.

ABSTRACT

Lately, sound quality systems have become central to the problem of sound evaluation. One of the features of sound quality systems is "loudness", which measures the sensation of sound intensity. The purpose of this paper is to present an artificial neural network technique to obtain the loudness index by Stevens' method. A multilayer neural network architecture is trained using the backpropagation technique. Central frequencies of 1/1 octave bands and the sound pressure levels are used as input information, and the Steven's loudness index is used as the output. In order to improve the training stage, a normalization process of input and output information is performed. After a training stage, the neural system is tested using input values in the same range as the values used during training process. Three different types of sounds are used in order to compare and validate the neural system. The first example is the sound of a human voice saying "shhhh", the second one is a signal of the noise inside the cabin of a car and the last one is the noise of a hermetic compressor measured in a low reverberation room. The comparison of the proposed method and the Stevens algorithm produced good results with errors below than 5% in all the cases. Besides, the proposed method provides the possibility to correct or add custom patterns, making it possible to use personalised patterns for specific groups.

1. INTRODUÇÃO

Ao longo do tempo tem havido uma preocupação crescente com relação ao nível de ruído emitido por equipamentos mecânicos, eletro-eletrônicos, automotivos e outros equipamentos que geram ruído. Neste sentido, o dB(A) – Nível de Pressão Sonora (NPS) ponderado pela curva A tem sido o parâmetro mais utilizado na quantificação da percepção da intensidade do ruído produzido por máquinas e veículos. Atualmente, na maioria dos países, existem normas que limitam os níveis de pressão sonora máxima e o tempo de exposição permitidos visando proteger a saúde das pessoas.

Entretanto, se inicialmente a preocupação era apenas em diminuir ou eliminar completamente qualquer ruído, observou-se que, em alguns casos, era importante a existência do ruído, porém com características agradáveis ao ouvido humano, já que o ruído é portador de informação. A partir destas premissas, surgiu um conjunto de parâmetros de percepção sonora que são conhecidos hoje como parâmetros de qualidade sonora. Esta nova filosofia levou ao desenvolvimento de padrões de análise, destacando-se os conceitos de “loudness” (sensação subjetiva de intensidade sonora), “sharpness” (sensação de equilíbrio espectral entre altas e baixas frequências), “roughness” (presença de frequências de modulação superiores a 20 Hz) e “fluctuation strength” (presença de frequências de modulação inferiores a 20 Hz). Através destes parâmetros é possível a quantificação objetiva de efeitos notados subjetivamente, que não são possíveis de serem detectados através da medida do nível de pressão sonora e de uma análise espectral simples. Dentre os parâmetros de qualidade mencionados, o “loudness” é provavelmente o parâmetro psicoacústico mais importante.

A relação entre estímulo e sensação não pode ser avaliada apenas em função das variações de pressão sonora diretamente, mas deve ser obtida como resultado de outros tipos de parâmetro obtidos pelo processamento dos sinais. Desta forma, em adição ao conceito de “loudness”, o nível de “loudness” também é importante. Ele não é somente um valor físico, mas sim a descrição de algo que está entre a sensação e o valor físico.

Comparações de “loudness” possibilitam obter resultados mais precisos que simples estimativas de magnitude. Por esta razão o nível de “loudness” foi criado para caracterizar a sensação “loudness” de qualquer som. Foi introduzido nos anos 20 por Barkhausen, pesquisador cujo nome foi abreviado para denominar a unidade de faixas de bandas críticas de frequência, o “bark” (Zwicker e Fastl, 1999).

2. MÉTODO DE STEVENS NO CÁLCULO DE “LOUDNESS”

Stevens (1955) definiu uma relação entre “loudness” e a intensidade do estímulo. Para ruídos com espectros contínuos, Stevens (1956), obteve empiricamente o cálculo de “loudness” a partir do espectro de bandas de 1/1 oitava usando a seguinte expressão,

$$S_t = S_m + 0.3 \left(\sum S - S_m \right) \quad (1)$$

onde, S_t é o “loudness” total, S_m é o “loudness” de maior valor, e $\sum S$ corresponde à soma de todos os “loudness” das outras bandas. O “loudness” de cada banda é determinado a partir das medidas de níveis de pressão sonora junto à tabela de equivalência que permite transformar estas medidas em níveis de “loudness”.

Para generalizar o procedimento anterior, Stevens (1961), propôs a mudança na equação (1) transformando-a em

$$S_t = S_m + F \left(\sum S - S_m \right) \quad (2)$$

onde ao coeficiente F é atribuído o valor de 0.3, 0.2, ou 0.15 que corresponde às bandas de 1/1 oitava, 1/2 de oitava ou 1/3 de oitava respectivamente, (ISO 532a, 1975).

O “loudness” total é convertido em nível de “loudness” total utilizando a relação $S_t = 2^{(P-40)/10}$, de onde é isolado P , de maneira que

$$P = 40 + 10 \log_{10}(S_t) \quad (3)$$

sendo S_t (em *sones*) o “loudness” total e P (em *phones*) o nível de “loudness” equivalente. O nível de “loudness” é o parâmetro psicoacústico que caracteriza a intensidade percebida pelo ouvido humano.

3. UTILIZAÇÃO DE UMA REDE NEURAL MULTICAMADA NA OBTENÇÃO DOS ÍNDICES DE STEVENS PARA O CALCULO DE “LOUDNESS”

Uma rede neural é um arranjo de unidades simples de processamento chamadas de neurônios cuja funcionalidade é inspirada no neurônio biológico. A capacidade de processamento da rede neural está associada aos pesos das interligações entre os neurônios obtidos através de um processo de adaptação ou aprendizagem, utilizando um conjunto de padrões de excitação e resposta conhecidos. Esta ferramenta é de reconhecida aplicação na identificação de sistemas permitindo produzir modelos explícitos (Haykin, 1999).

Neste trabalho é utilizada uma rede neural multicamada de 2 entradas, sendo uma das informações de entrada correspondente às frequências de bandas de 1/1 oitava e a outra entrada correspondente aos níveis de pressão sonora em *dB* para cada banda (**Figura 1**). A camada intermediária é composta de 15 neurônios com saídas do tipo tangente hiperbólico (equação (4)),

$$\begin{aligned} u_1 &= w_{11,P}P + w_{11,F}F + b_{11}, & v_1 &= \tanh(u_1) \\ &\vdots & & \\ u_{15} &= w_{115,P}P + w_{115,F}F + b_{115}, & v_{15} &= \tanh(u_{15}) \end{aligned} \quad (4)$$

e a camada de saída é composta por um único neurônio com saída do tipo linear (equação (5)),

$$y = \sum_{k=1}^{15} w_{2k} v_k + b_2 \quad (5)$$

o qual fornece uma única informação correspondente ao índice de “loudness” para cada par de informações (frequência e nível de pressão sonora) obtendo-se como resposta o valor do índice de “loudness”

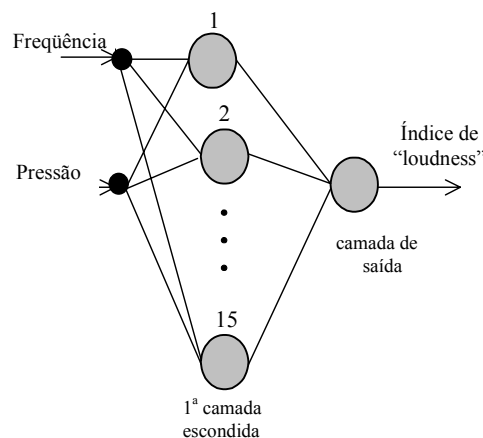


Figura 1. Rede neural utilizada para a obtenção do índice de “loudness”.

Para realizar o treinamento da rede neural, os valores dos índices foram normalizados utilizando primeiro a função logarítmica e depois uma expansão linear entre -1 e 1 (**Figura 2**). O número de padrões utilizados foi de 954 pontos que corresponde ao número total de índice de “loudness” utilizados para o treinamento da rede neural. Este número corresponde às 9 freqüências centrais de bandas de 1/1 oitava utilizadas (31.5 Hz a 8000 Hz) e aos 106 de níveis de pressão sonora (20 dB a 125 dB) a incrementos de 1dB.

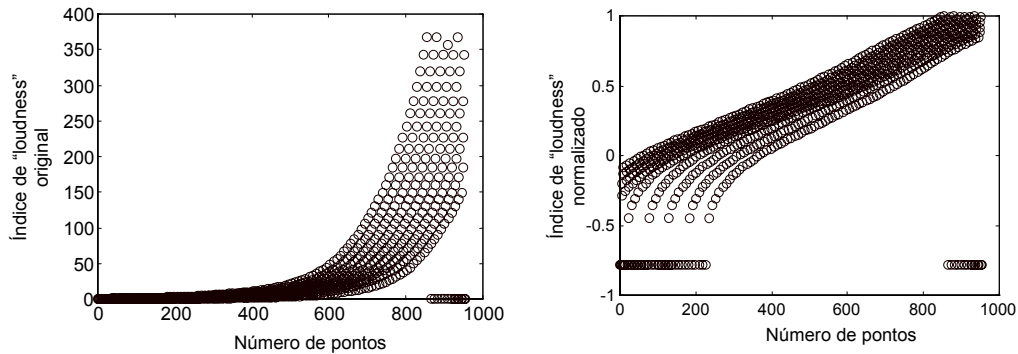


Figura 2. Normalização dos índices de “loudness”.

As informações de entrada também foram normalizadas entre os valores de -1 e 1 . Para as freqüências centrais das bandas de oitava foram utilizados os mesmos procedimentos usados na normalização do índice de “loudness”. Para os níveis de pressão sonora foi aplicada apenas uma expansão linear.

No processo de treinamento foi utilizado o método clássico padrão a padrão, isto é, para cada padrão entrada-saída são atualizados os pesos sinápticos das camadas de saída e escondida. O padrões de entrada e saída são mostrados na **Figura 3(a)**, onde no eixo das abscissas estão representados os níveis de pressão sonora em dB, no eixo das ordenadas estão representadas as freqüências centrais das bandas de 1/1 oitava e os valores da matriz correspondem aos índices de “loudness” utilizados no método de Stevens.

Uma vez realizado o treinamento da rede neural, ela foi submetida a testes com valores de freqüência e níveis de pressão sonora próximos aos padrões utilizados na fase de treinamento obtendo-se respostas coerentes, como mostra a **Figura 3(b)**.

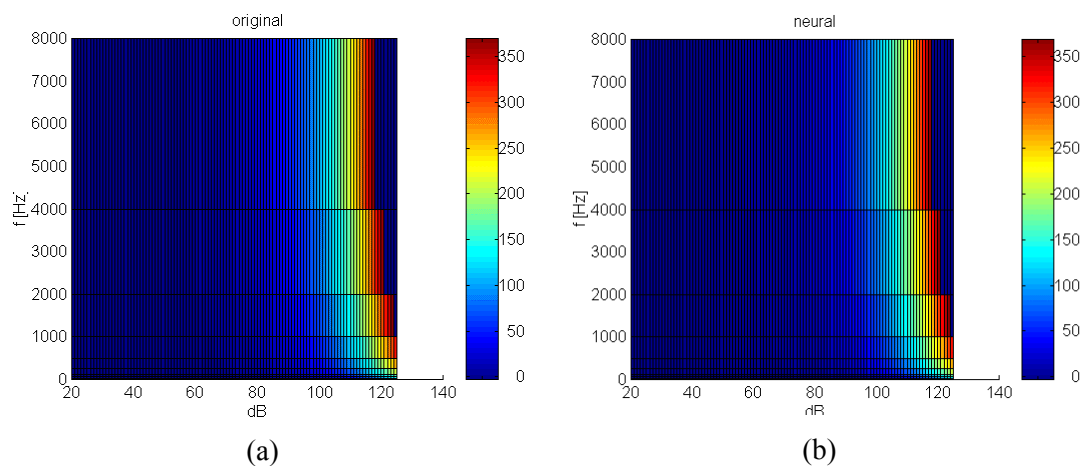


Figura 3. (a) Padrões de treinamento utilizados e (b) valores obtidos com a rede neural após o treinamento.

4. RESULTADOS

Como primeiro exemplo de teste foi utilizado o som produzido a partir do sinal de nível de pressão sonora de voz dizendo a nota musical “shhhh” medido com um microfone de eletreto acoplado ao sistema comercial SYMPHONIE da 01dB. A **Figura 4** mostra os valores do espectro em bandas de oitava (31.5 Hz a 2 kHz) sendo a referência de 20 μ Pa, os quais foram obtidos utilizando o equipamento comercial da 01dB.

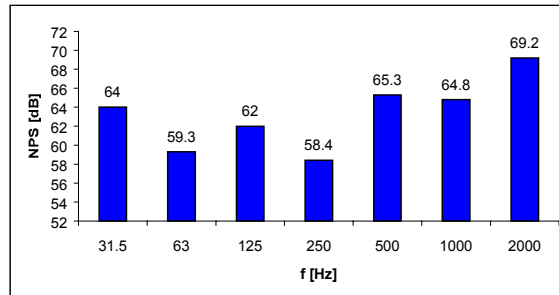


Figura 4. Valores do espectro do sinal de nível de pressão sonora de voz dizendo “shhhh” em bandas de oitava

Na **Tabela 1**, são mostrados os valores de “loudness” parciais (para cada banda) calculados utilizando o método de Stevens e os valores de “loudness” parciais obtidos com a rede neural.

Tabela 1. Valores de “loudness” parciais para o exemplo da Figura 4

f[Hz]	31.5	63	125	250	500	1000	2000
Stevens	1.33	1.47	2.81	3.08	5.59	6.52	10.02
Neural	1.33	1.56	2.81	3.01	5.62	6.52	10.27

O “loudness” global é 16.26 Sones segundo o método de Stevens e 16.54 Sones segundo a rede neural o qual gera um erro de 1.72%.

Como segundo exemplo, foi utilizada as informações de ruído gerados pelo motor de um caminhão pesado (19 ton.). O teste foi realizado com o veículo carregado em trecho asfaltado, plano, liso e reto, acelerando o motor lentamente de 1000 rpm até 2300 rpm (aproximadamente). As informações foram obtidas no habitáculo do caminhão através de um microfone que foi fixado no ouvido direito do motorista. As janelas estavam todas fechadas assim como todas as entradas de ar. As informações de níveis de pressão sonora foram obtidas para uma rotação de 1500 rpm.

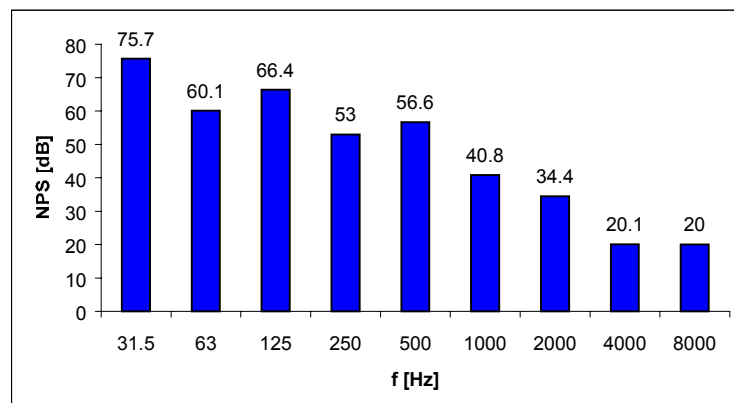


Figura 5. Valores do espectro do sinal de nível de pressão sonora de um motor de caminhão a 1500 rpm em bandas de 1/1 oitava

Os valores de “loudness” parciais calculados pelo método de Stevens e pela rede neural são mostrados na Tabela 2. Observa-se que apesar do valor do nível de pressão sonora ser maior em 31.5 Hz, o maior valor de “loudness” é atingido em 125 Hz cujo nível de pressão sonora é menor que o nível de pressão sonora em 31.5 Hz. Os valores de “loudness” global calculados são de 8.04 Sones para o método de Stevens e 7.96 Sones para a rede neural gerando um erro de 1.00 %.

Tabela 2. Valores de “loudness” parciais para o exemplo da **Figura 5**

f [Hz]	31.5	63	125	250	500	1000	2000	4000	8000
Stevens	3.14	1.573	3.82	2.24	3.32	1.52	1.216	0.455	0.61
Neural	3.103	1.666	3.705	2.193	3.326	1.554	1.212	0.486	0.643

Como terceiro exemplo foi utilizado o sinal de ruído emitido por um compressor hermético de geladeira (**Figura 6**).



Figura 6. Medidas de sinais sonoros emitidos por compressor hermético

As informações foram obtidas numa sala pouco reverberante (Laboratório de Vibroacústica da UNICAMP) utilizando o sistema comercial SYMPHONIE da 01dB. Os níveis de pressão sonora em bandas de oitavas estão na **Figura 7**

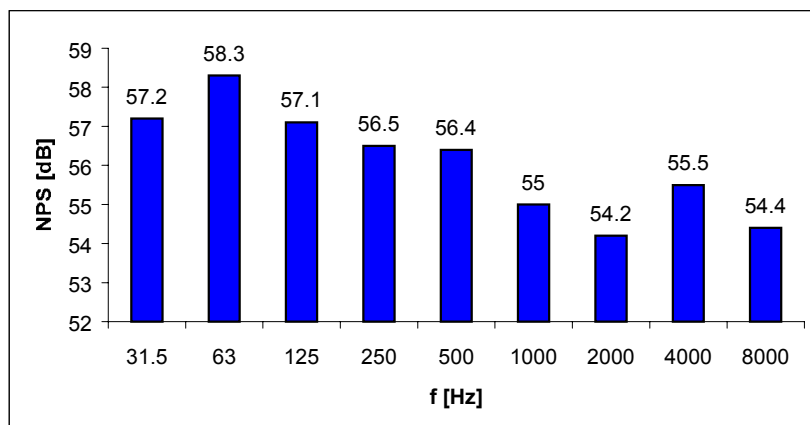


Figura 7. Valores do espectro do sinal de nível de pressão sonora emitidos por um compressor de geladeira em bandas de 1/1 oitava

Os valores de “loudness” parciais calculados pelo método de Stevens e pela rede neural são mostrados na **Tabela 3**. Os valores de “loudness” global calculados são de 12.91 Sones para o método de Stevens e 13.12 Sones para a rede neural produzindo um erro de 1.63 %.

Tabela 3. Valores de “loudness” parciais para o exemplo da Figura 7

f [Hz]	31.5	63	125	250	500	1000	2000	4000	8000
Stevens	0.706	1.363	1.975	2.76	3.28	3.6	4.14	5.35	5.96
Neural	0.6439	1.4414	2.0726	2.6959	3.2879	3.5855	4.0787	5.3905	6.1644

5. CONCLUSÕES

A rede neural multicamada utilizada neste trabalho mostra a capacidade de aprendizagem e também mostra a capacidade de atuar como um aproximador não linear, pois as variáveis utilizadas como frequências centrais das bandas de 1/1 oitava e os índices de “loudness” segundo o método de Stevens são não lineares. Nos três exemplos de aplicação, a rede neural mostrou erros abaixo dos 5% comparados com o modelo de Stevens no cálculo de índice de “loudness”. Além disso existe a possibilidade de corrigir ou adicionar padrões não utilizados no treinamento até o momento, e até pode-se pensar tornar o sistema de cálculo de índice de “loudness” personalizado.

6. AGRADECIMENTOS

Os autores gostariam de agradecer a FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo, Proc. 99/08862-2) pelo suporte financeiro.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- HAYKIN S. (1999), *Neural Networks: A comprehensive Foundation*, MC MILLAN PUBLISHING COMPANY, INC., USA.
- ISO 532^a (1975), Method for calculating “loudness” level, INTERNATIONAL ORGANIZATION FOR STANDARIZATION.
- STEVENS S.S. (1955), The Measurement of “loudness”, JOURNAL OF ACOUSTICAL SOCIETY OF AMERICA, v.27, p.815-829.
- STEVENS S.S. (1956), Calculation of the “loudness” of Complex noise, JOURNAL OF ACOUSTICAL SOCIETY OF AMERICA, v.28, p.807-829.
- STEVENS S.S. (1961), Procedure for Calculating “loudness”: Mark VI, JOURNAL OF ACOUSTICAL SOCIETY OF AMERICA, v.33, p.1577-1586.
- ZWICKER E., FASTL H. (1999), *Psychoacoustics – Facts and Models*, SPRINGER-VERLAG, HEIDELBERG.