

Aplicabilidade da Análise Fractal ao Reconhecimento de Locutor

Dirceu Gonzaga*

Rosângela Coelho

Instituto Militar de Engenharia
Departamento de Engenharia Elétrica
Praça General Tibúrcio, 80
Rio de Janeiro 22290-270
E-mail: coelho@ime.eb.br

Resumo— Este artigo apresenta os resultados de uma investigação sobre a aplicabilidade da análise fractal ao reconhecimento de locutor através de sinais de voz. A análise fractal foi realizada através do parâmetro de Hurst e, para isso, foram utilizados três algoritmos de estimação, estatística R/S, Higuchi e Wavelets. Em nossa análise, foram utilizados fonemas sonoros gravados por 4 locutores. Os resultados demonstraram que o parâmetro de Hurst pode representar uma importante característica da voz, podendo dessa forma ser utilizado no reconhecimento de locutor.

I. INTRODUÇÃO

As informações do locutor podem ser classificadas, através do sinal de voz, em dois grandes grupos: as características chamadas de alto nível e as fisiológicas. O primeiro grupo é composto de características tais como dialeto, contexto, estilo de fala, estado emocional, entre outras. Essas características são utilizadas para uma identificação “manual” ou através de um perito [9]. Todavia, em sistemas automáticos de verificação de locutor, as características de alto nível são difíceis de avaliar e quantificar, sendo necessárias a utilização de características baseadas na fisiologia e prosódia [6] tais como: coeficientes *cepstrum*, formante, magnitude espectral, *pitch*, energia, entre outras.

A análise fractal tem sido aplicada com sucesso em problemas que envolvem modelos complexos da natureza. Considerando que o sinal de voz é um sinal não linear e dinâmico, surge a necessidade de uma melhor modelagem da fala através da análise fractal [10] e de sistemas não lineares dinâmicos [5]. Na área de reconhecimento de voz, em [4] é feita uma análise fractal para a classificação de fonemas onde foram utilizadas a dimensão fractal em conjunto com os parâmetros do *iterated function system* (IFS)[2]. Em [11] é feita uma aplicação da dimensão fractal juntamente com os coeficientes LPC no reconhecimento do locutor. Esta combinação apresentou um considerável ganho na taxa de acerto do reconhecimento quando comparado ao LPC isolado, mostrando que a dimensão fractal possui importantes características do locutor.

O objetivo deste trabalho é fazer uma investigação da aplicação da análise fractal através da estimação do parâmetro de Hurst [8], no reconhecimento de locutor. Este

parâmetro tem um relação direta com a dimensão fractal [2] e possui um custo computacional menor do que as técnicas convencionais. Para esta análise foram utilizados três estimadores de Hurst.

Este trabalho está organizado da seguinte forma: na seção II é apresentada a definição de processos auto-similares, de dependência temporal e os estimadores de Hurst utilizados. Os resultados obtidos são discutidos na seção III. Finalmente, a seção IV apresenta as conclusões principais deste trabalho.

II. DEPENDÊNCIA TEMPORAL E ESTIMADORES DE HURST

É importante esclarecer a diferença existente entre processos com auto-similaridade e processos com dependência temporal. Embora sejam utilizados indistintamente em alguns trabalhos, os termos dependência temporal e auto-similaridade pertencem a contextos diferentes.

Um processo estocástico caracterizado como auto-similar, além da propriedade referente à invariância da distribuição marginal para quaisquer incrementos do processo, deve possuir um grau de dependência temporal. Por exemplo, o processo de movimento puro *browniano* é auto-similar e no entanto, não possui dependência ($H = 0.5$). Da mesma forma, um processo com dependência não necessariamente é auto-similar. Na literatura, tanto o grau de similaridade como o de dependência temporal são geralmente expressos pelo parâmetro de Hurst.

Considere um processo estocástico $X(t)$ com variância finita e função autocorrelação $\rho(k)$ definida por:

$$\rho(k) = \frac{Cov[X(t), X(t+k)]}{Var[X(t)]}, \quad k = 0, 1, 2, \dots \quad (1)$$

A função autocorrelação toma valores na faixa $[-1, 1]$ e $\lim_{k \rightarrow \infty} \rho(k) = 0$. A dependência temporal é definida pela taxa com que $\rho(k)$ decai a zero para $k \rightarrow \infty$.

Considerando que o comportamento assintótico da função $\rho(k)$ é dado por $\rho(k) \sim L(k)k^{2(H-1)}$. Define-se o grau de dependência temporal como valor do parâmetro de Hurst (H), onde $L(k)$ é uma função de variação lenta no infinito [3].

*O autor é aluno do programa de mestrado do Departamento de Engenharia Elétrica do IME.

Este decaimento hiperbólico da função autocorrelação difere do apresentado em processos tradicionais markovianos, onde o decaimento é exponencial.

O parâmetro de Hurst representa portanto, a forma de decaimento assintótico de $\rho(k)$. Considerando processos de variância finita o valor de H pertence a faixa $0 < H < 1$. Nesta faixa, podemos distinguir três diferentes processos. Estes processos são definidos como: i) $0 < H < \frac{1}{2}$ - processos que possuem anti-persistência ou dependência negativa. A função autocorrelação rapidamente cai a zero e em particular, $\sum_{k=-\infty}^{\infty} \rho(k) = 0$; ii) $H = \frac{1}{2}$ - nesta situação, os processos apresentam *Short Range Dependence* (SRD) e a função autocorrelação $\rho(k)$ apresenta decaimento exponencial para zero tal que $\sum_{k=-\infty}^{\infty} \rho(k) = c$, onde $c > 0$ é uma constante finita; iii) $\frac{1}{2} < H < 1$ - estes processos apresentam dependência positiva, *Long Range Dependence* (LRD) ou persistência. A função $\rho(k)$ apresenta decaimento lento indicando dependência mesmo entre amostras distantes. Neste caso, temos $\sum_{k=-\infty}^{\infty} \rho(k) = \infty$.

A. Estimadores de Hurst

Diversos métodos para realizar a estimação de Hurst tem sido apresentados na literatura [14]. A obtenção do parâmetro de Hurst é complexa visto que os efeitos da dependência se manifestam de maneira assintótica. Além disso, o método de coleta e tratamento de amostras pode influenciar bastante no processo de estimação [12].

Em nosso estudo serão utilizados três estimadores para a obtenção da dependência temporal, e conseqüentemente, a dimensão fractal do sinal de voz. Os estimadores são: a estatística R/S (*Rescaled adjusted range*) [8], o método de Higuchi [7] e o estimador Abry-Veitch (AV) [1], baseado em *wavelets*. Todos os estimadores utilizados fazem parte de um pacote HEP¹ desenvolvido em [12].

A seguir serão descritos, sucintamente, os estimadores R/S, Higuchi e AV.

A.1 Estatística R/S

A estatística R/S foi definida por Harold Edwin Hurst [8], para o dimensionamento de reservatórios de água no rio Nilo. O estimador R/S é o método mais utilizado para estimar o parâmetro de Hurst. A sua principal vantagem é a independência com relação a distribuição marginal das amostras do processo em análise.

Considere uma seqüência de amostras aleatórias X_i com somas parciais $Y(n) = \sum_{i=1}^n X_i$ e variância amostral $S^2(n) = (\frac{1}{n}) \sum_{i=1}^n X_i^2 - (\frac{1}{n})^2 Y(n)^2$. A estatística R/S é dada por:

$$\frac{R(n)}{S(n)} = \frac{1}{S(n)} \left[\max_{0 \leq t \leq n} (Y(t) - \frac{t}{n} Y(n)) - \min_{0 \leq t \leq n} (Y(t) - \frac{t}{n} Y(n)) \right] \quad (2)$$

Verifica-se que $E[R(n)/S(n)] \sim C_1 n^H$ para $n \rightarrow \infty$, onde C_1 é uma constante positiva. Considerando-se uma seqüência de tamanho N , devemos subdividi-la em K blocos de tamanho N/K . Para valores de $n < N/K$ são

¹O HEP está disponível e pode ser obtido por email: coelho@ime.eb.br

obtidas K estimativas de $R(n)/S(n)$ calculadas em cada bloco. Caso n seja maior que o tamanho do bloco ($n > N/K$) obtém-se tantos valores de $R(n)/S(n)$ quanto possível, partindo-se dos pontos iniciais de cada bloco. O valor do parâmetro H pode ser estimado através de regressão linear sobre os pontos obtidos, em um gráfico log-log de $R(n)/S(n)$ versus n .

O parâmetro n nas equações acima é referido como *lag* [5], [14]. Então para cada *lag* n , calcula-se $R(k_i, n)/S(k_i, n)$, começando-se no ponto $k_i = iN/K + 1$, $i = 1, 2, \dots$, de forma que $k_i + n \leq N$. Para realizar a estimação do parâmetro de Hurst, devemos escolher uma faixa de valores para n , ou seja, $n \in [n_{inf}, n_{sup}]$. Para $n < n_{inf}$, os resultados para a estatística R/S ainda não se manifestam, sendo afetados por efeitos de dependência em curta escala. Para $n > n_{sup}$ existem poucos pontos disponíveis para obter estimativas razoáveis de $E[R(n)/S(n)]$. A escolha dos limites n_{inf}, n_{sup} depende de uma análise visual dos pontos obtidos, conforme pode ser visto na Figura 1. Quando a dependência é examinada em curto intervalo de tempo, ocorre uma imprecisão no valor de H (por ex: $H = 0.9$), levando a interpretação incorreta do verdadeiro grau de dependência presente no sinal.

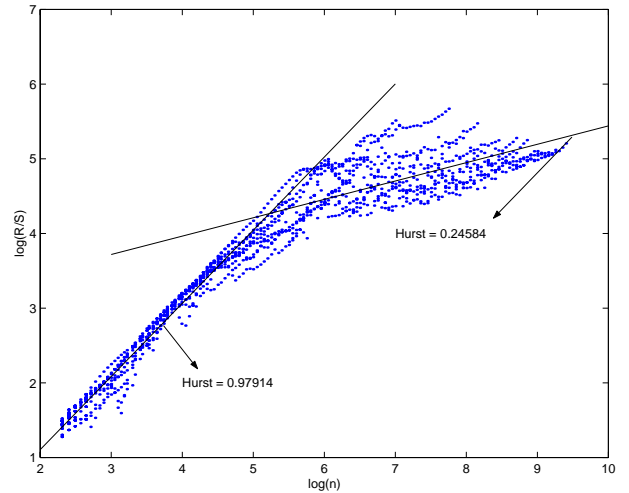


Fig. 1
ESTIMAÇÃO DO HURST ATRAVÉS DO MÉTODO R/S

A.2 Método Higuchi

O método Higuchi considera a dimensão fractal de uma série de amostras. O conceito de dimensão fractal está inserido no escopo da geometria fractal, desenvolvida por Mandelbrot nos anos 70/80 [2].

Uma das definições existentes para dimensão fractal é dada pela dimensão de Hausdorff (D_h) [2]. A dimensão de Hausdorff de um conjunto S em um espaço métrico é expressa por

$$D_h = \lim_{\epsilon \rightarrow 0} \frac{\ln[N(\epsilon)]}{\ln(\epsilon)} \quad (3)$$

onde $N(\epsilon)$ é o número de bolas abertas de raio ϵ necessário pra cobrir todo o conjunto S .

A dimensão fractal é diretamente relacionada ao parâmetro de Hurst pela relação $D_h = 2 - H$.

Para uma seqüência de amostras aleatórias $\{X_i\}$ com somas parciais $Y(n) = \sum_{i=1}^n X_i$, obtém-se k seqüências derivadas $Y_k^m = Y_m, Y_{m+k}, Y_{m+2k}, \dots, Y_{m+\lfloor \frac{N-m}{k} \rfloor k}$, onde $m = 1, 2, \dots, k$.

Para cada seqüência derivada Y_k^m , calcula-se o comprimento normalizado da curva:

$$L_m(k) = \frac{N-1}{K^2 \lfloor \frac{N-m}{k} \rfloor} \sum_{i=1}^{\lfloor \frac{N-m}{k} \rfloor} |Y_{m+ik} - Y_{m+(i-1)k}| \quad (4)$$

e definimos o comprimento da curva para cada intervalo k como

$$L(k) = \frac{1}{k} \sum_{m=1}^k L_m(k). \quad (5)$$

Verifica-se que $E[L(k)] \sim C_2 k^{-D_h}$ para $k \rightarrow \infty$, onde $D_h = 2 - H$. Novamente, o parâmetro H é estimado através de regressão em um gráfico de $\log L(k)$ por $\log(k)$, conforme visto na Figura 2 (exemplo de $H = 0.26$).

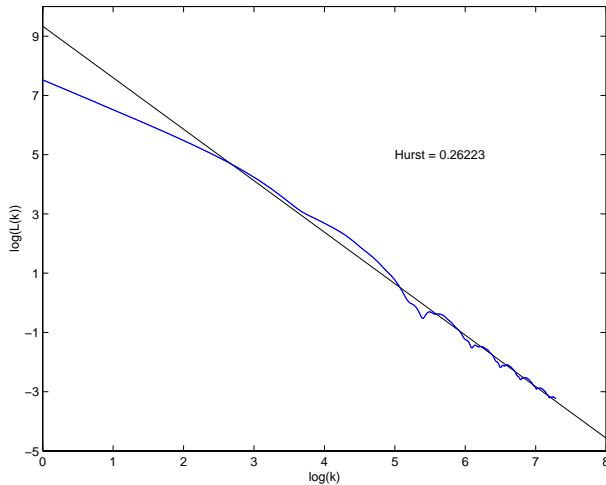


Fig. 2

ESTIMAÇÃO DO HURST ATRAVÉS DO MÉTODO HIGUCHI

A.3 Estimador AV

O estimador AV decompõe as amostras de um processo em seqüências de aproximação (passa-baixa) e detalhes (passa-alta) através de transformada discreta de *wavelet* (*DWT - discrete wavelet transform*).

Estas seqüências são obtidas através de filtros digitais especialmente projetados, cujos coeficientes são determinados pela *wavelet* utilizada. Não há restrição quanto à escolha das *wavelets* a serem utilizadas no estimador AV. Para manter compatibilidade com a proposta original do estimador apresentado em [1], utilizamos as *wavelets* de *Daubechies* com diferentes números de coeficientes.

Partindo de uma seqüência de amostras original, sucessivas seqüências de aproximação e detalhe são calculadas. Estas seqüências são obtidas aplicando-se filtragem digital de forma recursiva, ou seja, a saída de um estágio de filtragem é novamente aplicada ao estágio de filtragem e assim por diante. Este processo é ilustrado na Figura 3, para os três primeiros estágios de decomposição.

Cada nova seqüência é denominada $d(j, k)$, $k = 1, 2, \dots$, onde j representa a escala de decomposição no banco de filtros e k o índice de cada coeficiente gerado no banco de filtro em uma dada escala. Cada par de filtros é seguido de decimadores e portanto, a taxa de geração das seqüências $d(j, k)$ é reduzida à metade após um estágio de filtragem.

O estimador AV pode ser descrito, portanto, em três etapas:

1. Decomposição em *wavelets*: A DWT é aplicada nas amostras, gerando as seqüências de detalhes $d(j, k)$.

A DWT é obtida através do algoritmo piramidal (vide Figura 3). Este algoritmo é descrito detalhadamente em [13]. O algoritmo piramidal possui baixa complexidade computacional da ordem de $O(n)$, permitindo um cálculo rápido das seqüências de detalhes $d(j, k)$.

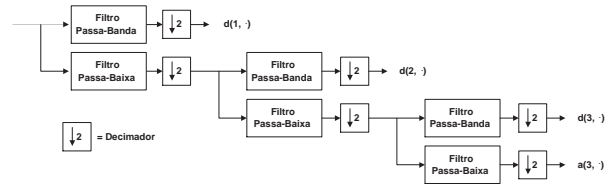


Fig. 3

BANCO DE FILTRO PIRAMIDAL PARA ESTIMADOR *wavelets*

2. Estimação da variância dos coeficientes de detalhes: Para cada escala j , obtém-se uma estimativa da variância dos coeficientes $d(j, k)$ denominada μ_j . Como estes coeficientes possuem média nula, tem-se que $\mu_j = (1/n_j \sum_k d(j, k)^2)$, onde n_j indica o número de coeficientes obtidos na escala j com expoente $\alpha = 2H - 1$, tal que $E[\mu_j] = c_\gamma j^{2H-1}$

3. Estimação do parâmetro H : Traça-se o gráfico de $y_j = \log_2(\mu_j)$ versus j , denominado diagrama log-escala. Através de regressão linear ponderado, obtém-se a inclinação α do gráfico e portanto, a estimativo do parâmetro $H = (1 + \alpha)/2$.

A etapa de regressão ponderada deve ser realizada somente nos pontos que apresentarem um alinhamento no diagrama log-escala. Esta região de alinhamento é denominada região de escalamento.

III. RESULTADOS E DISCUSSÃO

Para a análise foram utilizados os fonemas sonoros /a/, /é/, /ê/, /i/, /ó/, /ô/, /u/, gravados isoladamente de forma sustentada. Foram utilizadas cerca de 280 gravações, faladas por 4 locutores, 2 masculinos (LM) e 2 femininos (LF), sendo 10 repetições de cada fonema. Foi utilizada inicialmente a taxa de amostragem de 44.1 Khz. Taxas mais

baixas foram obtidas através de *downsample* com a finalidade de examinar os valores de H para diferentes taxas de amostragem. A duração média em *ms* dos fonemas, por locutor, está apresentado na Tabela I.

TABELA I
TEMPO MÉDIO EM *ms* DOS FONEMAS POR LOCUTOR

	/a/	/é/	/ê/	/i/	/ó/	/ô/	/u/
LM 1	355	351	352	334	337	326	307
LM 2	443	442	423	386	406	359	333
LF 1	505	515	492	483	506	495	529
LF 2	508	483	425	371	364	346	387

A seguir, serão apresentados os resultados da estimação do parâmetro de Hurst dos fonemas sonoros. A aplicabilidade do parâmetro de Hurst no reconhecimento de locutor também é examinada. A estimação do parâmetro de Hurst foi realizada utilizando-se os três estimadores descritos no item II-A.

A. Resultados

A fim de verificar a influência da taxa de amostragem na estimação de H , utilizamos os valores de 8, 11, 22 e 44 KHz para todas as gravações de cada locutor. Os resultados médios por locutor, estimados pelo algoritmo R/S, estão apresentados na Figura 4.

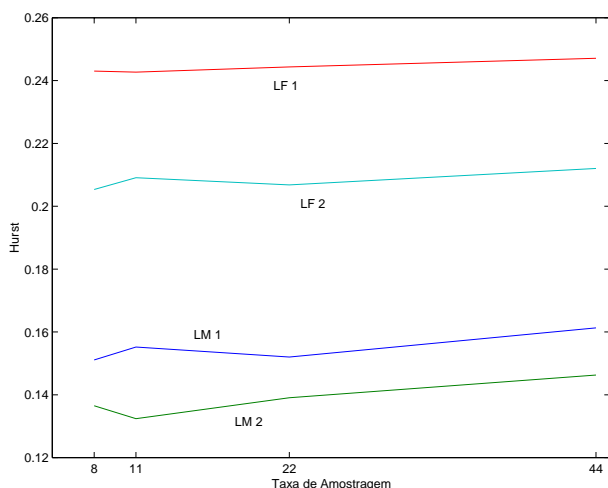


Fig. 4

HURST COM ESTIMADOR R/S VARIANDO-SE A TAXA DE AMOSTRAGEM PARA LOCUTORES

Como podemos observar, percebe-se uma separação bem nítida entre os locutores para os valores de H , principalmente para a taxa de 11 KHz, a qual foi a escolhida para o restante do trabalho. Além disso, o valor de H não sofreu variação brusca mudando-se a taxa de amostragem.

A Figura 5 mostra a estatística média de R/S do fonema /a/, para cada locutor. Nessa figura, pode-se examinar o comportamento do *lag* discutido no item II-A.1. Pode-se observar que a separação entre os locutores se inicia

para *lag* em torno de $\log(n) \approx 4$ o que equivale a cerca de 5ms para uma taxa de amostragem de 11 kHz. Observa-se também que para valores abaixo desse *lag* não há separação entre os quatro locutores.

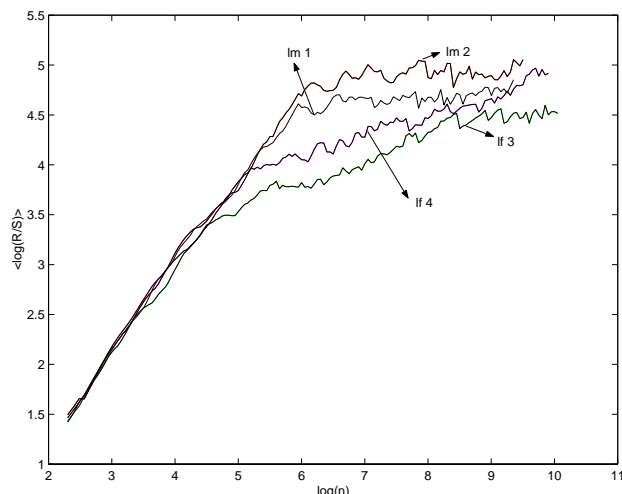


Fig. 5

ESTIMAÇÃO DO HURST MOSTRANDO AS MÉDIAS DE R/S POR LOCUTOR

A Tabela II, apresenta os valores limites de *lag*, para todos os fonemas, a partir dos quais começa a haver separação entre os locutores. Os valores do *lag* são dados em *ms*.

TABELA II
TEMPO MÉDIO EM *ms* DOS VALORES LIMITES DO *lag* PARA CADA FONEMA.

	/a/	/é/	/ê/	/i/	/ó/	/ô/	/u/
LM 1	3.61	5.39	6.01	6.79	3.83	5.41	7.03
LM 2	3.75	4.36	6.60	7.13	4.83	5.39	6.99
LF 1	1.92	2.06	2.68	3.33	2.47	3.22	3.04
LF 2	1.95	3.41	4.20	5.34	2.84	3.44	4.30

A Tabela III apresenta os valores médios de H obtidos com os estimadores R/S e Higuchi, discutidos anteriormente. Pode-se perceber que os estimadores seguem uma mesma tendência, não mostrando grandes discrepâncias de valores.

Com o estimador AV, não foi possível, até o momento, se chegar a resultados confiáveis, visto que a quantidade de amostras após a filtragem é pequena para se realizar uma estimação, ou seja, que fosse possível se observar uma tendência para a regressão linear. Novos estudos estão em andamento com a utilização de palavras e frases.

De nossos resultados concluímos ainda, que o estimador R/S fornece uma maior "separabilidade" de locutor que o estimador Higuchi. A fim de realizar outros experimentos com a finalidade de reconhecimento de locutor, escolhemos a estatística R/S média, conforme mostrado na Figura 6.

		/a/	/é/	/ê/	/i/	/ó/	/ô/	/u/
LM1	R/S	0.16	0.22	0.15	0.14	0.13	0.14	0.14
	Hig	0.16	0.23	0.19	0.20	0.13	0.18	0.20
LM2	R/S	0.14	0.14	0.12	0.14	0.12	0.12	0.14
	Hig	0.15	0.15	0.18	0.19	0.15	0.14	0.21
LF1	R/S	0.25	0.25	0.20	0.28	0.20	0.17	0.36
	Hig	0.30	0.24	0.21	0.33	0.23	0.21	0.37
LF2	R/S	0.27	0.25	0.18	0.23	0.18	0.15	0.20
	Hig	0.23	0.24	0.16	0.22	0.14	0.13	0.21

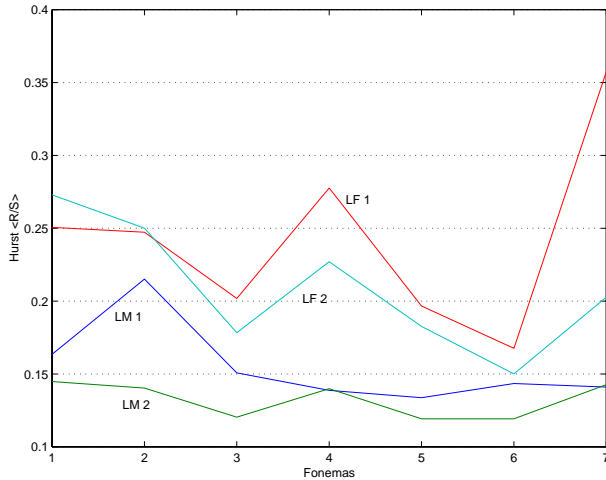


Fig. 6

VALORES DE HURST PARA A CLASSIFICAÇÃO DE LOCUTORES PELO ESTIMADOR R/S POR FONEMAS NA SEQÜÊNCIA /A/, /É/, /Ê/, /I/, /Ó/, /Ô/, /U/.

Pelos resultados apresentados nas Figuras 4 e 6, pode-se destacar o seguinte:

- Os fonemas apresentaram características de anti-persistência ($H < 0.5$) conforme visto no item II;
- Há uma separação direta entre os locutores femininos e masculinos;
- Para os locutores masculinos os fonemas /i/ e /u/ não apresentam discriminação. Para os locutores femininos o fonema /é/ é o menos discriminativo.
- Realizando-se uma média de todos os fonemas por locutor (vide Figura 4), percebe-se que há uma separação bem nítida entre os quatro locutores, com ênfase na taxa de amostragem de 11 kHz.

A Figura 7 apresenta um gráfico de Hurst médio por tamanho de janela de voz. Todo o sinal foi janelado sem superposição. Pode-se observar que janelas com duração em torno de 75 ms apresentam melhores condições para classificação. Janelas abaixo de 50 ms não apresentam separação entre os locutores masculinos.

Na Figura 8 pode-se observar a distribuição dos locutores por fonema, demonstrando que é possível realizar uma

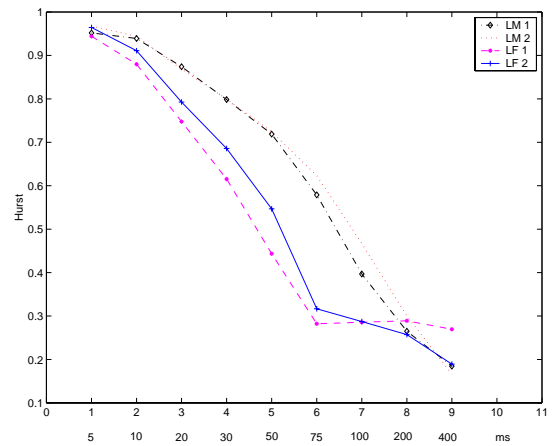


Fig. 7

VARIAÇÃO DO HURST PELO TAMANHO DA JANELA DO SINAL DE VOZ EM MS

interessante classificação, utilizando-se janelas de 75 ms, conforme mencionado anteriormente. Cabe ressaltar que os valores de Hurst inverteram comparados com o gráfico da Figura 6. Isso é devido ao maior grau de correlação entre as amostras, visto que estamos considerando janelas curtas.

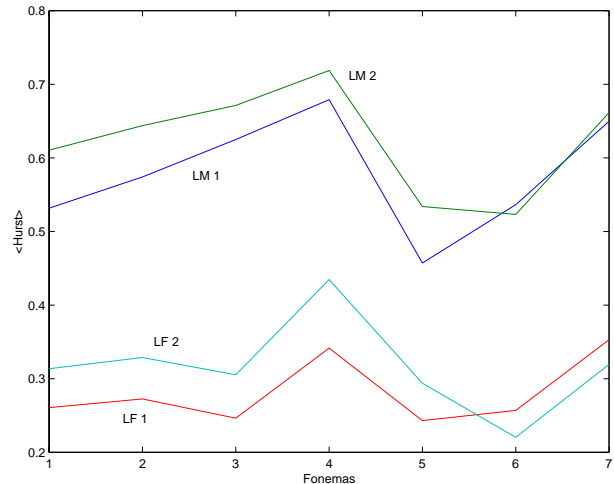


Fig. 8

VALORES DE HURST PARA A CLASSIFICAÇÃO DE LOCUTORES POR FONEMAS NA SEQÜÊNCIA /A/, /É/, /Ê/, /I/, /Ó/, /Ô/, /U/, UTILIZANDO JANELAS DE 75 MS.

A fim de avaliarmos a capacidade de classificação dos locutores utilizando a característica fractal (H), foi feita uma comparação utilizando-se a *pitch*. Esta foi extraída através do método AMDF [15]. A Figura 9 mostra os valores médios da *pitch* dos fonemas por locutor.

Para realizar a classificação, foi utilizada uma rede neural associativa [16] na qual o vetor de entrada era composto pelo Hurst de cada fonema, formando uma matriz de ve-

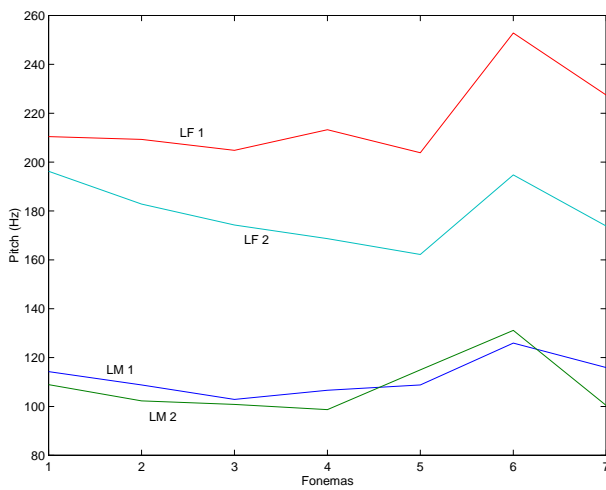


Fig. 9

VALORES DE PITCH POR FONEMAS NA SEQÜÊNCIA /A/, /É/, /Ê/, /I/, /Ó/, /Ô/, /U/.

tores de H dos 7 fonemas x 40 repetições, sendo 10 para cada locutor. A saída foi formada por 4 códigos ortogonais, onde cada um destes indica um locutor, o esquema geral dessa implementação pode ser visto na Figura 10. A matriz de entrada da *pitch* foi composta da mesma forma.

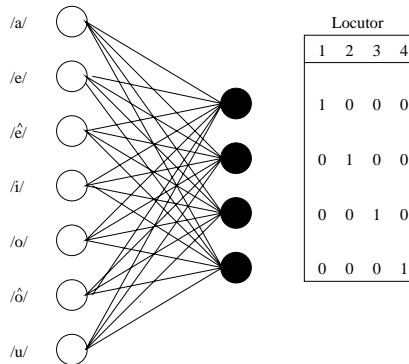


Fig. 10

ESQUEMA DE TREINAMENTO COM A REDE NEURAL ASSOCIATIVA.

A Tabela IV apresenta os resultados, para os dados de treinamento, da taxa de acerto utilizando o parâmetro de Hurst e a *pitch*. Pode-se observar que, para um classificador linear, o parâmetro de Hurst apresentou uma taxa de acerto maior que a *pitch* em cerca de 12%. A tabela também mostra que estimando-se o Hurst em janelas de 75 ms, consegue-se um desempenho de “separabilidade” melhor que quando estima-se H com o sinal completo. Este resultado confirma o que foi observado na Figura 7. Pode-se ver também que para um vetor composto de *pitch* mais Hurst, (matriz com entrada de 14x40 - 7 Hurst e 7 *pitch*), a taxa de acerto sobre consideravelmente (97.5%). Isto demonstra que o Hurst juntamente com outras caracte-

rísticas de voz pode apresentar excelentes resultados para o reconhecimento de locutor.

TABELA IV

TAXA DE ACERTOS UTILIZANDO UM CLASSIFICADOR LINEAR, EM %

Pitch	Hurst(75 ms)	Hurst (SC)	(Pitch + Hurst)
75	87,5	82,5	97,5

IV. CONCLUSÃO

Neste artigo foi apresentado um estudo sobre a aplicação da análise fractal, através do parâmetro de Hurst, ao reconhecimento de locutor. Foram estudados três estimadores para estimação de H , e observou-se que o R/S apresentou melhores resultados para um reconhecimento. Através de um classificador linear, baseado em redes neurais, foi possível perceber que a análise fractal pode se constituir numa importante característica de voz e portanto no reconhecimento do locutor.

REFERÊNCIAS

- [1] Abry, P., Veitch, D. *Wavelet analysis of long-range dependent traffic*. IEEE Transaction on Information Theory, 44(1), pp.2-15, jan. 1998.
- [2] Barnsley, M.F. *The Science of Fractal Images*. Springer-Verlag Inc, New York, USA, 1988.
- [3] Beran, J. *Statistics for Long-Memory Processes*. Chapman & Hall, 1994.
- [4] Bohez, E.L.J., Senevirathne, T.R., Van Winden, J.A. *Fractal Dimension and Iterated Function System (IFS) for Speech Recognition*. Electronics Letters, Vol 28, nr 15, pp. 1382-1384, jul. 1992.
- [5] Boshoff, H.F.V., Grotepass, M. *The Fractal Dimension of Fricative Speech Sounds*. IEEE Proceedings ICASSP 1991.
- [6] Deller, J.R. JR, Proakis, J.G., Hansen, J.H.L.. *Discret Time Processing of Speech Signals*. Macmillan Publishing Company, New York, 1993.
- [7] Higuchi T. *Approach to an Irregular Time Series on the Basis of the Fractal Theory*. Physica D., vol. 31, pp.277-283, 1988
- [8] Hurst, H.E. *Long-term storage capacity of reservoirs*. Transaction of American Society of Civil Engineers, 116, pp.770-799, 1951.
- [9] Jayant, M. Naik *Speaker Verification: A Tutorial*. IEEE Communication Magazine, pp. 42-48, jan. 1990
- [10] Mann, I. *An Investigation of Nonlinear Speech Syntesis and Pitch Modification Techiques*. Thesis PhD, The University of Edinburgh, set. 1999.
- [11] Petry, A., Barone, D.A.C. *Fractal Dimension Applied to Speaker Identification*. IEEE Proceedings ICASSP 2001.
- [12] Pontes, A.R.S. e Coelho R. *Controle de Admissão de Conexão Baseado em Medidas de Tráfego de Vídeo com Dependência Temporal*. Anais do XVIII Simpósio Brasileiro de Telecomunicações, 2000.
- [13] Press, W.H., Vetterling, W.T., Teukolsky, S.A., Flannery, B.P. *Numerical Recipes in C*. Cambridge University Press, 2 Ed, 1992.
- [14] Taquq, M.S., Teverovsky, V., Willinger, W. *Estimators for long-range dependence: an empirical study*. Fractals, Vol 3, nr 4, pp. 785-788, 1995.
- [15] Un, Chong K., Yang, S.C. *A Pitch Extraction Based on LPC Inverse Filtering and AMDF*. IEEE Transaction on Acoustic, Speech, and Signal Processing, vol. ASSP-25 n. 6, pp. 565-572, dez. 1977.
- [16] Zurada, J.M. *Introduction to Artificial Neural Systems*. West Publishing Company, 1992.