

# UTILIZAÇÃO DA ANÁLISE DE COMPONENTE PRINCIPAL NA REDUÇÃO DOS VETORES DE PARÂMETROS EM SISTEMAS DE RECONHECIMENTO DE FALA

Carlos Alberto Ynoguti<sup>(1)</sup> e Fábio Violaro<sup>(2)</sup>

<sup>(1)</sup> Grupo de Pesquisa em Processamento Digital de Sinais  
Instituto Nacional de Telecomunicações  
CP 05, CEP: 37540-000, Santa Rita do Sapucaí, MG  
ynoguti@inatel.br

<sup>(2)</sup> DECOM-FEEC-UNICAMP  
CP 6101, CEP: 13083-970, Campinas, SP  
fabio@decom.fee.unicamp.br

## RESUMO

Este trabalho tem por objetivo investigar a aplicação da técnica de Análise de Componente Principal [4] (*Principal Component Analysis*, PCA) na redução da dimensão dos vetores de parâmetros de um sistema de reconhecimento de fala contínua. Parâmetros de dimensão reduzida levam a uma economia tanto em termos de memória como em custo computacional. Nos testes realizados em um sistema baseado em HMM's contínuos conseguiu-se uma redução de quase 30 % na dimensão dos vetores de parâmetros, sem queda na taxa de acertos de palavras.

## 1. INTRODUÇÃO

Atualmente, grande parte dos sistemas de reconhecimento de fala emprega como parâmetros os coeficientes mel-cepstrais [2], juntamente com suas derivadas de primeira e de Segunda ordem. Se considerarmos que cada vetor de coeficientes mel tem dimensão 12, os parâmetros delta e delta-delta também terão dimensão 12. Desta forma, a cada quadro de análise do sinal e para cada estado do espaço de busca, tem-se 3 vetores de 12 componentes. Esses vetores podem ser modelados isoladamente ou agrupados em um único vetor de dimensão 36.

Considerando-se que esses vetores vão ser modelados estatisticamente por misturas de  $n$  gaussianas multidimensionais, e levando-se em conta adicionalmente o uso de matrizes de covariância diagonais (componentes incorrelatas), resulta que  $36n$  gaussianas unidimensionais são necessárias para o modelamento dos parâmetros em cada quadro do sinal de fala. Como para o caso de um sistema de reconhecimento de fala com vocabulário extenso tem-se vários milhares de estados, verifica-se que o número de gaussianas a ser calculado é bastante elevado. Além disso, o cálculo de cada uma destas gaussianas (ou

outras funções densidade de probabilidade) envolve um custo computacional relativamente alto, como pode ser visto pela expressão abaixo.

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) \quad (1)$$

Uma redução da dimensão do vetor de parâmetros utilizados no processo de reconhecimento diminui a carga computacional imposta ao sistema de reconhecimento, reduzindo conseqüente-mente o tempo de processamento.

Com este objetivo adotou-se a seguinte estratégia na elaboração do presente trabalho: inicialmente os parâmetros mel-cepstrais, delta mel e delta-delta mel foram agrupados em um único vetor. Desta forma o sistema passa a enxergar não três parâmetros de dimensão 12, mas apenas um parâmetro de dimensão 36. Obviamente este procedimento não reduz a carga computacional, pois ainda tem-se que calcular  $36n$  gaussianas para cada estado.

Aplicando-se agora a Análise de Componente Principal ao vetor composto, consegue-se reduzir a dimensão deste e forma que a perda de informação seja mínima.

Os testes realizados mostraram que, em alguns casos, é possível uma redução de até 30 % na dimensão dos vetores compostos, sem prejuízo na taxa de acertos do sistema.

Na seção seguinte será dada uma rápida visão sobre a teoria envolvida na Análise de Componente Principal e, nas demais seções, será descrita sua aplicação a um sistema de reconhecimento de fala contínua e os resultados obtidos.

## 2. ANÁLISE DE COMPONENTE PRINCIPAL

A análise de componente principal está relacionada à explicação da estrutura de variância-covariância através de algumas poucas combinações lineares das variáveis originais. Seus objetivos principais são (1) redução de dados, e (2) interpretação.

Embora sejam necessários  $p$  componentes para reproduzir toda a variabilidade do sistema, geralmente muito desta variabilidade pode ser avaliada por um número menor  $k$  de componentes principais. Se isto for verdade, existe quase tanta informação nas  $k$  componentes quanto nas  $p$  variáveis originais. As  $k$  componentes principais poderiam então substituir as  $p$  variáveis originais, e o conjunto de dados inicial, consistindo de  $n$  medidas em  $p$  variáveis, é reduzido para outro consistindo de  $n$  medidas em  $k$  componentes principais.

Algebricamente, componentes principais são combinações lineares particulares das  $p$  variáveis aleatórias  $X_1, X_2, \dots, X_p$ . Geometricamente, estas combinações lineares representam a seleção de um novo sistema de coordenadas obtido pela rotação do sistema original com  $X_1, X_2, \dots, X_p$  como eixos coordenados. Os novos eixos representam as direções com máxima variabilidade e fornecem uma descrição mais simples e parcimoniosa da estrutura de covariância.

Seja  $\Sigma$  a matriz de covariância associada ao vetor aleatório  $\mathbf{X}' = [X_1, X_2, \dots, X_p]$ . Se  $\Sigma$  tem pares autovalores/autovetores dados por  $(\lambda_1, \mathbf{e}_1), (\lambda_2, \mathbf{e}_2), \dots, (\lambda_p, \mathbf{e}_p)$ , onde  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ , então a  $i$ -ésima componente principal é dada por

$$Y_i = \mathbf{e}_i' \mathbf{X} = e_{i1} X_1 + e_{i2} X_2 + \dots + e_{ip} X_p, i = 1, 2, \dots, p \quad (2)$$

Este resultado mostra que as componentes principais são descorrelacionadas e têm variâncias iguais aos autovalores de  $\Sigma$ .

Sejam  $Y_1 = \mathbf{e}_1' \mathbf{X}$ ,  $Y_2 = \mathbf{e}_2' \mathbf{X}$ , ...,  $Y_p = \mathbf{e}_p' \mathbf{X}$  as componentes principais. Então

$$\begin{aligned} \sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} &= \sum_{i=1}^p \text{var}(X_i) = \\ \lambda_1 + \lambda_2 + \dots + \lambda_p &= \sum_{i=1}^p \text{var}(Y_i) \end{aligned} \quad (3)$$

Este resultado mostra que a variância total é dada por :

$$\text{Variância total} = \sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} = \lambda_1 + \lambda_2 + \dots + \lambda_p \quad (4)$$

e, conseqüentemente, a proporção da variância total devido à  $k$ -ésima componente principal é dada por

$$\left( \begin{array}{l} \text{Proporção da variância} \\ \text{total devido à } k \text{-ésima} \\ \text{componente principal} \end{array} \right) = \frac{\lambda_k}{\lambda_1 + \lambda_2 + \dots + \lambda_p}, k = 1, 2, \dots, p \quad (5)$$

Se a maior parte (por exemplo, 80 a 90%) da variância populacional total, para  $p$  grande, puder ser atribuída às primeiras uma, duas ou três componentes, então estas poderiam "substituir" as  $p$  variáveis originais sem muita perda de informação.

É importante observar que, embora os resultados acima tenham sido desenvolvidos a partir da matriz de covariância  $\Sigma$ , pode-se obter também as componentes principais a partir da matriz de correlação  $\mathbf{R}$  dos dados sob análise, e todos os resultados obtidos até aqui se aplicam integralmente.

## 3. APARATO EXPERIMENTAL

### 3.1 Sistema utilizado

Para os testes foi utilizado um software desenvolvido no LPDF [6], baseado em modelos ocultos de Markov contínuos.

Este sistema utiliza fones independentes de contexto como unidades fundamentais, o One Pass [3] como algoritmo de busca, e parâmetros mel-cepstrais [2] com suas respectivas derivadas de primeira e de segunda ordem (parâmetros delta mel e delta-delta mel) como a informação associada ao sinal de fala.

Para o cálculo dos autovetores e autovalores das matrizes de correlação e covariância, foi utilizado o algoritmo QL com deslocamentos implícitos em conjunto com o procedimento de Householder para tridiagonalização da matriz de covariância [5].

Os tempos de reconhecimento foram obtidos em um microcomputador com processador AMD-K6 350 MHz, 64 MB de RAM, sob a plataforma Windows 98.

### 3.2 Base de Dados [7]

As frases foram escolhidas segundo o trabalho realizado por Alcaim et. al. [1]. Neste, foram criadas 20 listas de 10 frases foneticamente balanceadas, segundo o português falado no Rio de Janeiro. Nestas listas, contou-se 694 palavras distintas.

Para as gravações foram selecionados 40 locutores adultos, sendo 20 homens e 20 mulheres. Para a formação do subconjunto de teste foram escolhidos, de forma aleatória, cinco locutores do sexo masculino e cinco do sexo feminino. Os demais locutores, 15 masculinos e 15 femininos, formam o subconjunto de treinamento. Nos

testes com dependência de sexo, os locutores de treinamento e teste são extraídos dos subconjuntos anteriores, resultando em 5 locutores de teste e 15 de treinamento.

Um locutor extra do sexo masculino completa a base de dados. Este locutor pronunciou todas as 200 frases, repetindo-as 4 vezes. Três repetições foram utilizadas para treinar o sistema e a última serviu como material de testes. Estas locuções foram utilizadas para testes com dependência de locutor.

As gravações foram realizadas em ambiente relativamente silencioso, com um microfone direcional de boa qualidade, utilizando uma placa de som SoundBlaster AWE 64. A taxa de amostragem utilizada foi de 11,025 kHz, com resolução de 16 bits por amostra. Os dados foram armazenados em formato Windows PCM (WAV).

A transcrição fonética foi feita manualmente para cada locução, utilizando programa de visualização gráfica do espectrograma e forma de onda do sinal, e fones de ouvido para audição da mesma.

## 4. TESTES

### 4.1 Descrição dos testes

Os testes foram realizados em quatro modos diferentes: dependente de locutor, apenas com locutores masculinos, apenas com locutores femininos e independente de locutor. Esta divisão foi realizada por dois motivos: a) a maioria dos sistemas de reconhecimento de fala comerciais pergunta inicialmente de que sexo é o locutor; b) como a análise PCA está intimamente relacionada à variância do espaço amostral, diferentes graus de variância entre os locutores podem gerar resultados diferentes (e conclusões diferentes também).

### 4.2 Testes iniciais

A primeira série de testes foi realizada utilizando-se os parâmetros mel-cepstrais, delta mel-cepstrais e delta-delta mel-cepstrais como três parâmetros independentes, cada qual modelado isoladamente com o seu conjunto de gaussianas. Esta série tem como objetivo estabelecer um desempenho base, com a qual os testes subsequentes serão comparados.

Foram feitos ensaios com misturas contendo de 4 a 7 gaussianas por estado, e os melhores resultados são mostrados na Tabela 1.

Nesta tabela e nas tabelas subsequentes, D, S e I correspondem à porcentagem de erros de deleção, substituição e inserção, respectivamente, e o tempo mínimo de reconhecimento das locuções é dado em minutos.

Tabela 1. Resultados dos testes iniciais.

Testes	D (%)	S (%)	I (%)	total (%)	número gauss.	tempo (min.)
Dep.	0,84	2,51	2,05	5,40	5	02:44
Masc.	3,88	8,22	2,21	14,31	5	02:06
Fem.	2,81	10,58	3,04	16,44	5	03:05
Indep.	3,69	11,68	2,82	18,19	5	02:43

### 4.3 Testes com parâmetros combinados

Em uma segunda etapa, os parâmetros mel-cepstrais, delta mel e delta-delta mel foram combinados em um único vetor. Desta forma, este novo vetor tem dimensões três vezes maior que os originais. Como neste trabalho foram utilizados parâmetros mel-cepstrais de ordem 12, este novo vetor combinado tem dimensão 36. Desta forma, trocam-se 3 parâmetros de ordem 12 por um parâmetro de ordem 36.

Agora, para cada estado, o sistema precisa calcular apenas uma mistura de gaussianas ao invés de 3, mas nesta única mistura as gaussianas tem dimensão 3 vezes maior. Em termos computacionais não há decréscimo de operações a serem realizadas.

Como no caso anterior, os testes foram realizados com misturas contendo de 4 a 7 gaussianas por estado. Os melhores resultados são mostrados na Tabela 2.

Tabela 2. Resultados dos testes com parâmetros combinados.

Testes	D (%)	S (%)	I (%)	total (%)	número gauss.	tempo (min.)
Dep.	1,67	2,51	1,07	5,25	4	03:11
Masc.	2,74	7,53	2,21	12,48	6	02:42
Fem.	3,04	12,33	3,80	19,18	5	02:56
Indep.	3,92	12,79	2,32	19,03	5	02:53

### 4.4 Testes utilizando a análise de componente principal

Na etapa final foi aplicada a análise de componente principal sobre os vetores combinados, testando-se vários níveis de redução da dimensão.

A partir da análise da proporção da variância total dada pela Equação (5), pode-se ter uma idéia de quanto é possível reduzir a dimensão dos vetores combinados sem perda significativa de informação. A título de ilustração, tem-se nas Tabelas 3 e 4 a explicabilidade do espaço amostral quando as variáveis originais são substituídas pelas componentes principais. Estas Tabelas foram obtidas a partir das locuções referentes ao modo independente de locutor.

Na Tabela 3 são apresentados os resultados para as componentes principais obtidas a partir da matriz de correlação e, na Tabela 4, os resultados para as componentes principais obtidas a partir da matriz de covariância.

Tabela 3. Proporção da variância total versus número de componentes utilizadas na representação das variáveis originais. Componentes principais obtidas a partir da matriz de correlação.

# comp.	% variância	# comp.	% variância
1	39.95%.	19	98.89%.
2	65.29%.	20	99.04%.
3	73.43%.	21	99.17%.
4	80.54%.	22	99.30%.
5	84.14%.	23	99.41%.
6	87.17%.	24	99.51%.
7	89.35%.	25	99.60%.
8	91.23%.	26	99.67%.
9	92.71%.	27	99.74%.
10	93.96%.	28	99.80%.
11	95.21%.	29	99.84%.
12	96.16%.	30	99.87%.
13	96.93%.	31	99.90%.
14	97.60%.	32	99.93%.
15	97.95%.	33	99.95%.
16	98.28%.	34	99.97%.
17	98.49%.	35	99.99%.
18	98.70%.	36	100.00%.

Tabela 4. Proporção da variância total versus número de componentes utilizadas na representação das variáveis originais. Componentes principais obtidas a partir da matriz de covariância.

# comp.	% variância	# comp.	% variância
1	41.40%.	19	98.74%.
2	61.85%.	20	98.91%.
3	71.13%.	21	99.06%.
4	78.45%.	22	99.20%.
5	82.41%.	23	99.33%.
6	85.51%.	24	99.45%.
7	87.99%.	25	99.55%.
8	90.10%.	26	99.63%.
9	91.74%.	27	99.70%.
10	93.18%.	28	99.77%.
11	94.54%.	29	99.82%.
12	95.62%.	30	99.86%.
13	96.50%.	31	99.89%.
14	97.26%.	32	99.92%.
15	97.67%.	33	99.94%.
16	98.04%.	34	99.97%.
17	98.28%.	35	99.98%.
18	98.52%.	36	100.00%.

A inspeção destas tabelas mostra que, para ambos os casos, uma redução para 20 componentes principais mantém aproximadamente 99 % da informação contida nas variáveis originais (vetor de parâmetros combinados), o que parece ser um limiar razoável de redução.

É interessante observar que este comportamento se repetiu também nos testes com locutores masculinos, locutores femininos e dependente de locutor, contrariando nossas expectativas de obtenção de uma possibilidade de redução maior nestes casos, devido à menor variabilidade entre locutores.

Qual é o efeito, entretanto, da redução de dimensionalidade dos coeficientes no desempenho do sistema de reconhecimento de fala? Pode-se assegurar que uma redução mínima da variância dos coeficientes representados num espaço de dimensão menor vai corresponder a uma degradação também mínima no desempenho do sistema de reconhecimento de fala? Para analisar o efeito líquido da redução da dimensionalidade no desempenho do sistema de reconhecimento de fala, foram realizados testes com diferentes níveis de redução.

Novamente, a título de ilustração, são apresentadas nas Tabelas 5 e 6 apenas os resultados para o modo independente de locutor, com misturas contendo 6 gaussianas

Tabela 5. Desempenho do sistema versus número de componentes principais para o modo independente de locutor. PCA sobre matriz de correlação.

número comp.	D (%)	S (%)	I (%)	total (%)	tempo (min.)
15	5,93	17,16	2,74	25,84	02:41
20	5,14	13,43	2,40	20,97	02:41
25	4,37	13,39	2,66	20,43	02:42
30	3,65	11,34	2,70	17,69	02:42
36	3,65	12,18	3,08	18,91	02:42

Tabela 6. Desempenho do sistema versus número de componentes principais para o modo independente de locutor. PCA sobre matriz de covariância.

número comp.	D (%)	S (%)	I (%)	total (%)	tempo (min.)
15	5,71	17,46	2,85	26,03	02:43
20	3,81	13,55	1,86	19,22	02:43
25	4,37	13,93	2,51	20,81	02:43
30	4,11	12,48	2,62	19,22	02:42
36	4,30	13,16	2,47	19,94	02:47

Analisando as tabelas acima, pode-se ver que a taxa de acertos permanece razoavelmente estável quando são utilizadas 20 ou mais componentes principais em lugar das

variáveis originais. Para uma redução maior, o desempenho cai de forma bastante acentuada, o que confirma as conclusões obtidas previamente pela análise das tabelas 3 e 4.

A seguir são mostrados os resultados dos testes efetuados utilizando 20 componentes principais para a representação dos coeficientes e 6 gaussianas por mistura. São considerados testes com independência e dependência de locutor.

Tabela 7. Resultados dos testes com PCA sobre matriz de correlação utilizando 20 componentes principais e 6 gaussianas por mistura.

testes	D (%)	S (%)	I (%)	total (%)	tempo (min.)
Dep.	1,37	2,44	0,68	4,49	02:16
Masc.	4,72	9,13	1,29	15,44	02:38
Fem.	3,58	14,15	2,97	20,70	02:51
Indep.	5,14	13,43	2,40	20,97	02:41

Tabela 8. Resultados dos testes com PCA sobre matriz de covariância, utilizando 20 componentes principais e 6 gaussianas por mistura.

testes	D (%)	S (%)	I (%)	total (%)	tempo (min.)
Dep.	1,90	3,04	1,41	6,09	03:14
Masc.	4,72	8,83	1,21	14,76	02:30
Fem.	3,12	15,37	3,73	22,22	02:42
Indep.	3,81	13,55	1,86	19,22	02:43

Verifica-se das tabelas acima que o desempenho é semelhante àquele obtido antes de se efetuar a PCA.

Tomando-se agora o número de componentes principais que resultou o melhor desempenho em termos de taxa de acertos, tem-se os resultados apresentados nas Tabelas 9 e 10.

Tabela 9. Resultados dos testes com PCA sobre matriz de correlação.

testes	núm comp.	D (%)	S (%)	I (%)	total (%)	tempo (min.)
Dep.	20	1,37	2,44	0,68	4,49	02:16
Masc.	30	3,27	6,77	1,52	11,57	02:37
Fem.	30	2,97	10,88	2,89	16,74	02:53
Indep.	30	3,65	11,34	2,70	17,69	02:42

## 5. ANÁLISE DOS RESULTADOS

O primeiro fato a ser notado é que a combinação dos parâmetros mel-cepstrais, delta mel-cepstrais e delta-delta mel-cepstrais em um único vetor produziu um aumento no desempenho para os testes com dependência de locutor e

com locutores masculinos, tendo efeito oposto para os testes com independência de locutor e para locutores femininos. Como a base de dados utilizada é relativamente pequena e os resultados foram bastante próximos (exceto no caso de locutores femininos), o agrupamento dos parâmetros em um único vetor parece não influenciar de forma significativa o desempenho do sistema em termos de taxa de acertos de palavras.

Tabela 10. Resultados dos testes com PCA sobre matriz de covariância.

Testes	número comp.	D (%)	S (%)	I (%)	total (%)	tempo (min.)
Dep.	30	1,75	2,66	0,76	5,18	02:15
Masc.	30	2,89	8,22	1,83	12,94	02:36
Fem.	30	3,80	11,64	3,20	18,65	02:52
Indep.	30	3,81	13,55	1,86	19,22	02:43

Em relação à análise de componente principal, verifica-se que a taxa de acertos é mais ou menos estável para os testes com 20 ou mais componentes principais. Os testes utilizando menos componentes resultaram em uma queda grande na taxa de acertos de palavras.

Os testes utilizando a matriz de correlação apresentaram resultados ligeiramente superiores àqueles utilizando a matriz de covariância. A grosso modo, entretanto, pode-se dizer que os desempenhos foram equivalentes.

Comparando-se os resultados da Tabela 1 com os das Tabelas 9 e 10, verifica-se que em quase todos os casos conseguiu-se até uma melhora na taxa de acertos do sistema (exceto para os testes com locutores femininos), além da redução da dimensão do espaço de busca, embora com uma taxa de compressão menor para os testes com locutores masculinos e independente de locutor (30 componentes principais).

Embora se tenha conseguido uma redução substancial na dimensão dos vetores de parâmetros, não foi conseguida uma redução no tempo de processamento. Isto parece indicar que existem outros fatores que apresentam uma influência maior no tempo de processamento do que o cálculo das gaussianas (por exemplo o tamanho do espaço de busca). De qualquer forma houve uma diminuição do espaço de memória que deve ser alocado para o armazenamento das gaussianas.

## 6. CONCLUSÕES

Neste trabalho foi feito um estudo teórico sobre a técnica de análise de componente principal, com a sua implementação prática sobre os vetores de parâmetros em um sistema de reconhecimento de fala contínua de vocabulário médio baseado em modelos ocultos de Markov contínuos.

Os testes realizados mostraram que é possível com esta técnica uma redução de aproximadamente 30 % na dimensão dos vetores de parâmetros sem uma queda significativa na taxa de acertos de palavras do sistema.

Na maioria dos casos conseguiu-se, além de uma redução na dimensão dos vetores de parâmetros, um aumento na taxa de acertos do sistema, o que parece ser um resultado realmente surpreendente.

De certa forma, ao se descorrelacionar as componentes dos vetores de parâmetros, a análise de componente principal propicia uma descrição mais rica do espaço amostral, ou seja, fornece mais informação com menos dimensões, o que poderia explicar a melhora de desempenho do sistema.

O tempo de processamento não caiu como era de se esperar, o que leva à conclusão de que existem outros fatores com uma influência maior no tempo de processamento (por exemplo o tamanho do espaço de busca). Houve, porém, uma diminuição substancial na quantidade de memória necessária ao armazenamento dos parâmetros das gaussianas.

É possível, entretanto, que tenha havido uma ligeira diminuição no tempo de processamento, mascarada pelo modo de operação do sistema operacional, pois o Windows é um sistema multitarefas e, neste caso, nem todo o tempo de processamento da CPU é utilizado apenas para o sistema de reconhecimento.

**Agradecimentos:** Este trabalho foi parcialmente financiado pela FAPESP, através de Bolsa de Pós-Doutoramento, processo 99/01241-2.

## 7. REFERÊNCIAS

- [1] Alcain, A., Solewicz, J. A., Moraes, J. A. "Frequência de ocorrência dos fonos e lista de frases foneticamente balanceadas no português falado no Rio de Janeiro". *Revista da Sociedade Brasileira de Telecomunicações*, 7(1):23-41, 1992.
- [2] Davis, S.B. and Mermelstein, P. "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences". *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASP-28(4):357-366, August, 1980.
- [3] Ney, H. "The use of a one-stage dynamic programming algorithm for connected word recognition". *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-32(2), April, 1984.
- [4] Johnson, R. and Wichern, D. W. *Applied Multivariate Statistical Analysis*. Prentice Hall, 1998.
- [5] Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T. *Numerical recipes – the art of scientific computing*. Cambridge University Press, Cambridge, 1987.
- [6] Ynoguty, C. A. e Violaro, F. "Um sistema de reconhecimento de fala contínua baseado em modelos de Markov contínuos". *Anais do XVIII Simpósio Brasileiro de Telecomunicações*, 3 a 6 de setembro de 2000, Gramado, RS.
- [7] Ynoguti, C. A. "Reconhecimento de fala contínua usando modelos ocultos de Markov". *Tese de Doutorado*, UNICAMP, Campinas, 1999.