

## 4

### **Resultados de Simulação e Avaliação das Técnicas Utilizadas para o Reconhecimento de Locutor Independente do Texto em Presença de Ruído**

Neste capítulo serão apresentadas, avaliadas e comparadas as técnicas utilizadas no capítulo 2 e 3 para o reconhecimento de locutor independente do texto .

#### 4.1

##### **Caracterização da Base de Dados**

No presente trabalho, é utilizada a base de dados TIMIT, a qual apresenta um alinhamento ortográfico, fonético, e transcrições das palavras de todas suas frases. Apresenta de 16 bits de resolução e uma taxa de amostragem de 16kHz. As frases encontram-se gravadas em formato wave. O projeto da base de dados foi um esforço conjunto entre o Instituto de Tecnologia de Massachusetts(MIT), SRI International(SRI) e Texas Instruments, Inc.(TI). As frases foram gravados no TI, transcritas no MIT e verificadas e preparadas para a produção de CD-ROM pelo National Institute of Standards and Technology(NIST). A base de dados é formado por 8 diretórios, "dr1" a "dr8". Em cada um desses diretórios há uma serie de subdiretórios para cada locutor, ou seja em cada um dos subdiretórios há 10 locuções foneticamente balanceadas de um locutor específico. Os diretórios foram divididos de acordo com a região em que foram extraídas as locuções nos Estados Unidos de America.

A tabela (4.1) amostra o número de subdiretórios disponíveis por diretório. Os arquivos presentes em cada um destes subdiretórios contém informações de um único locutor, fazendo um total de 630 locutores dos quais, 70% são homens (438) e 30% restantes são mulheres(192). Cada um dos 630 locutores lê um total de 10 frases que tem uma duração aproximada de 3 segundos e que são identificadas de uma forma lógica, ou seja seus nomes iniciam pelas siglas "sa", "sx" ou "si".

Tabela 4.1: Estruturas dos subdiretórios da base TIMIT.

Dialeto	Diretório	Número de Subdiretórios	Região
1	dr1	49	New England
2	dr2	102	Northern
3	dr3	102	North Midland
4	dr4	100	South Midland
5	dr5	98	Southern
6	dr6	46	New York City
7	dr7	100	Western
8	dr8	33	Army Brat

Há sempre 5 arquivos que iniciam por "sx", 2 por "sa" e 3 por "si", fazendo um total de 10 frases por locutor ou subdiretório. Para todos os testes realizados no presente trabalho utilizou-se 8 frases para o treinamento (fazendo-se uma concatenação de todas as frases "sa", "si" e de 3 frases "sx") e 2 frases ("sx") para teste, resultando em um total de 630 locuções de aproximadamente 24 segundos de duração para treinamento e 1260 locuções com 3 segundos de duração aproximada para teste.

## 4.2

### Base de Ruídos Noisex-92

A base Noisex-92 foi desenvolvida pelo "Speech Research Unit (RSRE)" do Reino Unido e o "Institute for Perception (TNO)" de Netherland. Nesta dissertação utilizou-se 4 tipos de ruído da base Noisex-92 com diferente relação sinal ruído (SNR): Ruído branco, ruído pink, ruído de fábrica e ruído babble ou falatório. Todos os ruídos tem uma duração de 235 segundos, e apresentam uma taxa de amostragem de 19.98KHz e 16 bits de resolução.

O ruído branco foi adquirido amostrando um gerador de ruído analógico de alta qualidade. Este ruído apresenta igual energia em todas as frequências do espectro. O ruído pink foi adquirido amostrando um gerador de ruído analógico de alta qualidade (Wandel & Goltermann), o qual apresenta energia igual por 1/3 de oitava. O ruído de fábrica foi adquirido através de amostras de gravação de um microfone 1/2" B&K em fita de audio digital (DAT). Este ruído foi gravado perto de equipamentos que cortam prata e de soldadura elétrica. Por ultimo, o ruído de falatório consiste de 100 pessoas falando numa cantina onde o nível de som durante a gravação foi de 88dBA [16].

O arquivo de voz com ruído foi gerada da seguinte forma: Para cada arquivo de voz que será avaliado um segmento de ruído de comprimento igual ao comprimento do arquivo de voz foi extraída, e somada com o arquivo de voz.

A figura 4.1 mostra a forma de onda dos diversos ruídos utilizados nesta dissertação.

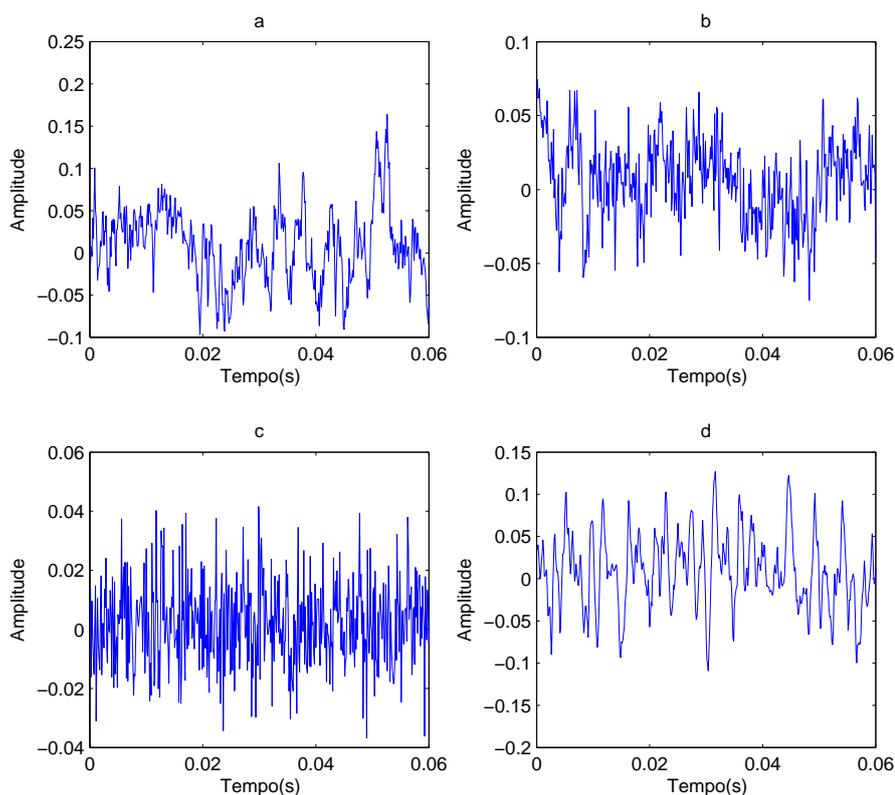


Figura 4.1: Forma de onda para 60 msec de ruído: a) fábrica, b) pink, c) branco, d) falatório

## 4.3

### Avaliação Experimental

#### 4.3.1

##### Otimizando os valores dos parâmetros

Todos os métodos de extração de características dependem de muitos parâmetros livres. Métodos convencionais têm sido avaliados em um grande número de diferentes tarefas para o reconhecimento automático de voz (ASR). Isto resultou no estabelecimento de parâmetros de valores padrão para esses

métodos. Esta seção discute a escolha dos parâmetros livres para os métodos SSCH, MFCC, TECCs, PAC-MFCC e PAC-MFCC com Classificador em Sub-bandas utilizados nesta dissertação para o reconhecimento automático de locutor .

Em cada um dos métodos utilizados, todas as locuções tanto do treinamento como de teste foram re-amostradas a 8000 kHz devido, à limitação da banda telefonica e, em seguida, passadas por um filtro de pré-ênfase de primeira ordem com coeficiente de filtro de 0.95, seguido de um janelamento Hamming. O comprimento do quadro que foi utilizado em cada um dos métodos foi de 16 ms, com 4 ms de incremento de quadro entre quadros sucessivos. Como a frequência de amostragem foi de 8 kHz, obteve-se 128 amostras de voz por quadro, e 32 amostras de incremento de quadro.

### 4.3.2

#### **Histogramas de Centróides de Sub-Bandas Espectrais (SSCH)**

O cálculo das características do SSCH depende da escolha particular do comprimento do quadro, da estimação espectral, do parâmetro espectral de faixa dinâmica, dos parâmetros do banco de filtros e da alocação das raias no histograma.

#### **Estimação do Espectro de Potência:**

Para o cálculo do do espectro de potência foi utilizada a transformada rápida de Fourier(FFT) de ordem 512, e os quadrados da magnitude dos coeficientes resultantes.

#### **Faixa dinâmica espectral:**

A faixa dinâmica do espectro de potência utilizada no cálculo do SSCH é controlada pelo parâmetro  $\gamma$  da equação 2-10. Se  $\gamma$  é muito pequeno (perto de 0), SSCH poderia aproximar-se ao centros de suas subbandas, e assim não conter nenhuma informação. Se  $\gamma$  fosse muito longe (perto de  $\infty$ ), SSCH corresponderia à localização dos valores de pico da subbanda do FFT baseado no espectro de potência, e seria, portanto, estimativas de ruído [9]. A tabela 4.2 ,mostra a característica do desempenho do reconhecimento do método SSCH para diferentes valores do parâmetro  $\gamma$  avaliados na base de dados TIMIT, em sinal limpo e em presença de ruído branco com várias SNR.

Tabela 4.2: Desempenho das características do SSCH para diferentes valores do parâmetro de faixa dinâmica  $\gamma$ . A avaliação foi feita com a base TIMIT tanto para o sinal limpa e para o sinal com ruído de fábrica com diferentes SNR.

$\gamma$	Porcentagem de Acerto[%]			
	Limpo	SNR [dB]		
		20	15	10
0.5	91.83	79.52	59.29	30.16
1	91.59	77.94	59.60	30.87
2	88.77	72.14	55.71	30.63

Para uma alta SNR, o melhor resultado foi obtido utilizando  $\gamma=0.5$  enquanto que o melhor desempenho a uma baixa SNR foi obtido utilizando um valor maior de  $\gamma$ . Estes resultados são razoáveis, já que incrementa o  $\gamma$  faz que a amplitude espectral seja mais eminente, e deste modo reduz o efeito do ruído aditivo.

No resto do trabalho foi utilizado  $\gamma=1$ . Os resultados mostrados na tabela 4.2, foram obtidos utilizando um banco de filtros constituído de 48 filtros com resposta em frequência retangular. Os centros de frequência foram uniformemente distribuídos na escala bark entre 60 e 3800Hz. O histograma é composto por 100 frequências bins uniformemente distribuído na escala bark entre 60Hz e 3800Hz.

#### Projeto do banco de filtros e construção do histograma:

O banco de filtros utilizados para obter as características do SSCH neste estudo consiste de filtros com resposta de frequência retangular e centros de frequências uniformemente distribuídos na escala Bark entre 60 e 3800Hz. A fim de capturar as informações sobre a distribuição da potência da sub-banda, a relação entre largura de faixa do filtro e a largura do bin no histograma deve ser escolhido para ser suficientemente grande.

Para as larguras de banda dos filtros, isso é conseguido através do aumento do número total de frequência bins. Neste trabalho, o histograma bins tem igual longitude na escala bark onde foram utilizados. Uma serie de experimento de reconhecimento foram realizados a fim de otimizar o número de bins no histograma e o número de filtros no banco de filtros. Os resultados são apresentados na tabela 4.3. Observa-se dessa tabela, que quando a largura de faixa BW aumenta, os porcentagem de acerto tendem a diminuir. Além

disso, pode-se observar que o número de filtros não afeta muito no reconhecimento. O número de bins no histograma faz que o reconhecimento melhore quando o número de bins é maior que 60.

No resto deste estudo, as características do SSCH foram calculadas utilizando 48 filtros com uma largura de faixa igual a 0.5 bark. O número de bins no histograma foi estabelecido em 100 devido aos resultados obtidos.

Tabela 4.3: Desempenho das características do SSCH para diferentes valores de numero de filtros e número de bins. A avaliação foi feita com a base de dados TIMIT para locuções limpas, com ruído branco e ruído de fábrica para diferente relação sinal ruído SNR.

Filtros BW [Bark]	Número de Filtros	Número de Bins	Porcentagem de Acerto[%]				
			Limpo	SNR [dB]			
				Fábrica		Branco	
				20	10	20	10
0.5	24	150	93.57	75.63	28.33	62.54	17.06
0.5	36	29	88.81	59.37	15.48	50.24	8.73
0.5	48	38	89.44	67.78	24.84	60.63	16.19
0.5	48	60	90.87	74.76	31.03	64.62	20.48
0.5	48	100	91.59	77.94	30.87	69.29	20.56
1	24	60	86.43	54.60	23.81	55.16	17.86
1	24	150	89.52	56.19	19.13	61.67	19.76
1	36	29	85.40	48.81	12.92	51.03	11.03
1	36	60	87.06	56.75	24.84	61.90	21.43
1	48	38	87.22	53.81	18.81	56.35	14.44
1	48	60	88.65	57.22	17.78	59.60	18.17
2	24	23	75	38.97	11.67	45.79	11.98
2	24	150	84.13	39.29	11.90	44.92	7.86
2	36	60	78.49	46.67	27.46	49.52	19.76
2	48	60	82.62	40.87	14.52	47.14	11.98
2	48	100	77.62	49.19	25.32	50.24	18.33
3	24	23	62.62	35.56	19.76	34.92	10.79
3	24	60	66.27	38.41	21.11	39.05	11.35
3	24	150	67.94	37.38	20.24	38.65	10.56
3	36	60	65.71	39.29	27.46	39.13	9.60
3	48	60	65.95	38.57	12.46	38.02	11.59

A tabela 4.4 mostra uma comparação com as porcentagens obtidas utilizando a técnica SSCH sem derivadas e utilizando a primeira e segunda derivada. Observa-se da tabela, que o melhor porcentagem de acerto foi obtido sem utilizar derivadas; pelo qual no que resta do trabalho para a técnica SSCH não foi utilizado nem a primeira nem a segunda derivada.

Tabela 4.4: Desempenho das características SSCH utilizando a primeira e segunda derivada (delta e delta-delta). A avaliação foi feita com a base de dados TIMIT para locuções limpas, com ruído pink e ruído de fábrica para diferente relação sinal ruído SNR.

Técnica SSCH Utilizada	Porcentagem de Acerto [%]				
	Limp	SNR [dB]			
		Pink		Fábrica	
		20	10	20	10
Sem Deltas	91.59	74.52	27.38	77.94	30.87
Delta	93.49	70.71	18.02	72.78	18.73
Delta-Delta	93.65	69.21	20.87	71.75	23.41

### 4.3.3

#### Coefficientes Cepstrais de Frequência Mel (MFCC)

Para o estudo do MFCC, utilizou-se o algoritmo desenvolvido pelo toolbox voicebox chamado melcepst, do qual amostra-se na tabela os resultados obtidos ao se utilizar diferentes números de filtros, além disso se faz uma comparação das porcentagens obtidas ao se utilizar coeficientes delta e delta delta. Os centros de frequência vão desde 60 até 3800Hz na escala mel.

A continuação a tabela 4.5 mostra a técnica MFCC com diferentes número de filtros e diferentes ruídos com relação sinal ruído de 10 e 20 dB.

Tabela 4.5: Desempenho das características MFCC para diferentes valores de número de filtros. A avaliação foi feita com a base de dados TIMIT para locuções limpas, com ruído pink e ruído de fábrica para diferente relação sinal ruído SNR.

Número de Filtros	Porcentagem de Acerto [%]				
	Limp	SNR [dB]			
		Fábrica		Pink	
		20	10	20	10
17	93.57	79.84	29.60	75.87	21.43
24	92.46	81.83	31.27	79.05	22.22
36	94.05	79.29	29.29	77.22	21.98
40	93.97	77.30	26.67	74.76	21.19

Da tabela 4.5 pode-se observar que o MFCC apresenta um melhor desempenho com um número de filtros igual a 24; pelo qual, no resto deste estudo, as características do MFCC foram calculadas utilizando 24 filtros.

A seguir a tabela 4.6 mostra o desempenho obtido de utilizar diferentes números de coeficientes cepstrais com a técnica MFCC. Dessa tabela pode-se observar que o melhor desempenho é obtido utilizando 20 coeficientes cepstrais.

Tabela 4.6: Comparação entre os distintos número de coeficientes cepstrais utilizados para a obtenção dos porcentagem de acertos da técnica MFCC. A avaliação foi feita com a base de dados TIMIT para locuções limpas, com ruído branco e ruído de fábrica para diferente relação sinal ruído SNR.

Número de Coeficientes Cepstrais	Porcentagem de Acerto [%]				
	Limp	SNR [dB]			
		Branco		Fábrica	
		20	10	20	10
12	91.19	55.48	9.13	72.86	18.57
15	92.46	65	13.33	78.89	25.71
20	92.46	67.38	15.95	81.83	31.27

A tabela 4.7 apresenta uma comparação com as porcentagens obtidas utilizando a técnica MFCC sem derivadas e utilizando a primeira e segunda derivada com um número de filtros igual a 24.

Tabela 4.7: Desempenho das características MFCC utilizando a primeira e segunda derivada (delta e delta-delta) com um número de filtros igual a 24. A avaliação foi feita com a base de dados TIMIT para locuções limpas, com ruído pink e ruído de fábrica para diferente relação sinal ruído SNR.

Técnica MFCC Utilizada	Porcentagem de Acerto [%]				
	Limp	SNR [dB]			
		Pink		Fábrica	
		20	10	20	10
Sem Deltas	92.46	79.05	22.22	81.83	31.27
Delta	93.17	76.59	18.96	78.65	26.19
Delta-Delta	93.97	76.35	20.63	78.84	25.16

Da tabela 4.7 pode-se observar que se obtém um melhor desempenho sem utilizar a primeira e segunda derivada.

#### 4.3.4 Coeficientes Cepstrais da Energia Teager

Para o estudo dos Coeficientes Cepstrais da Energia Teager (TECC), os centros de frequência também vão desde 60 até 3800Hz na escala bark. A tabela 4.8 mostra a técnica dos Coeficientes Cepstrais da Energia Teager com diferentes número de filtros e diferentes ruídos com relação sinal ruído de 10 e 20 dB.

Tabela 4.8: Desempenho das características da técnica TECC para diferentes valores de número de filtros. A avaliação foi feita com a base de dados TIMIT para locuções limpas, com ruído branco e ruído de fábrica para diferente relação sinal ruído SNR.

Número de Filtros	Porcentagem de Acerto [%]				
	Limpo	SNR [dB]			
		Fábrica		Branco	
		20	10	20	10
20	88.25	68.49	22.86	63.10	15.63
27	87.78	71.06	25.30	68.81	24.13
34	89.68	71.27	26.03	70.32	28.41
40	89.05	70.79	25.16	68.41	26.98

Da tabela 4.8 pode-se observar que se obtém um melhor desempenho ao utilizar 34 filtros para o reconhecimento de locutor utilizando a técnica TECC.

A tabela 4.9 faz uma comparação entre o número de coeficientes cepstrais utilizados para a obtenção dos porcentagem de acerto obtidos com a técnica TECC. Desta tabela pode-se observar que o melhor desempenho para a técnica TECC é conseguida utilizando 20 coeficientes cepstrais. Na tabela 4.10 mostra-se uma comparação das porcentagens obtidas ao se utilizar a técnica TECC sem derivadas e utilizando a primeira e segunda derivada (delta e delta-delta) com um número de filtros igual a 34.

Tabela 4.9: Comparação entre os distintos número de coeficientes cepstrais utilizados para a obtenção dos porcentagem de acertos da técnica TECC. A avaliação foi feita com a base de dados TIMIT para locuções limpas, com ruído branco e ruído pink para diferente relação sinal ruído SNR.

Número de Coeficientes Cepstrais	Porcentagem de Acerto [%]				
	Limpo	SNR [dB]			
		Branco		Pink	
		20	10	20	10
12	76.19	54.76	13.97	62.86	31.05
15	79.76	60.79	17.94	68.10	37.06
20	81.83	67.38	24.13	71.90	42.06

Tabela 4.10: Desempenho das características TECC utilizando a primeira e segunda derivada (delta e delta-delta) com um número de filtros igual a 34. A avaliação foi feita com a base de dados TIMIT para locuções limpas, com ruído branco e ruído de fábrica para diferente relação sinal ruído SNR.

Técnica TECCs Utilizada	Porcentagem de Acerto [%]				
	Limpo	SNR [dB]			
		Branco		Fábrica	
		20	10	20	10
Sem Deltas	89.68	70.32	28.41	71.27	26.03
Delta	89.92	68.10	21.03	67.30	18.81
Delta-Delta	90.63	67.62	22.86	68.89	20.63

Da tabela 4.10 observa-se que o melhor desempenho também é obtido sem utilizar a primeira e segunda derivadas.

#### 4.3.5

#### Fase AutoCorrelação com Coeficientes Cepstrais de Frequência Mel (PAC-MFCC)

Para a técnica PAC-MFCC, foram utilizadas o mesmos parâmetros do que o MFCC tanto no número de filtros (24 filtros) assim como nos centros de frequências na escala mel (de 60Hz a 3800 Hz). A tabela 4.11 faz uma comparação entre o número de coeficientes cepstrais utilizados para a obtenção dos porcentagem de acerto obtidos com a técnica PAC-MFCC.

Tabela 4.11: Comparação entre os distintos número de coeficientes cepstrais utilizados para a obtenção dos porcentagem de acertos da técnica PAC-MFCC. A avaliação foi feita com a base de dados TIMIT para locuções limpas, com ruído branco e ruído pink para diferente relação sinal ruído SNR.

Número de Coeficientes Cepstrais	Porcentagem de Acerto [%]				
	Limpo	SNR [dB]			
		Branco		Pink	
		20	10	20	10
12	76.19	54.76	13.97	62.86	31.05
15	79.76	60.79	17.94	68.10	37.06
20	81.83	67.38	24.13	71.90	42.06

Da tabela 4.11 pode-se observar que o melhor desempenho para a técnica PAC-MFCC é conseguida utilizando 20 coeficientes cepstrais. A seguir a tabela 4.12, faz uma comparação com as porcentagens obtidas de utilizar a técnica PAC-MFCC sem derivadas e utilizando a primeira e segunda derivada. Essa tabela mostra que se obtém um melhor desempenho no reconhecimento sem utilizar nem a primeira nem a segunda derivada (delta e delta-delta). Por esses motivos no resto do trabalho foi utilizado 20 coeficientes cepstrais sem deltas no reconhecimento de locutor utilizando a técnica PAC-MFCC.

Tabela 4.12: Desempenho das características PAC-MFCC utilizando a primeira e segunda derivada (delta e delta-delta). A avaliação foi feita com a base de dados TIMIT para locuções limpas, com ruído branco e ruído pink para diferente relação sinal ruído SNR.

Técnica PAC-MFCC Utilizada	Porcentagem de Acerto [%]				
	Limpo	SNR [dB]			
		Branco		Pink	
		20	10	20	10
Sem Deltas	81.83	67.38	24.13	71.90	42.06
Delta	79.52	57.70	17.14	64.92	32.78
Delta-Delta	79.05	59.21	17.54	66.59	31.75

### 4.3.6

#### PAC-MFCC com Classificador em Sub-bandas

Para o atributo PAC-MFCC com Classificador em Sub-bandas, foram utilizados as mesmas características do atributo PAC-MFCC, para o qual foram adotados como padrão 24 filtros e 20 coeficientes cepestrais sem deltas para cada um dos atributos PAC-MFCC utilizados por sub-banda. A tabela 4.13 faz uma comparação entre o número de sub-bandas (2 e 4 sub-bandas) utilizadas para a obtenção dos porcentagem de acerto obtidos com o atributo PAC-MFCC com Classificador em Sub-bandas para o reconhecimento de locutor independente do texto.

Tabela 4.13: Desempenho das características PAC-MFCC com Classificador em Sub-bandas para 2 e 4 sub-bandas. A avaliação foi feita com a base de dados TIMIT para locuções limpas, com ruído de fábrica e ruído pink para diferente relação sinal ruído SNR.

Número de Sub-bandas Utilizadas	Porcentagem de Acerto [%]				
	Limpo	SNR [dB]			
		Fábrica		Pink	
		20	10	20	10
2	85.55	74.60	44.76	74.44	46.66
4	79.68	64.68	27.70	64.36	29.76

Da tabela 4.13 verifica-se que o atributo PAC-MFCC com Classificador em Sub-bandas para 2 sub-bandas apresenta melhores resultados que utilizando 4 sub-bandas, observa-se também que, tanto para o ruído de fábrica como para o ruído pink o atributo PAC-MFCC com Classificador em Sub-bandas com 2 sub-bandas apresenta um porcentagem de acerto de 10% superior que o atributo PAC-MFCC com Classificador em Sub-bandas com 4 sub-bandas, em 20dB de RSR e 20% superior em 10dB de RSR.

### 4.4

#### Comparação dos porcentagem de acerto obtidos de cada um dos atributos utilizados

Nesta seção se faz uma comparação das diversas técnicas utilizadas para o reconhecimento de locutor independente do texto.

#### 4.4.1

#### Comparação das características do atributo MFCC com o atributo TECCs

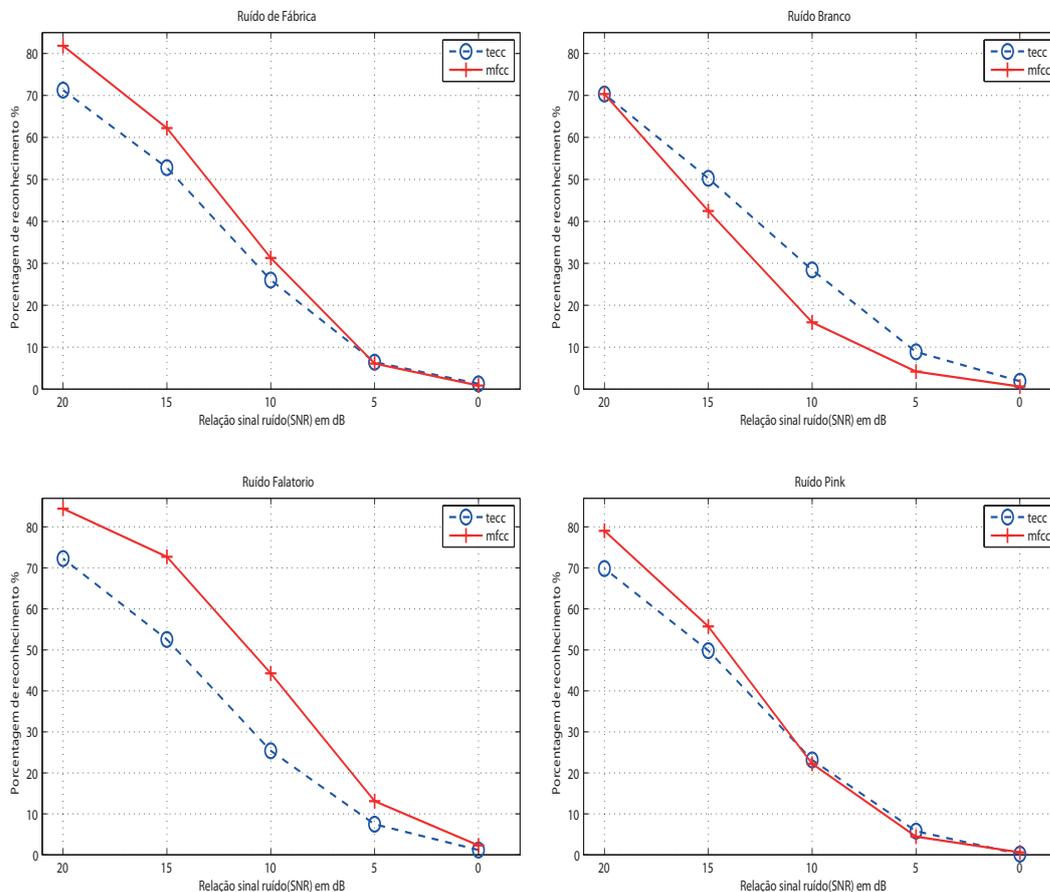


Figura 4.2: Comparação do atributo MFCC com o atributo TECCs em presença de diferentes tipos de ruído e de relação sinal ruído SNR

Da figura 4.2 nota-se que a técnica MFCC apresenta melhor desempenho do que TECC para o ruído de fábrica e ruído de falatório, porém a técnica TECC apresenta um melhor desempenho para ruído branco. No caso do ruído pink, pode-se observar que a técnica MFCC apresenta um melhor desempenho quando se tem uma relação sinal ruído maior a 15dB, além disso observa-se que a técnica TECC apresenta melhor desempenho para ruído pink menor a 10dB, porém desempenho da técnica MFCC é muito próxima.

#### 4.4.2

#### Comparação das características do atributo SSCH com o atributo MFCC

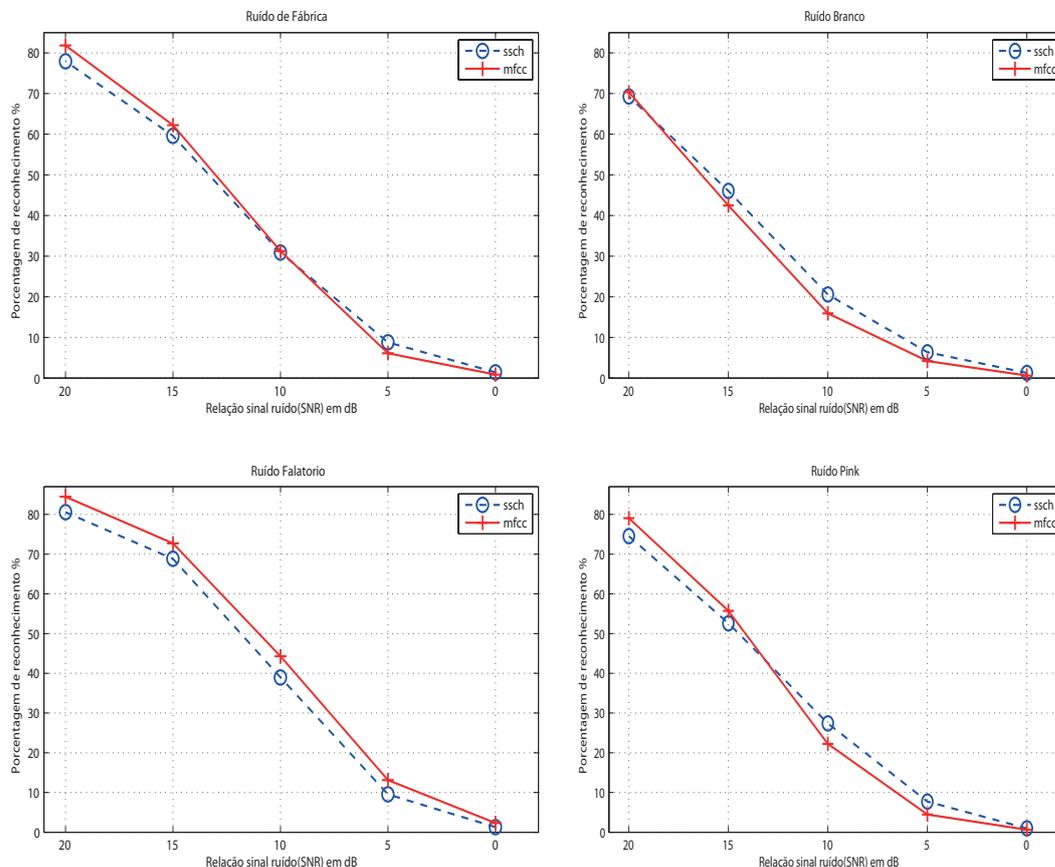


Figura 4.3: Comparação do atributo SSCH com o atributo MFCC em presença de diferentes tipos de ruído e de relação sinal ruído SNR

Da figura 4.3 pode-se observar que a técnica SSCH apresenta melhor desempenho do que MFCC para o ruído de branco quando tem-se uma relação sinal ruído menor a 15dB. No caso dos ruídos pink e de fábrica, a técnica SSCH apresenta um melhor desempenho para relação sinal ruído menor que 10dB. Para o ruído de falatório a técnica MFCC apresentou um melhor desempenho. Pode-se concluir dos gráficos apresentados na fig.4.3 que a técnica MFCC apresenta um melhor desempenho do que a técnica SSCH para os ruídos de fábrica, babble, pink e branco quando se tem uma relação sinal ruído maior que 15dB.

### 4.4.3

#### Comparação das características do atributo SSCH com o atributo TECC

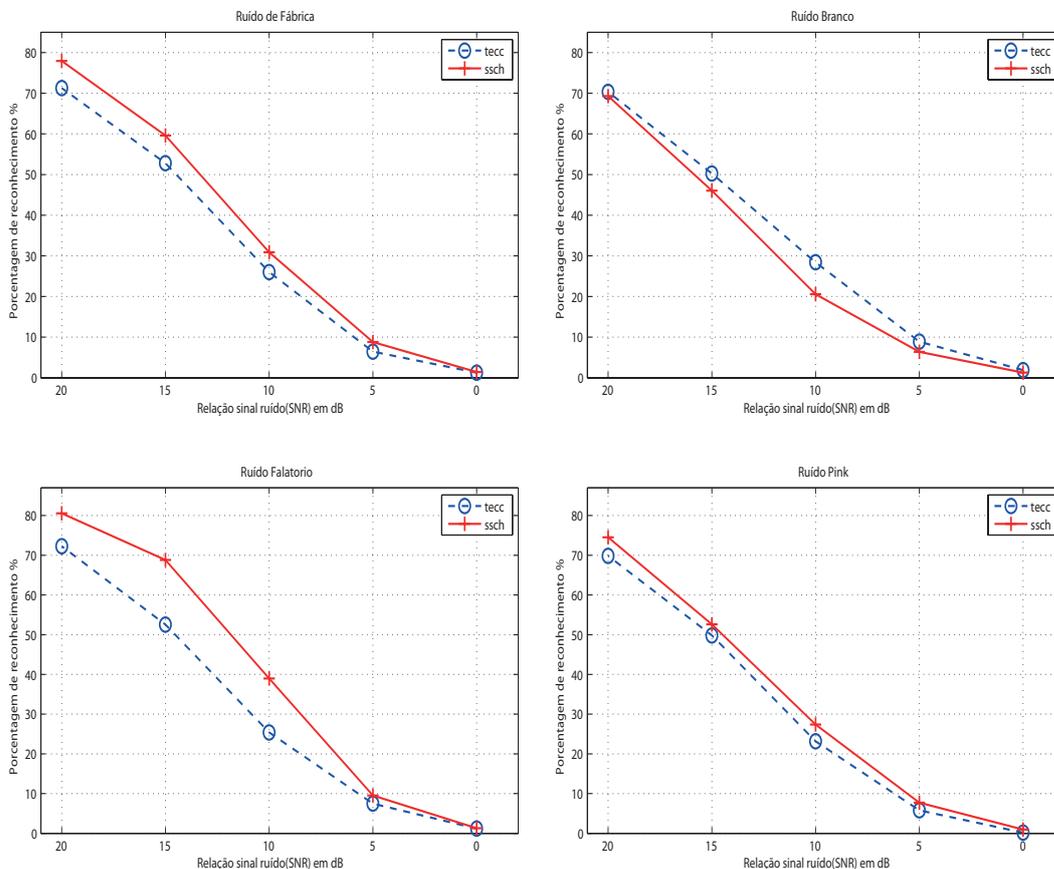


Figura 4.4: Comparação do atributo SSCH com o atributo TECC em presença de diferentes tipos de ruído e de relação sinal ruído SNR

Da figura 4.4 pode-se observar que a técnica SSCH apresenta melhor desempenho do que TECCs para os ruídos de fábrica, de falatório e pink; podendo-se observar que no caso do ruído pink que os resultados obtidos são muito próximos como por exemplo para 5dB de ruído pink obteve-se 7.7% para o SCCH e 5.79% para o TECC. No caso do ruído Branco, a técnica TECC apresentou um melhor desempenho do que SSCH para todos os casos de relação de sinal ruído encontradas.

#### 4.4.4

### Comparação das características do atributo PAC-MFCC com o atributo MFCC

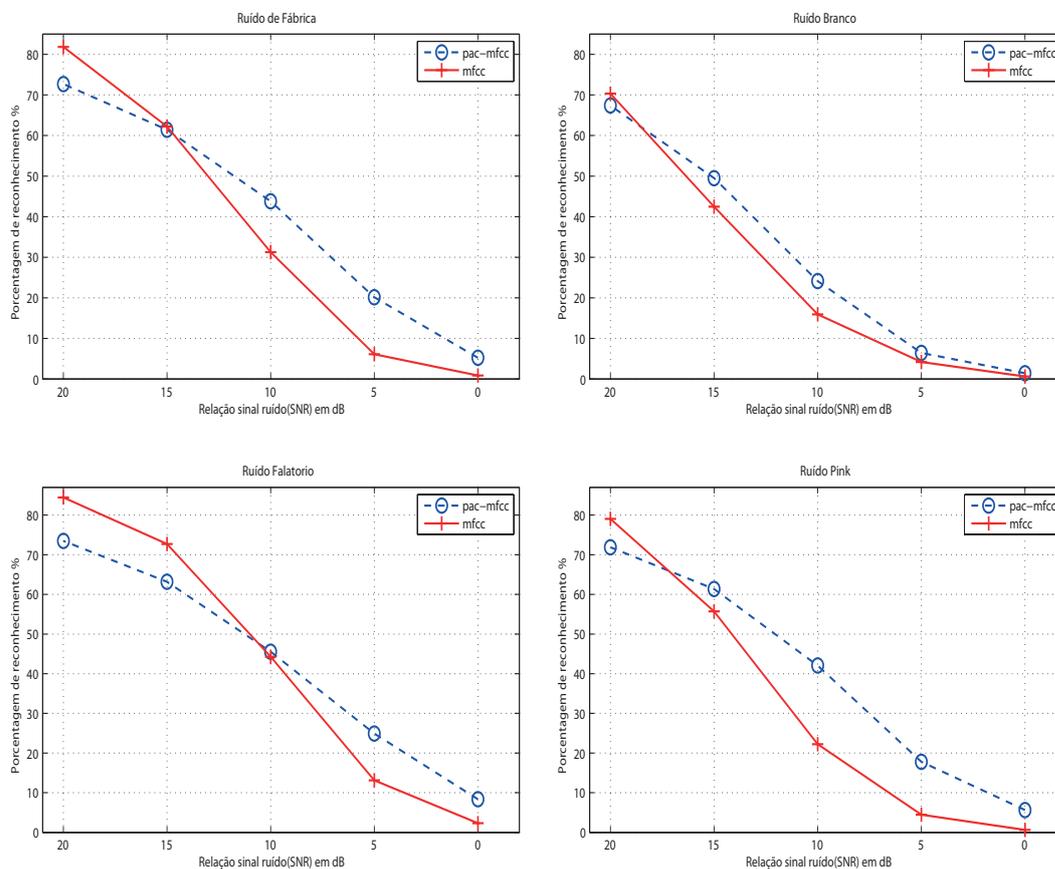


Figura 4.5: Comparação do atributo PAC-MFCC com o atributo MFCC em presença de diferentes tipos de ruído e de relação sinal ruído SNR

A técnica PAC-MFCC é muito melhor que a técnica MFCC para os ruídos branco, de fábrica e de pink em caso de ruído severo (RSN menor que 15dB), observa-se que a porcentagem de acerto no caso de ruído pink para 10db e 5 db a diferença é muito significativa ( obteve-se 42.06% de acerto com o PAC-MFCC e 22.22% para MFCC a 10dB), no caso do ruído de falatório, a técnica PAC-MFCC apresenta um melhor desempenho a partir de uma relação sinal ruído menor a 10dB.

#### 4.4.5

### Comparação das características do atributo PAC-MFCC com o atributo SSCH

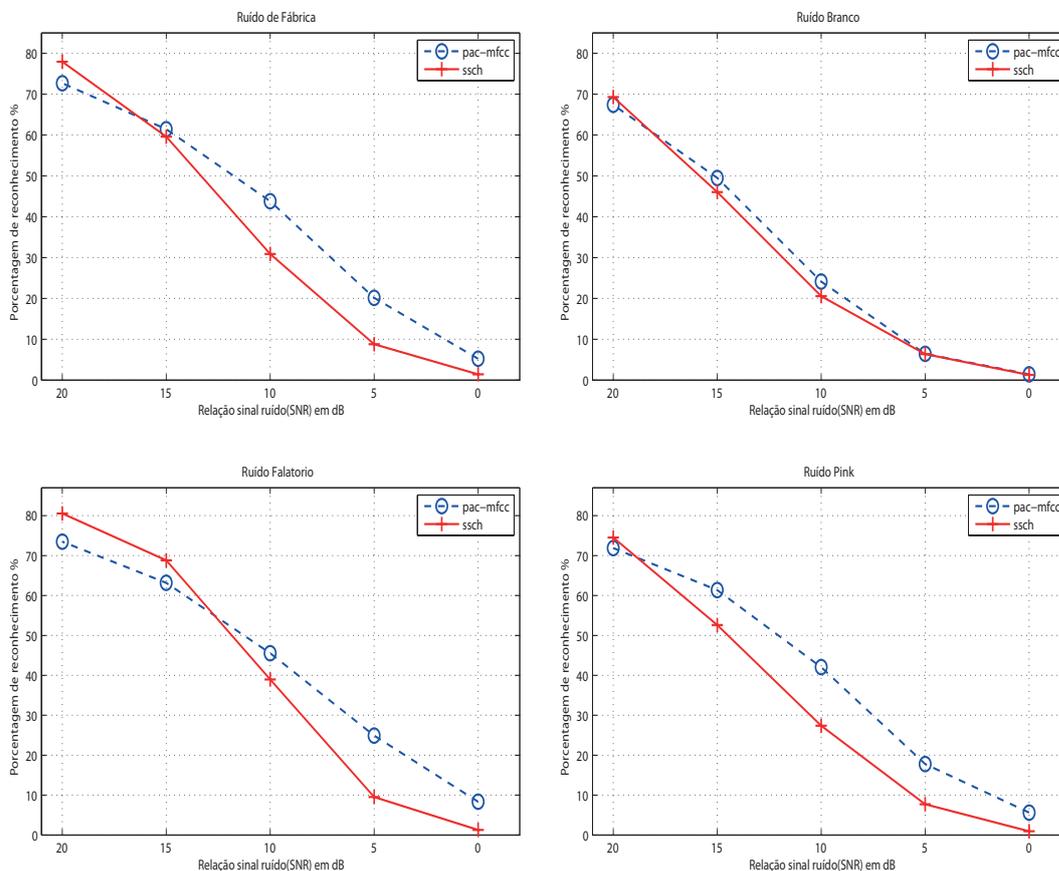


Figura 4.6: Comparação do atributo PAC-MFCC com o atributo SSCH em presença de diferentes tipos de ruído e de relação sinal ruído SNR

Da figura 4.6 verifica-se que a técnica PAC-MFCC apresenta um melhor desempenho que a técnica SSCH para os ruídos de fábrica, e pink quando estes tem uma relação sinal ruído menor que 15dB. No caso do ruído de falatório a técnica PAC-MFCC apresenta um melhor desempenho para uma relação sinal ruído menor que 10dB. Para o ruído branco, o PAC-MFCC apresentou um melhor desempenho que o SSCH, ainda que as porcentagens de acertos foram próximas (obteve-se 24.13% de acerto com o PAC-MFCC e 20.56% para SSCH a 10dB).

#### 4.4.6

### Comparação das características do atributo PAC-MFCC com o atributo TECC

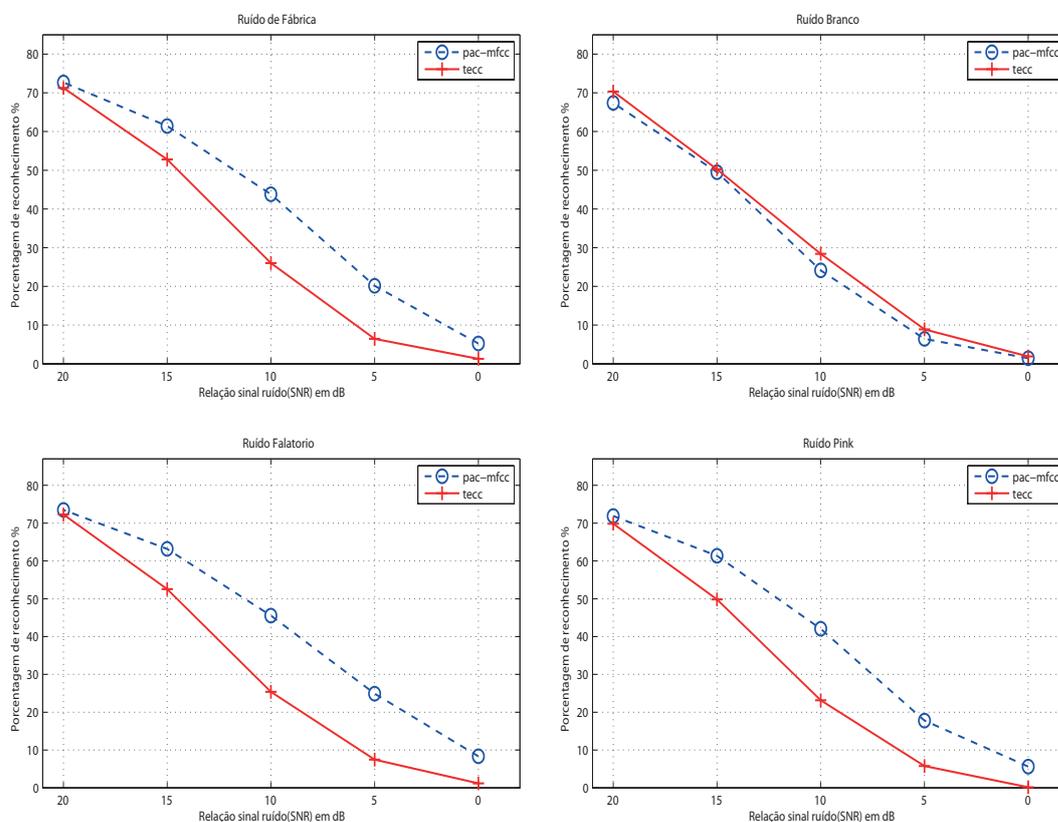


Figura 4.7: Comparação do atributo PAC-MFCC com o atributo TECC em presença de diferentes tipos de ruído e de relação sinal ruído SNR

Da figura 4.7 pode-se observar que a técnica PAC-MFCC apresenta um melhor desempenho que a técnica TECC para os ruídos de fábrica, pink e falatório, pode-se observar também que quando se tem uma relação de sinal ruído menor que 15dB, as porcentagens de acerto obtidas para estes tipos de ruído foram significativamente melhores utilizando a técnica PAC-MFCC (obteve-se 45.56% com ruído de falatório, 43.81% com ruído de fábrica e 42.06% com ruído pink com a técnica PAC-MFCC a 10dB de SNR, enquanto que com TECC obteve-se 25.40% para ruído de falatório, 25.03% para ruído de fábrica e 23.17% para ruído pink). No caso de ruído branco, a técnica TECC apresentou um melhor desempenho do que o PAC-MFCC mas as porcentagens de acertos foram próximas.

#### 4.4.7

### Comparação das características do atributo PAC-MFCC com Classificador em Sub-bandas com o atributo MFCC

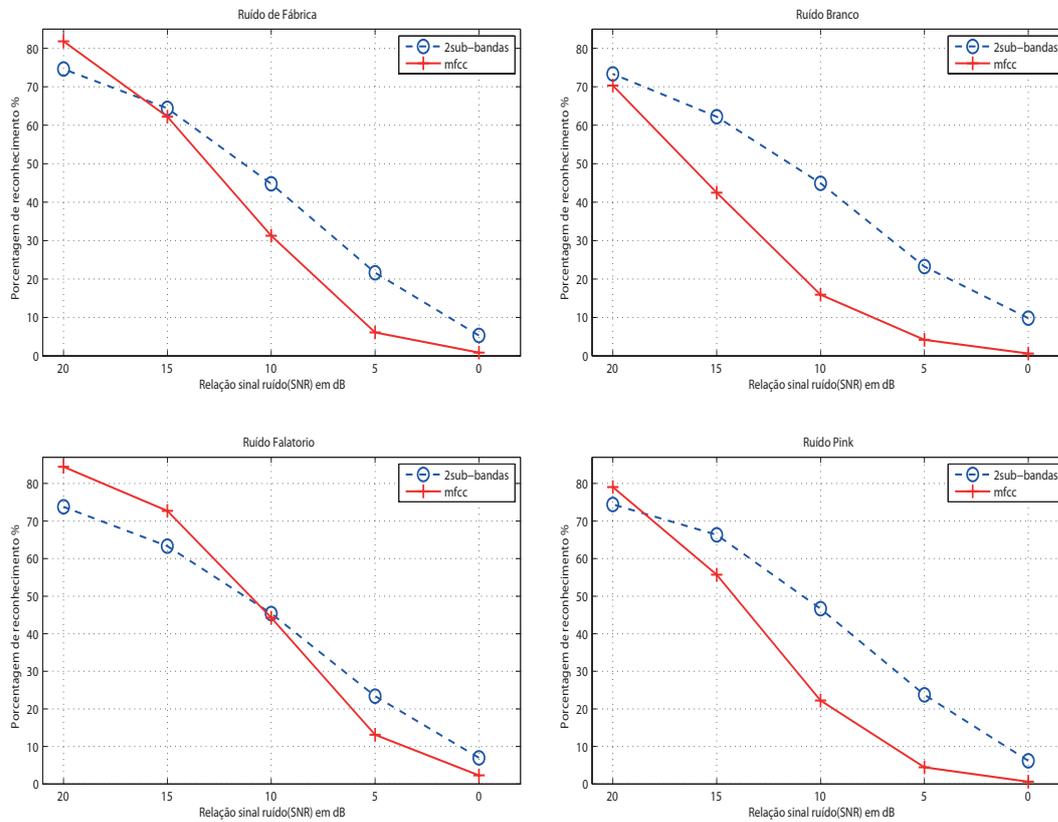


Figura 4.8: Comparação do atributo PAC-MFCC com Classificador em Sub-bandas com o atributo MFCC em presença de diferentes tipos de ruído e de relação sinal ruído SNR

A figura 4.8 mostra que a técnica PAC-MFCC com Classificador em Sub-bandas apresenta um melhor desempenho que a técnica MFCC para os ruídos de fábrica, branco, e pink. No caso do ruído de fábrica e de pink observa-se que os porcentagens de acerto são muitos melhores do que o MFCC para relações sinal ruído menores que 15db. No caso do ruído branco observa-se que as porcentagens obtidos superaram aos obtidos com a técnica MFCC inclusive para uma relação sinal ruído de 20dB. Além disso, as porcentagens obtidas para relações sinal ruído menores que 15db são muito melhores do que a técnica MFCC (obteve-se 44.92% com a técnica PAC-MFCC com Classificador em Sub-bandas enquanto a técnica MFCC obteve 20.56% para uma relação sinal ruído de 10dB). No caso do ruído falatório, a técnica PAC-MFCC com Classificador em Sub-bandas apresenta um melhor desempenho para uma relação sinal ruído menor a 10dB.

#### 4.4.8

### Comparação das características do atributo PAC-MFCC com Classificador em Sub-bandas com o atributo SSCH

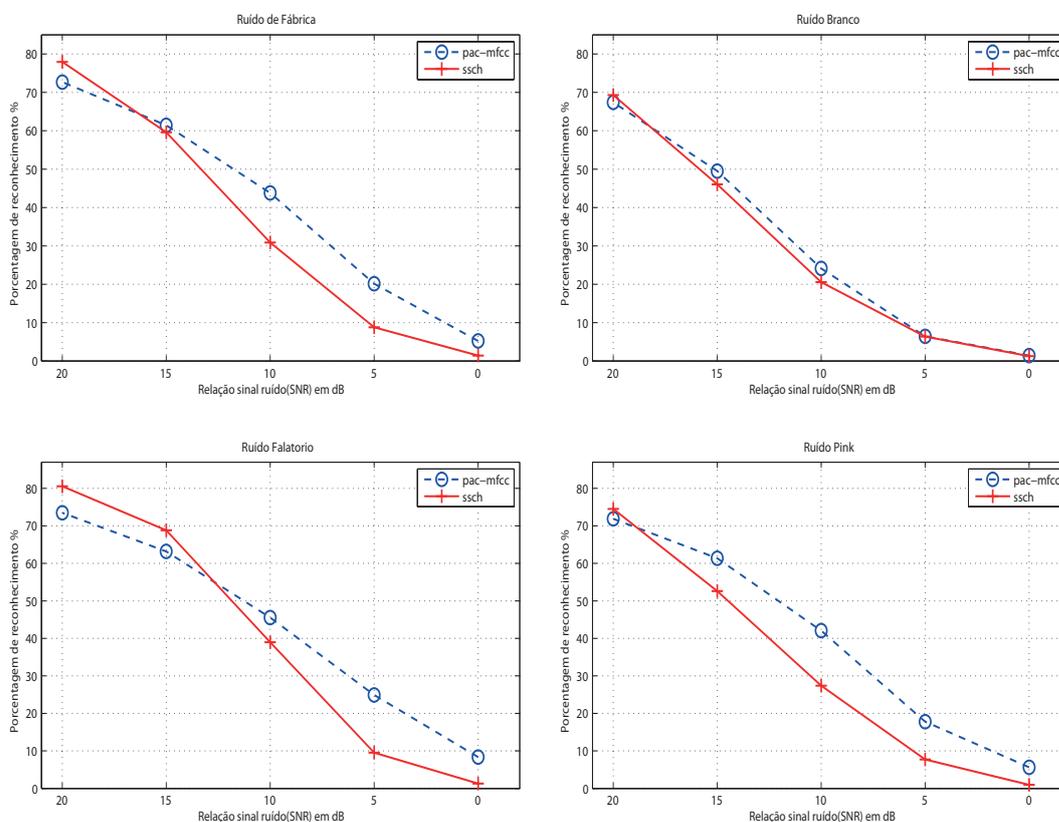


Figura 4.9: Comparação do atributo PAC-MFCC com Classificador em Sub-bandas com o atributo SSCH em presença de diferentes tipos de ruído e de relação sinal ruído SNR

A figura 4.9 mostra que a técnica PAC-MFCC com Classificador em Sub-bandas apresenta um melhor desempenho que a técnica SSCH para os ruídos de fábrica, branco, e pink. No caso do ruído de fábrica e de pink observa-se que as porcentagens de acerto são muito melhores que o atributo SSCH para relações sinal ruído menores que 15dB. No caso do ruído branco verifica-se que as porcentagens obtidas superaram as da técnica SSCH em todos os casos testados e, também, que para uma relação sinal ruído menor que 15 dB as porcentagens obtidas superaram significativamente as porcentagens obtidas com a técnica SSCH. No caso do ruído de falatório, a técnica PAC-MFCC com Classificador em Sub-bandas apresenta um melhor desempenho para uma relação sinal ruído menor que 10dB.

#### 4.4.9

### Comparação das características do atributo PAC-MFCC com Classificador em Sub-bandas com o atributo TECC

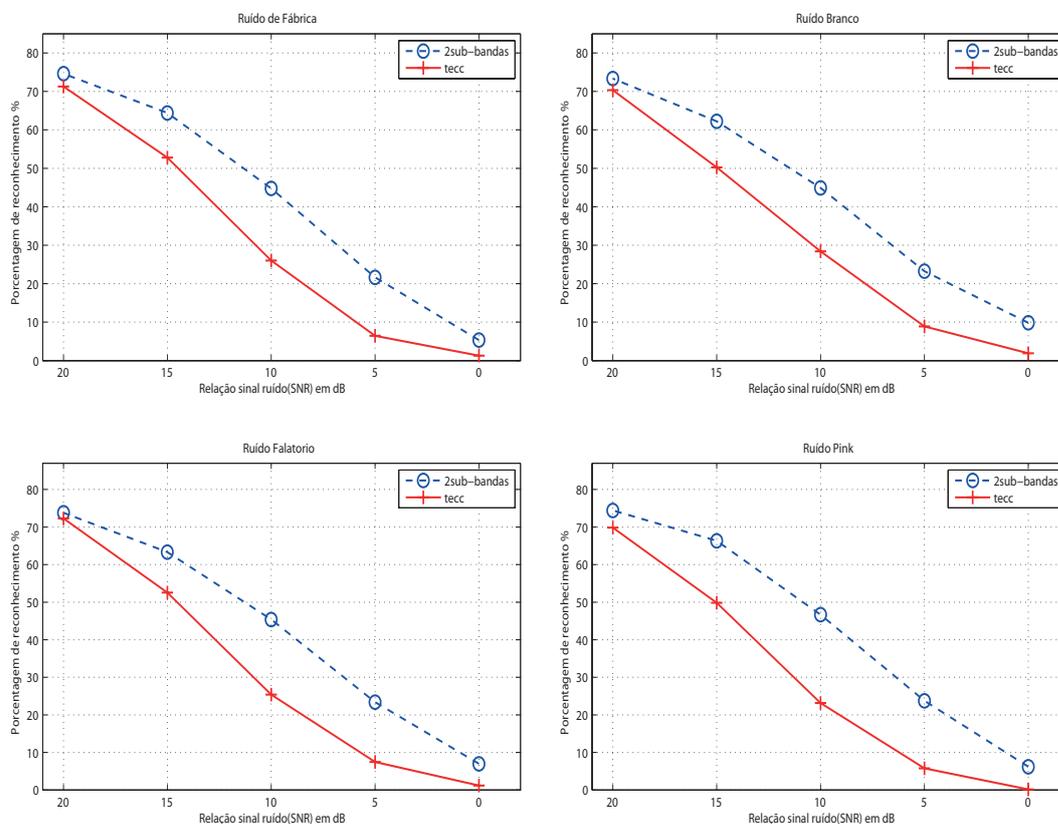


Figura 4.10: Comparação do atributo PAC-MFCC com Classificador em Sub-bandas com o atributo TECC em presença de diferentes tipos de ruído e de relação sinal ruído SNR

Da figura 4.10 pode-se verificar que a técnica PAC-MFCC com Classificador em Sub-bandas apresenta um desempenho significativamente superior à técnica TECC em todos os tipos de ruídos testados. Observa-se também que quando se tem uma relação sinal ruído menor que 15dB, as porcentagens obtidas pela técnica PAC-MFCC com Classificador em Sub-bandas são consideravelmente melhor que os obtidos com a técnica TECC (obteve-se 46.66% de acerto com a técnica PAC-MFCC com Classificador em Sub-bandas enquanto a técnica TECC obteve 23.17% para uma relação sinal ruído de 10dB para ruído pink).

#### 4.4.10

### Comparação das características do atributo PAC-MFCC com Classificador em Sub-bandas com o atributo PAC-MFCC

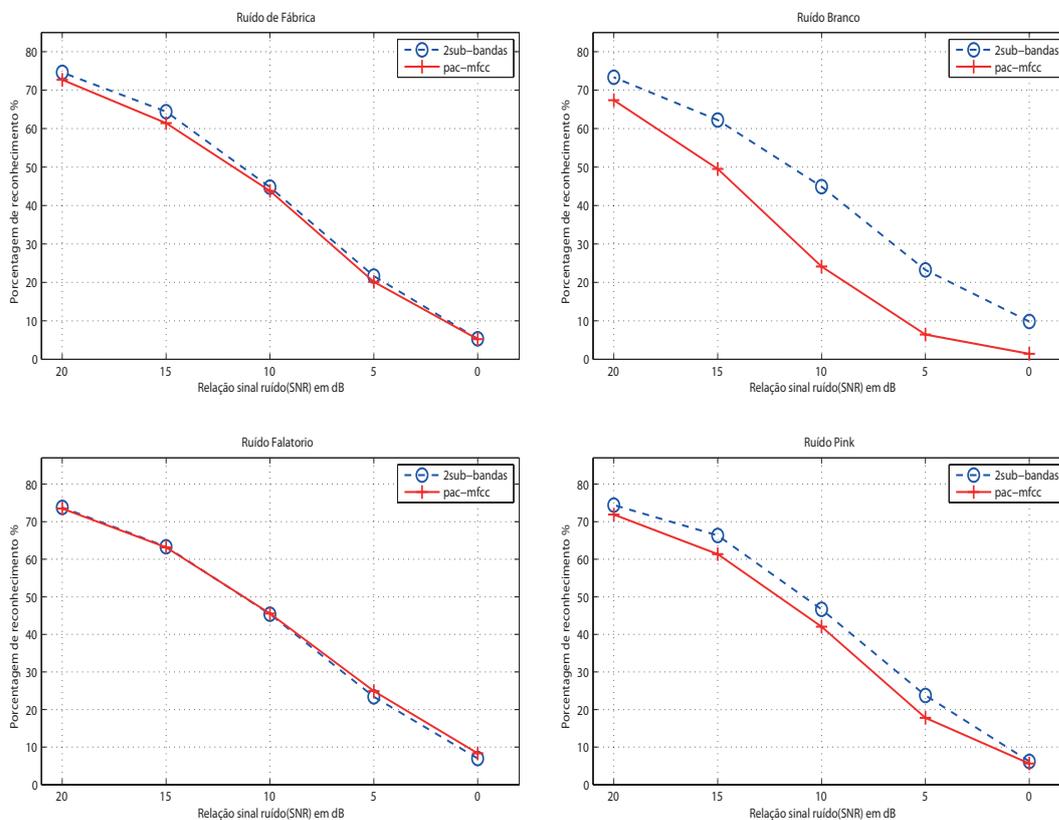


Figura 4.11: Comparação do atributo PAC-MFCC com Classificador em Sub-bandas com o atributo PAC-MFCC em presença de diferentes tipos de ruído e de relação sinal ruído SNR

Da figura 4.11 pode-se observar que a técnica PAC-MFCC com Classificador em Sub-bandas apresenta um melhor desempenho que a técnica PAC-MFCC. Vale salientar, porém, que no caso dos ruídos de fábrica e de falatório, as porcentagens obtidas são muito próximas (obteve-se com a técnica PAC-MFCC com Classificador em sub-bandas 44.76% enquanto que com a técnica PAC-MFCC obteve-se 43.81%, com uma relação sinal ruído de 10dB para ruído de fábrica). No caso do ruído branco, a diferença entre as 2 técnicas é muito significativa quando a relação sinal ruído é menor a 15 dB (obteve-se 23.25% com a técnica PAC-MFCC com Classificador em sub-bandas enquanto que com a técnica PAC-MFCC obteve-se 6.43% de acerto para uma relação sinal ruído de 5dB).

#### 4.4.11

#### Comparação de todas as técnicas utilizadas para os diferentes tipos de ruído

As figuras 4.12, 4.13, 4.14, 4.15 mostram o desempenho das diferentes técnicas utilizadas com os ruídos de fábrica, falatório, branco e pink, nas quais a técnica PAC-MFCC com Classificador em Sub-bandas, em geral, apresentou um melhor desempenho.

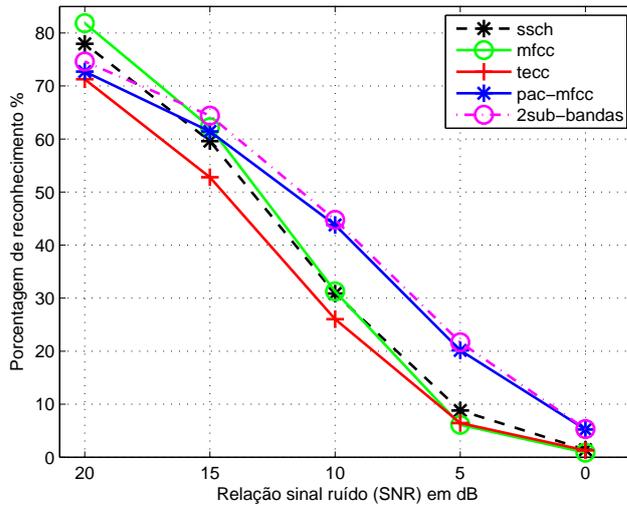


Figura 4.12: Comparação das técnicas MFCC, SSCH, TECC, PAC-MFCC e PAC-MFCC com Classificador em Sub-bandas em ruído de fábrica com diferentes relação sinal ruído SNR

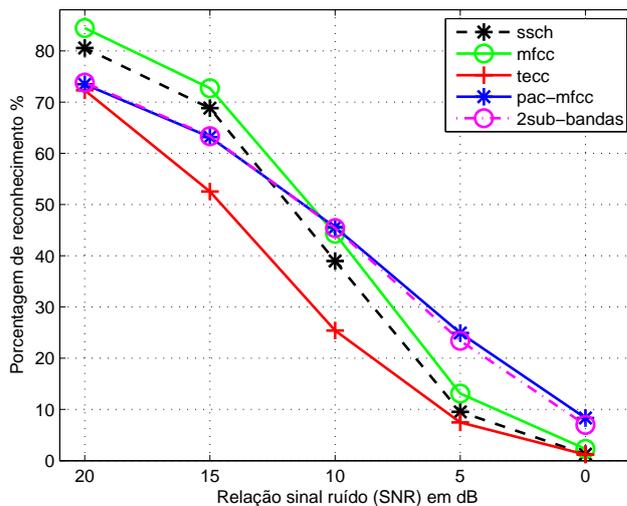


Figura 4.13: Comparação das técnicas MFCC, SSCH, TECC, PAC-MFCC e PAC-MFCC com Classificador em Sub-bandas em ruído falatório com diferente relação sinal ruído SNR

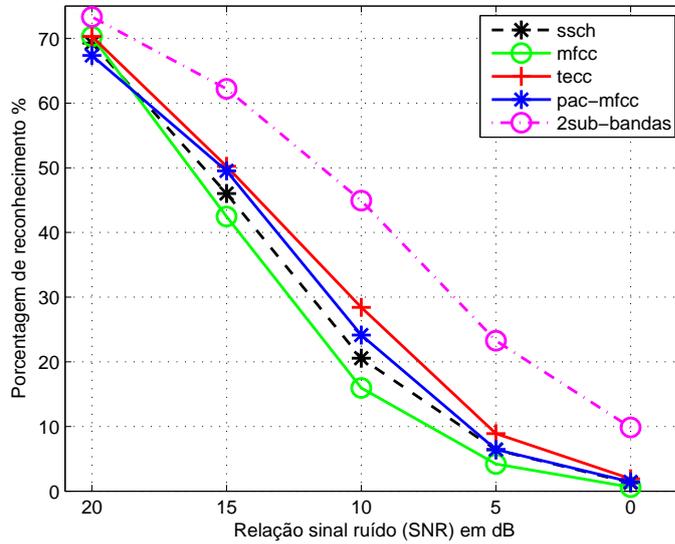


Figura 4.14: Comparação das técnicas MFCC, SSCH, TECC, PAC-MFCC e PAC-MFCC com Classificador em Sub-bandas em ruído de branco com diferentes relação sinal ruído SNR

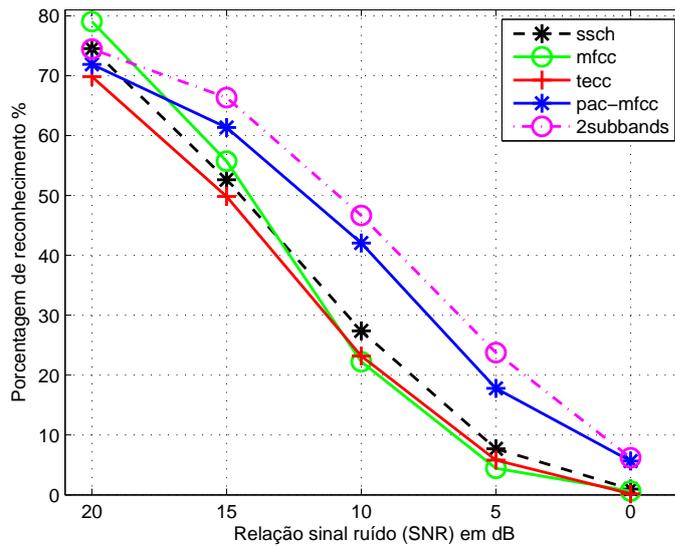


Figura 4.15: Comparação das técnicas MFCC, SSCH, TECC, PAC-MFCC e PAC-MFCC com Classificador em Sub-bandas em ruído pink com diferentes relação sinal ruído SNR