

Detecção e estimação de sinais

Reconhecimento de Padrões Utilizando Filtros Casados

Aline da Rocha Gesualdi Mello, José Manuel de Seixas,
Márcio Portes de Albuquerque, Eugênio Suares Caner,
Marcelo Portes de Albuquerque, Érica Marques da Silva

1. Objetivo:

Reconhecimento de letras e algarismos contidos em imagens de placas de automóveis utilizando filtros casados.

2. Introdução:

Os filtros casados são filtros lineares que têm a característica de apresentar uma saída máxima quando o sinal ao qual o filtro é associado se apresenta na sua entrada. Para qualquer outro sinal de mesma energia, a saída do filtro se apresenta com amplitude inferior ao sinal associado ao filtro. Dessa forma, se diz que o filtro é casado para esse último sinal.

Neste trabalho abordaremos a utilização destes filtros em um sistema de reconhecimento de caracteres contidos em placas de licença de veículos particulares brasileiros. As placas de licença são compostas por sete caracteres alfanuméricos, sendo os três primeiros letras, e os quatro restantes algarismos.

Um banco de dados com 360 imagens de placas foi construído objetivando treinar e testar os filtros casados dos algarismos e letras para a tarefa de reconhecimento. Devemos, a partir deste banco de dados, separar as imagens dos caracteres para a tarefa de reconhecimento.

A partir das imagens individualizadas dos algarismos e letras, foi feito um redimensionamento, reduzindo, desta forma, a dimensionalidade na entrada dos filtros. Cada amostra dos algarismos e letras contidos nas placas dos veículos foram manualmente recortados e redimensionados para o tamanho 5 x 6 pixels.

Nas sessões seguintes, serão apresentados os passos realizados na construção dos filtros. Iremos abordar as diferentes configurações dos padrões a serem casados. Em seguida, será feita uma comparação da performance do reconhecimento realizado pelos filtros casados e com os classificadores neurais artificiais.

3. A construção dos filtros casados

Inicialmente, cada imagem dos caracteres foi transformada em um vetor coluna de dimensão $N=R \times C$, onde R e C são, respectivamente, o número de linhas e colunas da imagem. As imagens foram transformadas em vetores coluna fazendo-se uma varredura linha por linha. Desta forma, uma seqüência de M imagens dão origem a uma matriz \mathbf{X} , como mostra a Equação 1.

$$\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_M], \quad \mathbf{X} \in \mathfrak{R}^{N \times M} \quad (1)$$

Cada coluna da matriz \mathbf{X} (isto é, os vetores \mathbf{x}_i) representa uma imagem de caractere. Estas imagens foram separadas manualmente de acordo com seu padrão e

representadas no conjunto de matrizes \mathbf{X}_{ij} , apresentado na Equação 2, sendo composto pelas matrizes dos caracteres que contêm imagens de um determinado padrão.

$$\mathbf{X}_{ij} = \{\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_9, \mathbf{X}_A, \mathbf{X}_B, \dots, \mathbf{X}_Z\} \quad (2)$$

Cada matriz do conjunto de \mathbf{X}_{ij} foi dividida em duas, sendo chamadas de matriz de treinamento e de teste. As amostras do conjunto de treinamento são utilizadas para extração de atributos, e as de teste para análise de performance dos filtros.

Dois bancos de filtros distintos foram utilizados. Um para o reconhecimento dos algarismos e outro para o reconhecimento das letras. A Figura 1 apresenta de forma esquemática a estrutura dos filtros de reconhecimento.

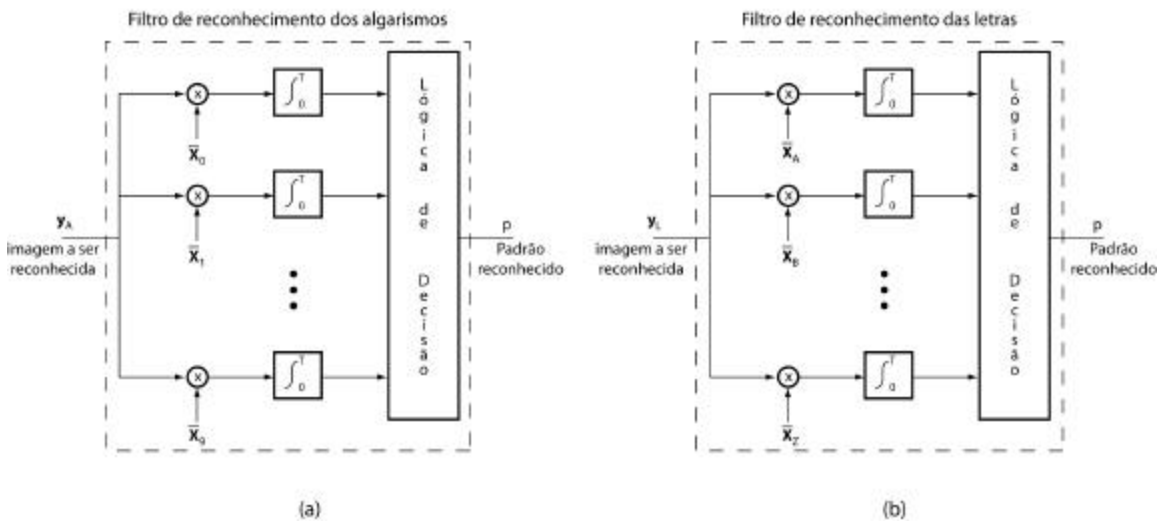


Figura 1 – Representação esquemática dos filtros casados. (a) Filtro casado para reconhecimento dos algarismos. (b) Filtro casado para reconhecimento das letras.

O vetor de entrada y_A é obtido de forma aleatória a partir das matrizes de teste dos algarismos e y_L obtido a partir das matrizes de teste das letras.

Os vetores $\bar{x}_0 \dots \bar{x}_Z$ são os padrões de entrada representativos de cada algarismo e letra, ou seja, são os sinais de associação dos vetores de entrada. Podemos representar estes sinais de associação de diferentes formas, como, por exemplo, assumir que estes são as médias aritméticas dos pixels das imagens da matriz de treinamento [1], conforme ilustra a Figura 2.

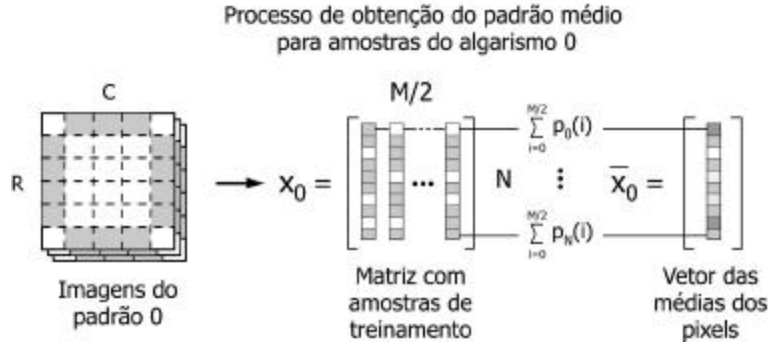


Figura 2 – Exemplo do processo de obtenção dos vetores de média aritmética dos pixels

A lógica de decisão consiste na verificação do maior valor de correlação do vetor de entrada \mathbf{y} com os vetores $\bar{x}_0 \dots \bar{x}_z$, também conhecida como *winner takes all*. Neste trabalho serão abordadas as duas técnicas de construção dos filtros, sendo exemplificados os passos de obtenção.

4. Estudo da influência da adição de ruído no cálculo da eficiência

O primeiro teste realizado, com os filtros casados no reconhecimento das imagens dos caracteres contidos nas placas de licença, foi a caracterização da eficiência de reconhecimento em função do aumento de ruído aplicado nas imagens médias dos padrões.

Inicialmente foi calculada, para as imagens do conjunto de treinamento, a imagem média de caracterização dos caracteres, conforme ilustrado na Figura 2. Na seqüência adicionei ruído branco gaussiano de média zero e desvio padrão variável, a cada uma das imagens médias deste conjunto de treino.

A Figura 3 apresenta os resultados obtidos da eficiência de classificação dos algoritmos (ϵ) em função da variação do desvio padrão (σ) do ruído gaussiano aplicado às imagens médias.

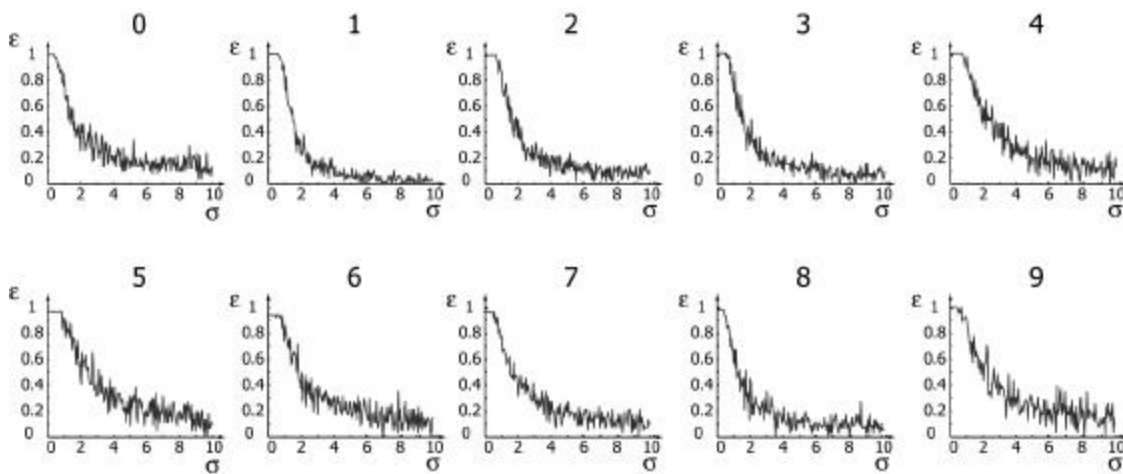


Figura 3 – Gráficos da eficiência de reconhecimento do algarismos em função da variação do desvio padrão do ruído gaussiano.

Pode ser verificado a partir dos gráficos da Figura 3 que todos os padrões possuem comportamento semelhante à adição de ruído. A partir de $\sigma > 2$ a eficiência de classificação é muito baixa, indicando que não há mais o reconhecimento dos caracteres.

5. Filtros casados utilizando imagens das médias dos caracteres

A partir da análise realizada na sessão anterior, foi iniciado o estudo dos filtros casados para as amostras do conjunto de teste tendo como sinal a ser casado as imagens médias dos caracteres no conjunto de treinamento.

Duas arquiteturas de construção foram utilizadas, sendo todas as imagens do conjunto de treinamento usadas no cálculo das imagens médias dos padrões. A diferença das arquiteturas dos filtros dos algarismos e das letras está na separação dos padrões em mais de uma representação, por exemplo, a utilização de mais de um padrão na representação de algarismo 6 (seis).

A primeira arquitetura de filtro não separa os algarismos e letras em diferentes representações. Desta forma, temos para o filtro dos algarismos 10 (dez) padrões a serem casados e, no filtro das letras 26 (vinte e seis) padrões.

Na segunda arquitetura foi utilizada mais de uma representação para os padrões que apresentavam amostras muito diferentes, como, por exemplo, os padrões ilustrados na Figura 4.



Figura 4 – Exemplo de representação em diferentes padrões do algarismo 9 e da letra M.

Para amostras com mais de uma representação foi criado um banco de filtros. A Figura 5 ilustra a arquitetura dos filtros dos algarismos na representação sem separação e com separação de padrões. O mesmo procedimento também foi realizado para o filtro das letras.

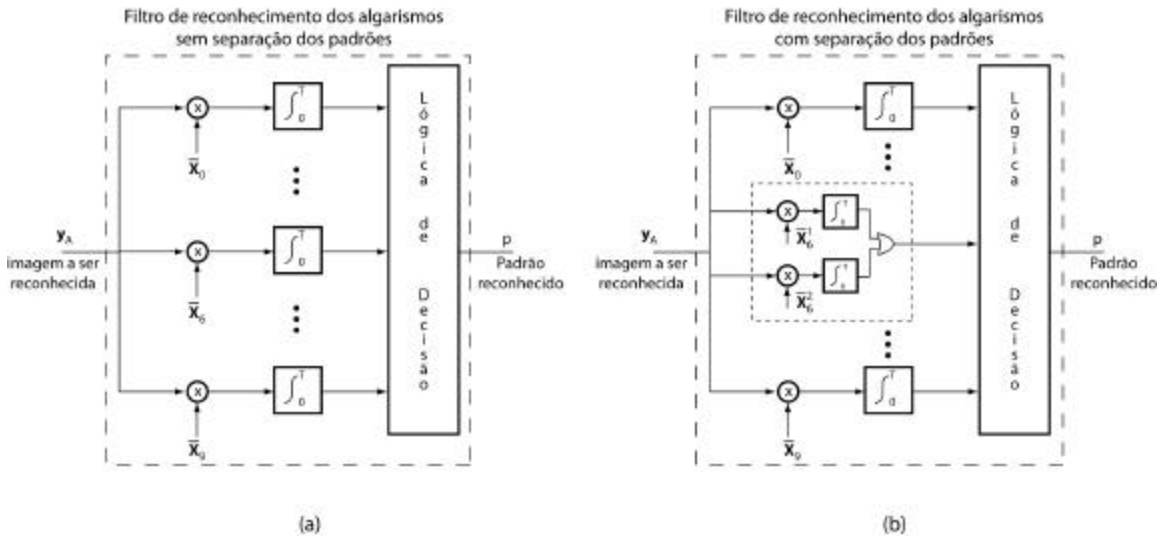
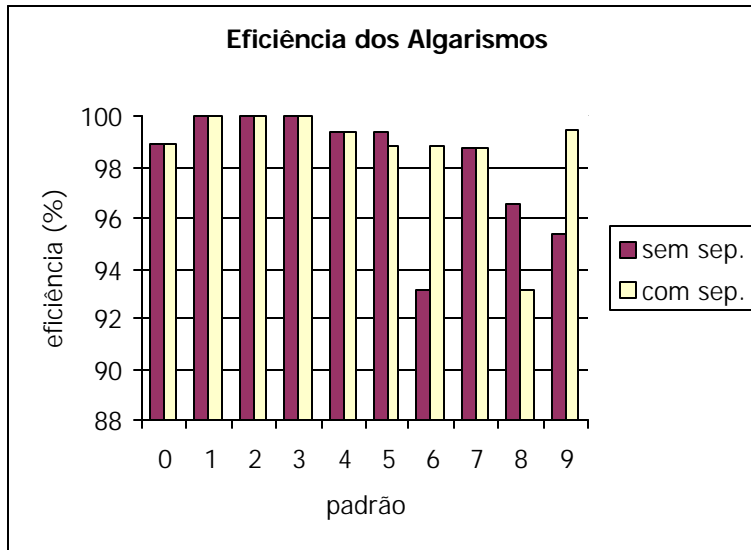


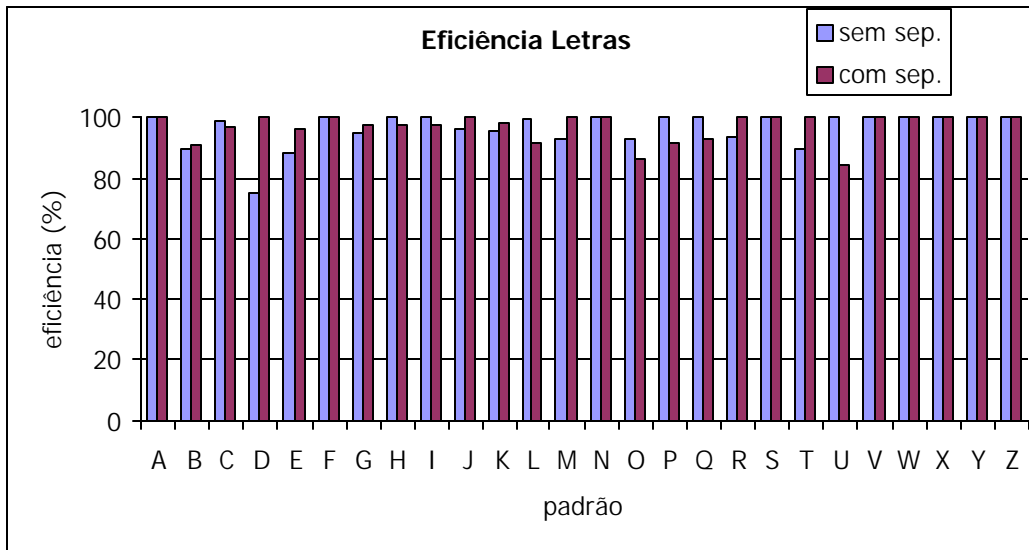
Figura 5 – Arquitetura dos filtros casados. (a) representação sem separação dos padrões e (b) representação com separação (utilização de banco de filtros).

O cálculo da eficiência do conjunto de teste foi feito para as duas arquiteturas de filtro criadas. A Figura 6 mostra os resultados obtidos para os filtros dos algarismos e letras sem separação de padrões e com a separação.

A separação de padrões para os algarismos foi feita para os números 6 e 9, onde foi observado duas formas diferentes de representação no universo de amostras. Um teste de separação de padrões para o algarismo 8 foi realizado, mas ao ser subdividido este prejudicou bastante o reconhecimento do algarismo 0, sendo desta forma descartada sua subdivisão. A separação para as letras foi feita para as classes B, D, K, M e R.



(a)



(b)

Figura 6 – Eficiência do conjunto de teste (a) dos algarismos e (b) das letras.

Analisando os gráficos da Figura 6 observamos que as classes que foram subdivididas obtiveram um aumento na eficiência de classificação, em compensação a subdivisão prejudicou a classificação de algumas outras classes. Na Tabela 1 são apresentadas as eficiências médias para os algarismos e para as letras. Notamos que há uma pequena melhoria na classificação, mas não é significativa.

Eficiência média de classificação		
	Sem separação	Com separação
Algarismos	98,14($\pm 2,37$)%	98,72($\pm 2,05$)%
Letras	96,36($\pm 5,85$)%	96,88($\pm 4,60$)%

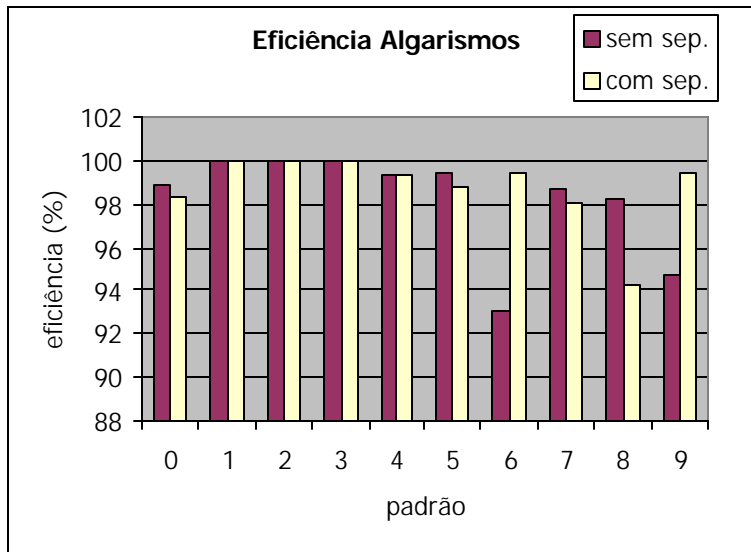
Tabela 1 – Eficiência média na classificação dos caracteres utilizando a imagem média como sinal casado.

6. Banco de filtros utilizando relação de assimetria

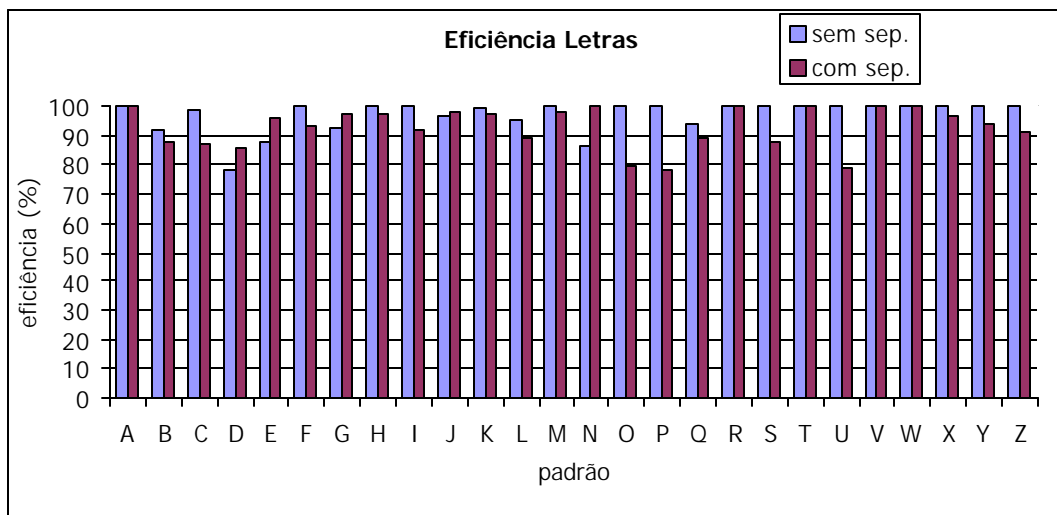
A relação de assimetria dos pixels da imagem do caractere, caracterizada em [2], distingue os pixels como pertencentes ao fundo ou ao objeto na imagem. Com base nesta informação podemos classificar as imagens do banco de dados de acordo com seus diferentes padrões.

Os vetores $\bar{x}_0 \dots \bar{x}_z$ (Figura 1) são compostos pelos valores de RA (relação de assimetria) de cada pixel da imagem. O vetor de entrada \mathbf{y}_A e \mathbf{y}_L , são as imagens dos caracteres, com os valores normalizados para a faixa de (-1,+1).

Os dois filtros, com e sem separação dos padrões (Figura 5), explicados na seção anterior, foram utilizados. A eficiência por caractere, apresentada nos gráficos da Figura 7, representa a comparação da execução destes filtros.



(a)



(b)

Figura 7 – Eficiência do conjunto de teste (a) dos algarismos e (b) das letras.

Os resultados obtidos pelos filtros casados com o método da relação de assimetria e com o método da média das imagens é semelhante. Podemos observar, na Figura 7, que os caracteres que foram separados em mais de uma representação obtiveram eficiências maiores que sem a separação, mas houve ligeiras quedas de eficiência nos demais padrões. A Tabela 2 apresenta a eficiência média dos algarismos e letras obtidos com os filtros casados pela relação de assimetria.

Eficiência média de classificação		
	Sem separação	Com separação
Algarismos	98,26($\pm 2,39$)%	98,77($\pm 1,73$)%
Letras	96,35($\pm 5,60$)%	92,79($\pm 6,69$)%

Tabela 2 – Eficiência média na classificação dos caracteres utilizando a relação de assimetria como sinal casado.

7. Comparação das eficiências dos filtros casados com os classificadores neurais

Em [3] é apresentado um método de reconhecimento dos algarismos e letras em imagens de placas de licença, baseado em classificação neuronal. Uma rede neuronal de 35x13x10 foi utilizada no reconhecimento dos algarismos e outra de 30x18x26 foi utilizada no reconhecimento das letras. A tabela 3 apresenta as eficiências médias para os algarismos e letras obtidas por estes classificadores neurais, bem como dos filtros casados.

	Filtro casado				Redes Neurais
	Média Aritmética		Relação Assimetria		
	sem sep.	com sep.	com sep.	com sep.	
Algarismos	98,14($\pm 2,37$)%	98,72($\pm 2,05$)%	98,26($\pm 2,39$)%	98,77($\pm 1,73$)%	99,19($\pm 0,23$)%
Letras	96,36($\pm 5,85$)%	96,88($\pm 4,60$)%	96,35($\pm 5,60$)%	92,79($\pm 6,69$)%	97,72($\pm 0,78$)%

Tabela 3 – Comparação das eficiências dos filtros casados com os classificadores neurais das letras e algarismos.

Analisando a Tabela 3 observamos que os filtros casados dos algarismos tiveram uma eficiência próxima da eficiência dos classificadores neurais, mas com um desvio padrão bem maior. Observamos também que a eficiência dos classificadores neurais foram superiores na classificação das letras.

8. Conclusão

O objetivo do trabalho foi alcançado. Foram construídos quatro filtros diferentes para o reconhecimento dos algarismos, bem como para o reconhecimento das letras. Pudemos notar que todos estes obtiveram eficiência de classificação superior a 92%. Os filtros casados com a média aritmética das amostras não levam em consideração a estocacidade do processo, enquanto que os filtros casados pela relação de assimetria representam o processo como estocástico. No entanto, não houve uma diferença significativa nas eficiências de ambos os filtros.

Podemos notar também, que os classificadores neurais são capazes de generalizar melhor o processo de reconhecimento que os filtros implementados. Uma possibilidade de melhoria na eficiência de classificação dos filtros é a utilização da expansão das amostras na série de Karhunen-Loeve, onde são extraídas as componentes principais de representação das amostras de treinamento.

9. Bibliografia

- [1] Shanmugan, K. S., Breipohl, A. M., *Random signal: detection, estimation and data analysis*, John Wiley & Sons, 1988.
- [2] Mello, A. G., *Estudo da representatividade estatística de algarismos e letras em imagens de placas de automóveis*, Relatório do 1º trabalho da disciplina Detecção e estimação de sinais, 2001.
- [3] Mello, A. G., *Sistema neuronal de reconhecimento de placas em veículos automotores particulares*, dissertação Ms.C., COPPE/UFRJ & CBPF, agosto, 2001.