

RECONHECIMENTO DE OBJETOS CONTIDOS EM IMAGENS ATRAVÉS DE REDES NEURAIAS

LUIZ EDUARDO NICOLINI DO PATROCÍNIO NUNES

Departamento de Engenharia Mecânica
Universidade de Taubaté

PEDRO PAULO LEITE DO PRADO

Departamento de Engenharia Elétrica
Universidade de Taubaté

RESUMO

Esse trabalho implementou um simulador de sistema de visão computacional, utilizando redes neurais artificiais para a classificação de objetos de formas geométricas (quadrado, círculo, triângulo e retângulo) contidos em imagens bidimensionais. As imagens para os testes foram adquiridas por dois equipamentos de captura: um scanner e uma câmera CCD. A tarefa de reconhecimento de imagens foi dividida em duas etapas: pré-processamento e classificação através de uma rede neural. Na etapa de pré-processamento, as imagens foram processadas por um algoritmo de detecção de bordas, utilizando o método de Sobel, que elimina o fundo contra o qual a imagem se encontra, deixando apenas o contorno do objeto a ser reconhecido. As imagens adquiridas pela câmera CCD sofreram binarização antes de serem processadas pelo algoritmo de detecção de bordas. As imagens binárias resultantes desse algoritmo foram, então, processadas pela Transformada de Fourier Log-polar, para tornar o sistema invariante a efeitos de translação, rotação e escala. Os resultados do algoritmo log-polar foram as entradas para a rede neural. Na etapa de classificação foi utilizada uma rede neural probabilística, por meio do software Matlab. Das imagens pré-processadas, duas amostras foram utilizadas para treinar a rede, e as restantes, para os testes de classificação. A rede classificou corretamente todas as imagens, sendo necessário, no caso das imagens adquiridas pela câmera, aumentar o número de amostras no vetor de treinamento. Os resultados do trabalho demonstraram que as redes neurais podem ser empregadas como ferramentas eficientes na tarefa de reconhecimento de formas geométricas.

PALAVRAS-CHAVE: redes neurais, reconhecimento de padrões, transformada log-polar, método de Sobel, formas geométricas.

INTRODUÇÃO

As redes neurais artificiais têm evoluído intensamente nos últimos anos em decorrência da sua constante aplicação nas mais diversas áreas e atividades. Como exemplo, as redes neurais artificiais têm sido utilizadas em sistemas de controle e otimização, análise de aplicações financeiras, reconhecimento de voz, controle de manipuladores robóticos, classificação, visão

artificial, reconhecimento de padrões em imagens, entre outros (Demuth e Beale, 1996).

Redes neurais artificiais são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes, e que adquirem conhecimento através da experiência. São compostas por diversas unidades computacionais (neurônios artificiais) operando em paralelo, interconectadas total ou parcialmente.

Cada neurônio artificial executa um certo número de operações simples e transmite seus resultados aos neurônios vizinhos, com os quais possui conexão (Perelmuter et al., 1995).

REVISÃO DA LITERATURA

Diversas pesquisas estão direcionadas à solução do problema de reconhecimento e classificação de objetos contidos numa imagem bidimensional.

He e Kundu (1991) usaram uma técnica de classificação baseada no modelo de Cadeias de Markov. Essa técnica não modela todo o padrão ou forma como um simples vetor de características, mas explora a relação entre segmentos consecutivos de um padrão a ser classificado.

Zwicke e Kiss (1983) e Perelmuter et al. (1995) usaram a Transformada de Mellin para tornar o sistema invariante a transformações de rotação, translação e escalonamento.

Moreira e Costa (1996) apresentaram um método para segmentação de imagens em cores que usa a classificação para agrupar pixels em regiões. A cromaticidade é usada como fonte de dados para o método, porque é normalizada e considera somente matiz e saturação, excluindo a componente luminância. A classificação é feita por uma rede neural do tipo Self-Organizing Map (SOM), que é empregada para obter as

cromaticidades principais apresentadas na imagem. Cada pixel é classificado então, de acordo com as classes identificadas. A priori, o número de classes é desconhecido e a rede neural artificial é usada para determinar as classes principais.

Huang e Liu (1997) propuseram um sistema de reconhecimento de padrão híbrido, baseado em um Algoritmo Genético (AG) com uma rede neural modelo Hopfield (HP), que pode reconhecer padrões modificados pela transformação causada pela rotação, escala ou translação. O método proposto se apoia na técnica de aproximação poligonal, que extrai os vetores de características apropriados de dimensões especificadas, caracterizando uma forma.

Eom (1998) apresentou a classificação de formas planas usando características espectrais obtidas por modelamento autoregressivo de uma função de contorno, representada por funções periódicas definidas pelo comprimento e ângulo de um vetor radial da centróide para um ponto de contorno movendo-se em uma velocidade constante.

Nos trabalhos de Loesch (1998) e Davoli et al. (1999) foi proposto um sistema neural híbrido projetado para o reconhecimento de objetos singulares em imagens em tons de cinza com a utilização de uma transformada discreta com

fundamentação matemática na transformação log-polar, que possui propriedades invariantes em relação à translação, rotação e escalonamento.

METODOLOGIA

A visão computacional é efetuada através de um conjunto de transformações, que podem ser em paralelo, permitindo a extração dos aspectos invariantes das imagens, independentemente do ambiente em que estejam situadas. Essas invariâncias possibilitam o reconhecimento da imagem, permitindo a interação do sistema com as mesmas. A visão computacional procura extrair informações contidas numa imagem com o objetivo de classificá-las, caracterizá-las e/ou reconstruí-las (Perelmuter et al., 1995).

O simulador do sistema de visão computacional por redes neurais foi implementado com a utilização do software Matlab, seu Toolbox de Processamento de Imagens, empregado no pré-processamento das imagens, e seu Toolbox de Redes Neurais, composto por uma coleção de funções para projetar e simular redes neurais.

Sistema de aquisição de imagens

O simulador do sistema de visão computacional tem o objetivo de reconhecer,

através de uma rede neural artificial, objetos de formas geométricas diversas, independentemente de suas variações em relação à rotação, translação e escalonamento. As classes de objetos utilizadas nesse trabalho foram o quadrado, o círculo, o triângulo e o retângulo. Para a realização de testes foram criadas duas redes neurais, sendo que, para cada uma, as imagens foram obtidas utilizando-se dois equipamentos distintos de captura: uma câmera CCD e um *scanner*. Tanto as imagens da câmera CCD, seis de cada classe, quanto às imagens do *scanner*, quarenta de cada classe, foram adquiridas com resolução de 256 x 256 *pixels* e em 256 níveis de cinza. O motivo da criação de redes neurais distintas para a realização dos testes deve-se ao fato das imagens adquiridas pela câmera apresentarem distorções em relação às imagens adquiridas pelo *scanner*, conforme ilustra a Figura 1.

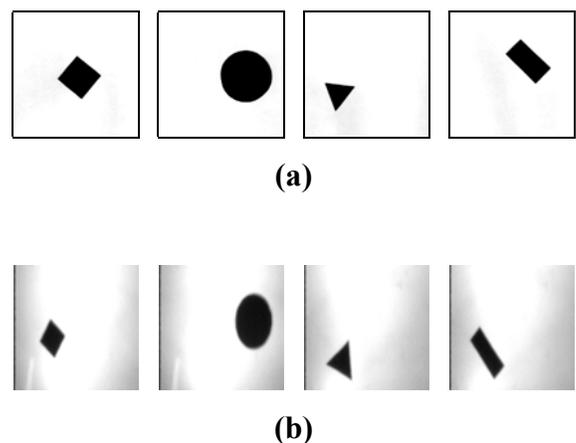


Figura 1. Variação entre as imagens adquiridas: (a) *scanner* e (b) câmera CCD.

Etapas da tarefa de reconhecimento

A tarefa de reconhecimento de objetos em imagens consistiu em duas etapas, conforme ilustra a Figura 2.

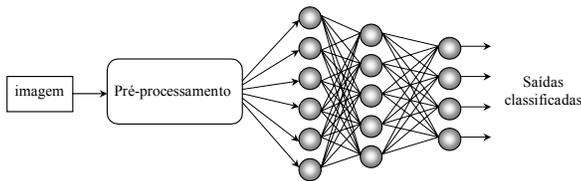


Figura 2. Etapas do reconhecimento.

Pré-processamento

A etapa de pré-processamento foi de grande importância nesse trabalho, pois executou uma série de transformações na imagem de modo a extrair suas características identificadoras, além de reduzir as dimensões dos dados de entrada da rede neural. Tais transformações foram obtidas com a utilização do método de Sobel e da Transformada Log-polar. O método de detecção de bordas (método de Sobel) foi aplicado como uma das fases da etapa do pré-processamento, visto que interessa ao sistema somente a geometria do objeto a ser reconhecido. Informações desnecessárias como, por exemplo, o fundo contra o qual a imagem se encontra, são descartadas. Um sistema de visão computacional busca a classificação correta do objeto independentemente de seu tamanho, posição e rotação. Por essa razão, a transformada Log-

polar também foi aplicada como uma das fases da etapa de pré-processamento. A transformada Log-polar torna o sistema invariante a efeitos de translação, rotação e escalonamento. O esquema ilustrado na Figura 3 descreve as fases da etapa de pré-processamento.

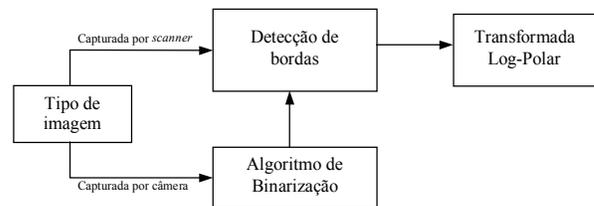


Figura 3. Fases do pré-processamento.

A etapa de pré-processamento foi iniciada com as imagens brutas de 256×256 pixels de resolução e com 256 níveis de cinza. Para eliminar o fundo contra o qual a imagem se encontra, aplicou-se em cada imagem, o algoritmo de detecção de bordas.

No caso das imagens capturadas pela câmera CCD, as imagens foram inicialmente processadas por um algoritmo de binarização para transformá-las em imagens binárias (branco e preto), para posterior aplicação do algoritmo de Sobel.

O algoritmo de binarização calcula um valor de limiar da imagem e compara esse valor com os valores dos pixels da imagem, transformando em branco os valores maiores que o limiar e, em preto, os valores menores.

Esse procedimento foi adotado devido aos resultados não-satisfatórios obtidos com a aplicação direta do algoritmo de detecção de bordas sobre as imagens.

Para terminar a etapa de pré-processamento foi aplicado a todas as imagens o algoritmo que executa a Transformada Log-polar (Loesch, 1998), para tornar o sistema invariante.

A entrada para o algoritmo Log-polar é a imagem binária resultante do processo de detecção de bordas. Antes de processar a transformada da imagem, o algoritmo procura por um ponto inicial que garanta a invariância às diferenças de escala.

Nesse algoritmo, inicialmente são tomados apenas os valores não-nulos (cores não-pretas), colocados num vetor, para serem processados, a fim de evitar o processamento desnecessário com zeros. Nesse caso, o ganho de tempo foi grande, visto que, como estão sendo processados apenas os contornos da imagem, a maior parte da imagem não carrega informação significativa, e aparece em preto.

A invariância em relação à rotação foi obtida pela aplicação da transformada de Fourier nas colunas da matriz resultante da operação log-polar. A imagem resultante da operação log-polar tinha dimensões de 64 linhas por 15 colunas, o que reduziu o tempo de processamento.

Classificação

Na tarefa de reconhecimento das imagens foi utilizada uma rede neural probabilística, que é um tipo de rede empregado em problemas de classificação e aproximação de funções.

Essa rede possui apenas três camadas: camada de entrada, camada escondida e camada de saída, conforme ilustra a Figura 4. A vantagem desse tipo de rede é ter um tempo de treinamento bastante reduzido, além de prescindir da determinação do número de neurônios na camada escondida.

O número de neurônios na camada escondida é definido de acordo com o número de vetores utilizados para o treinamento. Na camada de saída, a quantidade de neurônios é a mesma das classes a serem reconhecidas.

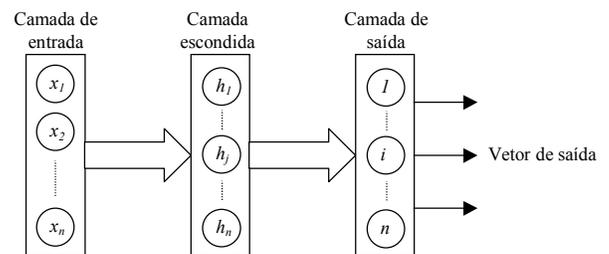


Figura 4. Camadas da rede neural probabilística.

Após a realização do pré-processamento de todas as imagens, os resultados da operação Log-polar foram utilizados como as entradas da rede neural. Inicialmente, esses resultados foram

importados para o Matlab e transformados em vetores coluna.

Para cada classe de objetos (quadrado, circunferência, triângulo e retângulo) foram utilizadas duas amostras para o treinamento da rede e, as outras, para a tarefa de classificação propriamente dita.

A composição do conjunto de treinamento teve como base a observação dos gráficos gerados pela transformada Log-polar: para cada classe foram escolhidas as duas amostras que apresentaram maior variação entre si.

Na fase de aprendizado da rede, foi fornecida a saída desejada (alvo), de acordo com o procedimento de aprendizado supervisionado.

A Tabela 1 mostra as classes de objetos com os respectivos vetores de saída desejada. Na fase de operação, somente as entradas são fornecidas para a rede, que deverá retornar uma classificação coerente.

Tabela 1. Vetores de saída desejada.

Classe	Vetor de saída desejada
Quadrado	[1 0 0 0]
Circunferência	[0 1 0 0]
Triângulo	[0 0 1 0]
Retângulo	[0 0 0 1]

RESULTADOS OBTIDOS PARA AS IMAGENS ADQUIRIDAS POR *SCANNER*

Devido à grande quantidade de imagens adquiridas, 160 no total, foram selecionadas apenas duas amostras de cada classe para a apresentação dos resultados obtidos na etapa de pré-processamento das imagens.

Pré-processamento

A Figura 5 apresenta as imagens capturadas por *scanner* e os resultados obtidos na etapa de pré-processamento, que inclui a detecção de bordas e a transformada log-polar.

Na fase de detecção das bordas, os resultados foram satisfatórios, dispensando a aplicação prévia de um algoritmo de binarização.

As imagens binárias geradas pelo algoritmo de detecção de bordas foram, então, processadas pela Transformada de Fourier Log-polar, para tornar o sistema invariante.

Como pode ser observado nas Figura 5, as imagens geradas pelo algoritmo log-polar são praticamente idênticas para cada classe de objeto.

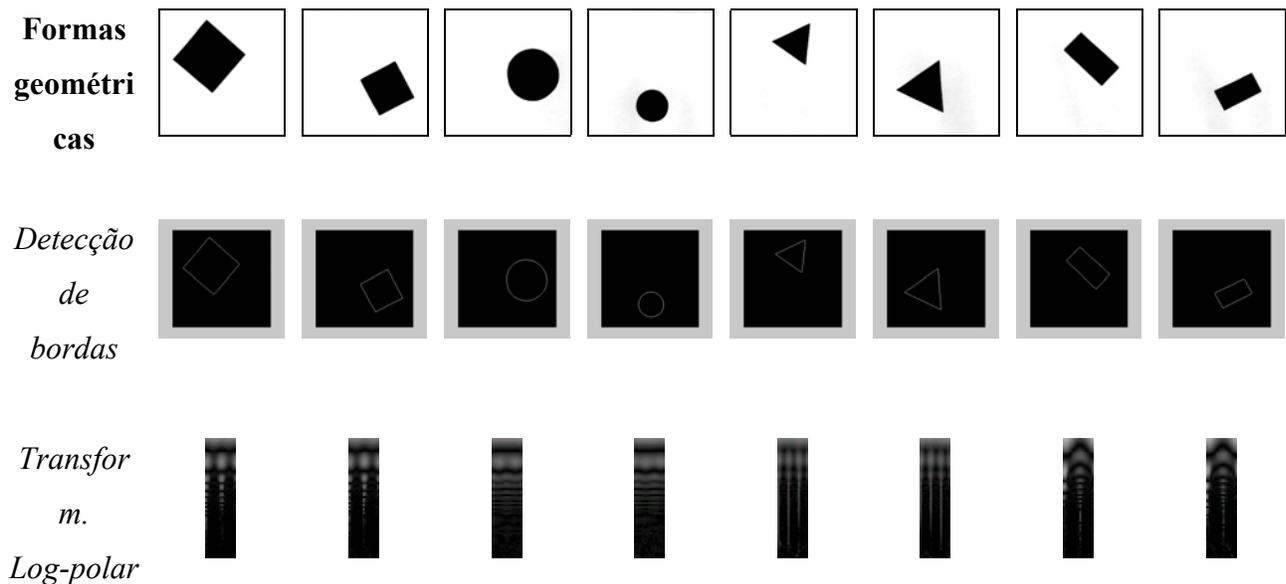


Figura 5. Imagens adquiridas por scanner e resultados do pré-processamento

Classificação

Na etapa de classificação foram utilizadas duas amostras de cada classe para treinar a rede neural e as 38 restantes para os testes. A rede apresentou bons resultados no reconhecimento dos objetos nas imagens adquiridas por *scanner*, classificando corretamente todos os objetos. A taxa de erro da rede foi estabelecida de acordo com o número ou porcentagem de objetos classificados errados. Os quadrados, circunferências, triângulos e retângulos, todos foram corretamente classificados, com porcentagem de acerto de 100%. Devido a esses resultados, não houve a necessidade de aumentar o número de amostras no vetor de treinamento da rede. A Tabela 2 apresenta a porcentagem de

acerto da rede na classificação dos objetos nas imagens.

Tabela 2. Porcentagem de acerto da rede.

Classe	Porcentagem de acerto
Quadrado	100%
Circunferência	100%
Triângulo	100%
Retângulo	100%

RESULTADOS OBTIDOS PARA AS IMAGENS ADQUIRIDAS PELA CÂMERA CCD

Esse trabalho propunha-se inicialmente a realizar testes com figuras capturadas por

scanner, mas também foram realizados testes com algumas imagens adquiridas por uma câmera CCD, para se verificar o comportamento da rede no processo de reconhecimento.

Pré-processamento

A Figura 6 apresenta as imagens capturadas pela câmera CCD e os resultados obtidos na etapa de pré-processamento. Pode-se observar que as bordas detectadas aparecem bem definidas, o que justificou a utilização do algoritmo de binarização antes da detecção de bordas.

As imagens geradas pelo algoritmo log-polar apresentam variações de maior ou menor intensidade conforme a classe de objetos, sendo o caso mais crítico a dos quadrados, que apresentaram três tipos de gráficos distintos entre si.

Tais variações se devem ao fato das imagens adquiridas pela câmera apresentarem distorções na geometria dos objetos, o mesmo não ocorrendo nas imagens adquiridas por scanner

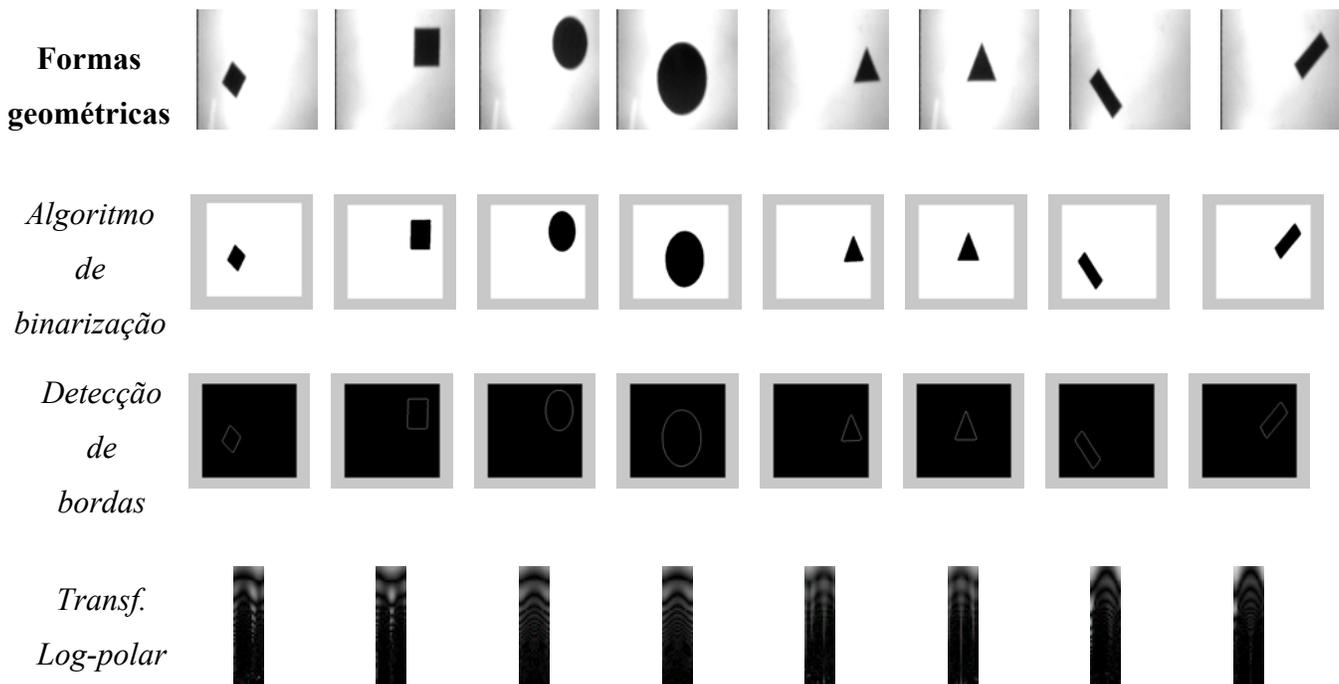


Figura 6. Imagens adquiridas pela câmera CCD e resultados do pré-processamento.

Classificação

Na etapa de classificação, foram utilizadas duas amostras de cada classe para treinar a rede neural, e as quatro restantes para os testes. A rede apresentou resultados razoáveis no reconhecimento dos objetos nas imagens adquiridas pela câmera. Os círculos e os triângulos foram classificados com 100% de acerto, três quadrados foram classificados

corretamente e, apenas dois triângulos obtiveram classificação correta.

Outras duas amostras de cada classe no vetor de treinamento foram utilizadas, a fim de se verificar uma possível melhora no desempenho da rede.

A Tabela 3 apresenta os diversos vetores utilizados no treinamento da rede e as respectivas porcentagens de acerto na classificação dos objetos nas imagens.

Tabela 3. Vetores de treinamento e porcentagem de acerto.

Vetor de Treinamento	Porcentagem de acerto			
	Quadrado	Círculo	Triângulo	Retângulo
1	75%	100%	100%	50%
2	100%	100%	100%	75%
3	100%	100%	75%	25%
4	75%	100%	100%	100%
5	50%	100%	100%	50%

De acordo com a Tabela 3, em todos os vetores de treinamento utilizados, os círculos foram corretamente classificados e os triângulos também obtiveram boa classificação em todos os vetores. Para as imagens utilizadas nesse teste, pode-se observar que a rede treinada com os vetores 2 e 4, apresentou os melhores resultados na tarefa de classificação.

Procurando melhorar o desempenho da rede na tarefa de reconhecimento, ela também foi

treinada com três amostras de cada classe no vetor de treinamento.

As três imagens restantes foram utilizadas na tarefa de classificação. Foram escolhidas para o vetor de treinamento as imagens geradas pelo algoritmo log-polar que apresentaram maiores diferenças entre si. Os resultados dos testes estão apresentados na Tabela 4.

Tabela 4. Resultados da rede treinada com três amostras de cada classe.

Classe	Porcentagem de acerto
Quadrado	100%
Circunferência	100%
Triângulo	100%
Retângulo	100%

De acordo com o observado na Tabela 4, o acréscimo do número de amostras no vetor de treinamento, possibilitou à rede reconhecer todos os objetos, melhorando, assim, o seu desempenho para 100% de acerto.

CONCLUSÕES

O objetivo proposto de implementação do simulador de um sistema de visão computacional em Matlab foi atingido, pois a rede neural utilizada classificou corretamente os objetos contidos nas imagens. O trabalho mostrou a viabilidade da utilização das redes neurais artificiais na tarefa de classificação de objetos em imagens. A importância desse trabalho se deve ao fato de ter apresentado um sistema de visão que pode ser implementado e utilizado em aplicações reais.

O Método de Sobel para a detecção das bordas dos objetos e a Transformada Log-Polar para extrair as variações de rotação, posição e escala foram considerados suficientes e eficientes como ferramentas na etapa de pré-processamento das imagens.

As imagens adquiridas por scanner se mostraram apropriadas para o teste dos algoritmos: tiveram um pré-processamento mais rápido e a escolha das matrizes resultantes da transformação log-polar para treinar a rede neural foi praticamente aleatória, visto que, a variação entre as mesmas era quase imperceptível. Essas imagens apresentaram ótimos resultados tanto na etapa de pré-processamento, quanto na etapa de classificação.

Para as imagens adquiridas pela câmera, foi necessário o emprego, na etapa de pré-processamento, da binarização, devido à distorção ocasionada pela câmera utilizada. As matrizes da transformação log-polar tiveram que ser escolhidas para compor o vetor de treinamento, visando a atingir melhores resultados da rede na tarefa de classificação. Na etapa de classificação, a rede teve também que ser treinada com um número maior de amostras no vetor de treinamento. Com o acréscimo de amostras de cada classe de objeto no vetor de treinamento, a rede foi capaz de reconhecer todos

os objetos, validando perfeitamente o sistema de visão computacional desenvolvido.

.ABSTRACT

This work developed a computational vision system simulator, using artificial neural networks for the geometric shapes classification (square, circle, triangle and rectangle) contained in bidimensional images. The images for the tests were acquired from two capture devices: a scanner and a CCD camera. The images recognition task was divided in two phases: pre-processing and classification through a neural network. In the pre-processing stage, the images were processed by a border detection algorithm, using the Sobel method, that eliminated the image background, just leaving the contour of the object to be recognized. The images acquired by the CCD camera suffered binarization before the processing by the border detection algorithm. The resulting binary images were, then, processed by the Fourier Log-polar transform, to make the system invariant to translation, rotation and scale effects. The results of the log-polar algorithm were the inputs for the neural network. In the classification phase, a probabilistic neural network was adopted, using the Matlab software. Two samples of the processed images were used to train the network, and the remaining ones, for the classification tests. The network classified

correctly all the images. In the case of the images acquired by the camera, the number of samples in the training vector had to be increased. The results of this work demonstrated that the neural networks can be used as efficient tools in the geometric shapes recognition task.

Key-words: neural networks, pattern recognition, log-polar transform, Sobel method, geometric shapes

REFERÊNCIAS BIBLIOGRÁFICAS

- Davoli, R., Tamburini, F. *Data Algorithm: A Numerical Method to Extract Shape Information from Gray Scale Images*. Technical Report UBLCS-93-15, University of Bologna, Italy, 1993.
- Demuth, H., Beale, M. *Neural Network Toolbox User's Guide*. The Mathworks Inc, 1998.
- Eom, K. B. *Shape Recognition Using Spectral Features*. Pattern Recognition Letters, No. 19, pp. 189-195, 1998.
- He, Y., Kundu, A. *2-D Shape Classification using Hidden Markov Model*. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13, No. 11, 1991.
- Huang, J. S., Liu, H. C. *Object Recognition using Genetic Algorithms with a Hopfield's Neural Network*. Expert Systems With Applications, Vol. 13, No. 3, pp. 191-199, 1997.

- Loesch, C. *Reconhecimento de Objetos em Imagens Bidimensionais*. Rev. Ciências Exatas, Taubaté, Vol. 4, No. 1/2, pp. 31-38, 1998.
- Moreira, J., Costa, L. F. *Neural-based Color Image Segmentation and Classification using Self-organizing Maps*. Anais do IX SIBGRAPI, pp. 47-54, 1996.
- Perelmuter, G., Carrera, E. V., Vellasco, M., Pacheco, A. *Reconhecimento de Imagens Bidimensionais utilizando Redes Neurais Artificiais*. Anais do VIII SIBGRAPI, pp. 197-203, 1995.
- Zwicke, P. E., Kiss, I. Jr. *A New Implementation of the Mellin Transform and its Application to Radar Classification of Ships*. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-5, No. 2, 1983.