

Classificação de Imagens Codificadas por Cadeias Direcionais Utilizando Redes Neurais Artificiais

Kleyton C. de Oliveira e Adolfo Bauchspiess
Departamento de Engenharia Elétrica
Universidade de Brasília – Brasília – DF/Brasil
kleyton@ieee.org, bauchspiess@ene.unb.br

Abstract: *This paper presents a procedure for a robust classification of chain-encoded images by using artificial neural networks to recognize sets of features based on normalized Fourier series representation. A particular string-encoding technique and its normalization are described and its results are used in a procedure to determine the best topology of neural networks to be used. The results presented by Radial Basis Function Networks and Multilayer Perceptrons Networks are compared in order to determine the most appropriate architecture for such an application.*

Keywords: *Pattern Classification, Chain-Encoded Images, Fourier Transform, Artificial Neural Networks*

1. Introdução

Uma das áreas que se ocupam de realizar análise de imagens é a chamada visão computacional, a qual tem por finalidade obter, a partir de uma imagem, informações geométricas, topológicas ou físicas sobre o ambiente representado em uma imagem [3].

Quando se trata da classificação de objetos, alguns aspectos, tais como orientação, magnitude, grau de luminosidade ou ruídos introduzidos no sinal de entrada, devem ser considerados. Tais aspectos geram restrições, como, por exemplo, uma quantidade finita de orientações ou faixas de valores de magnitude das imagens permitidas aos objetos sob inspeção. A robustez apresenta-se, assim, como um fator de grande importância a ser implementado no sistema.

O objetivo deste trabalho foi o desenvolvimento de um sistema de classificação de imagens robusto às restrições impostas pelos fatores acima relacionados. A técnica empregada consistiu na obtenção dos coeficientes normalizados das séries de Fourier das imagens codificadas por um método chamado de codificação por cadeias direcionais. Cadeias direcionais dos contornos de objetos permitem transformar a informação da imagem para um domínio unidimensional.

O processo teve continuidade com a apresentação dos coeficientes das séries de Fourier como pares de treinamento para duas arquiteturas de redes neurais artificiais adotadas neste trabalho.

2. Cadeias Direcionais

Uma forma de se efetuar o reconhecimento de imagens é através da comparação de palavras conhecidas como cadeias direcionais. Tais palavras são obtidas a partir da codificação de uma dada figura em um conjunto finito de informações [1].

A codificação por cadeia utiliza-se de um alfabeto A formado por 8 caracteres $A = \{0,1,\dots,7\}$. As infinitas possibilidades de seqüências que podem ser criadas a partir dos caracteres de A geram o espaço A^* . Cada caractere do alfabeto A representa um segmento orientado de reta de comprimento 1 ou $\sqrt{2}$ e de direção múltipla de 45° . O segmento associado a cada caractere a_i é dado por:

$$v_i = \left[1 + \frac{\sqrt{2}-1}{2} (1 - (-1)^{a_i}) \right] \angle \left(\frac{\pi}{4} \right) a_i \quad (1)$$

Uma cadeia direcional c é, então, uma palavra de comprimento m em A^* cujos elementos correspondem a caracteres de A organizados de modo que seus respectivos segmentos orientados estejam conectados de forma a reproduzir, da melhor maneira possível, os contornos de uma dada imagem [1].

A Figura 1 apresenta os segmentos orientados e seus respectivos caracteres, bem como a imagem codificada por $c = 0022746005443$.

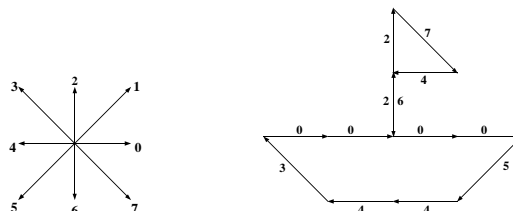


Figura 1 – Segmentos orientados do alfabeto A e imagem associada à palavra $c = 0022746005443$

A imagem é formada da seguinte maneira: a origem do segmento v_i , correspondente a c_i , é adotada como origem das coordenadas. O processo prossegue, então, com a origem de v_i sendo colocada na ponta de v_{i-1} , até que seja posicionado o último segmento, v_m , correspondente a c_m .

Nota-se que a imagem associada a c é uma imagem linear por partes e que existe apenas uma única imagem associada a uma determinada cadeia direcional. No entanto, uma mesma imagem pode ser representada por diversas cadeias direcionais diferentes, dependendo do ponto da imagem em que se quer iniciar a descrição da mesma. Se $c = c_1 c_2 \dots c_K$, for uma palavra em A^* , então a soma dos comprimentos dos segmentos de reta v_1, v_2, \dots, v_i , componentes de v , resulta em um comprimento de arco total, T, dado por:

$$T = \sum_{i=1}^K \left[1 + \frac{\sqrt{2}-1}{2} (1 - (-1)^{c_i}) \right] \quad (2)$$

Admitindo-se que t represente o parâmetro comprimento de arco, pode-se considerar v como uma curva fechada $v(t)$, apresentando componentes horizontais e verticais $x(t)$ e $y(t)$, respectivamente, com $0 \leq t \leq T$. As curvas formadas por $x(t)$ e $y(t)$ com $0 \leq t \leq T$ são chamadas projeções x - y .

Pode-se mostrar que as funções $x(t)$ e $y(t)$ são periódicas, de período igual ao comprimento de arco da imagem, T , contínuas, integráveis e absolutamente integráveis, atendendo, portanto, a todos os requisitos necessários para a sua representação em forma de séries de Fourier.

Com uma quantidade finita de coeficientes de Fourier, ou harmônicos, é possível reconstruir, de maneira satisfatória, as projeções da imagem original v , a qual pode ser constituída de uma quantidade muito grande de segmentos orientados v_i .

As representações das projeções x - y em forma de séries de Fourier são descritas por:

$$\begin{aligned} x(t) &= A_0 + \sum_{n=1}^{\infty} A_n \cos\left(\frac{2n\pi t}{T}\right) + B_n \sin\left(\frac{2n\pi t}{T}\right) \\ y(t) &= C_0 + \sum_{n=1}^{\infty} C_n \cos\left(\frac{2n\pi t}{T}\right) + D_n \sin\left(\frac{2n\pi t}{T}\right) \end{aligned} \quad (3)$$

onde A_n , B_n , C_n e D_n são os coeficientes não normalizados de Fourier.

A natureza contínua e linear por partes de $x(t)$ permite que os coeficientes A_n e B_n sejam calculados pela Equação (4).

$$\begin{aligned} A_n &= \frac{T}{2n^2\pi^2} \sum_{p=1}^K \frac{\Delta x_p}{\Delta t_p} \left[\cos\left(\frac{2n\pi t_p}{T}\right) - \cos\left(\frac{2n\pi t_{p-1}}{T}\right) \right] \\ B_n &= \frac{T}{2n^2\pi^2} \sum_{p=1}^K \frac{\Delta x_p}{\Delta t_p} \left[\sin\left(\frac{2n\pi t_p}{T}\right) - \sin\left(\frac{2n\pi t_{p-1}}{T}\right) \right] \end{aligned} \quad (4)$$

De maneira análoga, calcula-se os coeficientes C_n e D_n , de $y(t)$.

$$\begin{aligned} C_n &= \frac{T}{2n^2\pi^2} \sum_{p=1}^K \frac{\Delta y_p}{\Delta t_p} \left[\cos\left(\frac{2n\pi t_p}{T}\right) - \cos\left(\frac{2n\pi t_{p-1}}{T}\right) \right] \\ D_n &= \frac{T}{2n^2\pi^2} \sum_{p=1}^K \frac{\Delta y_p}{\Delta t_p} \left[\sin\left(\frac{2n\pi t_p}{T}\right) - \sin\left(\frac{2n\pi t_{p-1}}{T}\right) \right] \end{aligned} \quad (5)$$

Os componentes DC A_0 e C_0 são calculados pela Equação (6).

$$\begin{aligned} A_0 &= \frac{1}{T} \sum_{p=1}^K \frac{\Delta x_p}{2\Delta t_p} (t_p^2 - t_{p-1}^2) + \xi_p (t_p - t_{p-1}) \\ C_0 &= \frac{1}{T} \sum_{p=1}^K \frac{\Delta y_p}{2\Delta t_p} (t_p^2 - t_{p-1}^2) + \delta_p (t_p - t_{p-1}) \end{aligned} \quad (6)$$

onde

$$\begin{aligned} \xi_p &= \sum_{j=1}^{p-1} \Delta x_j - \frac{\Delta x_p}{\Delta t_p} \sum_{j=1}^{p-1} \Delta t_j \\ \delta_p &= \sum_{j=1}^{p-1} \Delta y_j - \frac{\Delta y_p}{\Delta t_p} \sum_{j=1}^{p-1} \Delta t_j \end{aligned} \quad (7)$$

e $\xi_j = \delta_j = 0$.

Adota-se, para a representação por séries de Fourier, um erro máximo admissível, ε_s , e calcula-se o número de harmônicos, N , necessários para que a diferença entre o sinal, s , e sua aproximação por séries de Fourier, \tilde{s} , seja menor que o erro ε_s .

Sejam $\tilde{X}_N(t)$ e $\tilde{Y}_N(t)$ as aproximações por séries de Fourier, com N harmônicos, das projeções $x(t)$ e $y(t)$, calculadas pela Equação (8).

$$\begin{aligned} \tilde{X}_N(t) &= A_0 + \sum_{n=1}^N A_n \cos\left(\frac{2n\pi t}{T}\right) + B_n \sin\left(\frac{2n\pi t}{T}\right) \\ \tilde{Y}_N(t) &= C_0 + \sum_{n=1}^N C_n \cos\left(\frac{2n\pi t}{T}\right) + D_n \sin\left(\frac{2n\pi t}{T}\right) \end{aligned} \quad (8)$$

A Figura 2 mostra a imagem da Figura 1 – Segmentos orientados do alfabeto A e imagem associada à palavra $c = 0022746005443$

e suas aproximações por séries de Fourier para $N = 1, 2, 4$ e 15 harmônicos.

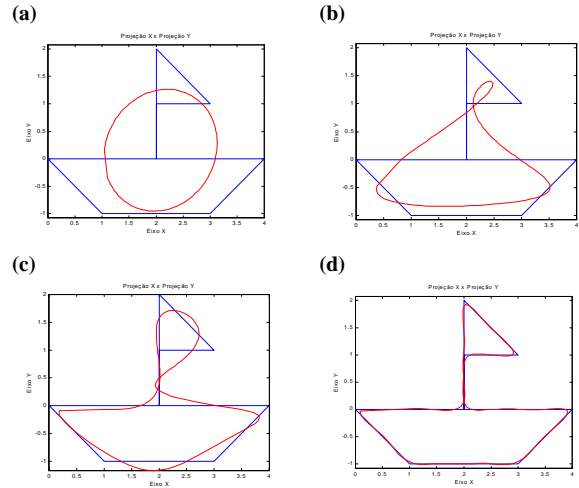


Figura 2 (a) a (d) - Imagem gerada por uma cadeia direcional e harmônicos $N = 1, 2, 4$, e 15

Define-se o erro acumulado de reconstituição ε como:

$$\varepsilon = \max \left(\sup_t |x(t) - \tilde{X}_N(t)|, \sup_t |y(t) - \tilde{Y}_N(t)| \right) \quad (9)$$

Para obter uma informação mais precisa sobre a diferença entre a imagem v e a imagem reconstruída a partir das aproximações de x - y , utiliza-se o erro quadrático médio ε_{qm} dado por:

$$\varepsilon_{qm}(N) = \frac{1}{K} \sum_{p=1}^K \sqrt{(x(t_p) - \tilde{X}_N(t_p))^2 + (y(t_p) - \tilde{Y}_N(t_p))^2} \quad (10)$$

Há, neste ponto, duas observações a serem feitas:

1. uma mesma imagem em \mathcal{R}^2 possui várias representações em A^* ;
2. uma imagem v' obtida a partir de uma rotação, translação ou multiplicação de v por um fator possui uma representação c' em A^* diferente da palavra c , correspondente a v .

A fim de solucionar este problema, recorre-se a uma normalização das aproximações por séries de Fourier pelas chamadas elipses harmônicas.

Pode-se reescrever a Equação (8) como:

$$\begin{aligned}\tilde{X}_N(t) &= A_0 + \sum_{n=1}^N X_n(t) \\ \tilde{Y}_N(t) &= C_0 + \sum_{n=1}^N Y_n(t)\end{aligned}\quad (11)$$

onde, para $1 \leq n \leq N$, tem-se:

$$\begin{aligned}X_n(t) &= A_n \cos\left(\frac{2n\pi t}{T}\right) + B_n \operatorname{sen}\left(\frac{2n\pi t}{T}\right) \\ Y_n(t) &= C_n \cos\left(\frac{2n\pi t}{T}\right) + D_n \operatorname{sen}\left(\frac{2n\pi t}{T}\right)\end{aligned}\quad (12)$$

Da mesma forma, tem-se $v_n(t) = (X_n(t), Y_n(t))$, para todo n . Para cada n , quando t varia de 0 a T , $v_n(t)$ descreve uma elipse no plano. De fato, partindo da Equação (12) e considerando $A_n D_n - B_n C_n$ diferente de zero, obtém-se o seguinte resultado:

$$\frac{(C_n^2 + D_n^2)X_n^2 + (A_n^2 + B_n^2)Y_n^2 - 2(A_n C_n + B_n D_n)X_n Y_n}{(A_n D_n - B_n C_n)^2} = 1 \quad (13)$$

A Equação (13) é a equação de uma elipse e é válida para todo n . Para o caso em que $A_n D_n - B_n C_n = 0$, a Equação (13) representa uma elipse degenerada: uma reta. Vale ressaltar que a mesma elipse $v_n = (X_n, Y_n)$ será encontrada para o n -ésimo harmônico, não importando o ponto da imagem adotada como origem para a codificação em cadeia direcional.

A relação direta entre os coeficientes originais A_n , B_n , C_n , e D_n e os coeficientes após as operações conjugadas de rotação da imagem por um ângulo ψ e de deslocamento de ponto inicial por λ unidades, A_n^{**} , B_n^{**} , C_n^{**} e D_n^{**} , pode ser escrita na forma:

$$\begin{pmatrix} A_n^{**} & B_n^{**} \\ C_n^{**} & D_n^{**} \end{pmatrix} = \begin{pmatrix} \cos\psi & \operatorname{sen}\psi \\ -\operatorname{sen}\psi & \cos\psi \end{pmatrix} \begin{pmatrix} A_n & B_n \\ C_n & D_n \end{pmatrix} \begin{pmatrix} \cos\frac{2n\pi\lambda}{T} & -\operatorname{sen}\frac{2n\pi\lambda}{T} \\ \operatorname{sen}\frac{2n\pi\lambda}{T} & \cos\frac{2n\pi\lambda}{T} \end{pmatrix} \quad (14)$$

O mapeamento $\begin{pmatrix} A_n \\ B_n \\ C_n \\ D_n \end{pmatrix} \rightarrow \begin{pmatrix} A_n^{**} \\ B_n^{**} \\ C_n^{**} \\ D_n^{**} \end{pmatrix}$, denominado Λ ,

normaliza os coeficientes de Fourier originais com relação a uma rotação da imagem original e de uma mudança de ponto inicial. Resta, no entanto, decidir qual o ponto inicial e a orientação padrão para gerar representações canônicas das imagens a serem examinadas por cadeias direcionais, ou seja, dadas duas imagens distintas que representam um mesmo objeto, quaisquer que sejam os pontos iniciais ou as orientações das imagens apresentadas, a aplicação de Λ sobre os coeficientes de Fourier deverá resultar em um mesmo conjunto de coeficientes de Fourier, denominado de representação canônica.

Este problema é resolvido tomando-se a elipse dos harmônicos de ordem 1. O valor de λ_1 é tomado como sendo o valor necessário para que o ponto inicial da imagem v' esteja sobre o eixo semi-maior da elipse de primeiro harmônico. Este eixo determina, também, a rotação ψ_1 da imagem, de modo que a orientação da imagem seja igual à orientação do eixo. É importante ressaltar que, uma vez que eixo semi-maior cruza a elipse harmônica de ordem 1 em dois pontos distintos, existem sempre duas possibilidades de representações canônicas, distintas por uma rotação de 180° , resultando em dois conjuntos de coeficientes de Fourier com sinais opostos. Matematicamente, o que foi apresentado é expresso como:

$$\lambda_1 = \frac{T}{4\pi} \arctan\left(\frac{2(A_1 B_1 + C_1 D_1)}{A_1^2 + C_1^2 - B_1^2 - D_1^2}\right) \quad (15)$$

e

$$\psi_1 = \arctan\left(\frac{C_1^*}{A_1^*}\right) \quad (16)$$

A divisão de todos os coeficientes pelo comprimento do eixo semi-maior, $E^* = \sqrt{(A_1^*)^2 + (C_1^*)^2}$, gera uma normalização em relação ao tamanho da imagem codificada.

Por fim, para obter uma normalização em relação a uma translação da imagem, basta desconsiderar os componentes DC, para efeitos de comparação.

3. Codificação por Cadeias Direcionais

Para a execução da fase experimental do projeto, foi criada uma base de dados composta de dez imagens gravadas em padrão JPEG. Foi elaborado um programa o qual gera, a partir de uma dada imagem qualquer, uma imagem linear por partes. O interesse, em cada imagem da base de dados, está nas bordas do objeto apresentado em primeiro plano. A partir de tais bordas, foram obtidas imagens lineares por partes semelhantes à imagem apresentada na Figura 1. Em seguida, as imagens lineares por partes foram codificadas em cadeias direcionais.

O processo adotado para a codificação das imagens consiste de duas etapas: uma referente à detecção das bordas e linearização por partes da imagem digitalizada em estudo e a outra que codifica a imagem linear por partes em cadeias direcionais.

De modo a se obter maior facilidade de implementação, este trabalho foi desenvolvido com base em imagens no formato “escala de cinza” com 256 valores de intensidade luminosa.

Considere-se uma imagem em que se tem um fundo claro e um contorno escuro qualquer. Neste caso, tomando-se o centro geométrico de um pixel escuro como referência, as direções que podem ser seguidas para encontrar o centro geométrico do próximo pixel escuro coincidem com as direções do alfabeto A, definido na seção 2.

O algoritmo de detecção de bordas foi implementado em duas etapas:

1. Varredura da imagem em busca do pixel P_0 que serviria de ponto de início da codificação;
2. Determinação das bordas do objeto desejado.

O critério de otimização de busca empregado neste trabalho baseia-se no conhecimento prévio do pixel P_k e nas direções dos caracteres do alfabeto A, definido na seção 2, para determinação do próximo pixel P_{k+1} da borda do objeto.

O processo de detecção de borda e linearização por partes utilizado permite que se obtenha o código em cadeia direcional, c_i , correspondente ao objeto apresentado na imagem I .

4. Obtenção dos Coeficientes de Fourier

Uma vez de posse das palavras em cadeias direcionais das imagens componentes da base de dados, procedeu-se à obtenção dos coeficientes normalizados de Fourier das projeções x - y das imagens lineares por parte v induzidas por c . Vale lembrar que os coeficientes normalizados de Fourier de cada imagem compõem a chamada representação canônica, definida na seção 2.

Neste ponto, as relações apresentadas nas Equações (1) a (16) mostraram-se extremamente úteis.

A truncagem das séries de Fourier das projeções x - y permite que seja armazenada, em um arquivo relativamente pequeno, uma quantidade de informações suficiente para a recuperação satisfatória das imagens que se deseja classificar. Além disso, a normalização torna os coeficientes válidos para qualquer orientação e magnitude das imagens.

É interessante, no que diz respeito à truncagem das séries de Fourier, evitar a armazenagem de uma quantidade muito grande de harmônicos, visto que desta forma não haveria nenhuma vantagem na aplicação deste método. Também deve ser tomado cuidado quanto a uma sub-representação da imagem face a uma quantidade muito pequena de harmônicos.

Em atenção a estes detalhes, foram efetuados testes comparando o erro de reconstituição obtido na comparação da imagem original e a imagem reconstruída a partir das séries de Fourier truncadas em N , onde N variou entre 10 e 25 harmônicos. Notou-se que uma imagem reconstruída a partir de séries truncadas no 25º harmônico apresentavam erros acumulados desprezíveis, tendo sido esta a razão de ser escolhida a truncagem neste harmônico. Notou-se, ainda, que abaixo de 10 harmônicos a reconstituição das imagens tornava-se prejudicada.

5. Classificação com Redes Neurais Artificiais

A truncagem das séries de Fourier das projeções x - y constituiu um avanço no que se refere à representação das imagens, visto as dimensões das palavras que descrevem as imagens terem sido reduzidas de centenas de elementos, como era o caso das cadeias direcionais, a algumas dezenas de elementos. Restava, porém, determinar a quantidade otimizada de harmônicos a serem utilizados, além de implementar uma rede neural

que apresentasse a maior robustez a distorções geradas sobre os coeficientes em questão.

A tarefa de determinação da quantidade otimizada de coeficientes de Fourier foi executada por dois programas de classificação por redes neurais artificiais implementados neste trabalho.

Estes programas são bastante semelhantes entre si, diferindo apenas no tipo de arquitetura empregada: Perceptron Multicamadas ou Rede Neural de Base Radial.

O processo de projeto das redes neurais artificiais ocorreu segundo o descrito a seguir. Primeiramente, os módulos dos coeficientes de Fourier das projeções x - y , truncados no harmônico N , com N variando de 10 a 25 foram organizados em palavras de entrada e de saída desejada para as redes neurais. A apresentação dos módulos dos coeficientes de Fourier confere a robustez final quanto à orientação da figura, resolvendo o problema das duas possibilidades de orientação de representações canônica de cada imagem. Uma vez que se tem 4 coeficientes para cada harmônico N , as palavras de entrada das redes neurais artificiais tinham dimensão $4N$. Devido ao fato de terem sido empregadas 10 imagens como Base de Dados, a saída das redes neurais artificiais eram palavras de dimensão 10, sendo que a cada saída era associada uma imagem. A imagem cujo neurônio correspondente apresentava maior valor de saída era dita a imagem reconhecida. As redes foram treinadas, então, sem ruído algum, a fim de se obter uma primeira aproximação dos pesos das conexões. Em seguida, o grau de dificuldade do mapeamento foi sendo aumentado gradativamente, treinando-se as redes com ruídos sobre os coeficientes do Fourier de 5%, 10%,... até 75% dos valores dos mesmos. Os resultados destes treinamentos foram armazenados e os erros de classificação das redes neurais para diferentes valores de N foram comparados.

Uma vez treinada cada topologia, apresentaram-se, para fins de avaliação e em seqüência aleatória, todas as imagens da Base de Dados como entrada das redes neurais artificiais. O processo foi repetido 100 vezes para cada valor de nível máximo de ruído sobre a imagem, R_{max} , o qual variou de 0% a 100%, em passos de 5%, do valor dos coeficientes normalizados de Fourier.

Por fim, de posse dos resultados das comparações dos erros de classificação para várias quantidades de harmônicos e das duas arquiteturas diferentes, bastou escolher a rede de melhor desempenho como a rede projetada para o problema de classificação de imagens codificadas por cadeias direcionais.

6. Resultados

As bordas das imagens da Base de Dados, convertidas para o padrão escala de cinza, foram detectadas segundo o critério de otimização de busca citado na seção 3. Vale lembrar que o processo de detecção de bordas empregado neste trabalho permite a obtenção simultânea das cadeias direcionais que descrevem as imagens lineares por partes representadas pelas bordas dos objetos.

6.1. Resultados da Aplicação das Séries de Fourier

A partir das cadeias direcionais da Base de Dados, foram determinadas as projeções x - y de cada imagem. Obtiveram-se, então as representações canônicas das imagens da Base de Dados, sendo que os módulos dos primeiros 25 harmônicos de $\tilde{X}(t)$ e $\tilde{Y}(t)$ de cada imagem foram colhidos para servir de banco de dados, de onde foram criados os pares de treinamento das redes neurais.

As representações canônicas são conjuntos de coeficientes de Fourier resultantes da aplicação do mapeamento Λ , definido na seção 2, sobre os coeficientes de Fourier de duas imagens distintas que representam um mesmo objeto sob diferentes orientações e magnitudes. Os resultados da aplicação deste mapeamento são ilustrados na Figura 3, onde se tem a sobreposição de 4 imagens distintas referentes a um mesmo objeto sobre as respectivas imagens reconstituídas a partir dos coeficientes normalizados de Fourier obtidos para cada imagem. Nota-se que existem apenas 2 representações canônicas possíveis, diferentes por uma rotação de 180°, conforme previsto na seção 2. A apresentação dos módulos dos coeficientes de Fourier como entrada das redes neurais garante o reconhecimento destas duas representações possíveis como um único padrão.

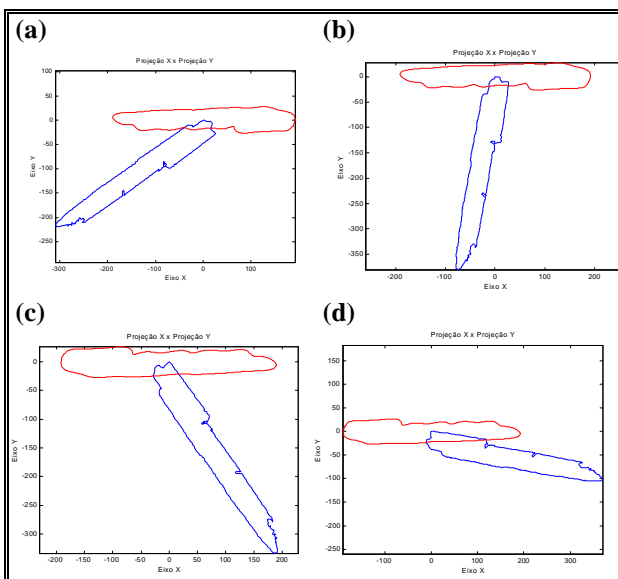


Figura 3 (a) a (d) – Comparações de várias imagens distintas referentes a um mesmo objeto em orientações diferentes com as imagens resultantes das representações canônicas

A quantidade de 25 harmônicos, como sendo uma quantidade suficiente para garantir uma boa reconstrução das imagens, foi determinada a partir de um gráfico de erros médios.

A Figura 4 mostra uma imagem v_i , induzida por uma cadeia direcional c_i e a imagem \tilde{v}_i , reconstrução, a partir dos coeficientes normalizados de Fourier, também representados nestas figuras.

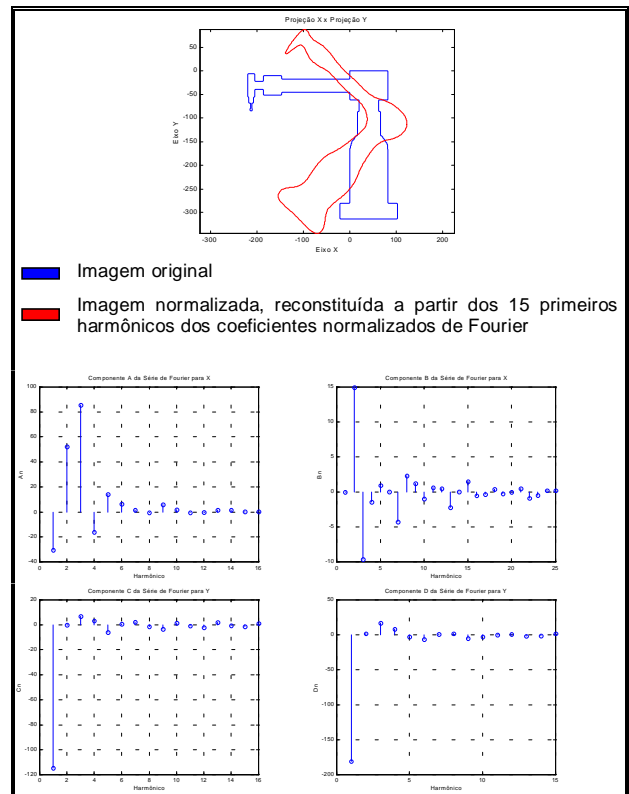
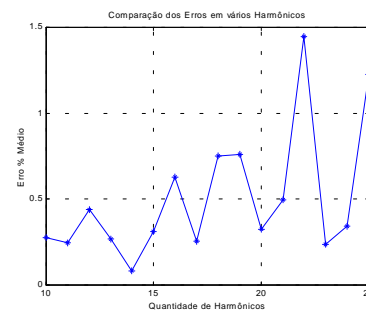


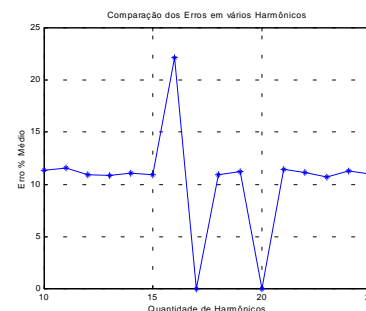
Figura 4 – Robô de soldagem e seus coeficientes normalizados de Fourier

6.2. Resultados das Redes Neurais Artificiais

A Figura 5 mostra uma comparação entre os resultados estatísticos obtidos através das Redes Neurais de Base Radial e Redes Perceptron Multicamadas.



(a) Redes Neurais de Base Radial



(b) Redes Perceptron Multicamadas

Figura 5 – Erro Percentual médio em relação à quantidade de harmônicos

Segundo os resultados obtidos, os percentuais de Erro médio das Redes Neurais de Base Radial encontram-se todos dentro de uma margem compreendida entre os valores de 0% a 1,5%, com um erro médio global de 0,51%, enquanto, para as Redes Perceptron Multicamadas, a margem vai de 0% a 25%, com um erro geral de 10,41%.

Os erros mínimos apresentados pelas redes neurais artificiais foram 0,086%, para o 14º harmônico na classificação com Redes Neurais de Base Radial, e 0% para o 17º e o 20º harmônico na classificação com Redes Perceptron Multicamadas.

Foi determinada, então, como sendo a rede neural projetada para o sistema de classificação de imagens codificadas por cadeias direcionais, a Rede Neural de Base Radial com 56 entradas, 20 neurônios na camada escondida e 10 neurônios na camada de saída, ilustrada na Figura 6.

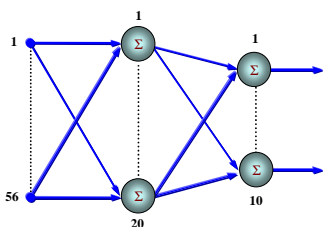


Figura 6 – Topologia da Rede Neural adotada para o Sistema de classificação de imagens codificadas por cadeias direcionais

As figuras 7 e 8 ilustram o resultado de uma classificação de imagem pelo método descrito neste trabalho. Para efetuar esta classificação, foi adotado um ruído de no máximo 15% sobre a imagem a ser classificada, de modo a simular uma possível situação real, como impurezas nas lentes da câmera ou variações de luminosidade, por exemplo.



Figura 7 - Imagem com 15% de ruído e a borda detectada

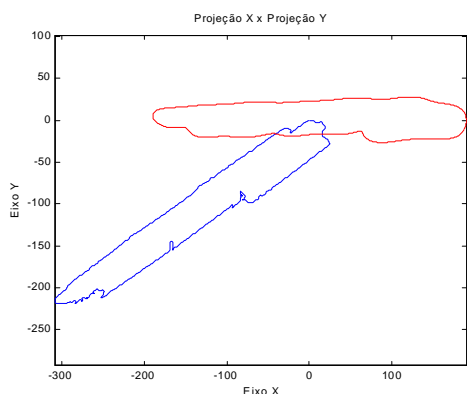


Figura 8 – Imagem codificada e sua representação canônica

7. Conclusões e Perspectivas

Este trabalho apresentou um método de compressão de imagens baseado nas chamadas cadeias direcionais. Este método mostrou-se relativamente simples e eficiente, sendo indicado para aplicações em que haja interesse em descrever as características de um objeto pelas suas bordas.

Em relação aos resultados das redes neurais artificiais, percebe-se que as Redes Neurais de Base Radial são mais adequadas, de um modo geral, para o projeto de um sistema robusto de classificação de imagens codificadas do que as Redes Perceptron Multicamadas, segundo uma série de critérios.

O primeiro fator a favor das Redes Neurais de Base Radial é a maior velocidade de treinamento. Isto significa que, caso seja necessário ensinar novos padrões a uma rede com esta arquitetura, a resposta será mais rápida do que seria caso a rede possuísse arquitetura Perceptron Multicamadas.

O segundo ponto em que as Redes Neurais de Base Radial mostram-se vantajosas em relação às Redes Perceptron Multicamadas está nos seus erros de classificação, os quais foram baixos, com uma média variando entre 0,08% e 1,4% do total de classificações, dependendo da quantidade de harmônicos empregados na descrição das características das imagens. Os erros de classificação das Redes Perceptron Multicamadas, por sua vez, variaram entre 0% e 25,8% da quantidade total das classificações efetuadas.

Algumas possibilidades de continuidade do trabalho ora apresentado são: a expansão do mesmo no sentido de permitir a determinação de um volume de contorno dos objetos de interesse; uma elaboração mais cuidadosa do algoritmo quanto ao tempo de processamento, de forma que se torne viável a implementação “on-line” do algoritmo como um todo em processos de produção; e a determinação das bordas de um objeto não levando em conta apenas a intensidade luminosa dos pixels, mas sim as cores dos objetos.

8. Referências Bibliográficas

- [1] DOUGHERTY, R. E.; GIARDINA, R. C.: *Mathematical Methods for Artificial Intelligence and Autonomous Systems*. Prentice-Hall International Editions, Englewood Cliffs, New Jersey, U.S.A., 1988.
- [2] KASABOV, N. K.: *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering*. Massachusetts Institute of Technology Press, U.S.A., 1996.
- [3] Gomes, J.; Velho, L.: *Computação Gráfica: Imagem* – IMPA/SBM, Rio de Janeiro, Brasil, 1994.
- [4] HALL, E. L.: *Computer Image Processing and Recognition*. Academic Press, Inc. New York, U.S.A., 1979.
- [5] SERRA, J.: *Image Analysis and Mathematical Morphology*. Vol.I. Academic Press, Inc. New York, U.S.A., 1982.
- [6] *Neural Network Toolbox User's Guide*. The MathWorks, January 1994.