

RECONHECIMENTO DE PADRÕES INVARIANTE A ROTAÇÃO UTILIZANDO UMA REDE NEURAL MORFOLÓGICA NÃO SUPERVISIONADA

Erico L. Rempel¹, Francisco A. T. F. Silva²

¹Instituto Nacional de Pesquisas Espaciais-LAC-CAP

²Instituto Nacional de Pesquisas Espaciais-CRN-ROEN

E-mails: erico@nucleo.inpe.br, tavares@roen.inpe.br

Abstract

This work presents the development of a rotation and translation invariant 2-dimensional image detection system using an unsupervised morphological neural network. Utilizing the translation invariance property of the operators of Mathematical Morphology, it is suggested a morphological pattern-matching mechanism based on the comparison between the input image and a set of rotated structuring elements, making possible rotation and translation invariant pattern detection by use of rotational weights applied to morphological neurons. A series of experiments were done using binary and gray scale images.

1. Introdução

O processamento de imagens digitais é uma área de pesquisa que tem experimentado um grande crescimento na última década. Dentro desta área, grandes esforços têm sido aplicados em tarefas de reconhecimento de padrões visuais, onde uma série de dificuldades são encontradas. Além de os padrões de entrada a serem detectados poderem estar corrompidos por ruído, muitas vezes se encontram conectados com outros objetos, necessitando de uma etapa de segmentação como pré-processamento. Além disso, os objetos podem se encontrar em diversos tamanhos e posições, podendo haver necessidade de o reconhecimento ser invariante a transformações de escala, translação e rotação. Dentre estes problemas, o mais complicado, segundo Deseilligny et al. [1], é o reconhecimento invariante a rotação.

Dentre as técnicas mais utilizadas para detecção de imagens invariantes a rotação encontram-se as expansões circulares harmônicas [2], as transformadas Fourier-Mellin [3] e Radon [4], momentos geométricos [5] e abordagens baseadas em Redes Neurais Artificiais (RNAs) [6-8]. Tais técnicas geralmente requerem que o padrão a ser detectado seja segmentado previamente, de maneira a ser separado do contexto.

O modelo de detecção invariante a rotação apresentado neste trabalho baseia-se no conceito de rotação e replicação de pesos de neurônios em diversos ângulos. O mecanismo de replicação de pesos, em

específico, foi inspirado em outros trabalhos [6,7], onde pode-se observar processos semelhantes. O modelo atual faz uso da robustez e flexibilidade da Rede Morfológica Não Supervisionada (RMNS), concebida por Silva e Banon [9].

Na seção 2 é comentada brevemente a estrutura e o funcionamento da RMNS. Na seção 3 é descrita a forma para se implementar invariância a rotação na rede. Na seção 4 são discutidos alguns resultados obtidos. As conclusões são dadas na seção 5.

2. Rede Morfológica Não Supervisionada

A RMNS é uma rede multicamadas não-supervisionada capaz de detectar padrões binários ou em escala de cinza por meio de operadores da Morfologia Matemática [10]. Dentre as principais características da RMNS encontram-se: reconhecimento invariante a translação; certo grau de insensibilidade ao contexto; capacidade de rejeitar padrões não treinados; aprendizado de novos padrões sem alterar padrões já aprendidos e robustez a ruídos controlada pelo usuário através do ajuste de parâmetros da rede. A Figura 1 ilustra a estrutura da rede, que combina uma camada de entrada com uma camada competitiva, mais duas camadas para recuperação e saída do padrão.

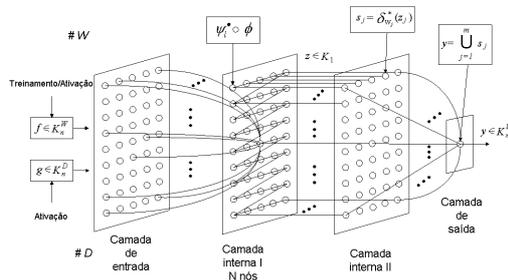


Figura 1: Rede Morfológica Não Supervisionada
FONTE: (Silva e Banon [9], p. 402.)

A forma de aprendizagem é bastante semelhante ao modelo de competição criado por Kohonen [11]. No entanto, na RMNS é feita uma pré-ativação antes do treinamento, de maneira a determinar se o padrão apresentado já foi anteriormente aprendido. Em caso afirmativo, a fase de treinamento é desnecessária. Mais

detalhes sobre o algoritmo de treinamento podem ser observados em [9].

Para uma análise comparativa sobre as modificações efetuadas, para implementação de invariância a rotação, apresenta-se o algoritmo original de ativação da RMNS a seguir:

Passo 1. Inicializar pesos e parâmetros;

Passo 2. Apresentar imagem à camada de entrada;

Passo 3. Ativar primeira camada interna com o operador de detecção com folga [9], sobre a imagem de entrada com os pesos dos neurônios da camada morfológica;

Passo 4. Ativar segunda camada interna, responsável pela recuperação de cada padrão detectado pela primeira camada interna;

Passo 5. Apresentar o primeiro padrão detectado na camada de saída;

Passo 6. Se existe alguma outra imagem, voltar ao passo 1;

Passo 7. Terminar ativação.

Um sistema de inibição, semelhante ao da rede ART (*Adaptive Resonance Theory* [12]) impede que mais de um neurônio da camada morfológica produza saída positiva após o passo 3.

3. O Modelo Proposto

A RMNS permite a recuperação de padrões que sejam semelhantes, dentro de uma faixa de similaridade estabelecida pelo usuário. Esta característica permite algum nível de robustez a pequenas rotações [9]. Para permitir reconhecimento invariante a qualquer ângulo de rotação, propõe-se um esquema de rotação de pesos dos neurônios da primeira camada interna.

Na RMNS os padrões aprendidos são armazenados nas matrizes de pesos dos neurônios da camada morfológica. A RMNS é treinada com os padrões que representam os elementos estruturantes (“pesos”) com orientação 0° conforme já descrito em [9]. Após o treinamento, os pesos dos neurônios morfológicos são rotacionados e replicados em outros neurônios, não utilizados na fase de treinamento, formando diversos agrupamentos invariantes a rotação (agrupamentos de neurônios direcionais). O conjunto de todos os agrupamentos compõe o mapa morfológico invariante a rotação.

Seja W a matriz de pesos de um neurônio treinado da camada morfológica. Um agrupamento de neurônios direcionais pode ser representado por:

$$[W \quad r_{1\alpha}(W) \quad r_{2\alpha}(W) \quad \dots \quad r_{(n-1)\alpha}(W)] \quad (1)$$

onde $r_{i\alpha}(W)$ denota a rotação de W por um ângulo $i \cdot \alpha$, com $i = 1, 2, \dots, n-1$. Dado que α é o intervalo de rotação entre uma cópia e outra, o número de elementos do agrupamento é dado por $n = \lceil 360/\alpha \rceil$.

A Figura 2 ilustra detalhes do mapa morfológico invariante a rotação na arquitetura da RMNS. As operações de união após a camada interna I aplicam-se sobre as saídas de elementos de um mesmo agrupamento de neurônios direcionais. O sistema de inibição da rede garante que somente um agrupamento produza saída positiva. Cada elemento da segunda camada interna recebe a saída de um agrupamento invariante a rotação. Se este valor for positivo, é exibida na camada de saída o padrão detectado com inclinação 0° .

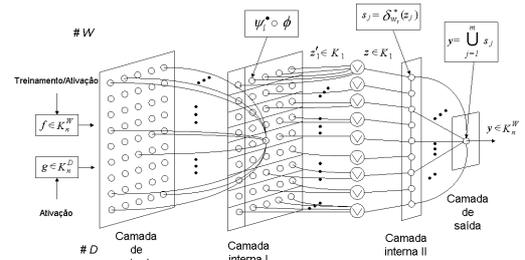


Figura 2: Detalhe da introdução dos neurônios especiais, com replicação de pesos, aplicados ao mapa morfológico

O algoritmo de ativação da RMNS, refletindo a inclusão de invariância a rotação, segue os seguintes passos:

Passo 1. Inicializar pesos e parâmetros;

Passo 2. Apresentar imagem à camada de entrada;

Passo 3. Aplicar ativação de todos os neurônios, da Camada interna I, utilizando-se o operador de detecção com folga como detalhado em [9], sobre a imagem de entrada com os pesos dos neurônios da camada morfológica e os respectivos pesos replicados;

Passo 4. Produzir uma única saída para cada agrupamento através da união das saídas dos seus neurônios;

Passo 5. Ativar segunda camada interna, responsável pela recuperação de cada padrão detectado pela primeira camada interna. As entradas dos elementos desta camada são as saídas dos agrupamentos da camada anterior;

Passo 7. Executar passos 5 e 6 do algoritmo de ativação da RMNS;

Passo 8. Terminar a ativação.

O passo 4 garante que seja produzida uma saída positiva de um agrupamento sempre que um de seus neurônios detectar um padrão conhecido na imagem de entrada. O sistema de inibição da RMNS atua sobre as saídas dos agrupamentos de maneira a impedir que mais de um agrupamento produza saída positiva.

O treinamento é obtido acrescentando-se a replicação dos pesos, nos neurônios direcionais, ao algoritmo original apresentado em [9]. Isto pode ser feito após o treinamento de cada padrão “on-line”, ou

após todo o processo de treinamento do lote original, isto é, em “batch”.

4. Resultados

Alguns resultados preliminares obtidos com o modelo desenvolvido são mostrados a seguir. Inicialmente foram realizados testes para verificar a robustez da rede com relação a entradas ruidosas e rotacionadas. A Figura 3 exibe algumas saídas obtidas para uma rede invariante a rotações de 90 graus treinada com 4 letras de tamanho 8x8. Os padrões de treinamento e ativação são os mesmos utilizados por Srinivasa e Jouaneh [9] para testes com a rede ART1. A RMNS foi capaz de classificar corretamente as entradas apresentadas.

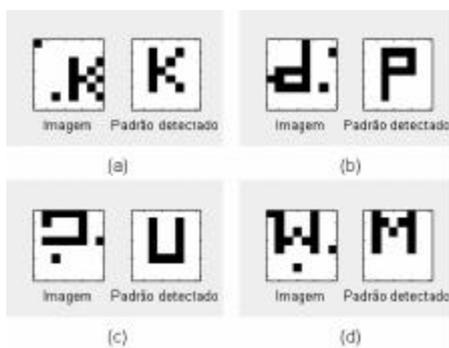


Figura 3: Exemplo de reconhecimento com invariância a rotação para entradas ruidosas e rotacionadas

Com o intuito de demonstrar a capacidade da RMNS de lidar com problemas de contexto, foram realizados testes com imagens binárias de ferramentas. A Figura 4 ilustra a detecção de um padrão (alicate) que encontra-se rotacionado e em contato com outras ferramentas. O padrão treinado é recuperado e apresentado à direita.

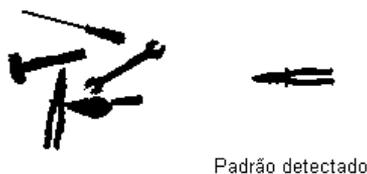


Figura 4: Saída da RMNS treinada com um alicate.

É importante destacar que a RMNS detecta os padrões treinados sem uma etapa de pré-processamento para segmentação da imagem. A detecção do alicate foi possível por meio da sondagem da imagem de entrada em busca do padrão desejado, aproveitando-se da capacidade da RMNS de produzir saída nula caso o padrão não se encontre presente na entrada. Esta característica da rede é importante porque em muitas aplicações é difícil, ou mesmo impossível, conseguir heurísticas robustas de segmentação [13].

As Figuras 5 e 6 mostram a aplicação da RMNS em um problema mais complexo. Neste caso utilizou-se

como padrões de treinamento 4 imagens de faces em posição central (0 grau de inclinação) extraídas do MIT Face Database¹, com resolução de 50x50 pontos. Após o treinamento, os pesos dos neurônios morfológicos foram replicados com inclinações entre 20° e 30°, e entre 340° e 330° com intervalos de 2 graus entre uma cópia e outra. Foram feitos testes com 8 imagens não treinadas em que as pessoas aparecem com o rosto inclinado para a esquerda ou para a direita. A rede foi capaz de classificar corretamente 7 das entradas apresentadas, mesmo com inclinação desconhecida. Uma das entradas produziu saída nula, tendo havido erro de exclusão.



Figura 5: Conjunto de padrões de treinamento

Na Figura 5 tem-se o conjunto de imagens usado no treinamento. A Figura 6 mostra um exemplo de ativação. A rede foi capaz de classificar corretamente o rosto apresentado, contendo inclinações em torno de 25 graus. A capacidade de generalização da rede permitiu a identificação da pessoa mesmo tendo sido apresentada uma imagem diferente daquela utilizada no treinamento. Estes são apenas resultados iniciais visando ilustrar o funcionamento da rede. Para um melhor desempenho sugere-se o uso de imagens de maior resolução.



Figura 6: Reconhecimento de face.

5. Conclusões

Neste trabalho foi apresentada uma forma de implementação de invariância a rotação inserindo-se neurônios com pesos replicados na RMNS. O modelo implementado constitui um sistema de detecção invariante a translação e rotação, robusto a ruídos como apresentado em [9], capaz de lidar com problemas de contexto da cena, sem alterar as características e o funcionamento básico da RMNS. Mais estudos devem

¹ <ftp://whitechapel.media.mit.edu/pub/Images/>

ser efetuados, visando determinar de maneira mais precisa o número mínimo de neurônios direcionais de cada agrupamento, sem perda da eficiência de reconhecimento. Trabalhos futuros também serão realizados para implementação de mecanismos necessários ao reconhecimento invariante a escala.

networks. Cambridge, Massachusetts: The MIT Press, p. 711-715, 1995.

Agradecimentos

Agradecemos a valiosa colaboração do Dr. Gerald J. F. Banon referentes aos diversos comentários fornecidos ao longo do desenvolvimento deste trabalho. Este trabalho teve apoio parcial da FAPESP.

Referências

- [1] Deseilligny, M.; Men, H.; Stanon, G. Character string recognition on maps, a rotation-invariant recognition method. *Pattern Recognition Letters*. v. 16, n. 12, p. 1297-1310, Dec. 1995.
- [2] Chen, C. H.; Wu, R. Study of a class of invariant recognition for machine vision. *Proceedings of SPIE – Visual Communications and Image Processing*. Cambridge, Massachusetts: SPIE. v. 707, p.17-22, Sep. 1986.
- [3] Wechsler, H.; Zimmerman, G. 2-D invariant object recognition using distributed associative memory. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. v. 10, n. 6, p. 811-821, Nov. 1988.
- [4] Simpson, M.; Ervin, P.; Snyder, M. Radon transform applications in optical inspection. *Optical Engineering*, v. 27, n. 2, p. 164-171, Feb. 1998.
- [5] Wong, W.; Siu, W.; Lam, K. Generation of moment invariants and their uses for character recognition. *Pattern Recognition Letters*, v. 16, n. 2, p. 115-123, Feb. 1995.
- [6] Widrow, B.; Winter, R.; Baxter, R. Layered neural nets for pattern recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*. v. 36, n. 7, p. 1109-1118, July 1988.
- [7] Srinivasa, N.; Jouaneh, M. An invariant pattern recognition machine using a modified ART architecture. *IEEE Transactions on System, Man, and Cybernetics*. v. 23, n. 5, p. 1432-1437, Sep./Oct. 1993.
- [8] Fukumi, M.; Omatu, S.; Nishikawa, Y. Rotation-invariant neural pattern recognition system estimating a rotation angle. *IEEE Transactions on Neural Networks*, v. 8, n. 3, p. 568-581, May 1997.
- [9] Silva, F. A. T. F.; Banon, G. J. F. Rede Morfológica Não Supervisionada-RMNS. In: *Proceedings of the IV Brazilian Conference on Neural Networks*, São José dos Campos, p. 400-405, July 1999.
- [10] Serra, J. *Image Analysis and Mathematical Morphology*. Volume 1. London: Academic Press, 1982.
- [11] Kohonen, T. The self-organizing map. *Proceedings of the IEEE*, v. 78, n. 9, p. 1464-1480, Sep. 1990.
- [12] Carpenter, G.; Grossberg, S. The ART of adaptive pattern recognition by a self-organizing neural network. *Computer*, v. 21, n. 3, p. 77-88, Mar. 1988.
- [13] LeCun, Y.; Bengio, Y. Pattern recognition. In: Arbib, M. A. ed. *The handbook of brain theory and neural*