

Reconhecimento de Gestos da Mão para Movimentação de Objetos Virtuais

Pedro Henrique L. M. Silva, Luis Rivera

Laboratório de Ciências Matemática (LCMAT)
Universidade Estadual do Norte Fluminense (UENF)
Av. Alberto Lamego 2000, CEP 28015-620, Campos dos Goytacazes, RJ-Brasil
pedrolmota@gmail.com, rivera@uenf.br

***Abstract.** In this paper, an interactive virtual environment based on immersion principle which consists in moving around virtual objects with the hand to fit these objects in their respective holes is presented. To recognize the hand gesture, the Hu's invariant moments are used to characterize the images and a neural network (MPL) is used for classification.*

1. Introdução

O uso de gestos na comunicação é uma forma comum e umas das mais utilizadas para a comunicação entre humanos. Com a busca por meios mais eficazes de comunicação entre homem/máquina, o estudo do reconhecimento de gestos da mão através de câmeras ganhou muita atenção por se tratar do desenvolvimento de uma forma mais interativa e natural de promover essa comunicação, sem uso de outros elementos intrusivos que limitam a realização de movimentos e gestos em forma natural.

Existem diversos trabalhos que abordam reconhecimento gestos da mão para comunicação humano-computador. Por exemplo, Chen et al. (2007) criaram uma ferramenta para a detecção de quatro posições da mão usando características Haar-like baseado no algoritmo desenvolvido por Viola e Jones (2001). Khotanzad e Lu (1990) desenvolveram um sistema de reconhecimento de linguagem de sinal americano dos 26 caracteres do alfabeto inglês a partir dos gestos de mão capturados por uma câmera, classificadas por rede neural Perceptron de múltiplas camadas. Os gestos foram caracterizados usando momentos invariantes de Hu e os momentos de Zernick. A tecnologia de sensores combinadas com câmeras, como a utilizada por Kinect (Kinect), também são usadas eficientemente, para reconhecimentos de posturas e posições. No sentido de interação natural, o ideal é reconhecer posturas e posições sem a ajuda de sensores.

Neste trabalho pretende-se construir uma ferramenta para reconhecer um conjunto de gestos da mão com princípios de imersão de forma a pegar um objeto, movimentar e encaixar no buraco que melhor ajuste geometricamente. Para isto, tem-se um ambiente virtual com um conjunto de diferentes objetos geométricos na mesa, uma parede com buracos para cada um dos objetos e uma pinça para segurar e mover os objetos.

2. Reconhecimento e movimentação de objetos

O sistema proposto consiste de vários módulos. Um módulo gerador do ambiente virtual. Um módulo reconhecedor do gesto baseado em rede neural que analisa cada frame binarizado da imagem capturada pela webcam. A imagem binarizada deve ser caracterizada para alimentar a rede neural tanto na aprendizagem quanto na classificação. Um terceiro módulo é o gerador de ações da pinça, como consequência do gesto reconhecido, movimentando e atuando sobre o objeto, em consonância com a

movimentação da mão. Um módulo de proximidade entre o objeto movimentado e a geometria do buraco. Finalmente, um módulo de encaixe com princípio de encaixe automático quando o objeto o buraco correspondente estão bem próximos.

2.1. Reconhecimento de gestos

O fato de a mão humana variar grandemente nas posições que pode assumir e movimentos que pode realizar é a principal razão pela qual a escolha de um bom conjunto de características é importante. Nesse trabalho, para se extrair a mão a partir de uma imagem, foi implementado, devido a sua simplicidade, o método usado por Peer et al. (2003), que detecta a mão a partir de valores correspondentes ao tom de pele definidos explicitamente em um determinado espaço de cores no formato RGB, como $R > 95$, $G > 40$ e $B > 20$, $\text{Max}\{R,G,B\} - \text{Min}\{R,G,B\} > 15$, $R - G > 15$, $R > G$ e $R > B$. A maior dificuldade nesse tipo de abordagem entretanto, é encontrar o espaço de cores adequado e o conjunto de regras de decisão que deve ser determinado de forma empírica. Nesse contexto, a partir da imagem da mão é extraída as características utilizando os momentos invariantes de Hu (1968). Esses momentos são um conjunto de funções não lineares que são invariantes a escala, translação e rotação e são definidos a partir dos momentos geométricos da imagem. Os momentos invariantes de Hu fornecem 7 valores de ponto flutuante que formam parte do conjunto de dados de entrada para uma rede neural no processo de aprendizado, como também no processo de classificação para reconhecimento de uma postura da mão. Neste caso se usam as 6 primeiras funções significativas, sendo descartada a sétima.

Modela-se a rede neuronal Perceptron de múltiplas camadas (MPL), em particular de três camadas, sendo a primeira camada de 6 neurônios, uma para cada valor de entrada, e as camadas ocultas definidas por testes empíricos.

3. Conclusões e trabalhos pendentes

Os quatro gestos (eleger, segurar, mover, soltar), caracterizados pelos momentos invariantes, permitiram treinar a MPL para o reconhecimento do gesto. O comando é transferido ao ambiente virtual para acionar a pinça nos objetos. Ainda faltam integração dos módulos e os testes. Quando concluído, o usuário interagirá com o ambiente de forma direta e interativa.

Referências

- Chen, Q.; Georganas, N.; Petriu, M. (2007), Real-time Vision-based Hand Gesture Recognition Using Haar-like Features. Instrumentation and Measurement Technology Conference – IMTC, Polônia.
- Khotanzad, A.; Lu, J. (1990), Classification of Invariant Image Representations Using a Neural Network. IEEE Transaction on acoustics, speech and signal processing. Vol 38. No 6. pp. 1028 – 1038
- Kinect, <http://www.engadget.com/2010/11/04/kinect-for-xbox-360-review/>.
- Hu, M. (1962), Visual pattern recognition by moments invariants. IRE Trans. Inform. Theory. Vol. IT-8. pp. 179-187.
- Peer, P., Kovac, J., and Solina, F. (2003), Human skin colour clustering for face detection. In submitted to EUROCON 2003 – International Conference on Computer as a Tool.
- Viola, P.; Jones, M. (2001), Robust real-time object detection. Cambridge Research Laboratory Technical Report Series CRL2001/01, pp. 1-24.