

Universidade de São Paulo  
Instituto de Ciências Matemáticas e Computação

# Aprendizado por imitação usando Redes Neurais

Aluna: Valéria de Carvalho Santos  
Prof<sup>a</sup>. Dra. Roseli A. F. Romero

# Sumário

- Introdução
- Modelos de RN para aprendizado por imitação
- Considerações finais
- Referências

# Introdução

- Aprendizado por Imitação
  - O objetivo do mecanismo de imitação é reproduzir precisamente um subconjunto pré-definido de tarefas e, assim, aprender por meio de exemplos [Billard *et al*, 2006]



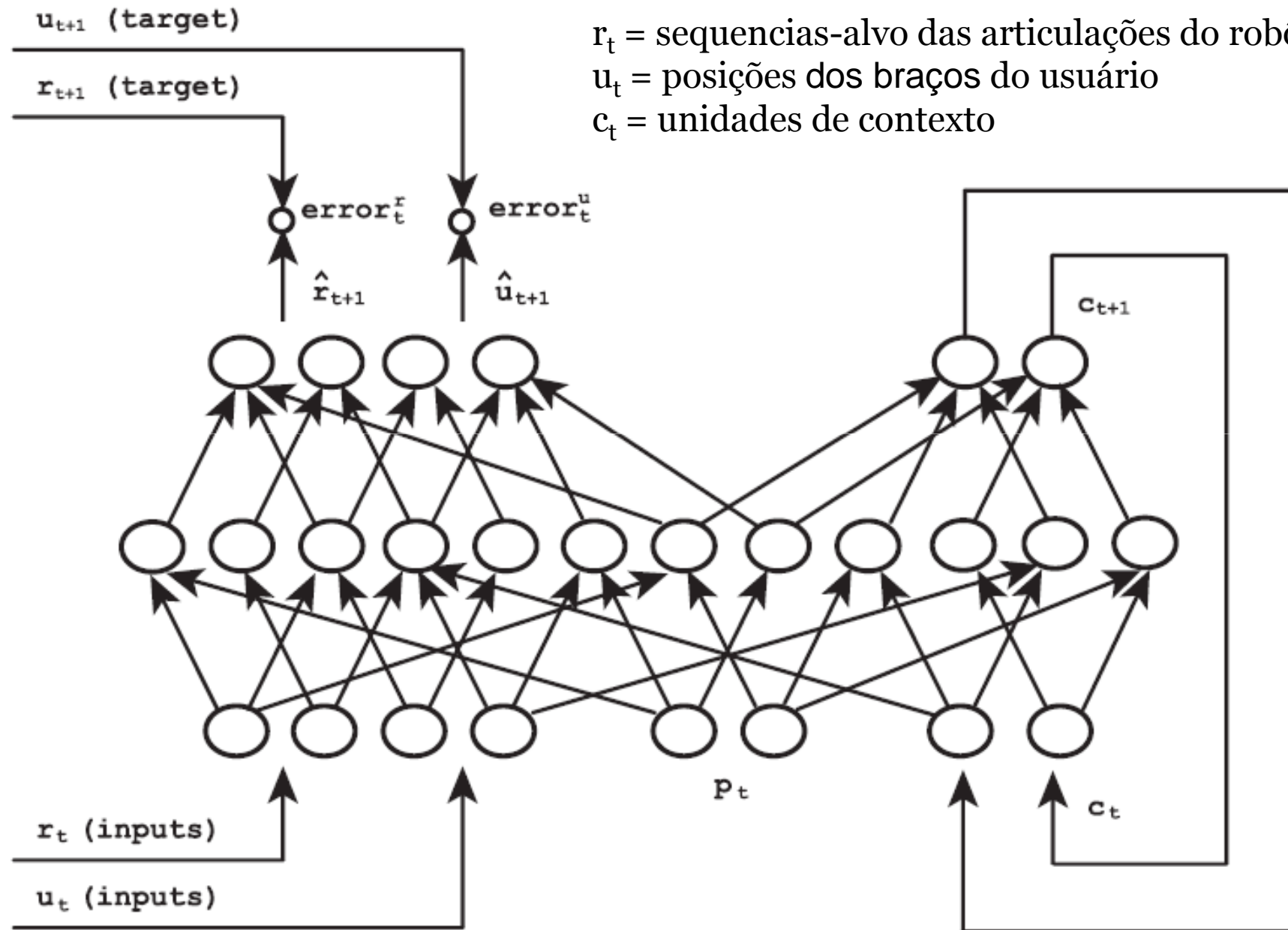
## Trabalhos Relacionados [Ito, 2004]

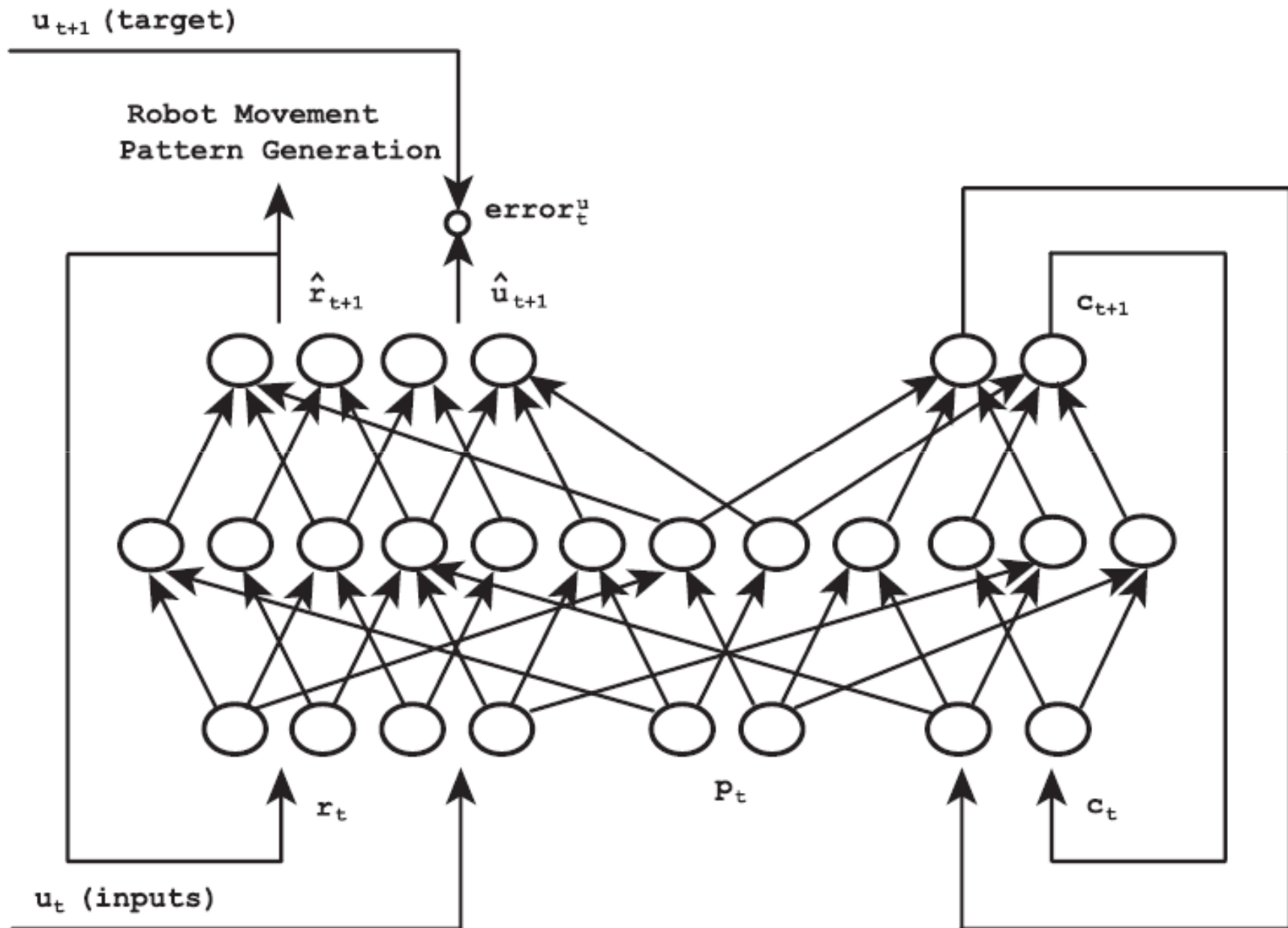
- Experimentos em interações imitativas entre um robô humanóide e um humano
- Modelo de rede neural dinâmica
  - Recurrent Neural Network Model with Parametric Bias
- QRIO (SDR-4X II)



## Trabalhos Relacionados [Ito, 2004]

- Apenas movimentos padrões dos braços foram considerados
- A tarefa do robô consiste de duas fases:
  - **Aprendizado**
    - Um conjunto de movimentos padrões com diferentes perfis é aprendido e associado aos correspondentes movimentos padrões visualmente percebidos dos braços do usuário
  - **Interação**
    - o robô percebe visualmente os movimentos padrões dos braços do usuário e gera seus movimentos padrões correspondentes.





## Trabalhos Relacionados [Ito, 2004]

- Os padrões de movimento do robô são gerados com base nos valores PB, enquanto esses valores são adaptados ao perceberem os padrões de movimento da mão
- Uma característica interessante desse modelo é que a geração e a percepção são executadas simultaneamente em sistema dinâmico neural



## Trabalhos Relacionados [Gotoh *et al*, 2007]

- Projeto de pesquisa conjunto entre universidade e indústria (Masayoshi Kanoh, 2004; Kato et al., 2004)
- Conversa com uma pessoa por um motor de síntese e reconhecimento de voz fundamental.
- Pode se comunicar com uma pessoa transmitindo suas emoções através de Mecanismos de gestos e expressões faciais.



## Trabalhos Relacionados [Gotoh *et al*, 2007]

- Aprender o mapeamento de expressão facial entre um humano e um robô Ifbot, pelo qual o robô pode imitar uma expressão facial arbitrária
- Utiliza o FACS (Facial Action Coding System) [Ekman e Friesen, 1978]
  - Decompõe uma expressão facial observada em uma ou mais de 44 unidades de ação, que representa a atividade muscular que produz mudanças momentâneas na aparência facial

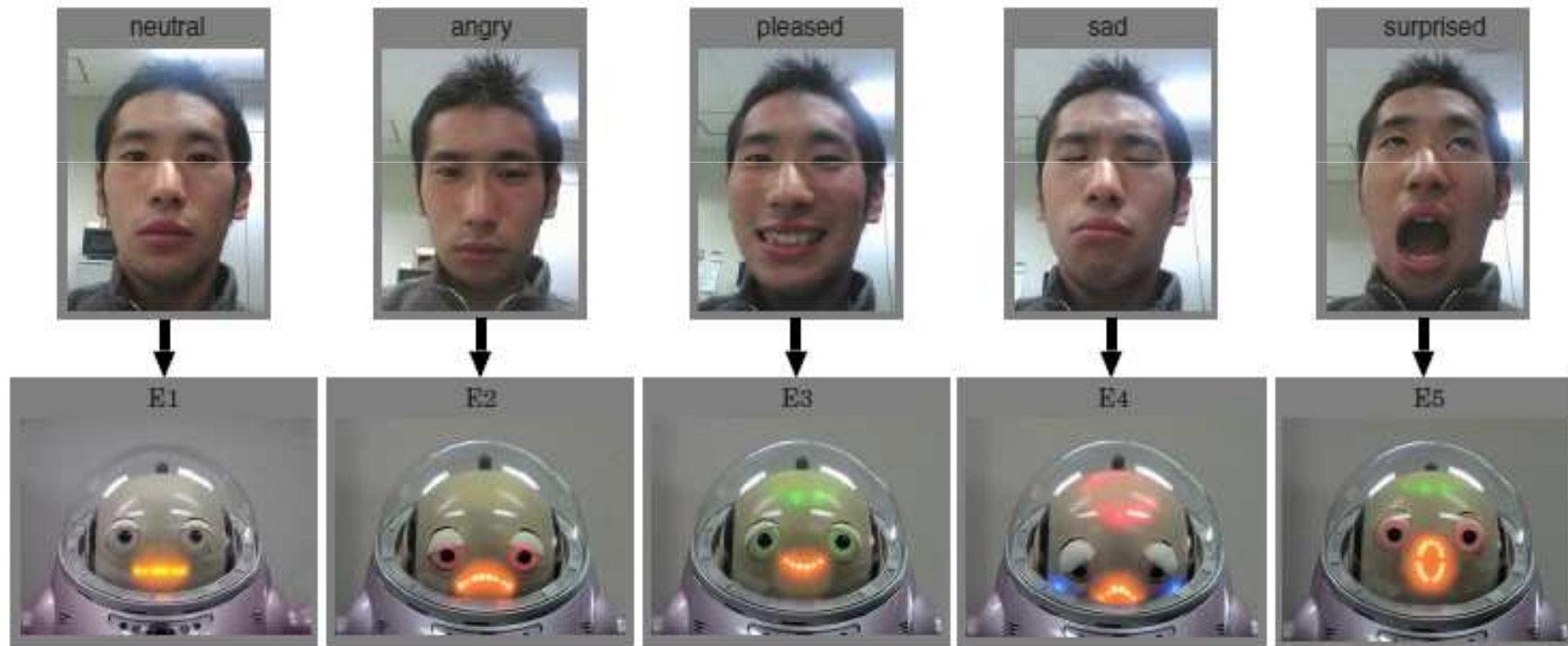
## Trabalhos Relacionados [Gotoh *et al*, 2007]

- Rede neural perceptron com três camadas para mapear expressões faciais entre um humano e o Ifbot
- As entradas da rede são os parâmetros das unidades de ação de faces humanas
- A rede é treinada com aprendizado supervisionado por retropropagação de erro, usando erro entre os dados de saída da rede e os parâmetros de expressão facial do Ifbot
- Foram definidas 16, 48 e 16 unidades para a primeira, segunda e terceira camada respectivamente

## Trabalhos Relacionados [Gotoh *et al*, 2007]

- Treinamento da rede:
  - quatro imagens faciais expressivas (raiva, satisfação, tristeza e surpresa) e uma neutra de quatro pessoas e as correspondentes expressões faciais do Ifbot.
- Para avaliar o método, foram apresentadas à rede imagens de três pessoas

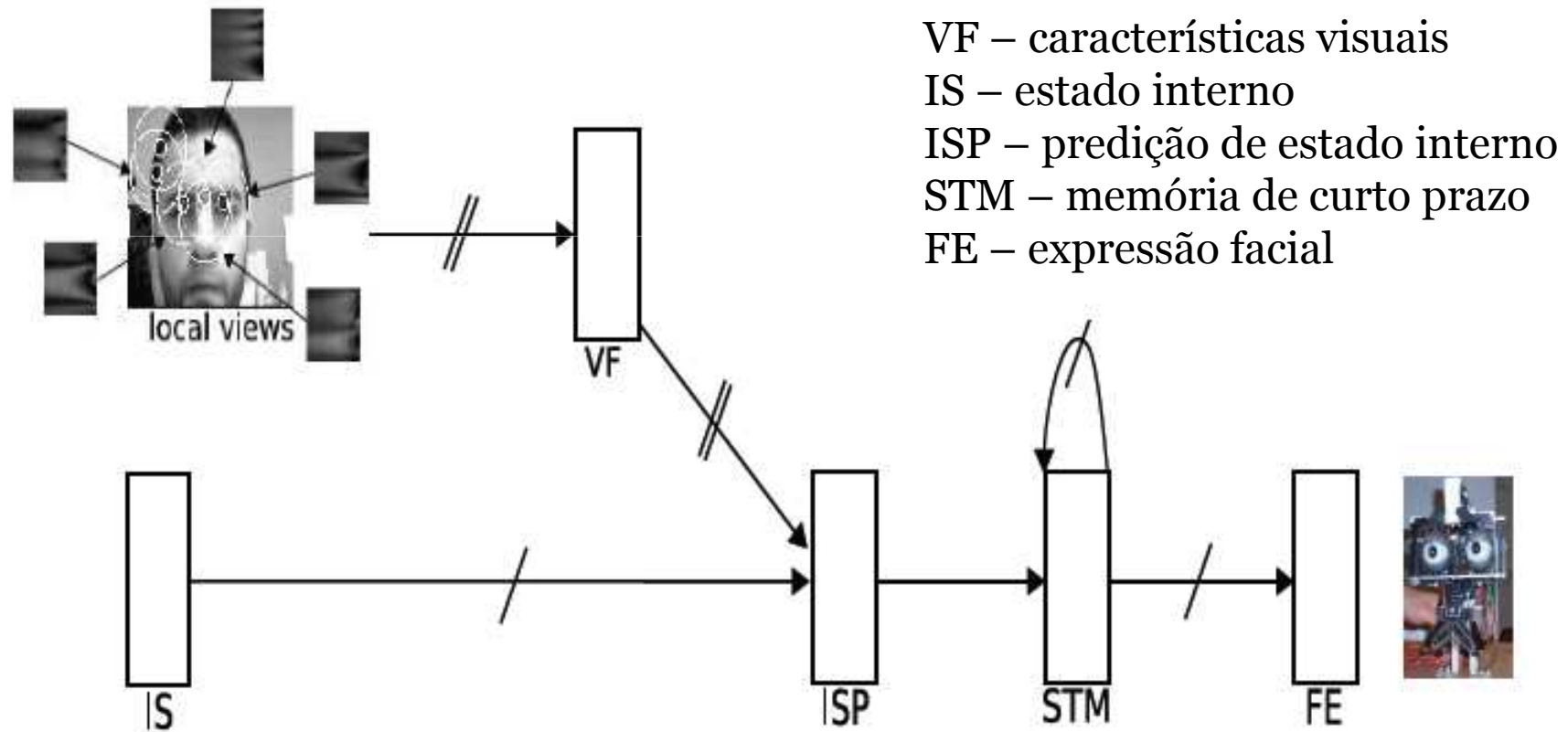
# Trabalhos Relacionados [Gotoh *et al*, 2007]



## Trabalhos Relacionados [Boucenna *et al*, 2010]

- Objetivo
  - Entender como os bebês aprendem a reconhecer expressões faciais sem ter um sinal de ensino
- No início o robô conhece nada do ambiente

## Trabalhos Relacionados [Boucenna *et al*, 2010]



IS, ISP, STM e FE são grupos de neurônios que contêm cinco neurônios correspondendo a quatro expressões faciais mais a expressão neutra.

## Trabalhos Relacionados [Boucenna *et al*, 2010]

- O conjunto de visões locais é aprendido pelo recrutamento de novos neurônios no grupo VF

$$VF_j = net_j \cdot H_{max(\gamma, \overline{net} + \sigma_{net})}(net_j)$$

$$net_j = 1 - \frac{1}{N} \sum_{i=1}^N |W_{ij} - I_i|$$

$$H_{\theta}(x) = \begin{cases} 1, & \theta < x \\ 0, & \text{caso contrário} \end{cases}$$

- $VF_j$  é a atividade do neurônio  $j$  no grupo VF
- $I$  é uma entrada visual
- $\gamma$  é um limiar de reconhecimento



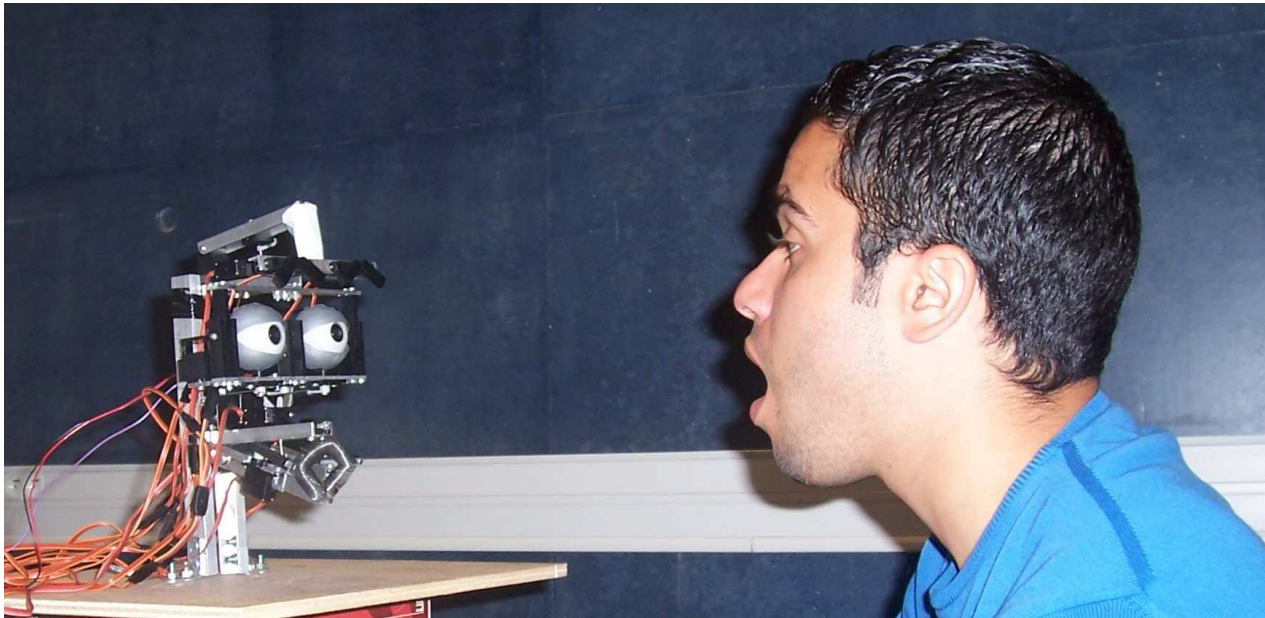
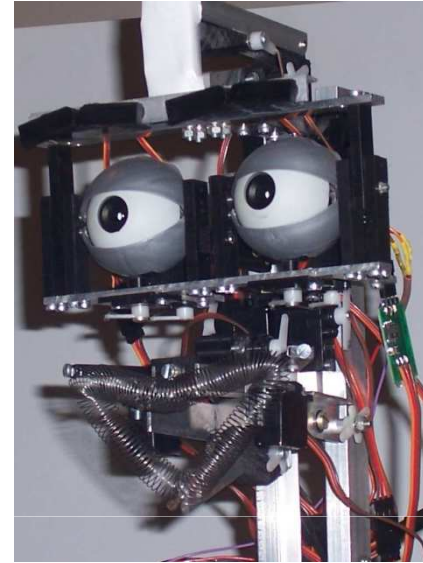
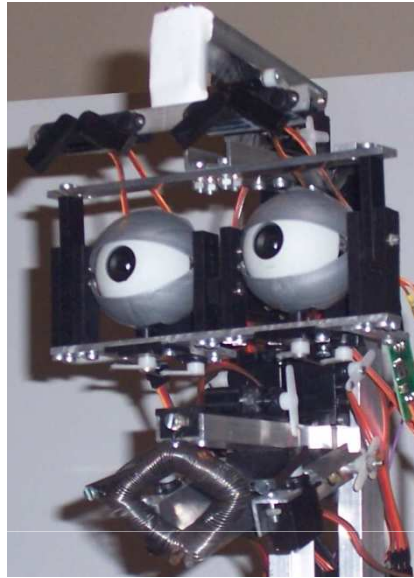
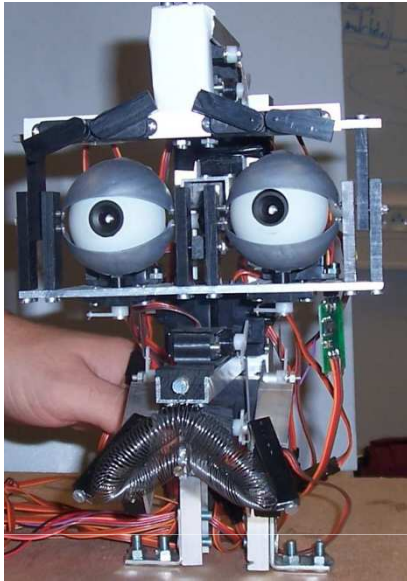
## Trabalhos Relacionados [Boucenna *et al*, 2010]

- Ajuste dos pesos:

$$\Delta W_{ij} = \delta_j^k (a_j(t) I_i + \epsilon (I_i - W_{ij}) (1 - V F_j))$$

$$\delta_j^k = \begin{cases} 1, & j = k \\ 0, & \text{caso contrário} \end{cases}$$

- $K = \text{argMax}(a_j)$ 
  - $a_j(t) = 1$  quando um novo neurônio é recrutado
  - $a_j(t) = 0$ , caso contrário
- A outra parte da regra pondera os protótipos já aprendidos



# Considerações Finais

- Observa-se que tanto modelos mais simples quanto modelos mais complexos reproduzem esse tipo de aprendizado

# Referências

- Bombini, G., Mauro, G. D., Basile, N., Ferilli, T. M. A., e Esposito, S. (2009). Relational learning by imitation. *Agent and Multi-Agent Systems: Technologies and Applications*, páginas 273–282.
- Boucenna, S., Gaussier, P., Andry, P., e Hafemeister, L. (2010). Imitation as a communication tool for online facial expression learning and recognition. *International Conference on Intelligent Robots and Systems*.
- Ekman, P. e Friesen, W. V. (1978). *Facial action coding system*. Consulting Psychologists Press.

# Referências

- Gotoh, M., Kanoh, M., Kato, S., e Itoh, H. (2007). A neural-based approach to facial expression mapping between human and robot. In Proceedings of the 11th international conference, KES 2007 and XVII Italian workshop on neural networks conference on Knowledge-based intelligent information and engineering systems: Part III, KES'07/WIRN'07, páginas 194–201, Berlin, Heidelberg. Springer-Verlag.
- Ito, M. (2004). On-line imitative interaction with a humanoid robot using a dynamic neural network model of a mirror system. *Adaptive Behavior*, 12:93–115.
- Kato, S., Ohshiro, S., Itoh, H., e Kimura, K. (2004). Development of a communication robot ifbot. In *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*.

# Referências

- Masayoshi Kanoh, Shohei Kato, H. I. (2004). Facial expressions using emotional space in sensitivity communication robot "ifbot". In IEEE/RSJ International Conference on Intelligent Robots and Systems, página 1586U" 1591, Sendai, Japan.
- Billard, A., Calinon, S., e Guenter, F. (2006). Discriminative and adaptive imitation in uni-manual and bi-manual tasks. *Robotics and Autonomous Systems*, 54(5):370–384.