



UNIVERSIDADE TÉCNICA DE LISBOA

INSTITUTO SUPERIOR TÉCNICO

# CONTROLO DE UM MANIPULADOR ROBÓTICO USANDO VISÃO

**Maria da Graça Vieira de Brito Almeida**

(Licenciada)

Dissertação para obtenção do grau de Mestre em Engenharia Electrotécnica e de Computadores

**Orientador:** Doutor Pedro Manuel Urbano de Almeida Lima

**Co-Orientador:** Doutor José Alberto Rosado dos Santos Vitor

**Júri:**

**Presidente:** Doutor João Rogério Caldas Pinto

**Vogais:** Doutor José Alberto Rosado dos Santos Vitor

Doutor Paulo José Cerqueira Gomes da Costa

Doutor Pedro Manuel Urbano de Almeida Lima

Agosto de 2004

## Resumo

O sistema de visão tem sido um aliado natural dos manipuladores. Com efeito, quando se procura dotar os manipuladores de autonomia, cada vez mais se alia aos manipuladores a capacidade de visão. O problema essencial reside na necessidade de conhecer a correspondência de um movimento no plano da imagem para o correspondente movimento no espaço tridimensional (3D). Tradicionalmente essa correspondência é fornecida ao sistema de controlo dos manipuladores por um tipo de procedimento que é designado por calibração.

Esta Tese utiliza uma arquitectura de controlo “dinâmico vê-e-move” (*dynamic look-and-move*), na qual a relação entre os movimentos do manipulador e o respectivo movimento na imagem é calculada por uma Rede Neuronal que intervém na cadeia de controlo. Esta Rede substitui o Jacobiano completo (que inclui o Jacobiano do Manipulador e o da Imagem), evitando deste modo a necessidade de conhecer a cinemática do manipulador e proceder à calibração do sistema de visão.

Apresentam-se resultados experimentais relativos à utilização da instalação de ensaio constituída por um manipulador, com 5 graus de liberdade, da Mistubishi, modelo Movemaster E3J, e um sistema de visão constituído por uma câmara analógica e uma placa de aquisição de vídeo.

**Palavras-chave:** manipulação robótica, visão robótica, seguimento visual, captura baseada em visão, controlo em tempo real, seguimento visual sem calibração.

## Abstract

The vision system has been a natural allied of the robot arm. In fact, it is commonly accepted that one of the ways to increase the manipulator autonomy is to add visual capacity. The key issue relies on the relation of a bi-dimensional movement with the correspondent three-dimensional space movement. This correspondence is traditionally given by a procedure designated as calibration.

This work makes use of a visual servo control (dynamic look-and-move). A Neural Network takes part on control chain imposing the relationship between the robot arm movements and the corresponding movements in the image. The network simulates the Jacobian of the robot arm and image avoiding the need to know the kinematics and vision system calibration. Experimental results are shown relating the use of a *Mitsubishi* robot arm with a vision system composed by a camera and a video frame grabber.

**KEY-WORDS:** Robotic manipulation, robotic vision, visual servoing, capture based on vision, real time control, uncalibrated visual servoing.

## **Agradecimentos**

Quero expressar em primeiro lugar os meus sinceros agradecimentos ao Prof. Doutor Pedro Manuel Urbano de Almeida Lima e Prof. Doutor José Alberto Rosado dos Santos Victor, respectivamente, na qualidade de Orientador e Co-orientador científicos, pelo empenho, incentivo e disponibilidade mostrada na orientação deste trabalho.

Agradeço ao Prof. Doutor João Carlos Pires da Palma, as ajudas prestadas no projecto, pela sua disponibilidade, e pela utilização do equipamento da Secção de Automação e Robótica do DEEA/ISEL de que é responsável, e da qual sou docente, compatibilizando o desempenho da minha actividade enquanto docente.

Agradeço, igualmente, a todas aquelas pessoas que, de uma forma ou de outra, deram o seu contributo para a realização deste trabalho, e que não foram aqui referidas.

Agradeço ao Instituto Politécnico de Lisboa, pelo financiamento do projecto de investigação I&D N° 3, “Controlo de um Braço de Robô Através de Visão”.

Por último, para a minha família a quem a realização deste trabalho tirou grande parte da minha atenção e disponibilidade, aqui fica um eterno e grato reconhecimento.

# Índice

<b>1</b>	<b>Introdução.....</b>	<b>1</b>
1.1	Motivação.....	1
1.2	Perspectiva Histórica da Auto-calibração.....	2
1.3	Objectivos.....	4
1.4	Contribuições Originais.....	5
1.5	Organização da Tese.....	6
<b>2</b>	<b>Controlo em Espaço Operacional Baseado em Visão.....</b>	<b>8</b>
2.1	Sistema de Controlo do Manipulador.....	9
2.1.1	Espaço Junta e Espaço Operacional.....	9
2.1.2	Jacobiano do Manipulador.....	10
2.2	Sistema de Visão.....	12
2.2.1	Modelo de Formação da Imagem.....	12
2.2.2	Jacobiano da Imagem.....	15
2.2.3	Jacobiano Completo.....	16
2.2.4	Características da Imagem.....	18
2.2.5	Segmentação de Imagens.....	20
2.3	Arquitectura de Seguimento Visual.....	20
2.4	Cinemática do erro.....	22
<b>3</b>	<b>Estimação do Jacobiano através da Rede Neuronal.....</b>	<b>25</b>
3.1	Caracterização das Redes Neurais.....	25
3.2	Algoritmo de Retropropagação do Erro.....	28
3.3	Jacobiano Completo através da Rede Neuronal.....	30
<b>4</b>	<b>Métodos e Algoritmos.....</b>	<b>33</b>
4.1	Descrição da Instalação.....	33
4.2	Sistema de Controlo Utilizado.....	34
4.3	O Manipulador Mitsubishi Movemaster RV- E3J.....	37
4.4	Sistema de Visão.....	38
<b>5</b>	<b>Resultados.....</b>	<b>40</b>
5.1	Simulação.....	40
5.1.1	Estudo da Função Jacobiano.....	41
5.1.2	Treino da Rede Neuronal.....	46
5.1.3	Testes.....	48
5.2	Resultados Experimentais com o Manipulador Real.....	51
<b>6</b>	<b>Conclusões e Trabalho Futuro.....</b>	<b>60</b>

<b>7</b>	<b>Referências Bibliográficas .....</b>	<b>63</b>
	<b>Anexo A – Manipulador Mitsubishi RV-E3J.....</b>	<b>65</b>
	<b>Anexo B - Parâmetros Denavit-Hartenberg.....</b>	<b>67</b>
	<b>Anexo C - Cinemática Directa.....</b>	<b>68</b>
	<b>Anexo D - Cinemática Inversa.....</b>	<b>69</b>
	<b>Anexo E - Jacobiano do Manipulador .....</b>	<b>71</b>
	<b>Anexo F - Jacobiano Completo .....</b>	<b>72</b>
	<b>Tabela 2 – Características do Movemaster RV-E3J.....</b>	<b>73</b>

# 1 Introdução

## 1.1 Motivação

A maioria dos manipuladores industriais necessita de ambientes de grande exactidão do ponto de vista do posicionamento dos objectos envolventes e adaptados às suas características. Com a utilização comum dos manipuladores, surge a necessidade de desenvolver aplicações onde a localização dos objectos a manipular, no ambiente de trabalho, não seja tão rigorosa, criando deste modo um sistema flexível. Tal objectivo pode ser conseguido aliando capacidade de visão ao sistema de controlo do manipulador, através da recolha de imagens por câmara de vídeo e do seu tratamento em computador.

Conhecem-se diversos tipos de estratégias para seguir um objecto usando visão por computador, que necessitam dos modelos das características do manipulador e da câmara. Estes modelos só se mantêm correctos se não se fizer nenhuma alteração das condições iniciais, nomeadamente o posicionamento relativo da câmara, manipulador e objectos a manipular. Manter estes modelos torna-se difícil, limitativo e, por vezes, oneroso.

Quando se procura dotar os manipuladores de autonomia, cada vez mais se alia aos manipuladores a capacidade de visão. Contudo, ainda que tecnologicamente a captação de imagens não coloque dificuldades especiais com equipamento de custo moderado, há problemas relativos ao seu tratamento que terão de ser resolvidos. O problema essencial reside na necessidade de conhecer a correspondência de um movimento no plano da imagem para o correspondente movimento no espaço tridimensional (3D). Tradicionalmente essa relação é fornecida ao sistema de controlo dos manipuladores por um tipo de procedimento que é designado por calibração.

O procedimento tradicional de calibração consiste em proceder à relação entre pontos do espaço cartesiano para pontos do plano de imagem. Para estabelecer esta correspondência é necessário conhecer as coordenadas de um conjunto de pontos do espaço tridimensional e conhecer as coordenadas destes mesmos pontos no plano de imagem, para depois, através de cálculos apropriados, se obter uma matriz, denominada

matriz de calibração, que estabelece a correspondência entre as coordenadas. Esta matriz de calibração já incorpora os parâmetros intrínsecos e extrínsecos da câmara, pelo que um novo posicionamento desta, e/ou uma mudança de lente, implicará um novo processo de calibração. Para além deste processo ser efectivamente moroso, a sua validação estará dependente da precisão que se obtiver das coordenadas dos pontos atrás citados.

Uma alternativa ao processo da calibração consiste na aprendizagem da relação entre pontos do espaço cartesiano para pontos do plano de imagem, através de uma rede Neuronal.

Esta Tese utiliza uma arquitectura de controlo “dinâmico “vê-e-move” (*dynamic look-and-move*)[3], na qual a relação entre os movimentos do manipulador e o respectivo movimento na imagem é calculada por uma Rede Neuronal que intervém na cadeia de controlo. Esta Rede substitui o Jacobiano completo (que inclui o Jacobiano do Manipulador e o da Imagem), evitando deste modo a necessidade de conhecer a cinemática do manipulador e proceder à calibração do sistema de visão.

## 1.2 Perspectiva Histórica da Auto-calibração

Desde 1973, com o trabalho de Shirai e Inoue [9], que se fazem estudos com base no uso de uma cadeia de retroacção visual no sentido de corrigir o movimento do braço manipulador. Referências importantes a este respeito são as de Corke [3] e Graefe [17].

A obtenção de modelos independentes, sem valores predefinidos, traz grande vantagem em termos de robustez e custo, e é objecto de estudo de grande parte dos investigadores. Os investigadores pretendem construir modelos que não dependam da cinemática do manipulador nem da geometria dos objectos a manipular.

Relativamente aos métodos até agora implementados, de auto-calibração, merecem destaque algumas abordagens que a seguir se mencionam:

- Modelo da cinemática e os parâmetros intrínsecos da câmara [25]

Zhuang *et al.* em 1995 apresentaram um sistema simultâneo de calibração do manipulador e de uma câmara acoplada no effector terminal do manipulador. Neste artigo, os autores realizaram um estudo do modelo da cinemática e das funções de custo. Adicionaram dois parâmetros ao modelo de calibração do manipulador que representam a geometria da câmara, sendo possível realizar



níveis diferentes de calibração. No modelo de calibração mais simples, identificam-se os parâmetros da câmara com outros que especificam a transformação câmara/efector terminal, e assume-se que é conhecida a relação entre os sistemas de coordenadas do mundo e do efector terminal; no segundo nível de calibração identificam-se os parâmetros da câmara que relacionam a transformação câmara/efector terminal e o sistema da base, e assume-se que a geometria do manipulador é conhecida. O sistema, na sua versão completa, calibra o sistema manipulador-câmara na totalidade. Assume-se, para os três níveis, que os parâmetros intrínsecos da câmara, o centro da imagem, os factores de escala e a distância focal são conhecidos. Utiliza-se o Jacobiano no bloco de calibração manipulador-câmara.

- Mapas de memória [14]

Desde 1986, a equipa de Ritter *et al.* estuda o tema de controlo utilizando mapas de memória Kohonen. O sistema utilizado compõe-se de duas câmaras e um manipulador com três graus de liberdade, cujo objectivo é apanhar objectos num espaço 3D. Ao longo deste estudo foi verificado que é possível obter bons resultados mesmo com um pequeno número de treinos. O processo de aprendizagem é dividido em diversos estágios de modo a ter redes mais pequenas e simples.

- Duas câmaras e restrição dos graus de liberdade do manipulador [16][17]

Este grupo, Graefe e Nguyen, iniciou o seu trabalho nesta área em 1995. O manipulador utilizado tem 5 graus de liberdade, mas neste sistema de controlo só se consideram 3 juntas; a 5ª junta só tem a ver com a orientação e não é utilizada, e a 4ª junta está sempre colocada perpendicularmente ao plano de trabalho. O sistema de visão é composto por duas câmaras solidárias com o movimento da 1ª junta. O primeiro passo consiste em mover a 1ª junta até encontrar o objecto no plano da imagem, e no passo seguinte controlam-se as juntas 2 e 3 de modo a aproximar o manipulador ao objecto.

- Restrição a um plano [24]

Este sistema utiliza uma câmara montada no efector terminal do manipulador e admite que a distância  $z$  do manipulador ao plano de trabalho é fixa e conhecida. Conhecendo o movimento da câmara é possível retirar

informação útil exclusivamente da imagem. A câmara está montada no efector terminal e solidária com o movimento da última junta, observando a área de trabalho por baixo do efector terminal, onde se encontra um objecto. À medida que o sistema da câmara roda em torno do eixo rotacional do efector terminal, o caminho desenhado pelo ponto característico forma uma elipse no plano de imagem. Observa-se que a distância do eixo rotacional ao ponto característico, no espaço cartesiano, é proporcional ao tamanho da elipse gerada, no plano de imagem. Quando o eixo rotacional está sobre o objecto obtém-se um simples ponto na imagem. O ponto representa o alinhamento do sistema de visão com o objecto seguido.

### 1.3 Objectivos

O principal objectivo desta Tese é desenvolver uma metodologia, que permita aproximar um manipulador dotado com sistema de visão monocular de um objecto, sem a necessidade de efectuar qualquer tipo de calibração da câmara. Não é necessário conhecer os parâmetros intrínsecos da câmara nem efectuar a sua calibração extrínseca. O manipulador é controlado pela informação que retira da imagem e não necessita de perspectiva inversa ou qualquer outra transformação de cinemática. A única restrição é, o objecto a capturar estar localizado num plano.

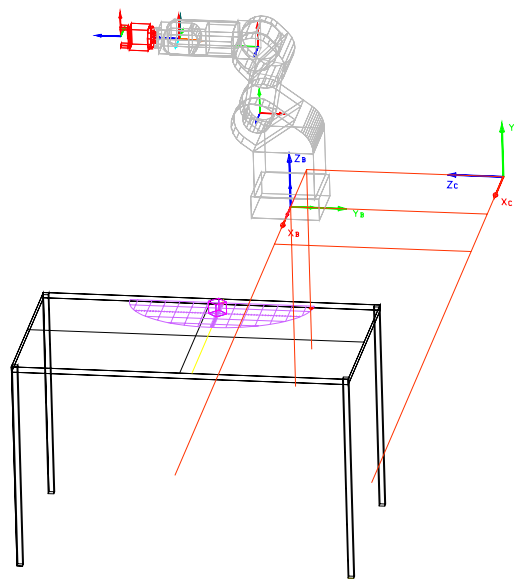
Propõe-se a concepção e desenvolvimento de uma nova técnica de aprendizagem da correspondência de deslocamentos do plano de imagem no espaço cartesiano, à custa de informação de visão recolhida através de câmara de vídeo. Subsequentemente procuram-se soluções de controlo de um manipulador, baseadas em visão utilizando esta técnica. Pretende-se que o processo de calibração seja feito de um modo natural e integrado, sem ser necessário efectuar medições por parte do operador. De igual modo não se pretende conhecer a posição da câmara, nem estar limitado a um tipo de câmara, lente, ou distância focal. O objectivo não é conhecer medidas exactas do plano cartesiano para o plano da imagem, mas sim encontrar a relação aproximada entre um deslocamento efectuado pelo manipulador, no plano cartesiano, e o respectivo deslocamento no plano de imagem.

Para se conhecer esta relação poder-se-ia optar por determinar a cinemática inversa ou o Jacobiano completo (Jacobiano da Imagem e Jacobiano do Manipulador); ambas as soluções implicariam conhecer as características físicas do manipulador, e como tal seria um trabalho limitado a um dado dispositivo.

Neste trabalho, utiliza-se uma rede neuronal na cadeia de controlo, sendo ela a responsável por efectuar o Jacobiano inverso, não havendo assim necessidade de cálculos matemáticos prévios nem a necessidade de se conhecer as características do manipulador.

Pretende-se sustentar o estudo com a indispensável componente experimental, utilizando equipamento com características idênticas às dos sistemas industriais.

Neste estudo será utilizado um manipulador, com 5 graus de liberdade, da Mistsubishi, modelo Movemaster E3J e um sistema de visão constituído por uma câmara e uma placa de aquisição. O manipulador será controlado por comandos elaborados a partir do sistema de visão e enviados directamente para o controlador do manipulador. Um desenho esquemático dos diversos elementos envolventes no sistema poderá ser visto na Figura 1.1. De salientar que neste trabalho o espaço de trabalho foi reduzido ao plano da mesa. A posição e orientação da câmara não são fixas nem conhecidas a priori.



*Figura 1.1– Desenho esquemático dos diversos elementos envolventes no sistema implementado*

## **1.4 Contribuições Originais**

Com a utilização de redes neuronais criou-se um sistema simples de controlo, com a restrição a um plano de trabalho. Houve a preocupação de não impor limitações ao sistema e de não assumir nada sobre as características do objecto observado, ou sobre o posicionamento da câmara. Optou-se por usar uma só câmara, em lugar das duas câmaras que a maior parte dos artigos sugere, numa primeira abordagem, para melhor compreender

as variáveis em causa e também devido ao alargamento do espaço de trabalho que uma só câmara permite. Por se utilizar somente uma câmara, houve a necessidade de aplicar um método que conseguisse, de algum modo, estimar a profundidade do objecto a capturar relativamente ao plano de imagem.

## 1.5 Organização da Tese

A Figura 1.2 representa, num diagrama de blocos, os diversos temas abordados nesta Tese e a sua organização.

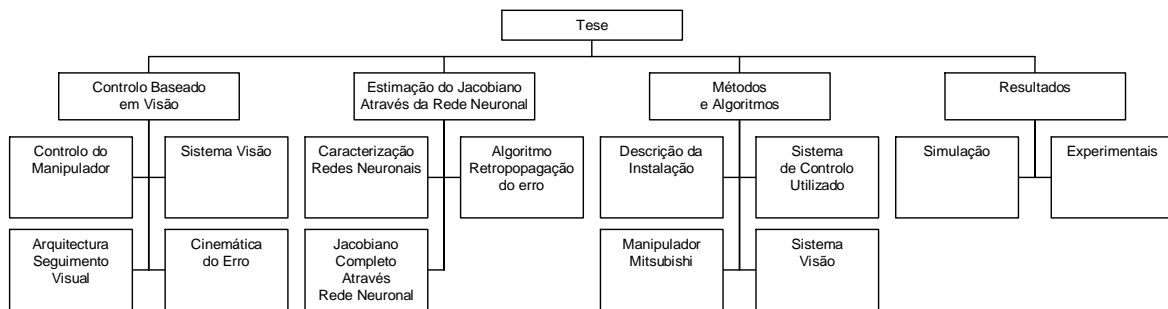


Figura 1.2- Organização da Tese

No Capítulo 2 introduzem-se os diversos aspectos que caracterizam um sistema de seguimento visual. Apresenta-se alguma teoria relevante, e introduz-se a terminologia e a notação utilizadas. No que concerne ao manipulador são introduzidas as noções de espaço junta e espaço operacional, a noção de Jacobiano, remetendo-se para o anexo as expressões concretas relacionadas com o manipulador utilizado. Em relação ao sistema de visão são introduzidas questões relativas ao modelo de formação de imagem, modelo da câmara, funções características da imagem e Jacobiano da Imagem. São introduzidas as arquitecturas de seguimento visual, dando relevo à implementada – estrutura de controlo de sistema dinâmico “vê-e-move”. Conclui-se o Capítulo fazendo referência à cinemática do erro.

A estimação do Jacobiano completo, que engloba o Jacobiano do Manipulador e o da Imagem, é apresentada no Capítulo 3. Aqui começa-se por fazer uma caracterização das redes neuronais, explicando de seguida o algoritmo de retropropagação de erro utilizado. Termina-se explicando a aplicação das redes neuronais à estimação do Jacobiano completo.

A descrição da instalação e o sistema de controlo implementado é apresentado no Capítulo 4, terminando este com a apresentação das características do manipulador RV-E3J da Mitsubishi.

No Capítulo 5 apresentam-se os resultados da simulação em Matlab. Mostram-se o estudo da função Jacobiano e o treino da rede neuronal. Finalmente indicam-se os resultados experimentais, relativos à extracção de parâmetros característicos de imagem, bem como às experiências realizadas de forma a aproximar o manipulador de um objecto em forma de cubo.

No Capítulo 6 apresenta-se as conclusões retiradas deste trabalho e estabelece-se as direcções principais de investigação para trabalho futuro.

## 2 Controlo em Espaço Operacional Baseado em Visão

Neste Capítulo introduzem-se os diversos aspectos que caracterizam um sistema de seguimento visual, e apresentam-se as diversas arquitecturas de seguimento visual. Aborda-se alguma teoria relevante e introduz-se a terminologia e a notação utilizadas.

Um sistema de seguimento visual, em termos genéricos, resulta na integração de dois subsistemas:

- Sistema de visão – com capacidade para fazer o seguimento dinâmico do manipulador e que faça a computação, em tempo real, dos parâmetros que caracterizam o seu movimento;
- Sistema de controlo do manipulador.

Em relação ao sistema de visão é necessário saber qual a arquitectura de seguimento visual a utilizar, assim como os algoritmos para a extracção dos parâmetros característicos de imagem.

A movimentação de um manipulador requer a aplicação de forças e binários. Trata-se de um sistema com um comportamento dinâmico complexo. No entanto a abordagem destes aspectos apenas se torna imprescindível nos casos em que se procura explorar intensivamente a rapidez de actuação. No presente estudo admite-se que a produção de força, bem como o controlo dinâmico dos movimentos do manipulador, ficam a cargo do controlador. Pretende-se realizar movimentos sem requisitos de evolução temporal restritivos, podendo admitir-se que a operação decorre em modo quase estacionário.

Por esse facto proceder-se-á apenas ao estudo da cinemática e do Jacobiano do Manipulador.

## 2.1 Sistema de Controlo do Manipulador

### 2.1.1 Espaço Junta e Espaço Operacional

Para descrever a trajectória do efector terminal do manipulador é necessário especificar a sua posição e orientação.

$$\mathbf{P} = \begin{bmatrix} \mathbf{p} \\ \phi \end{bmatrix} \quad (2.1)$$

onde  $\mathbf{p}$  representa a posição e  $\phi$  representa a orientação.

O vector  $\mathbf{p}^1$  é definido no espaço no qual a tarefa é especificada, sendo este espaço usualmente designado por *espaço operacional*.

O *espaço junta* denota o espaço no qual é definido o vector das variáveis de junta. Para um manipulador com  $n$  graus de liberdade e apenas com juntas de revolução, esse vector é dado por:

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \dots \\ \theta_n \end{bmatrix} \quad (2.2)$$

onde  $\theta_i$ , com  $i = 1..n$ , representa a posição angular da junta  $i$ .

As equações da cinemática directa [15] [23] permitem calcular a posição do efector terminal em função das variáveis de junta:

$$\mathbf{P} = k(\boldsymbol{\theta}) \quad (2.3)$$

As equações da cinemática directa, para o manipulador aqui utilizado, encontram-se no anexo C.

O problema da cinemática inversa consiste na determinação dos valores de junta, correspondentes a uma dada pose do efector terminal. Nem sempre é possível encontrar uma única solução. No anexo D encontram-se as equações que permitem resolver o problema da cinemática inversa para o manipulador utilizado.

---

<sup>1</sup> Vectores e Matrizes serão representados em **Negrito**

Na Figura 2.1 pode-se observar os diversos referenciais do espaço de trabalho;  $\{B\}$  referencial da base do manipulador,  $\{G\}$  referencial do efector terminal e  $\{C\}$  o referencial da câmara.

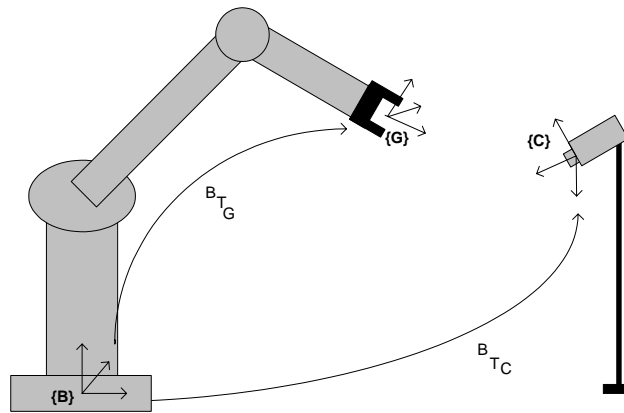


Figura 2.1- Sistemas de Coordenadas relevantes; Mundo( $\{B\}$ ), Efector terminal( $\{G\}$ ) e Câmara( $\{C\}$ )

A notação utilizada foi a seguinte:  ${}^xT_y$  representa a matriz de transformação através do qual é possível representar o sistema de coordenadas  $\{y\}$  no referencial do sistema de coordenadas  $\{x\}$ .

## 2.1.2 Jacobiano do Manipulador

Nos sistemas de controlo de movimento, é necessário conhecer como a variação infinitesimal das juntas afecta o movimento infinitesimal do efector terminal. A vantagem de trabalhar com variações pequenas é a relação linear entre o espaço infinitesimal das juntas e o espaço infinitesimal do efector terminal.

O objectivo da cinemática diferencial é estimar a relação entre a velocidade das juntas e a velocidade de translação e angular do efector terminal. Consideremos que  $\mathbf{\Omega}(t) = [\omega_x \ \omega_y \ \omega_z]^T$  representa a velocidade angular e  $\mathbf{T}(t) = [T_x \ T_y \ T_z]^T$  a velocidade de translação e que o manipulador tem  $n$  graus de liberdade [5][6][20].

Pretende-se exprimir a velocidade linear do efector terminal  $\dot{\mathbf{p}}$  e a velocidade angular  $\mathbf{\Omega}$  em função das velocidades de junta  $\dot{\mathbf{\theta}}$ , através da seguinte relação:

$$\begin{aligned} \dot{\mathbf{p}} &= \mathbf{J}_p \dot{\mathbf{\theta}} \\ \mathbf{\Omega} &= \mathbf{J}_\omega \dot{\mathbf{\theta}} \end{aligned} \tag{2.4}$$



onde  $\mathbf{J}_P$  é a matriz  $(3 \times n)$  relativa à contribuição das juntas de velocidade  $\dot{\boldsymbol{\theta}}$  para a velocidade linear do efector terminal  $\dot{\mathbf{p}}$ , enquanto  $\mathbf{J}_O$  é a matriz  $(3 \times n)$  relativa à contribuição das juntas de velocidade  $\dot{\boldsymbol{\theta}}$  para a velocidade angular do efector terminal  $\boldsymbol{\Omega}$ .

A matriz  $(6 \times n)$   $\mathbf{J}_M$  é o Jacobiano do manipulador.

$$\mathbf{J}_M = \begin{bmatrix} \mathbf{J}_P \\ \mathbf{J}_O \end{bmatrix} \quad (2.5)$$

Considerando o Jacobiano do manipulador,  $\mathbf{J}_M$  [10][20], dividido em vectores de coluna  $(3 \times 1)$ ; calcula-se  $\mathbf{J}_M$  através das matrizes de transformação de cada um dos elos do manipulador, tal como se indica em (2.6),

$$\mathbf{J}_M = \begin{bmatrix} J_{P1} \dots J_{Pn} \\ J_{O1} \dots J_{On} \end{bmatrix} \quad (2.6)$$

$$\begin{bmatrix} J_{Pi} \\ J_{Oi} \end{bmatrix} = \begin{bmatrix} Z_{i-1} \times (P - P_{i-1}) \\ Z_{i-1} \end{bmatrix}$$

Sendo  $P_0 = [0 \ 0 \ 0 \ 1]^T$  e  $Z_0 = [0 \ 0 \ 1]^T$ ,  $P_{i-1}$  é dado pela 3ª coluna da matriz de Translação e  $Z_{i-1}$  é dado por 3ª coluna da matriz de Rotação.

As equações da cinemática diferencial, podem ser escritas, numa forma compacta por:

$$\begin{bmatrix} T_X \\ T_Y \\ T_Z \\ \omega_X \\ \omega_Y \\ \omega_Z \end{bmatrix} = \mathbf{J}_M \dot{\boldsymbol{\theta}} \quad (2.7)$$

onde  $\boldsymbol{\Omega}(t) = [\omega_x \ \omega_y \ \omega_z]^T$  representa a velocidade angular e  $\mathbf{T}(t) = [T_x \ T_y \ T_z]^T$  a velocidade de translação.

No Anexo E apresenta-se as equações do Jacobiano para o manipulador utilizado neste trabalho.

## 2.2 Sistema de Visão

O objectivo do sistema de visão é extrair, em tempo real, informação da imagem de modo a fazer a detecção, o reconhecimento, e ainda o seguimento dos objectos relevantes.

Para controlar um manipulador, utilizando informação proveniente de um sistema de visão, é necessário compreender os aspectos geométricos do processo de formação da imagem. Cada câmara forma uma projecção bidimensional (2D), da cena, no plano de imagem onde o sensor está colocado. Pode-se obter a informação tridimensional utilizando várias vistas de uma só câmara, ou utilizando várias câmaras (*stereo*), ou conhecendo a relação geométrica entre pontos extraídos da imagem e o objecto real, ou limitando o espaço de trabalho. Nesta Tese utiliza-se uma só câmara e limita-se o espaço de trabalho a um plano. Pode-se utilizar vários modelos de projecção da imagem. Neste trabalho será adoptado o modelo de projecção perspectiva.

### 2.2.1 Modelo de Formação da Imagem

Um ponto 3D do mundo,  $\mathbf{p} = {}^C[X \ Y \ Z]^T$ , cujas coordenadas são expressas no sistema de coordenadas da câmara,  $\{C\}$ , projecta um ponto,  $\mathbf{c} = [u \ v]^T$ , no plano de imagem, tal como se indica na Figura 2.2 (assumindo uma lente *pinhole*), onde  $f$  é a distância entre o plano de imagem e o centro de projecção, também denominado de distância focal.

A notação utilizada foi a seguinte:  ${}^x\mathbf{p}$  representa um ponto  $\mathbf{p}$  no referencial de coordenadas  $\{x\}$

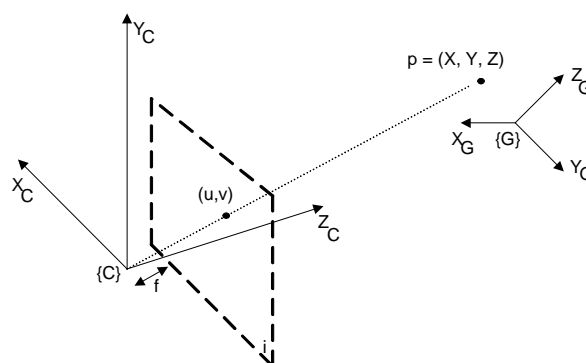


Figura 2.2 – Projecção Perspectiva de um ponto 3D no plano da imagem

A relação entre os dois pontos, segundo o modelo de projecção perspectiva [3][20], é dado por:

$$\begin{aligned} u &= f \frac{X}{Z} \\ v &= f \frac{Y}{Z} \end{aligned} \quad (2.8)$$

Ou ainda em coordenadas homogéneas, onde  $\lambda$  é um escalar diferente de 0,

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} {}^c \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.9)$$

O modelo da câmara é constituído pelos seus parâmetros extrínsecos e intrínsecos.

Os parâmetros intrínsecos da câmara descrevem:

- A dimensão dos pixels (*picture elements*) da imagem, escalamento vertical e horizontal;
- A posição do ponto principal (intersecção do eixo óptico com o plano da imagem);
- Distância focal.

Os parâmetros extrínsecos descrevem:

- A orientação e posição da câmara.

Nas equações (2.8) e (2.9) os pontos da imagem estão expressos em unidades métricas e não no índice dos pixels (*picture element*). Através dos parâmetros intrínsecos da câmara, é possível relacionar a representação de um ponto da imagem em pixels, com a representação do mesmo ponto em unidades métricas.

Utilizando os parâmetros intrínsecos da câmara, dimensão do pixel/metro da imagem ( $K_x, K_y$ ) e a posição do centro da imagem ( $C_x, C_y$ ) é possível reescrever (2.9),

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} fK_x & 0 & C_x & 0 \\ 0 & fK_y & C_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} {}^c \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.10)$$

$$\text{sendo } \begin{cases} u = (fK_x X + C_x Z) / Z \\ v = (fK_y Y + C_y Z) / Z \end{cases} \quad (2.11)$$

Através de transformação de coordenadas é possível exprimir um ponto  $\mathbf{P}$ , com coordenadas conhecidas no referencial da câmara ( $\{C\}$ ), no referencial do mundo ( $\{B\}$ ). Neste trabalho vamos considerar o referencial do mundo como sendo o sistema de coordenadas da base do manipulador,  $\{B\}$ . A relação entre os diferentes referenciais está representada na Figura 2.3.

A notação utilizada foi a seguinte:  ${}^xT_y$  representa a matriz de transformação através da qual é possível representar o sistema de coordenadas  $\{y\}$  no referencial do sistema de coordenadas  $\{x\}$ ;  ${}^x\mathbf{p}$  representa um ponto  $\mathbf{p}$  no referencial de coordenadas  $\{x\}$ .

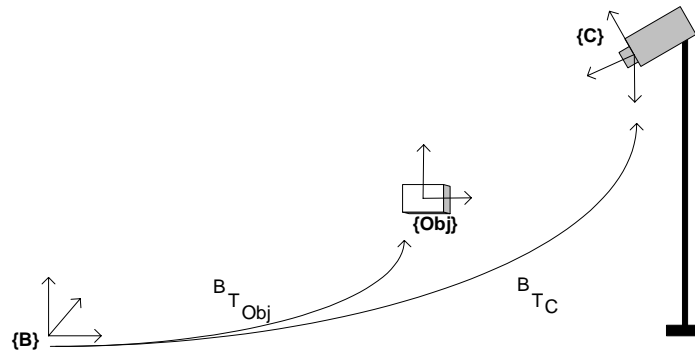


Figura 2.3 - Sistemas de Coordenadas relevantes; Mundo ( $\{B\}$ ), Objecto ( $\{Obj\}$ ) e Câmara ( $\{C\}$ )

Juntando as equações anteriores obtém-se o modelo completo da câmara que relaciona um ponto  ${}^B\mathbf{p} = [X \ Y \ Z]^T$ , expresso em coordenadas do mundo (referencial zero do manipulador), com a sua projecção no plano da imagem  $[u \ v]^T$ , expressa em pixels.

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} fK_x & 0 & C_x & 0 \\ 0 & fK_y & C_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \times ({}^C T_B) \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.12)$$

sendo  $({}^C T_B)$  a matriz de transformação através da qual é possível representar o sistema de coordenadas do mundo, ( $\{B\}$ ) no sistema de coordenadas da câmara, ( $\{C\}$ )

Supondo que o modelo de formação da imagem é a projecção perspectiva, o vector de parâmetros característicos da imagem é dado por  $\mathbf{c} = [u \ v]^T$ , onde  $u$  e  $v$  são dados por (2.12).

### 2.2.2 Jacobiano da Imagem

Nas aplicações de controlo por visão, é necessário relacionar os movimentos diferenciais dos parâmetros característicos da imagem com movimentos diferenciais do objecto a seguir visualmente.

Uma ferramenta útil para transformar o erro, no espaço de imagem (tipicamente expresso em pixel) no equivalente espaço operacional é o chamado Jacobiano de Imagem. O Jacobiano da imagem,  $\mathbf{J}_{img}$ , encarrega-se de estabelecer esta relação.

Supondo que o efector terminal do manipulador está a mover-se no espaço de trabalho, este movimento será descrito, no referencial da câmara ( $\{C\}$ ), por uma velocidade angular  $\boldsymbol{\Omega}(t) = [\omega_x \ \omega_y \ \omega_z]^T$  e por uma velocidade de translação  $\mathbf{T}(t) = [T_x \ T_y \ T_z]^T$ .

A velocidade de um ponto  $\mathbf{P}$  no efector terminal, expresso no referencial da câmara, pode ser escrito em notação vectorial como:

$$\dot{\mathbf{P}} = \boldsymbol{\Omega} \times \mathbf{P} + \mathbf{T} \quad (2.13)$$

Suponhamos o ponto  $\mathbf{P} = [X \ Y \ Z]^T$ . A velocidade de  $\mathbf{P}$ , pode ser dada por [3]:

$$\begin{aligned} \dot{X} &= Z\omega_y - Y\omega_z + T_x \\ \dot{Y} &= X\omega_z - Z\omega_x + T_y \\ \dot{Z} &= Y\omega_x - X\omega_y + T_z \end{aligned} \quad (2.14)$$

sendo  $\boldsymbol{\Omega}(t) = [\omega_x \ \omega_y \ \omega_z]^T$  velocidade angular e  $\mathbf{T}(t) = [T_x \ T_y \ T_z]^T$  a velocidade de translação.

Substituindo (2.8) em (2.14) obtém-se:

$$\begin{aligned}
\dot{X} &= Z\omega_y - \frac{vZ}{f}\omega_z + T_x \\
\dot{Y} &= \frac{uZ}{f}\omega_z - Z\omega_x + T_y \\
\dot{Z} &= \frac{vZ}{f}\omega_x - \frac{uZ}{f}\omega_y + T_z
\end{aligned} \tag{2.15}$$

Seja  $\mathbf{P}$  um ponto no efector terminal e  $\dot{\mathbf{P}}$  a sua velocidade instantânea. Se  $\mathbf{P}$  projectar em  $\mathbf{c} = [u \ v]^T$  então será induzida na imagem uma velocidade  $\dot{\mathbf{c}}$  que pode ser determinado por (2.16).

$$\dot{\mathbf{c}} = \begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = \mathbf{J}_{\text{img}} \begin{bmatrix} T_x \\ T_y \\ T_z \\ \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \tag{2.16}$$

Utilizando (2.11) e (2.15) e considerando  $f_x = f K_x$ ,  $f_y = f K_y$ ,  $u_0 = \frac{u - C_x}{f_x}$  e  $v_0 = \frac{v - C_y}{f_y}$  onde  $(K_x, K_y)$  é a dimensão dos pixel/metro da imagem e  $(C_x, C_y)$  a posição do centro da imagem, obtém-se a matriz que representa o Jacobiano da imagem,  $\mathbf{J}_{\text{img}}$ :

$$\mathbf{J}_{\text{img}} = \begin{bmatrix} \frac{f_x}{Z} & 0 & -\frac{u_0}{Z} & -\frac{u_0 v_0}{f_y} & \frac{f_x^2 + u_0^2}{f_x} & -v_0 \frac{f_x}{f_y} \\ 0 & \frac{f_y}{Z} & -\frac{v_0}{Z} & -\frac{f_y^2 + v_0^2}{f_y} & \frac{v_0 v_0}{f_x} & u_0 \frac{f_y}{f_x} \end{bmatrix} \tag{2.17}$$

### 2.2.3 Jacobiano Completo

A multiplicação do Jacobiano da Imagem pelo Jacobiano do Manipulador (Jacobiano completo) permite relacionar coordenadas de junta com os parâmetros característicos da imagem. A relação entre os diferentes Jacobianos poderá ser observado na Figura 2.4.

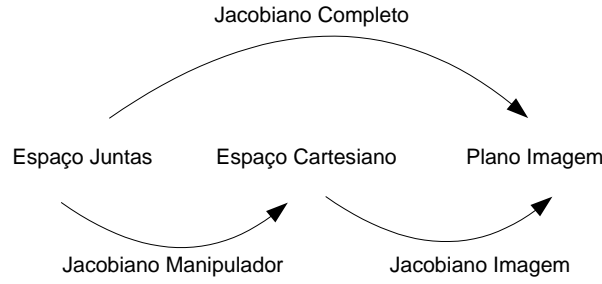


Figura 2.4- Relação entre os diversos Jacobianos: Manipulador, Imagem e Completo

Utilizando as equações (2.7) e (2.16) é possível obter a relação do Jacobiano completo através de:

$$\begin{bmatrix} \dot{\cdot} \\ \dot{\mathbf{u}} \\ \dot{\cdot} \\ \dot{\mathbf{v}} \end{bmatrix} = \mathbf{J}_{\text{img}} \mathbf{J}_M \dot{\boldsymbol{\theta}} = \mathbf{J} \dot{\boldsymbol{\theta}} \quad (2.18)$$

onde  $\mathbf{J}$  é o Jacobiano completo.

Como o nosso objectivo é calcular a cinemática inversa, o que se pretende é utilizar a seguinte equação:

$$\dot{\boldsymbol{\theta}} = (\mathbf{J})^{-1} \begin{bmatrix} \dot{\cdot} \\ \dot{\mathbf{u}} \\ \dot{\cdot} \\ \dot{\mathbf{v}} \end{bmatrix} \quad (2.19)$$

Só é possível calcular a inversa de uma matriz,  $\mathbf{J}$ , se ela for uma matriz quadrada e não singular.

Se  $\mathbf{J}$  for uma matriz não quadrada, de dimensões  $n \times m$ ,  $\mathbf{J}^{-1}$  não existe e através da computação da solução de erro mínimos quadráticos é possível obter uma solução genérica dada por:

$$\dot{\boldsymbol{\theta}} = \mathbf{J}^+ \begin{bmatrix} \dot{\cdot} \\ \dot{\mathbf{u}} \\ \dot{\cdot} \\ \dot{\mathbf{v}} \end{bmatrix} = \mathbf{J}^+ \begin{bmatrix} \dot{\cdot} \\ \dot{\mathbf{u}} \\ \dot{\cdot} \\ \dot{\mathbf{v}} \end{bmatrix} + (\mathbf{I} - \mathbf{J}^+ \mathbf{J}) \mathbf{b} \quad (2.20)$$

sendo  $\mathbf{J}^+$  a pseudo inversa de  $\mathbf{J}$  e  $\mathbf{b}$  um vector arbitrário ( $m \times 1$ ). A solução que minimiza o

erro quadrático resulta num valor para  $\dot{\boldsymbol{\theta}}$  que minimiza a norma  $\left\| \begin{bmatrix} \dot{\cdot} \\ \dot{\mathbf{u}} \\ \dot{\cdot} \\ \dot{\mathbf{v}} \end{bmatrix} - \mathbf{J} \dot{\boldsymbol{\theta}} \right\|$ .

Consoante a relação entre  $m$  e  $n$ , a pseudo inversa terá as seguintes particularidades[20]:

- $m > n$ : é possível definir a matriz pseudo-inversa à esquerda,  $\mathbf{J}_L^+$ ,

$$\mathbf{J}_L^+ = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \quad (2.21)$$

- $n > m$ : é possível definir a matriz pseudo-inversa à direita,  $\mathbf{J}_R^+$ ,

$$\mathbf{J}_R^+ = \mathbf{J} (\mathbf{J} \mathbf{J}^T)^{-1} \quad (2.22)$$

Vamos primeiro considerar o caso  $m > n$ , ou seja admitindo que há mais parâmetros característicos de imagem do que graus de liberdade. Neste caso existem  $m-n$  parâmetros característicos de imagem redundantes, o que resulta numa série de equações inconsistentes. Neste caso a pseudo inversa apropriada é dada por  $\mathbf{J}^+ = \mathbf{J}_L^+$ , através da equação (2.21). O termo  $(\mathbf{I} - \mathbf{J}_L^+ \mathbf{J})\mathbf{b}$  é nulo e a solução aproximada é dada por:

$$\dot{\boldsymbol{\theta}} = \mathbf{J}_L^+ \begin{bmatrix} \cdot \\ \mathbf{u} \\ \cdot \\ \mathbf{v} \end{bmatrix} \quad (2.23)$$

Se  $m < n$ , não se observa suficientes parâmetros característicos de imagem de modo a observar a velocidade do objecto, isto é, existem componentes de velocidade do objecto que não podem ser observadas. Neste caso a pseudo inversa apropriada é dada por  $\mathbf{J}^+ = \mathbf{J}_R^+$  e calculada através da equação (2.22). O termo  $(\mathbf{I} - \mathbf{J}_R^+ \mathbf{J})\mathbf{b}$  não é nulo e todos os vectores do tipo  $(\mathbf{I} - \mathbf{J}_R^+ \mathbf{J})\mathbf{b}$  caem no espaço nulo de  $\mathbf{J}$  e correspondem aos componentes em que a velocidade do objecto não é observável [3] [21].

## 2.2.4 Características da Imagem

Uma característica da imagem, (*image feature*), é qualquer característica estrutural que possa ser extraída da imagem, como por exemplo um contorno. Tipicamente corresponde à projecção de algum objecto físico no plano da imagem da câmara.

Define-se parâmetro característico da imagem como qualquer quantidade real que possa ser calculada a partir de uma ou mais características de uma imagem. Alguns



exemplos de parâmetros característicos da imagem, usados no seguimento visual, são a distância entre dois pontos presentes na imagem, a orientação da linha que une estes dois pontos, a área de uma superfície projectada na imagem, e o centro de área de uma superfície projectada na imagem [2].

Uma boa característica da imagem é aquela que pode ser localizada sem ambiguidades em diferentes vistas da mesma cena. O mais comum é usar as coordenadas de um ponto característico ou o centro de área [4].

Se utilizarmos imagens binárias, cada pixel (*picture element*) admite apenas dois valores possíveis, zero ou um. A vantagem de se trabalhar com imagens binárias tem a ver com o seu armazenamento ser mais leve e o seu processamento ser mais fácil.

Considerando imagens binárias, a função característica  $b(x,y)$  toma o valor 1 para pontos do objecto e zero para os restantes (*background*). Uma imagem binária pode ser obtida através de um valor de limiar (*threshold*); a operação de limiar (*thresholding*) define o valor 0 para a função característica onde a intensidade for maior que o limiar e 1 para os restantes casos.

A área de um objecto pode ser obtida por:

$$A = \iint_I b(x, y) dx dy \quad (2.24)$$

onde a integração é feita ao longo de toda a imagem I [11]. Se a imagem possuir mais que um objecto, o valor obtido por (2.24) será a área total.

A questão que se põe de seguida é a de como determinar a posição do objecto na imagem. Como o objecto não é dado por um simples ponto, é usual escolher-se o centro da área como representativo da posição desse objecto.

O centro de área é o centro de massa da figura com a mesma forma e com massa constante por unidade de área. O centro de massa é o ponto onde toda a massa do objecto pode ser concentrada, sem que com isso se altere o momento de primeira ordem relativamente a qualquer dos eixos.

Numa imagem 2D, o momento em torno de x é dado por [11][18]:

$$\bar{x} \iint_I b(x, y) dx dy = \iint_I x b(x, y) dx dy \quad (2.25)$$

e o momento em torno de y é dado por:

$$\bar{y} \iint_I b(x, y) dx dy = \iint_I y b(x, y) dx dy \quad (2.26)$$

onde  $(\bar{x}, \bar{y})$  é a posição do centro de área. Os integrais que aparecem à esquerda das equações determinam a área, tal como referido em (2.24) pelo que para calcular  $(\bar{x}, \bar{y})$  se fará:

$$\bar{x} = \frac{\iint_I x b(x, y) dx dy}{A} \quad (2.27)$$

$$\bar{y} = \frac{\iint_I y b(x, y) dx dy}{A}$$

onde  $A$  simboliza o momento de ordem zero de  $b(x, y)$ .

### 2.2.5 Segmentação de Imagens

O processo de segmentação de imagens consiste em separar a imagem nos seus diversos componentes, e assim separar os diversos objectos que nela existem. Esta separação consiste em agrupar todos os pixels, que pertencem a um dado objecto ou região, através de um critério de semelhança como por exemplo uma medida, brilho, etc.

Regra geral as imagens possuem ruído do tipo impulsivo (*salt & pepper*), sendo necessário assegurar a coerência espacial das regiões. Através de filtros de Max/Mínimo, ou da média, ou da mediana, é possível eliminar esse ruído [18].

## 2.3 Arquitectura de Seguimento Visual

As arquitecturas de controlo podem ser divididas em dois métodos: “sistema dinâmico vê-e-move” (*dynamic look-and-move*) e “sistema de seguimento visual directo” (*direct visual servo*) [3][4]. O primeiro utiliza o controlador do manipulador numa cadeia de controlo interna e, na cadeia de controlo exterior, utiliza informação visual. O segundo não utiliza o controlador do manipulador, e os valores dos binários a aplicar a cada uma das juntas são calculados a partir, exclusivamente, das poses estimadas visualmente e da pose (posição e orientação) do manipulador. A maior parte dos sistemas adoptam o primeiro método, isto porque muitos manipuladores já têm uma *interface* que permite dar comandos incrementais de posição ou velocidade cartesiana, o que simplifica a construção do sistema de controlo por visão.

O sistema de controlo pode ainda ser *baseado na posição* ou *baseado na imagem* [3].

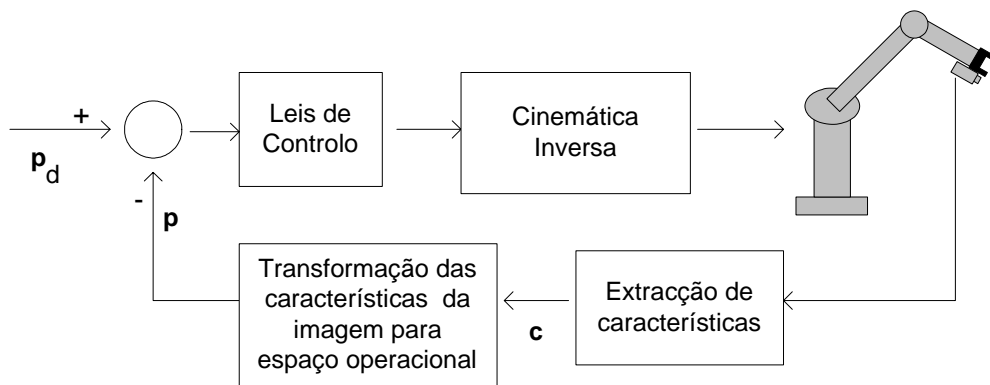


Figura 2.5 - Sistema de Controlo baseado na posição

No sistema de controlo baseado na posição, representado na Figura 2.5, as características são extraídas da imagem e usadas em conjunção com o modelo geométrico do objecto a manipular e o modelo da câmara, de modo a calcular a distância, no plano de imagem, entre o efector terminal e o objecto. Usando esses dados, o erro entre a posição corrente ( $\mathbf{p}$ ) e a posição desejada ( $\mathbf{p}_d$ ) para o manipulador é definido no espaço operacional. Este sistema apresenta uma desvantagem, que é estar muito dependente dos parâmetros resultantes da calibração e da cinemática do sistema. A precisão do movimento depende directamente da qualidade da calibração da câmara e da precisão do manipulador. Como vantagem, este sistema apresenta a possibilidade de descrever as tarefas de posição em coordenadas cartesianas.

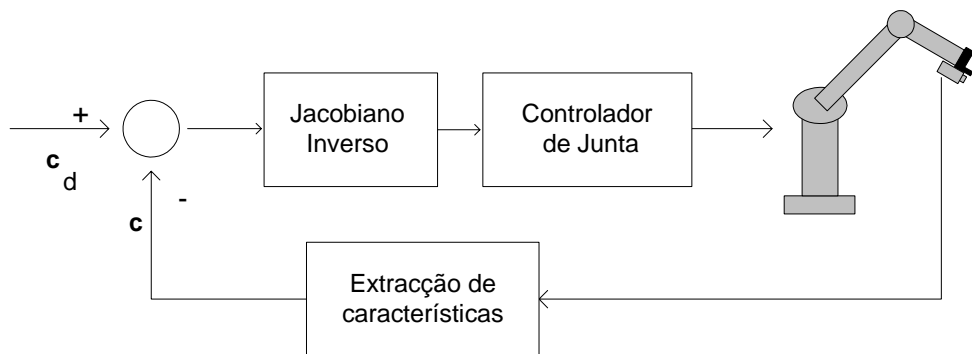


Figura 2.6 - Sistema de Controlo baseado na imagem

A tarefa do seguimento visual baseado na imagem, apresentada na Figura 2.6, consiste em controlar a posição e orientação do braço manipulador, utilizando informação visual

obtida através das características da imagem,  $\mathbf{c}_d$  e  $\mathbf{c}$ , respectivamente vector de parâmetros característicos da posição desejada e da posição actual do efector terminal.

O modelo baseado na imagem é mais apropriado nas tarefas em que não se dispõe de nenhum modelo geométrico do objecto a manipular. Os movimentos do manipulador causam mudanças na imagem observada pelo sistema de visão e é essa informação que intervém na cadeia de controlo.

Este modelo necessita de cálculos mais simples de computação e elimina a necessidade do modelo de calibração da câmara; contudo o sistema não é linear e é altamente acoplado.

Se o sistema observar em simultâneo o objecto e o efector terminal, diremos que estamos em presença de um sistema com o efector terminal em cadeia fechada (*ECL – endpoint closed-loop*), caso contrário o sistema é denominado de sistema com o efector terminal em cadeia aberta (*EOL – endpoint open-loop*)[4].

## 2.4 Cinemática do erro

O sistema pode utilizar uma câmara fixa ou uma câmara montada no efector terminal. Em qualquer dos casos, o movimento do manipulador causa mudanças na imagem, sendo estas observadas pelo sistema de visão.

Num sistema de controlo baseado na imagem, o sinal de erro é definido directamente em termos da informação visual obtida através das características da imagem[3], tal como se pode observar na Figura 2.7.

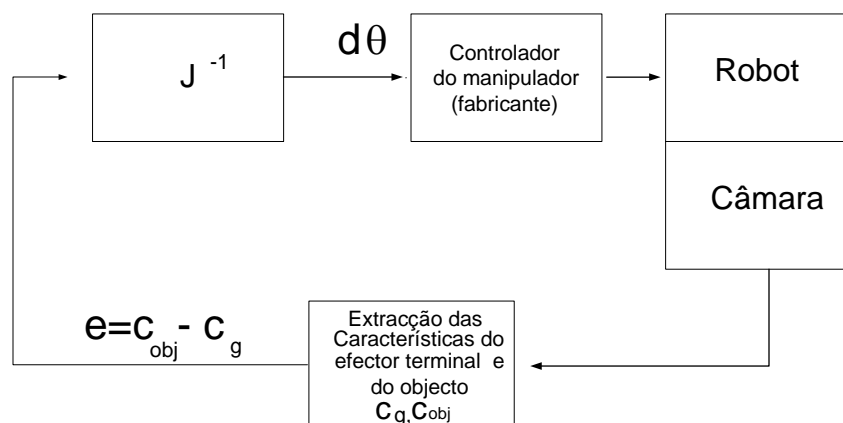


Figura 2.7 – Sistema de Controlo baseado na imagem

O erro é definido no espaço de imagem, e o controlo do manipulador é definido no espaço de coordenadas de juntas. É necessário estabelecer a relação entre as mudanças observadas no espaço de imagem e as respectivas mudanças de posição do manipulador. O Jacobiano completo encarrega-se desta relação;

A posição do objecto a ser manipulado, na imagem, corresponde ao vector de parâmetros característicos da imagem  $\mathbf{c}_{obj}=[u_{obj} \ v_{obj}]^T$ , a que corresponde  $\mathbf{p}_{obj}$ , nas coordenadas do mundo. A posição do efector terminal na imagem corresponde ao vector de parâmetros característicos da imagem  $\mathbf{c}_g=[u_g \ v_g]^T$  e a  $\mathbf{p}_g$ , nas coordenadas do mundo. O erro é dado por  $\mathbf{e} = \mathbf{c}_{obj} - \mathbf{c}_g$ , tal como se pode ver na Figura 2.8. O objectivo do sistema de controlo é tornar  $\mathbf{e} = 0$ .

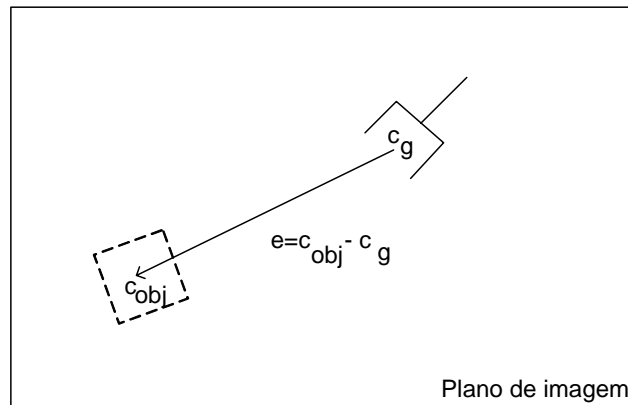


Figura 2.8 – Diagrama do plano de imagem representando o objecto, o efector terminal e a respectiva distância

Através do Jacobiano da Imagem,  $\dot{\mathbf{c}} = \mathbf{J}_{img} \begin{bmatrix} \mathbf{T} \\ \mathbf{\Omega} \end{bmatrix}$ , é possível descrever como os parâmetros característicos da imagem mudam com respeito à pose do manipulador. E, através do Jacobiano do Manipulador,  $\begin{bmatrix} \mathbf{T} \\ \mathbf{\Omega} \end{bmatrix} = \mathbf{J}_M \dot{\boldsymbol{\theta}}$ , é possível estabelecer a relação entre a variação infinitesimal das juntas e o movimento infinitesimal do efector terminal. Relacionando os dois jacobianos, obtém-se o jacobiano completo,  $\dot{\mathbf{c}} = \mathbf{J}_{img} \mathbf{J}_M \dot{\boldsymbol{\theta}} = \mathbf{J} \dot{\boldsymbol{\theta}}$ .

O Jacobiano completo (que inclui os Jacobianos do manipulador e da imagem) permite estabelecer a relação entre os diferenciais necessários a dar a cada uma das variáveis de junta e a velocidade de um ponto na imagem. O que se pretende calcular neste trabalho é o inverso do Jacobiano Completo.

O sistema de controlo, baseado na imagem, observa as variações da velocidade de um ponto na imagem e pretende-se saber quais os diferenciais a dar a cada uma das variáveis de junta para conseguir deslocar o manipulador para o ponto pretendido, utilizando para tal o jacobiano inverso completo,  $\dot{\boldsymbol{\theta}} = (\mathbf{J})^{-1} \dot{\mathbf{c}}$ .

Se a função de erro for dada por  $\mathbf{e} = \mathbf{c}_{obj} - \mathbf{c}_g$ , a lei de controlo proporcional é dada por,

$$\dot{\boldsymbol{\theta}} = \mathbf{K} (\mathbf{J})^{-1} \mathbf{e} \quad (2.28)$$

onde  $\mathbf{K}$ , é uma constante de ganho, de dimensões apropriadas[3]. Para as situações em que a matriz do Jacobiano é não quadrada, utilizar-se-á as técnicas descritas em 2.2.3.

Nesta Tese utiliza-se redes neuronais que aprendem a correspondência de deslocamentos do plano de imagem no espaço de juntas, evitando deste modo o processo tradicional de calibração da câmara e a necessidade de conhecer a priori a cinemática do manipulador.

Na próxima Secção proceder-se-á ao estudo teórico das redes neuronais e como utilizá-las no exemplo concreto do cálculo do Jacobiano completo inverso.

### **3 Estimação do Jacobiano através da Rede Neuronal**

Para que a Rede Neuronal aprenda o Jacobiano completo é necessário fornecer um conjunto de dados de treino. Após proceder-se ao treino, a rede está pronta a ser utilizada e, perante um novo conjunto de dados de entrada, irá produzir os valores de saída adequados.

Ao introduzir a rede neuronal na cadeia de controlo, evita-se proceder ao processo tradicional de calibração da câmara, processo este responsável por estabelecer a correspondência entre pontos do espaço cartesiano para pontos do plano de imagem. Neste processo tradicional de calibração a matriz é calculada para uma dada posição e orientação da câmara. Quando se altera a colocação da câmara ter-se-á de proceder a novos cálculos, resultando o processo moroso. Ao utilizar a Rede Neuronal, pode-se colocar a câmara numa qualquer posição e orientação, porque o primeiro passo a dar é treinar a rede. Neste conjunto de dados de treino já está encapsulada a informação da localização e orientação da câmara, a cinemática inversa do manipulador, e o Jacobiano de imagem.

Este tema será de novo discutido nos Capítulos 4 e 5 quando se analisar em detalhe a arquitectura de seguimento visual implementada.

#### **3.1 Caracterização das Redes Neurais**

As Redes Neurais Artificiais (RNA) têm sido utilizadas, por exemplo, na modelação de memória associativa, reconhecimento de padrões, representação de funções, previsão de séries temporais, optimização, etc. Têm sido aplicadas em áreas muito diversas, como por exemplo: física, controlo de processos industriais, robótica, indústria aeroespacial, telecomunicações, electrónica, indústria automóvel, medicina, reconhecimento de texto, imagem e voz, etc [7].

As redes neuronais são compostas por elementos simples que operam em paralelo. Estes elementos foram inspirados pelo sistema nervoso biológico.

O córtex cerebral é constituído por unidades celulares independentes designadas por neurónios, que podem comunicar entre si através das ligações sinápticas ou sinapses. A comunicação é unidireccional, através de sinais eléctricos e químicos, e de forma descontínua (impulsos). Sempre que um certo potencial de activação é ultrapassado em resultado da recepção de um ou mais sinais nervosos nas sinapses, desencadeia-se a comunicação. Estes sinais nervosos são, em geral, amplificados (ou pesados) de forma diferenciada, ao atravessar as sinapses do neurónio.

As Redes Neurais Artificiais (RNA) são, basicamente, modelos matemáticos para processar a informação; estas têm características comuns com as redes neuronais biológicas. As RNA foram desenvolvidas como uma generalização do modelo matemático da cognição humana, baseado na admissão de que [8]:

- A informação é processada nos elementos mais simples denominados neurónios.
- Os sinais são passados entre os neurónios, através das conexões.
- Cada conexão tem associado um peso, que numa rede neuronal típica multiplica o sinal a transmitir.
- Cada neurónio faz apelo a uma função de activação, ao sinal a transmitir, de modo a determinar o sinal de saída.

O primeiro modelo matemático de um neurónio foi desenvolvido em 1943 pelo neurofisiologista Warren McCulloch e pelo matemático Walter Pitts, baseado no conhecimento da neurologia. Em 1958, Rosenbaltt desenvolveu o perceptrão (é uma unidade que calcula o sinal de saída através de uma combinação linear das variáveis de entrada). Em 1974, Paul Werbos desenvolveu o método de retropropagação do erro. Muitos anos se passaram até que este método fosse popularizado; contudo, nos dias de hoje, é o método mais utilizado e conhecido[8].

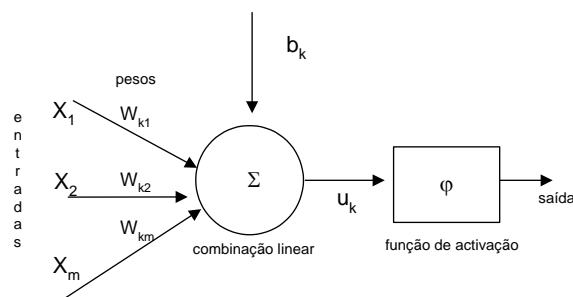


Figura 3.1 – Modelo genérico de um neurónio artificial



De um modo mais formal, as redes neuronais podem ser apresentadas por um grafo orientado, tal como se apresenta na Figura 3.1, que possui os seguintes componentes básicos:

- Um conjunto de sinapses ou ligações caracterizadas por um peso  $w_{kj}$  que se multiplica pelo sinal  $x_j$  de entrada na sinapse  $j$  do neurónio  $k$ ;
- Um módulo somador que soma as entradas pesadas;
- Um valor real  $b_k$  designado por desvio (*bias*);
- Uma função de activação que modula a amplitude de saída do neurónio.

Os nós são os neurónios e os arcos as sinapses. Neurónios de entrada são aqueles que não têm sinapses dirigidas para os mesmos, e neurónios de saída são aqueles que não possuem sinapses de saída.

As funções de activação mais comuns são:

- degrau,
- lineares,
- gaussianas
- trigonométricas.

As redes neuronais artificiais podem ser constituídas por diversas camadas. A primeira é designada por camada de entrada e a última por camada de saída. Os neurónios constituintes destas duas camadas designam-se por neurónios periféricos. As camadas intermédias designam-se por camadas escondidas.

Se a informação se dirige num só sentido, da entrada para a saída, sem retroacção (*feedback*), i.e. a saída de uma camada não afecta a mesma, a rede designa-se por rede de alimentação directa (*feedforward network*) [12]. Se a informação se desloca nas duas direcções, introduzindo retroacções, então temos as redes recorrentes (*feedback networks*).

Uma vez definida a rede, esta tem que ser treinada mediante o processo de aprendizagem. Normalmente isso significa modificar os pesos de modo a obter a solução óptima.

Todos os métodos de aprendizagem de redes adaptativas, que permitem a alteração dos seus pesos, podem ser classificados em duas categorias [8]:

- **Aprendizagem com supervisão** – é o caso mais típico. O treino é feito através de um conjunto de dados de entrada do qual se conhece o respectivo conjunto de saída. Os pesos são ajustados de acordo com os algoritmos de aprendizagem, com o objectivo da convergência do erro, i.e. a minimização do erro entre o valor desejado e o valor calculado. O objectivo desta aprendizagem consiste em determinar os pesos que minimizam este erro. Um dos algoritmos mais utilizados é baseado nos erros mínimos quadráticos (*least mean square*), sendo conhecida a sua generalização como retropropagação do erro (*error backpropagation*)[26].
- **Aprendizagem sem supervisão** – não utiliza nenhum método externo de ensinamento e é baseado exclusivamente na informação local. Alguns algoritmos de aprendizagem sem supervisão são: de aprendizagem por reforço (*reinforcement learning*), de aprendizagem associativa (*associative learning*), e de aprendizagem competitiva (*competitive learning*).

### 3.2 Algoritmo de Retropropagação do Erro

Quando se pretende treinar a rede neuronal para desempenhar uma tarefa, deve-se ajustar os pesos de cada nó de forma que o erro, entre o valor desejado e o valor actual, seja reduzido. O objectivo é treinar a rede de forma a atingir um equilíbrio entre a capacidade de responder correctamente a um padrão de entrada que foi usado no treino, isto é, a capacidade de *memorizar*, e a aptidão de dar uma resposta razoável a um outro qualquer conjunto de dados, ou seja, ser capaz de *generalizar*.

A rede neuronal de Retropropagação do Erro é a mais generalista e a mais utilizada. Utiliza o gradiente descendente para treinar a rede. A técnica do gradiente descendente é análoga ao processo de minimização do erro. O Algoritmo de Retropropagação do Erro aprende a gerar um mapa do espaço de entrada para o espaço de saída, através da minimização do erro entre a actual saída produzida pela rede e a saída desejada. O processo de aprendizagem começa pela apresentação dos dados de entrada à rede. Estes dados são propagados através de toda a rede até que seja produzida uma saída. O Algoritmo de Retropropagação do Erro determina o erro para cada uma das unidades, pelo que cada unidade modifica os seus pesos de conexão na direcção de redução do sinal de erro, e o processo repete-se para o próximo dado [22].

O número de camadas escondidas tem a ver com a aplicação mas, regra geral, só uma camada é suficiente para a maior parte dos casos [26]. Na Figura 3.2, visualiza-se a arquitectura de uma rede com uma camada escondida.

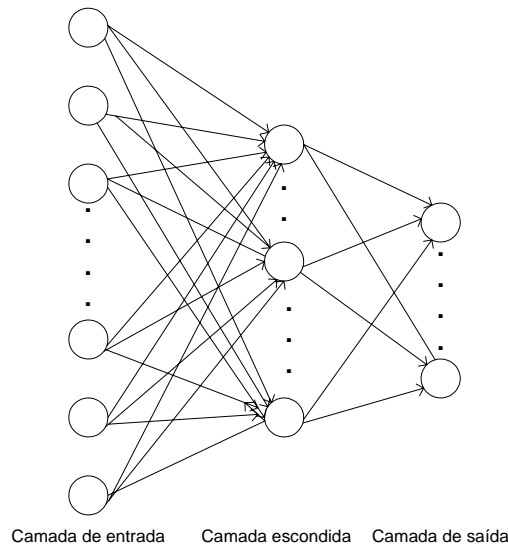


Figura 3.2 – Arquitectura de uma rede de alimentação directa

Cada unidade recebe estimulação de uma série de outras unidades, com excepção da camada de entrada que só recebe de uma fonte. O sinal recebido, por cada unidade, é multiplicado (modulado) por um peso de conexão; cada sinal modulado é integrado (soma algébrica) de modo a produzir o sinal de estimulação da unidade [1] [13].

O treino da rede envolve três passos:

- a rede de alimentação directa (*feedforward network*) para o conjunto de dados de entrada,
- a retropropagação do erro,
- e o ajuste dos pesos.

O valor do nó na camada de saída passa pelos seguintes passos:

- a) Cálculo do total dos pesos associados à entrada  $X_k$ , utilizando para isso a fórmula:  $X_k = \sum_i y_i W_{ik}$ , onde  $y_i$  é o nível de actividade do nó  $k$  na camada anterior.
- b) Cálculo da actividade  $y_i$  usando a soma de todos os pesos. Tipicamente utiliza-se uma função sinusoidal:  $y_k = \frac{1}{1 + e^{-x_k}}$ .

c) Cálculo do erro da rede através da seguinte expressão:  $E = \frac{1}{2} \sum_k (y_k - d_k)^2$ , onde  $d_k$  é o valor desejado.

Depois do treino, a aplicação da rede envolve exclusivamente a fase de alimentação directa. Mesmo que o treino tenha sido lento, uma rede treinada produz rapidamente uma saída. Variações do algoritmo da retropropagação do erro foram desenvolvidas de modo a melhorar a velocidade do processo do treino.

A rede neuronal requer que os dados de entrada e de saída sejam normalizados de modo a ter a mesma ordem de grandeza. Se não houvesse essa normalização dos dados, haveria certas grandezas que teriam mais importância que outras.

### 3.3 Jacobiano Completo através da Rede Neuronal

Neste trabalho, o sistema de controlo dinâmico do tipo “vê-e-move”, baseado na imagem, implementa a lei de controlo estabelecida por (2.28); sendo assim pretende-se observar velocidades do vector de imagem ( $du, dv$ ) e obter os respectivos diferenciais de cada uma das variáveis de junta ( $d\theta$ ).

Neste trabalho, a aproximação da função Jacobiano completo será feito à custa de uma Rede Neuronal, obviando assim toda a problemática estudada no Capítulo 2 relativamente ao cálculo e identificação de parâmetros do Jacobiano e à inversão de matrizes não quadradas. Outra vantagem da utilização da rede neuronal reside na total ausência de conhecimentos da cinemática do manipulador.

Na Figura 3.3 apresenta-se um diagrama de blocos da rede capaz de aprender o diferencial a ser dado a cada uma das variáveis de junta. A rede neuronal, para cada uma das juntas, é constituída por uma camada de entrada com 2 nós ( $u, v$ ) e uma camada de saída com 1 nó ( $d\theta$ ). O Jacobiano completo será constituído por quatro redes idênticas, uma para cada variável de junta.

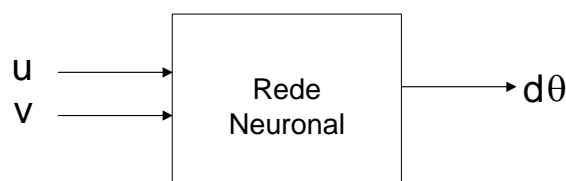
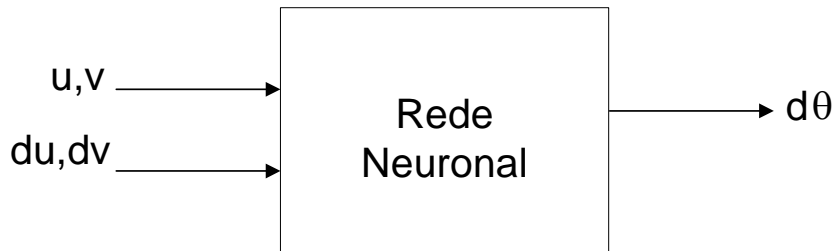


Figura 3.3- Diagrama de Blocos da Rede Neuronal

Houve a necessidade de introduzir a posição actual do manipulador, em termos de coordenadas de imagem, além do vector de velocidade de imagem, de forma a Rede Neuronal conseguir aprender melhor a função Jacobiano completo. Antes de se proceder a esta introdução, a RNA aprendia dados inconsistentes, isto porque para o mesmo vector de entrada havia várias possibilidades para o vector de saída e a RNA tinha dificuldade em aproximar a função pretendida.

O bloco de dados da rede neuronal implementado apresenta-se na Figura 3.4.



*Figura 3.4 – Diagrama de Blocos da Rede Neuronal implementada*

Para que a Rede Neuronal aprenda o Jacobiano completo é necessário fornecer um conjunto de dados de treino. Estes serão constituídos por um conjunto de dados de entrada do qual se conhece o respectivo conjunto de saída. De acordo com o algoritmo de aprendizagem da retropropagação do erro, os pesos são ajustados com vista à minimização do erro entre o valor desejado e o calculado. Após proceder-se ao treino, a rede está pronta a ser utilizada e, para um novo dado de entrada, irá produzir os valores de saída apropriados.

Ao introduzirmos esta rede neuronal na cadeia de controlo, evitamos proceder ao processo tradicional de calibração da câmara, processo este responsável por estabelecer a correspondência entre pontos do espaço cartesiano e pontos do plano de imagem. O processo tradicional de calibração implicaria conhecer um conjunto de pontos do espaço 3D e os respectivos pontos no plano de imagem. Com estes dados obteríamos uma matriz de calibração que seria a responsável no estabelecimento da correspondência do espaço 3D para o plano de imagem. Neste processo tradicional de calibração a matriz é calculada para uma dada posição e orientação da câmara. Quando se alterasse a colocação da câmara ter-se-ia de proceder a novos cálculos, resultando o processo moroso. Ao utilizar-se a RNA, pode-se colocar a câmara numa qualquer posição/orientação porque o primeiro passo a dar é treinar a rede. Neste conjunto de dados de treino já está encapsulada a informação da

localização e orientação da câmara, a cinemática inversa do manipulador e o Jacobiano de imagem.

No Capítulo 4, apresentar-se-á a inserção da rede neuronal, para a estimação do Jacobiano Completo, na arquitectura de seguimento visual implementada.

## 4 Métodos e Algoritmos

No âmbito deste trabalho foi implementado um sistema de controlo do tipo “vê-e-move” baseado na imagem.

Neste Capítulo é realizada a descrição do sistema, sendo apresentado em detalhe cada um dos seus subsistemas, nomeadamente a Rede Neuronal e o sistema de visão responsável pela extracção das características de imagem. Também são aqui fundamentadas as decisões e soluções efectuadas. Por fim é descrito o método experimental utilizado na obtenção dos resultados.

### 4.1 Descrição da Instalação

A arquitectura de seguimento adoptada neste estudo pode ser representada pela Figura 4.1, baseada no esquema da Figura 2.6.

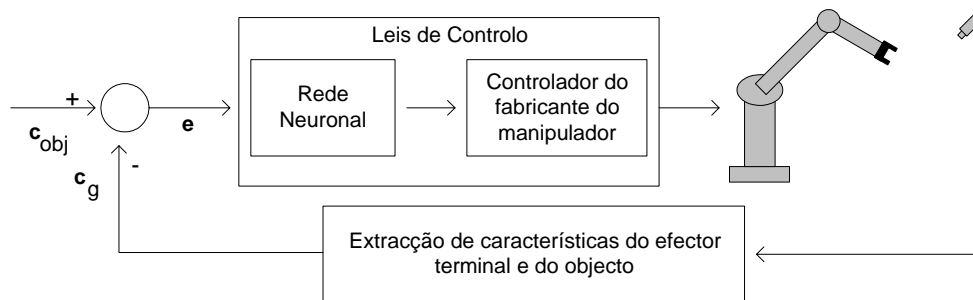


Figura 4.1 - Sistema de Controlo Implementado

O sistema utilizado é composto por 3 subsistemas:

- Sistema de visão constituído por uma câmara Sony XC-ES50, 1 placa de aquisição de imagem, um computador pessoal (PC) com microprocessador Pentium 1.6GHz e 512Mb de memória RAM, responsável pelos algoritmos de processamento e de extracção de características da imagem.

- Sistema de seguimento do terminal do manipulador. Este algoritmo de seguimento também se encontra no PC pertencente ao sistema de visão, e consiste na utilização da rede neuronal, que foi previamente treinada, em lugar do Jacobiano completo.
- Sistema de manipulação constituído por um manipulador Mitsubishi Movemaster RV-E3J e respectivo controlador. A partir do PC do sistema de visão, é possível especificar comandos incrementais directamente ao manipulador.

## 4.2 Sistema de Controlo Utilizado

O sistema de controlo implementado foi o sistema de controlo dinâmico do tipo “vê-e-move” baseado na imagem, tal como indicado no Capítulo 2, e que se apresenta agora em detalhe.

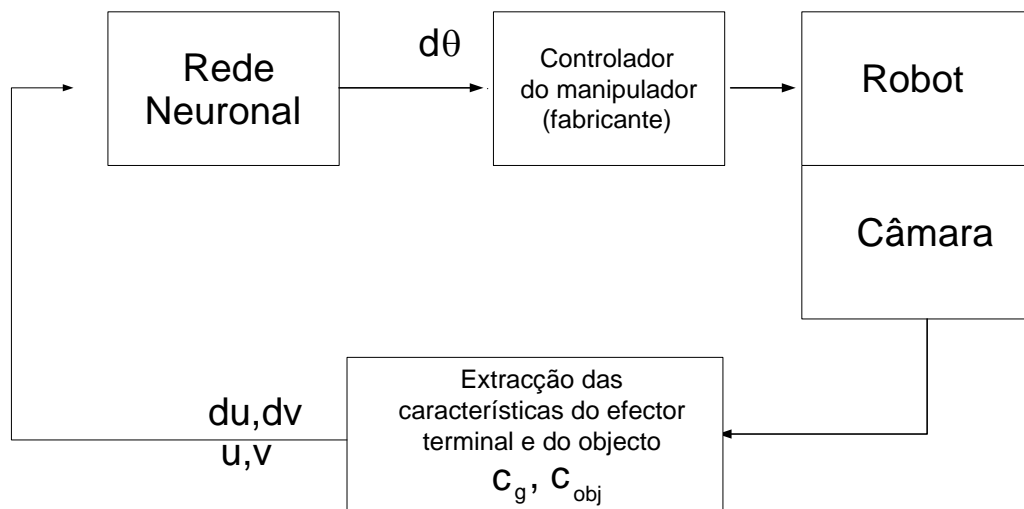


Figura 4.2 – Diagrama de blocos da arquitectura de seguimento visual

O controlo do manipulador é realizado no espaço junta, utilizando para isso o controlador de posição fornecido pelo construtor. O movimento do manipulador causa mudanças na imagem, sendo estas observadas pelo sistema de visão e transmitidas à Rede Neuronal, que se encarregará de dar directamente as referências ao controlador do manipulador.



O erro é definido directamente em termos de características da imagem. A especificação de uma arquitectura de seguimento visual baseado na imagem, envolve a determinação de uma função de erro,  $e$ , de tal modo que quando a tarefa é atingida,  $e = 0$ , o que se obtém utilizando para tal a lei de controlo apresentada na Secção 2.4. Se a posição do efector terminal na imagem corresponder ao vector de parâmetros característicos da imagem  $c_g$  e se o vector de parâmetros característicos da imagem desejado corresponder a  $c_{obj}$ , a função de erro apropriada é  $e = c_{obj} - c_g$ . A lei de controlo é dado por (2.28).

O Jacobiano completo, estabelece a relação entre a derivada dos parâmetros característicos e a velocidade do manipulador.

Devido à dificuldade do cálculo do Jacobiano e da sua inversa, já anteriormente referido, utilizar-se-á uma rede neuronal que será responsável por este mesmo controlo. Para o efeito recorreu-se à *Neural Networks Toolbox* do programa MatLab da MathWorks.

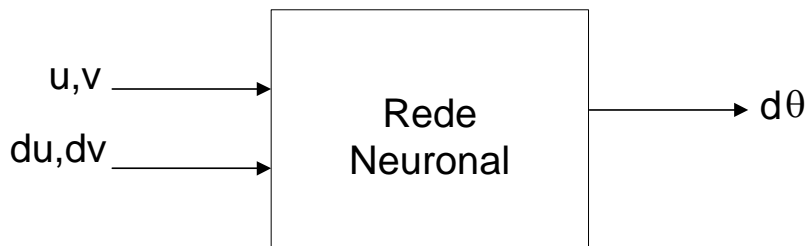


Figura 4.3 – Diagrama de blocos da Rede Neuronal

A RNA utilizada, representada na Figura 4.3, tem como entradas a posição actual do manipulador em termos do vector de parâmetros característicos da imagem ( $u, v$ ) e a velocidade do vector de imagem ( $du, dv$ ) sendo a saída o diferencial de movimento que é necessário realizar em cada uma das juntas de modo a deslocar o manipulador ao objecto ( $d\theta$ ). Os valores de saída do sistema de controlo do manipulador serão aplicados directamente ao controlador do fabricante do manipulador. Este bloco simula o Jacobiano completo apresentado na Secção 2.2.3, evitando deste modo quer o processo de calibração do manipulador quer o da câmara.

Através do estudo da função Jacobiano, apresentado no Capítulo 5, concluiu-se da necessidade de separar o espaço de trabalho por zonas, devido à complexidade do sistema e da sua não linearidade. Para cada zona existe uma rede neuronal.

A RNA utilizada tem uma camada de entrada com 4 nós, uma camada escondida com 3 nós e uma de saída com um nó. Este modelo repete-se para cada uma das variáveis de junta. As funções utilizadas nas duas primeiras camadas são funções trigonométricas e na

camada de saída foi aplicada uma função linear. A função de treino utilizada foi a *trainlm* (algoritmo de otimização Levenberg-Marquardt) para a retropropagação do erro baseado nas segundas derivadas. O treino consiste em 200 épocas para um parâmetro final de erro de 0.01.

O método de aprendizagem aqui utilizado foi aprendizagem com supervisão. Neste método existe a necessidade de se obter um conjunto de dados de entrada do qual se conhece o respectivo conjunto de saída. Os dados foram divididos em três subconjuntos, tal como se pode ver na Figura 4.4:

- treino ( $\frac{2}{4}$  dos dados de entrada),
- validação ( $\frac{1}{4}$  dos dados de entrada) ,
- teste ( $\frac{1}{4}$  dos dados de entrada +  $\frac{1}{4}$  dos dados de treino);

Os dados de treino e de validação servem para a aprendizagem da RNA e o conjunto de teste serve para verificar o funcionamento da rede treinada.

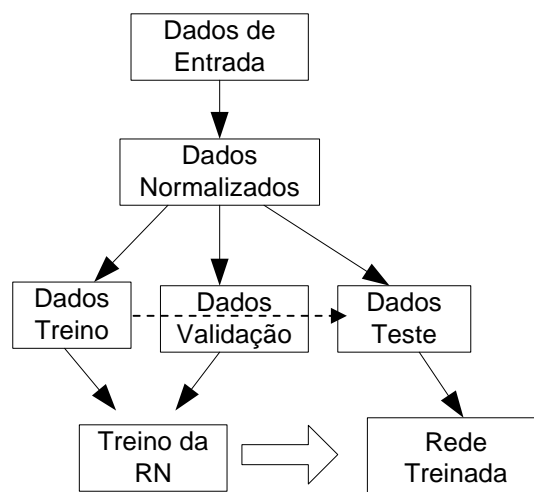


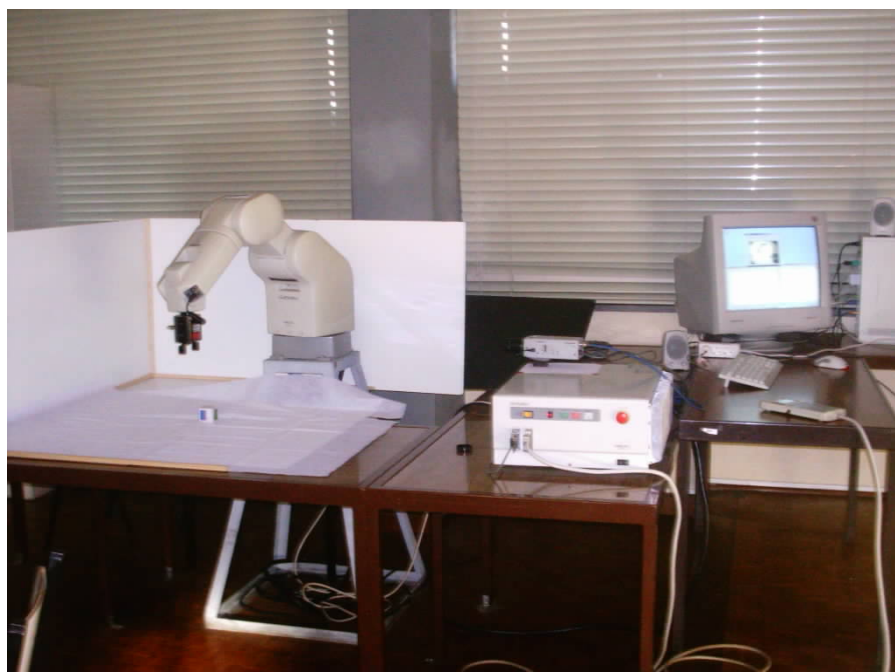
Figura 4.4- Diagrama de Blocos do Conjunto de Dados para treino da Rede Neuronal

Os dados foram retirados dos dados originais com espaçamentos iguais, tendo sido normalizados no intervalo  $[-1,1]$  com a função *premnmx* [7]. O subconjunto de treino é utilizado para calcular os gradientes e actualizar os pesos e os desvios. O erro no subconjunto da validação é monitorizado durante o processo de treino; o erro da validação na fase inicial de treino, normalmente, decresce, assim como o erro do conjunto de treino. Contudo, à medida que as redes se sob reajustam aos dados, o erro no conjunto de validação começa a crescer; quando este erro cresce para além de um número específico de

iterações, então o treino pára. Através desta divisão de dados é possível parar o treino mais cedo. O subconjunto de teste é útil para comparar modelos diferentes, verificar se houve memorização da rede, etc.

### 4.3 O Manipulador Mitsubishi Movemaster RV- E3J

O manipulador utilizado é o apresentado na Figura 4.5. A comunicação com o manipulador faz-se através da porta série do controlador, ligada à porta série do computador de processamento de imagem e cálculos das redes neuronais. O manipulador da Mitsubishi modelo Movemaster RV-E3J tem 5 graus de liberdade e uma capacidade de carga de 3Kgf.



*Figura 4.5- Manipulador Mitsubishi RV- E3J e a câmara utilizada*

Encontram-se no Anexo A os referenciais para este manipulador, assim como os alçados principais. No Anexo B, os parâmetros da matriz de Denavit-Hartenberg. No Anexo C, as matrizes necessárias para o cálculo da Cinemática Directa. No Anexo D, todos os cálculos para a Cinemática Inversa e no Anexo E, os Jacobianos.

Apresenta-se na Tabela 1, em anexo, os dados mais relevantes deste manipulador.

O tipo de linguagem utilizada é específico do MOVEMASTER e possui cerca de 94 comandos. Os comandos mais utilizados indicam-se a seguir.

MJ a b c d e – se assumirmos que as juntas estão nas coordenadas A, B, C, D, E. Este comando vai provocar uma mudança de posição para os valores A+a, B+b, C+c, D+d, E+e, utilizando para isso a interpolação de juntas.

DJ n,a – n representa o número da junta à qual se vai dar um incremento de a graus.

ER n – Através do comando ER é possível inquirir ao controlador a natureza do erro.

HE n – Permite memorizar a posição corrente através de um número n.

MO n – Permite colocar o manipulador numa posição previamente memorizada em n.

GO – Permite abrir o efector terminal.

GC – Permite fechar o efector terminal.

## 4.4 Sistema de Visão

O sistema de visão é constituído por uma única câmara, sendo esta colocada de um modo aleatório, mas de forma a permitir visualizar a mesa de trabalho e o manipulador.

Para facilitar a detecção pelo sistema de visão, colocou-se no terminal do manipulador uma fita colorida vermelha. No objecto foi colocado uma fita verde para facilitar a sua detecção. Uma vez que só se utiliza uma câmara, foi ainda utilizado uma marca fixa, de cor vermelha, de forma a conhecer a profundidade do sistema.

Este sistema é o responsável por extrair as características de imagem; utilizando para tal a segmentação de imagem através da luminosidade e calculando de seguida o centro de área, já discutido na Secção 2.2.4.

As imagens recolhidas pela câmara a cores, estão no formato *RGB (Red, Blue, Green)*, que é constituído por três matrizes em que cada uma delas tem uma componente de cor. Após se ter adquirido a imagem, converte-se as coordenadas de cor da imagem adquirida para o formato *HSV (Hue Saturation Value)* [19], facilitando deste modo a segmentação.

É possível obter uma imagem binária através da operação de limiar. Considerando que se tem uma imagem  $g(x,y)$ , é possível obter a imagem binária,  $b(x,y)$ , para um dado limiar T, através de [10][19][23]:

$$b(x,y) = \begin{cases} 1 & \text{se } g(x,y) > T \\ 0 & \text{se } g(x,y) \leq T \end{cases} \quad (4.1)$$

Através de um valor de limiar apropriado e utilizando matriz de *Hue*, é possível detectar os diversos objectos essenciais a este estudo. Com a operação de limiar obtém-se uma imagem a preto e branco (BW), sendo deste modo possível aplicar as equações (2.27) para o cálculo dos centros de massa, aos diversos elementos essenciais ao controlo deste sistema.

## 5 Resultados

Neste Capítulo serão apresentados e comentados os resultados, simulados e experimentais, relativos a cada um dos subsistemas que compõem o sistema de seguimento visual “vê-e-move”.

A primeira abordagem é o estudo da função Jacobiano completo. Para tal elaboraram-se diversos testes para observar o comportamento da função, função essa com alguma complexidade. Foi através deste estudo que se chegou à conclusão da necessidade da separação das redes neurais pelas variáveis de junta e da separação do espaço de trabalho, por três zonas.

Através de um programa realizado no Matlab, simula-se o sistema “vê-e-move”. A implementação deste sistema consiste na aprendizagem da função Jacobiano, e por fim no sistema de controlo propriamente dito.

Finalmente apresenta-se os resultados obtidos com o manipulador real.

### 5.1 Simulação

Para a extracção dos dados de simulação colocou-se a câmara, relativamente à base do manipulador, nas coordenadas  ${}^B[350\ 700\ 0]^T$  (medidas em mm), e de seguida procedeu-se a uma rotação de  $90^\circ$  em torno de  $X$ . A matriz que relaciona a câmara com o sistema da base é pois dado por:

$${}^C T_B = \begin{pmatrix} 1 & 0 & 0 & -350 \\ 0 & 0 & 1 & 700 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Através da Figura 5.1, é possível visualizar os referenciais do mundo (base do manipulador) e da câmara.

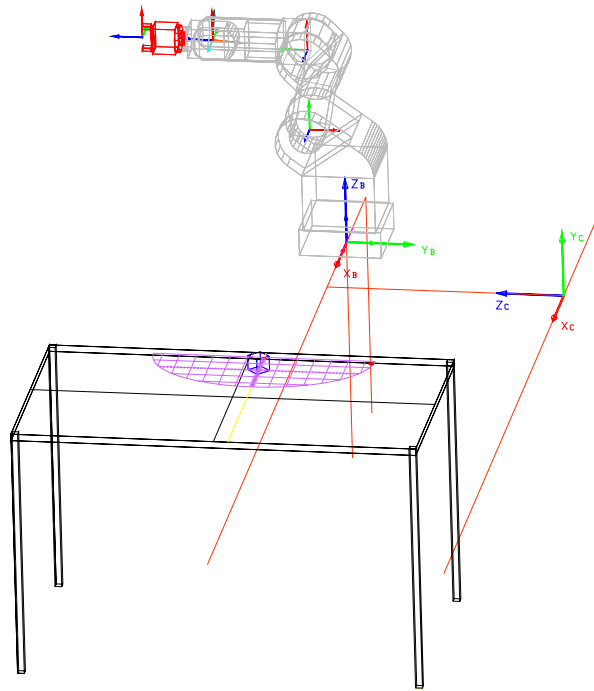


Figura 5.1 – Referencial da Câmera e do Mundo

### 5.1.1 Estudo da Função Jacobiano

Para se realizar o estudo da função Jacobiano completo, procedeu-se à elaboração de diversos testes. Com estes testes pretendia-se analisar a variação das coordenadas de imagem face às variáveis de junta.

Fez-se o estudo da função Jacobiano:

- A. Localmente num único ponto;
- B. Numa zona de trabalho, sendo a junta 1 constante;
- C. Na totalidade da área útil de trabalho.

#### A. Estudo da função Jacobiano localmente num único ponto

Para o teste da Figura 5.2, utilizou-se apenas um ponto do espaço em torno do qual se procedeu a pequenas variações das juntas 2, 3 e 4. A variação em torno do ponto foi de  $\pm 1^\circ$  nas diversas variáveis de junta.

Observa-se que o jacobiano em torno de um ponto apresenta uma superfície regular em qualquer das juntas. É possível verificar a elevada sensibilidade da junta 4 face a variações de  $u$  e de  $v$ .

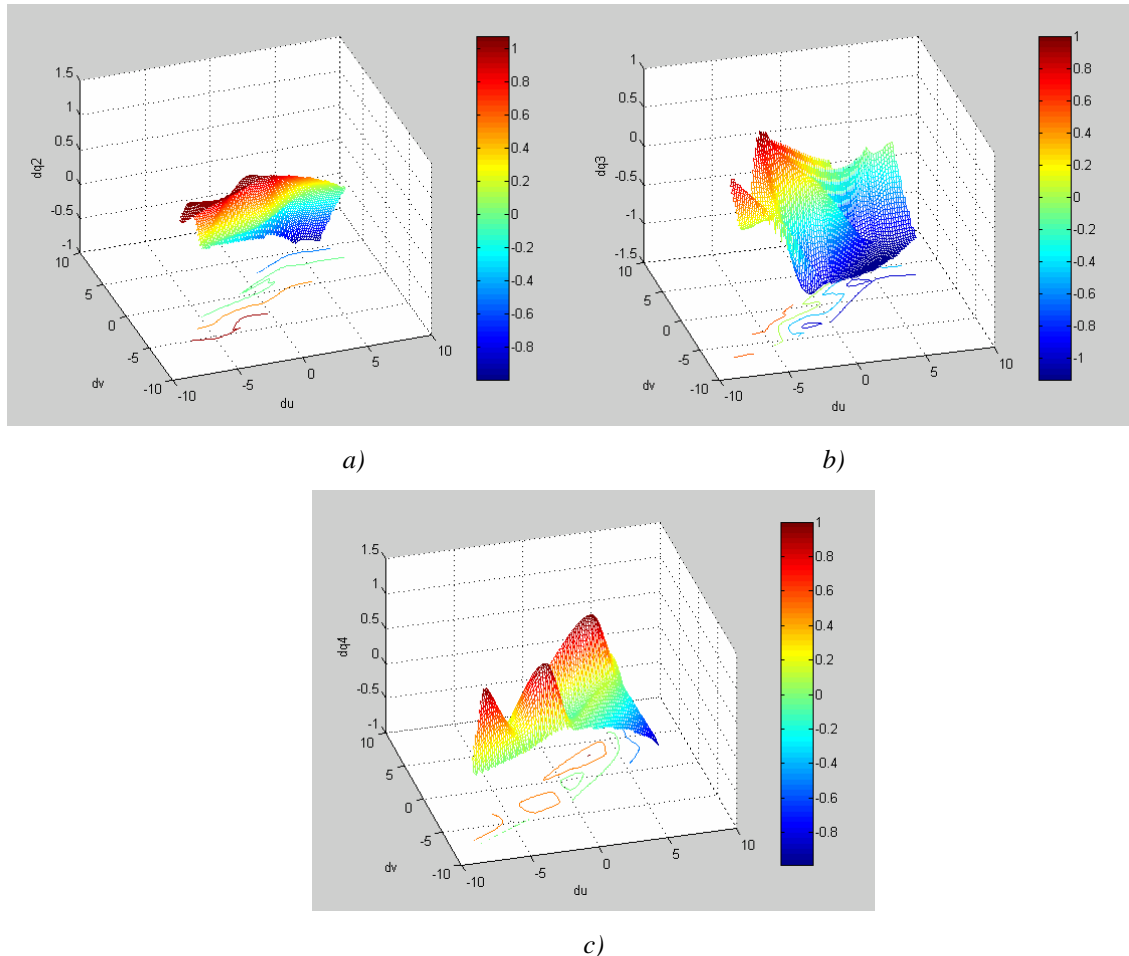


Figura 5.2 – Variações dos ângulos de junta, para um ponto do espaço, em função dos deslocamentos na imagem resultantes; a) Junta 2; b) Junta 3; c) Junta 4

## B. Estudo da função Jacobiano numa zona do espaço de trabalho

Neste teste, utilizou-se vários pontos, em torno dos quais se fez pequenas variações. Todos esses pontos têm em comum o facto do valor da junta 1 ser o mesmo.

No teste da Figura 5.3, utilizou-se uma zona do espaço de trabalho, em torno do qual se procedeu a variações das juntas 2, 3 e 4. A variação em torno do ponto foi de  $\pm 1^\circ$  nas diversas variáveis de junta. Observa-se que o jacobiano em torno de uma zona de trabalho, onde se manteve o valor da junta 1 constante, apresenta uma superfície regular.



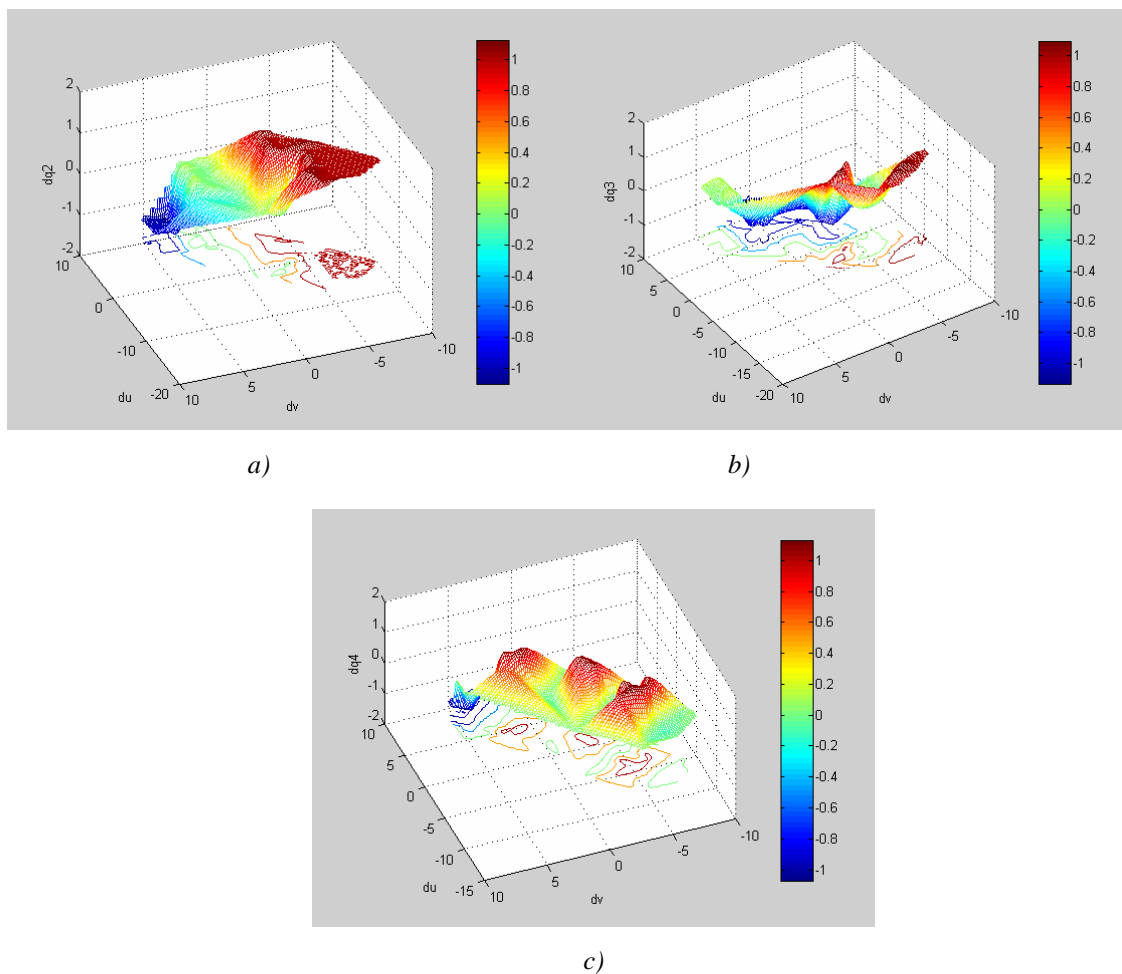


Figura 5.3 – Variações do ângulos de junta em função dos deslocamentos na imagem resultantes, para uma zona do espaço de trabalho; a) Junta 2; b) Junta 3; c) Junta 4

É possível verificar a elevada sensibilidade da junta 4 face a variações de  $u$  e de  $v$ . Sendo assim, num trabalho futuro, poder-se-ia trabalhar com a junta 4 e 5, juntas de orientação, numa única rede onde as variações seriam menores que as restantes juntas. É possível observar a forte dependência da função Jacobiano face ao ponto de funcionamento escolhido.

### C. Estudo da função Jacobiano

O último teste consistiu em colocar o manipulador em diversos pontos do espaço operacional, fazendo de seguida pequenas variações das variáveis de junta 1, 2, 3 e 4 em torno desses pontos. As imagens que se seguem mostram as variações das diversas juntas em função dos deslocamentos na imagem resultantes.

A variação em torno dos diversos pontos foi de  $\pm 1^\circ$  nas diversas variáveis de junta, podendo-se observar os resultados na Figura 5.4. Da esquerda para a direita e de cima para baixo, as imagens correspondem, sucessivamente, às variações efectuadas nas juntas 1, 2, 3 e 4.

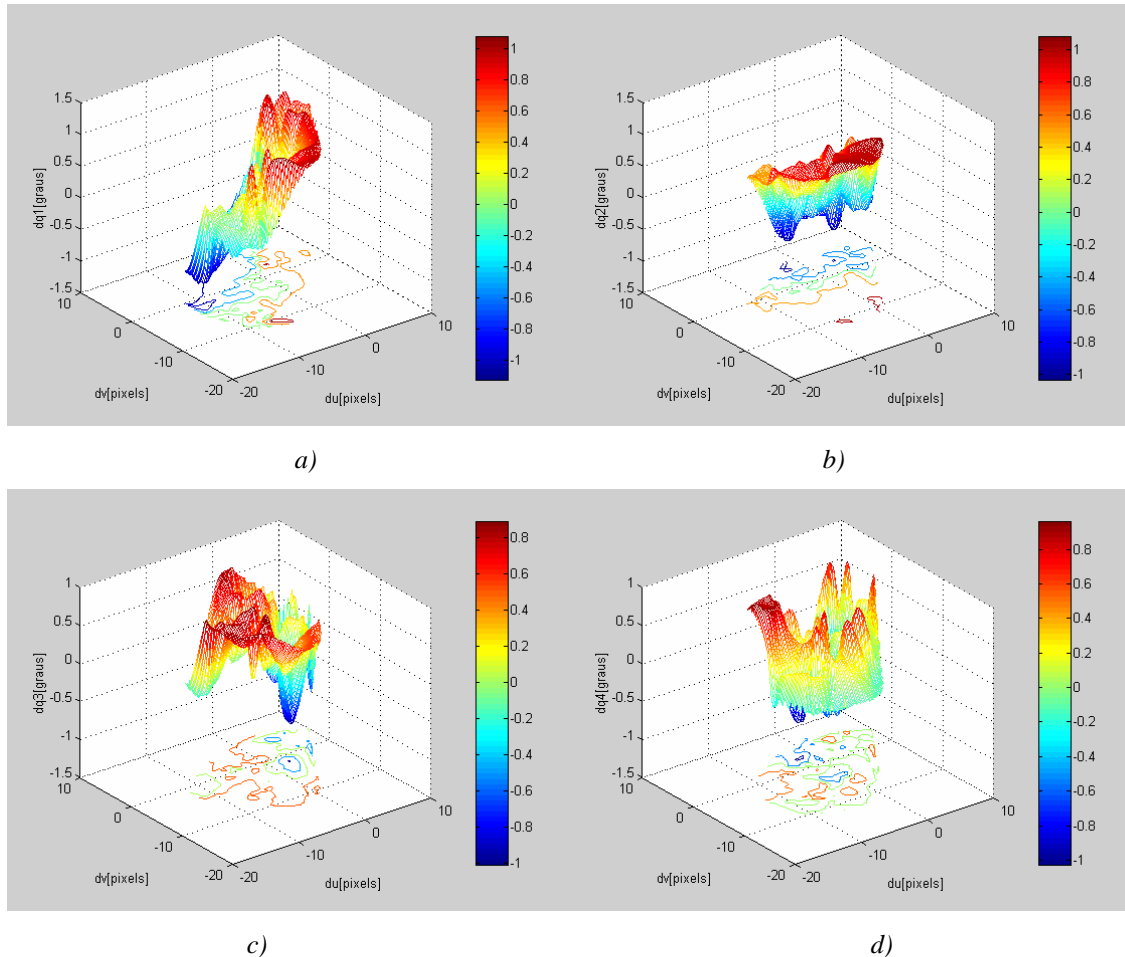


Figura 5.4 – Variações dos ângulos de junta, para diversos pontos, em função dos deslocamentos na imagem resultantes; a) Junta 1; b) Junta 2; c) Junta 3; d) Junta 4

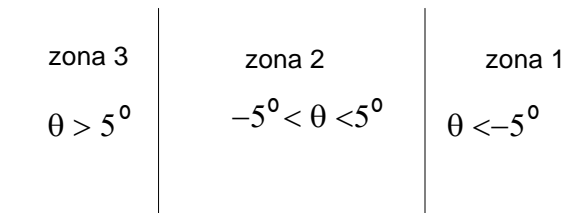
Destas quatro figuras, Figura 5.4, observa-se que o gráfico respeitante à junta 4 não é regular, ao passo que nos gráficos das outras juntas se consegue encontrar zonas de alguma regularidade. A junta 4 é a junta de orientação e observa-se que é muito sensível quer em  $u$  quer em  $v$ , ou seja, uma pequena variação do ângulo de junta 4 provoca uma grande variação na posição em termos da imagem.

Com estes testes chegou-se à conclusão da elevada irregularidade, em geral, das coordenadas de junta face às coordenadas de imagem.

Também se verificou a perda de informação da profundidade, pelo facto de se estar a trabalhar com uma só câmara. E, tal como seria de esperar, constatou-se em certos movimentos do manipulador, tais como os efectuados ao longo do raio de projecção, assim como as rotações efectuadas em torno de um ponto, que não se observa nenhuma variação na imagem. Estas situações ocorrem todas no espaço nulo do Jacobiano.

Através deste estudo se infere também, da necessidade de separar o espaço de trabalho por zonas, cada uma das quais dispendo de uma rede neuronal, devido à complexidade do sistema e da sua não linearidade.

O espaço de trabalho foi dividido em três zonas consoante o valor da variável de junta 1, como se mostra na Figura 5.5; se esse valor for inferior a  $-5^\circ$ , está na zona 1; para o intervalo de  $-5^\circ$  a  $5^\circ$ , a zona 2 e se for superior a  $5^\circ$ , a zona 3.



*Figura 5.5 – Divisão do espaço de trabalho em três zonas distintas*

Como o sistema pretende simular o Jacobiano completo, optou-se por fazer os cálculos teóricos das expressões do Jacobiano do Manipulador e do Jacobiano de Imagem. Pretendia-se, com este estudo, verificar a complexidade do sistema, na sua totalidade, e a sua não linearidade.

Devido ao facto de, neste caso particular, o manipulador ter 5 graus de liberdade, as matrizes Jacobiano não são quadradas, e como tal ter-se-ia, para controlar o sistema, que trabalhar com a pseudoinversa, tal como se explicou no Capítulo 2.

Pela observação das figuras, é possível constatar a complexidade da função que se pretende calcular. Para este caso específico a aproximação da função faz-se através de redes neuronais e não por via analítica. Um aspecto importante neste trabalho, incidiu na verificação da possibilidade da utilização de uma única rede neuronal, ou então em dividir o espaço de trabalho em diversas zonas, cada uma como sendo a aproximação local da função.

Assim o sistema aqui implementado resulta numa mais valia, devido à sua flexibilidade, uma vez que em lugar de se empregar a via analítica, opta-se pelas redes neuronais para aprenderem a função Jacobiano. Por não se seguir a via analítica, não é necessário conhecer a cinemática do manipulador, nem proceder à calibração da câmara.

### 5.1.2 Treino da Rede Neuronal

Terminado este estudo da função Jacobiano, passou-se para outra fase que consistiu em simular o sistema indicado na Figura 4.2. Para simular o Jacobiano completo é necessário conhecer diversos pares de dados (velocidade das coordenadas de junta, velocidade de deslocamento do ponto na imagem). Obtidos os diversos valores, procedeu-se ao treino da rede neuronal.

Foi efectuado um teste em que se provocou perturbações do efector terminal de modo a obter os dados de treino da rede. Estes dados são obtidos através da cinemática directa e da transformada de perspectiva.

Estes dados foram obtidos através da colocação do manipulador em 27 pontos diferentes do espaço operacional e em torno de cada um desses pontos procedeu-se a dez perturbações aleatórias com uma distribuição uniforme. Para cada uma dessas perturbações, nas variáveis de junta, procedeu-se ao cálculo resultante no deslocamento efectuado na imagem. Desde modo construiu-se um ficheiro com os dados de entrada e um outro com os dados de saída da rede.

Na Figura 5.6 observa-se a imagem obtida através do simulador; a mancha azul representa os diversos pontos utilizados no treino.

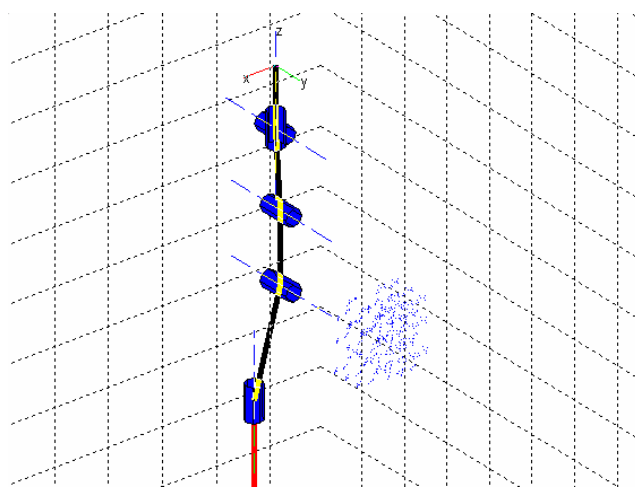


Figura 5.6 – Dados de treino da rede, os pontos marcados a azul representam os pontos utilizados no conjunto de treino da rede

Após a obtenção dos dados, é necessário proceder-se ao treino da rede.

Sendo comuns às quatro redes, os dados de entrada são, respectivamente, as coordenadas do ponto da imagem,  $u$  e  $v$ , e os diferenciais que se fazem em torno desse ponto em coordenadas de imagem,  $du$  e  $dv$ . O dado de saída da rede é o diferencial que é necessário provocar em cada uma das variáveis de juntas,  $d\theta_i$ , de modo a se conseguir o deslocamento na imagem pretendido.

Ilustra-se na Figura 5.7, um exemplo concreto das redes neuronais as variáveis de junta 2 e 4. A junta 5, junta de orientação, não é considerada neste trabalho.

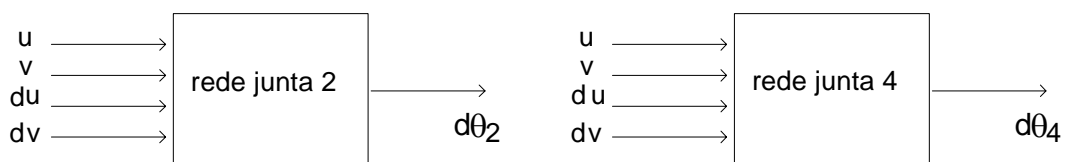


Figura 5.7 – Exemplo do diagrama de Blocos das Redes Neuronais para a junta 2 e 4

Um exemplo de treino da rede, para a junta 2 e 4 mostra-se na Figura 5.8. Foi utilizado o conjunto de validação de modo a parar o treino da rede mais cedo e deste modo não se correr o risco de a rede memorizar a função em vez de a aprender.

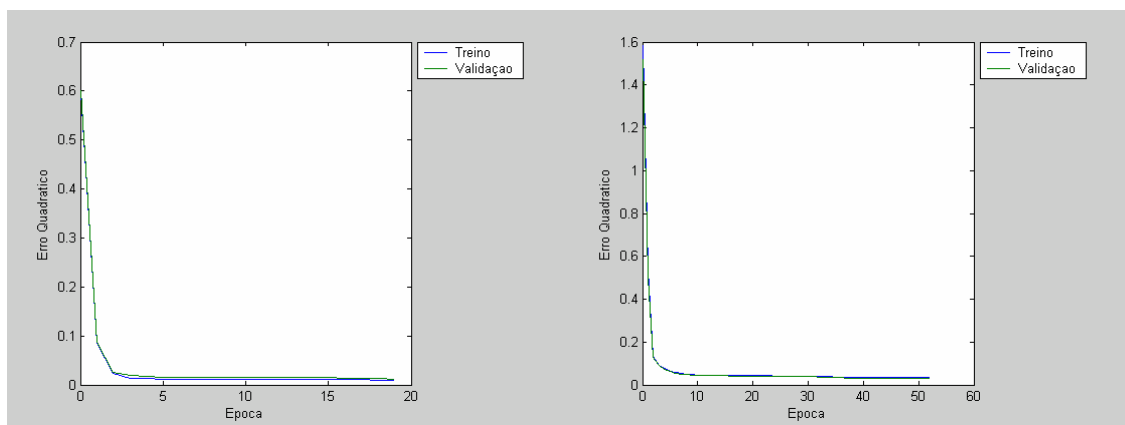


Figura 5.8 – Exemplo de Treino das redes para as variáveis de junta 2 e 4

Esta rede foi treinada com os seguintes dados:

- uma camada de entrada com 4 nós e uma camada escondida com 3 nós, ambas com funções de activação sinusoidal;

- uma camada de saída com um nó, com uma função de activação linear;
- o parâmetro de aprendizagem foi de 0.3 e o momento de 0.8;
- 200 épocas e um erro final de 0.01.

### 5.1.3 Testes

Após o treino da rede e utilizando o conjunto de teste, é importante verificar como é que a rede se comporta com novos dados. Para verificar o funcionamento da rede treinada, é-lhe fornecida o subconjunto de dados de teste, e verifica-se qual a resposta que apresenta. Estes valores são depois comparados com os valores esperados. O conjunto de teste é formado com valores do conjunto de treino e com valores nunca vistos pela rede.

Na Figura 5.9, apresenta-se os resultados obtidos, com o conjunto de teste, para as juntas 2 e 4. E na Tabela 1 apresenta-se os valores dos erros obtidos para a rede da junta 2 e 4.

Nesta figura representou-se por um círculo o valor correcto e por um \*, o valor devolvido pela rede. A junta 2 apresentou um erro absoluto médio de  $0.2616^\circ$  com uma variância de  $0.1670^\circ$  e a junta 4, um erro absoluto médio de  $0.6075^\circ$  com uma variância de  $0.8471^\circ$ .

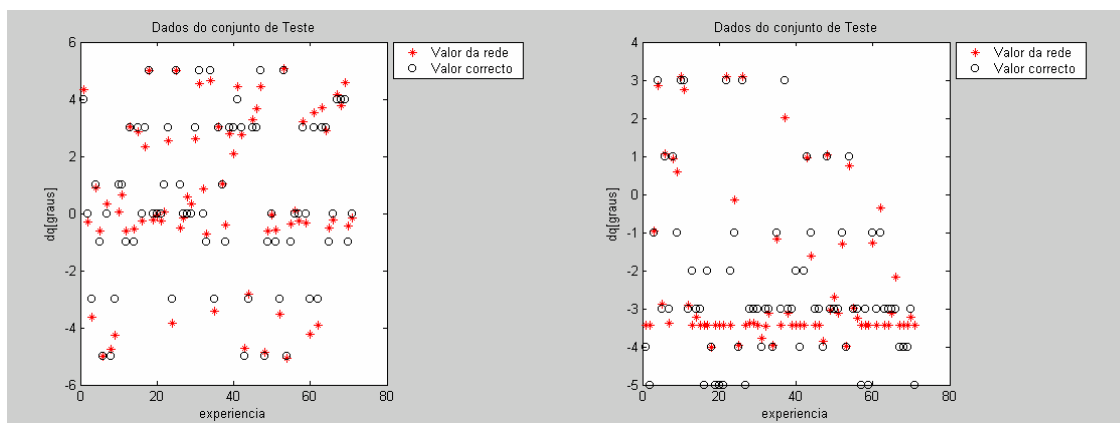


Figura 5.9 – Simulação da rede treinada com os dados do conjunto de teste da junta 2 e 4

Junta	Erro Absoluto médio	Variância
2	$0.2616^\circ$	$0.1670^\circ$
4	$0.6075^\circ$	$0.8471^\circ$

Tabela 1 – Resultados dos valores dos erros para as redes das juntas 2 e 4

O passo seguinte consistiu em verificar, com a rede treinada, como o manipulador se aproximaria de um cubo, colocado sobre uma mesa.

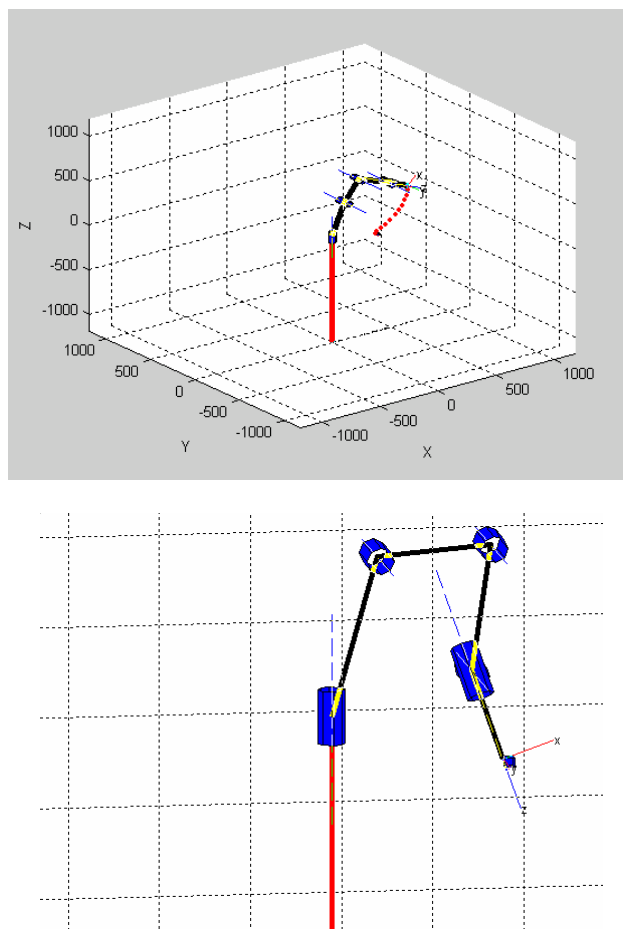
Nas figuras que se apresentam de seguida, os diversos pontos a vermelho na imagem pretendem simular as posições do manipulador ao aproximar-se do cubo.

Este teste foi efectuado para o posicionamento do cubo em duas posições distintas, referidas nas alíneas a) e b).

Em ambas as experiências a posição inicial do manipulador corresponde aos seguintes valores de juntas, em graus,  $\theta = [0 \ 30 \ 90 \ 0 \ 0]^T$ .

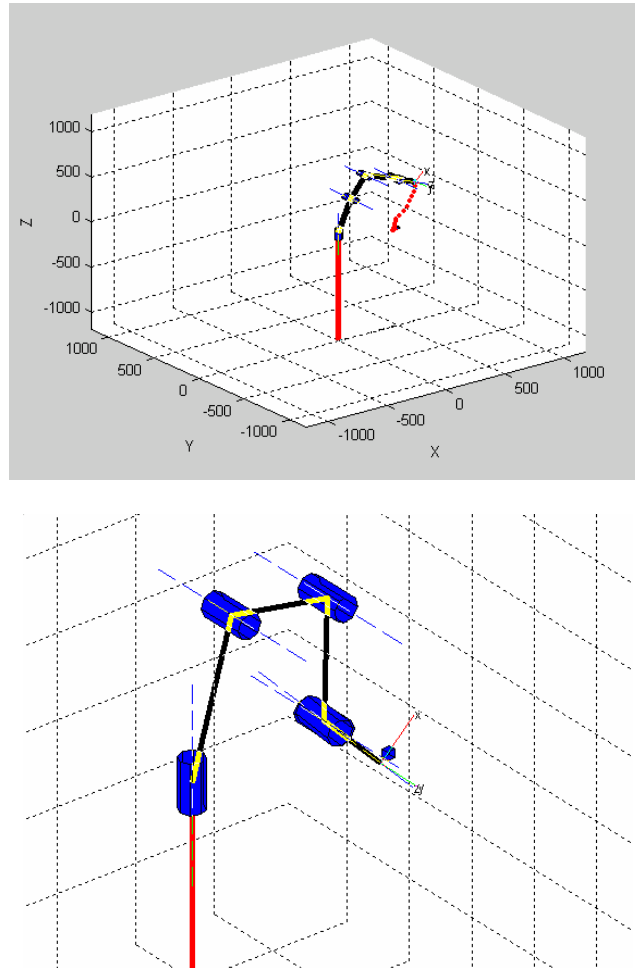
Procedeu-se à experiência para duas posições distintas:

- a) O centro do cubo situa-se nas coordenadas  $\mathbf{p} = {}^B[390 \ 0 \ -126]^T$  [mm]. Ao fim de 11 iterações o manipulador encontra-se a uma distância do cubo, em termos de coordenadas de imagem,  $\mathbf{c} = [11 \ -18]^T$  [pixels], tal como se pode observar na Figura 5.10.



*Figura 5.10 – Simulação do seguimento do efector terminar; os pontos a vermelho representam os diversos pontos que o manipulador vai passando até aproximar-se do cubo*

- b) O centro do cubo situa-se nas coordenadas do  $\mathbf{p}^B = [500 \ 0 \ -126]^T$  [mm]. Ao fim de 16 iterações o manipulador encontra-se a uma distância do cubo, em termos de coordenadas de imagem,  $\mathbf{c} = [14 \ 0]^T$  [pixels], tal como se pode observar na Figura 5.11



*Figura 5.11 – Simulação do seguimento do efector terminar; os pontos a vermelho representam os diversos pontos que o manipulador vai passando até aproximar-se do cubo*

Nas diversas experiências efectuadas obteve-se resultados satisfatórios. Conseguindo o manipulador aproximar-se do ponto desejado.

Após a realização dos testes simulados, deu-se início à realização do mesmo tipo de experiência, mas agora com o manipulador real. Nos resultados experimentais houve a necessidade de introduzir a aquisição e segmentação da imagem de modo a extrair os diversos pontos característicos de imagem. Um dos grandes problemas encontrados foi o controlo da luminosidade no espaço laboratorial.



## 5.2 Resultados Experimentais com o Manipulador Real

Tal como se verificou na Secção anterior, devido à não regularidade do sistema, não era possível utilizar apenas uma rede neuronal para todo o espaço de trabalho, havendo a necessidade de dividir este espaço em subzonas.

Pelo simples facto de se utilizar apenas uma câmara, o sistema perde informação da profundidade. A projecção vai causar essa perda de informação, dado que cada ponto do plano de imagem corresponde a uma linha no espaço tridimensional (3D) [3]. Informação adicional é necessária para fazer a correspondência de pontos 3D em pontos 2D. Surgiu assim a necessidade de colocar em cena uma marca fixa (de cor vermelha) com o objectivo de se conhecer a profundidade a que o manipulador está a operar.

Na Figura 5.12 é possível observar os diversos elementos relevantes ao sistema: cubo, o effector terminal do manipulador e marca fixa.

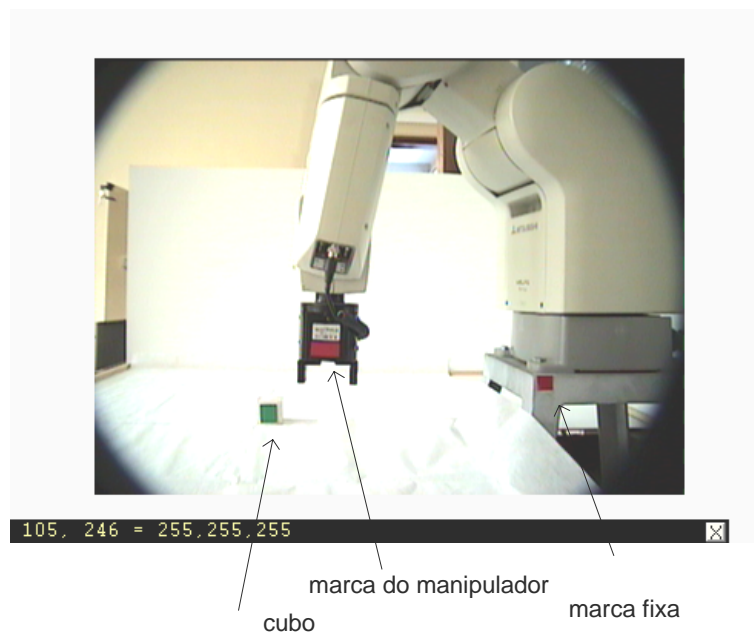


Figura 5.12 – Imagem do Espaço de Trabalho com os elementos mais relevantes ao Sistema de Visão

A rede da junta 1, Figura 5.13 a), tem uma construção diferente das outras porque o seu objectivo é simplesmente o de conhecer o plano de trabalho em que o manipulador vai funcionar, ou seja escolher a zona de trabalho.

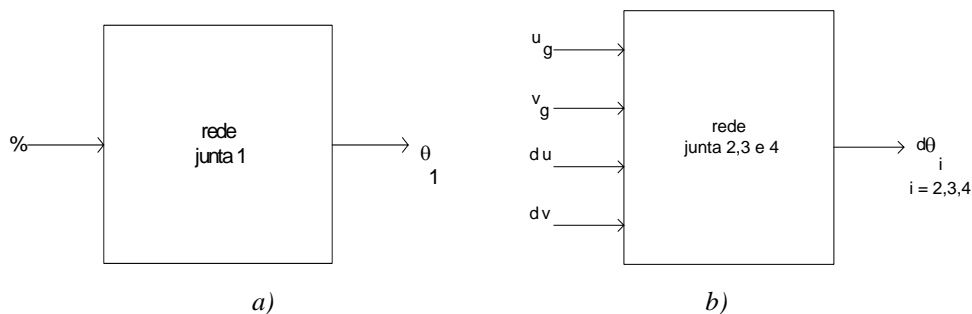


Figura 5.13 – Desenho esquemáticos do fluxo dos dados das redes; a) junta 1; b) restantes juntas

Cada uma das redes utilizadas é constituída por:

- Rede da Junta 1 – que tem como objectivo a escolha da zona de trabalho:
  - 1 nó de entrada – percentagem da área do cubo face à área da marca fixa;
  - 1 nó de saída – valor da junta 1.
- Redes das restantes juntas (constituição idêntica para cada uma das variáveis de junta) – que tem como objectivo aproximar o efector terminal ao objecto:
  - 4 nós de entrada – posição actual do efector terminal e o diferencial a efectuar em torno desse ponto;
  - 1 parâmetro de saída – diferencial a efectuar à variável de junta.

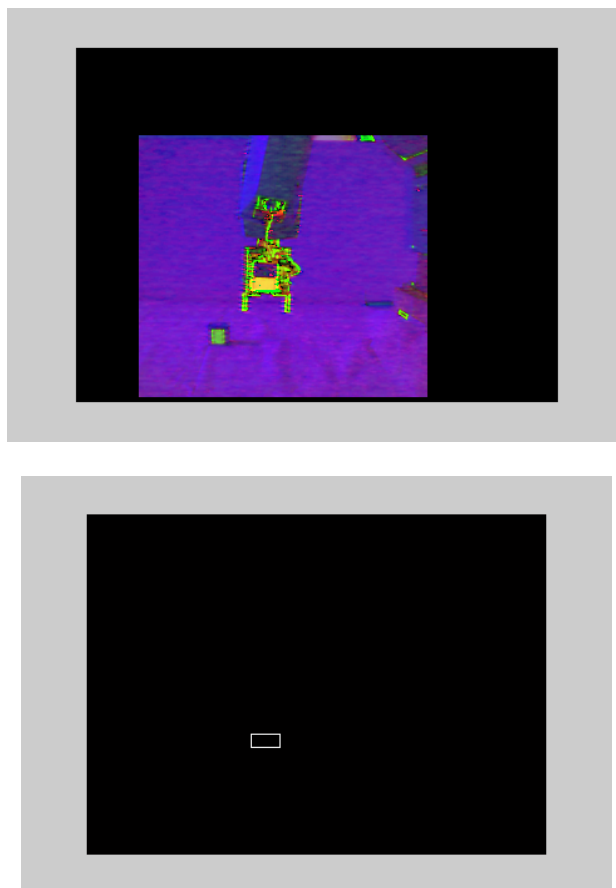
A obtenção dos dados da rede neuronal para a junta 1, consiste em colocar o cubo no efector terminal e depois efectuar pequenos deslocamentos na variável da junta 1 e as medições, na imagem, das áreas do cubo e da marca fixa. O par de dados que se armazena, no ficheiro de dados de entrada da rede, é a percentagem da área do cubo face à marca fixa, e no ficheiro de dados de saída da rede armazena-se o valor da variável da junta 1.

No caso da rede respeitante à junta 1, são feitos diversos movimentos da junta 1, mantendo-se as outras juntas estáticas, e calcula-se, em cada um desses pontos, a percentagem da área do cubo relativamente a uma referência fixa, colocada junto à base do manipulador. Deste modo é possível a RNA estabelecer uma relação entre o valor da variável da junta 1 e a distância do objecto à câmara.

Através da rede da junta 1, escolhe-se qual a zona de trabalho, tal como foi indicado na Figura 5.5; depois através das redes das variáveis das juntas 2, 3 e 4, calcula-se o

diferencial a dar a cada uma das variáveis de junta, de modo a aproximar o manipulador ao objecto.

Em cada uma das situações é necessário proceder à extracção dos pontos característicos de imagem. Apresenta-se na Figura 5.14 um exemplo do resultado da segmentação utilizada para o manipulador. Antes de se proceder à segmentação converte-se as coordenadas de cor da imagem adquirida para *Hue Saturation Value* (HSV), de modo a facilitar a segmentação [19].



*Figura 5.14 – Imagem convertida em HSV, e após segmentação, o contorno da imagem binária utilizada no cálculo do centro de massa do manipulador*

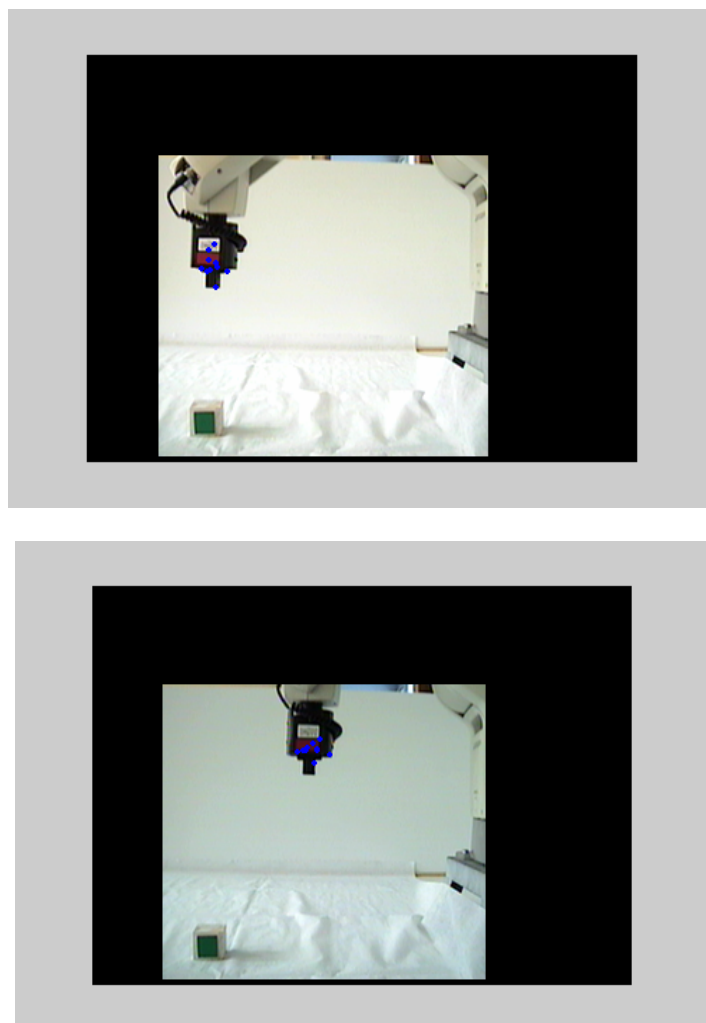
Os testes efectuados, cujas imagens se apresentam de seguida, consistiram em:

- a) Recolha do conjunto de treino da rede,
- b) Treino das Redes Neuronais Artificiais (RNA),
- c) Escolha correcta do plano de trabalho, através da RNA da junta 1,
- d) Aproximação do manipulador ao cubo, para duas posições distintas, através das RNA das juntas 2, 3 e 4.

Nos testes aqui apresentados posicionou-se a câmara aproximadamente a 90cm da base do manipulador, num plano perpendicular ao plano XZ do manipulador.

a) Recolha do conjunto de treino da rede

Na Figura 5.15 mostra-se as imagens recolhidas pela câmara em duas posições de aquisição de dados de treino. Os pontos marcados a azul representam os pontos adquiridos.



*Figura 5.15 – Figuras reais obtidas com o manipulador durante a aquisição de dados. Os pontos representados a azul são os utilizados na rede neuronal. A marca a vermelho representa a região utilizada pela segmentação para seguir o manipulador*

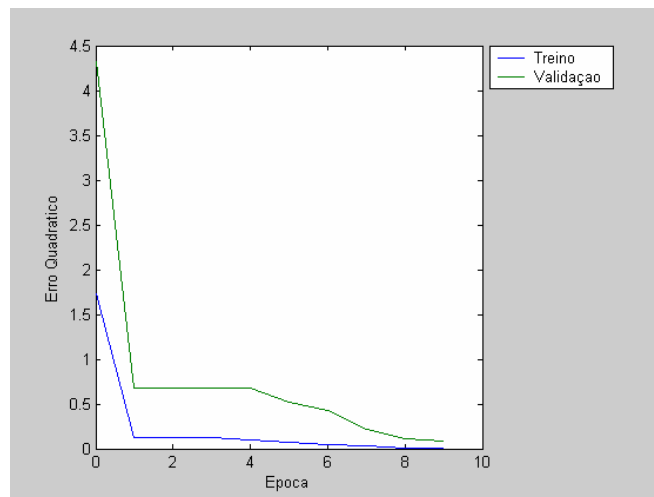
b) Treino das Redes Neurais Artificiais

Após recolher os dados de treino, procede-se ao treino da rede. A junta 1 é treinada à parte, numa rede em que o dado de entrada é a percentagem da área do cubo relativamente

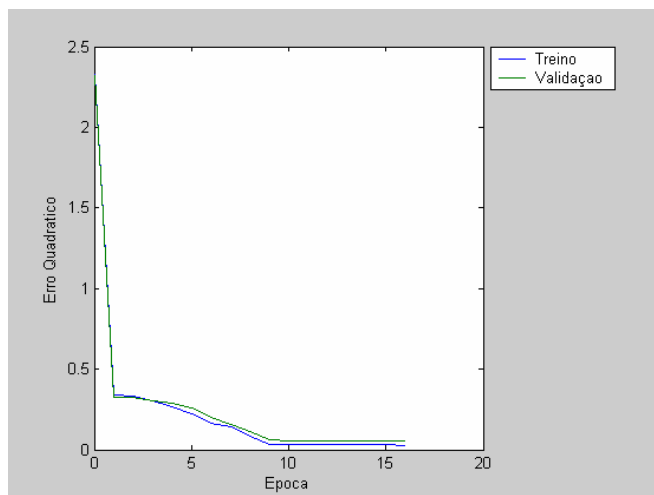
à área da marca fixa e como parâmetro de saída temos o valor da junta 1. Deste modo é possível saber em que plano de imagem está o cubo.

As juntas 2, 3 e 4 foram treinadas por outras redes em que os dados de entrada são as coordenadas do ponto da imagem,  $u$  e  $v$ , os diferenciais do movimento em torno desse ponto em coordenadas de imagem,  $du$  e  $dv$ , e como saída da rede o movimento diferencial que é necessário provocar em cada uma das variáveis de junta,  $d\theta_i$ , de modo a se conseguir o deslocamento na imagem pretendido.

Na Figura 5.16, apresenta-se um exemplo de treino para as juntas 1 e 2, utilizando valores reais obtidos com a câmara.



a)



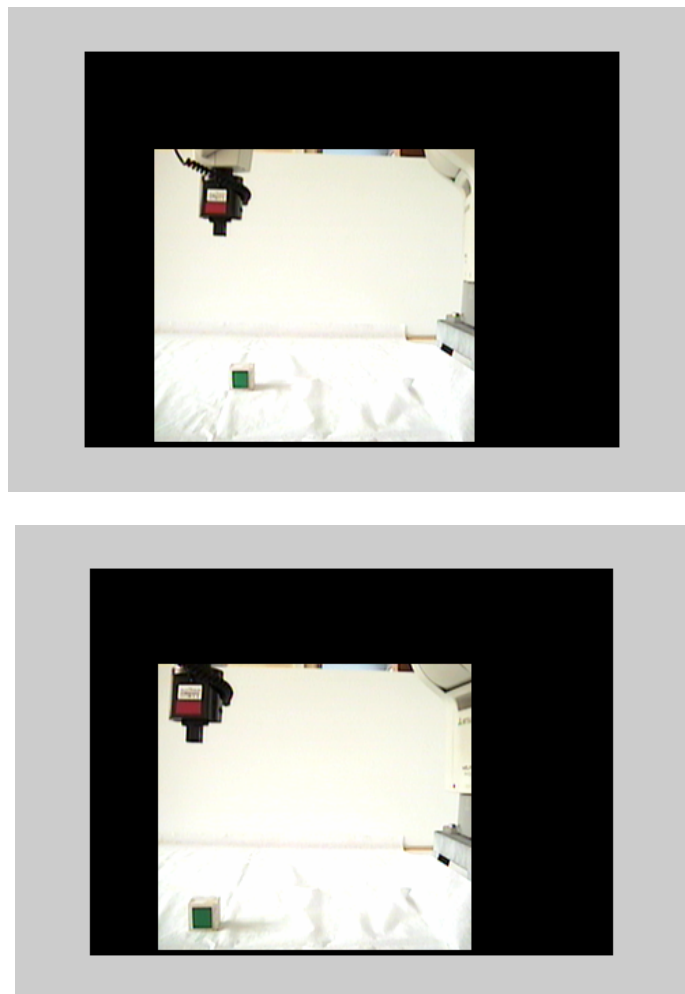
b)

Figura 5.16 – Treino das juntas 1 e 2 ; a) junta 1; b) junta2

Após se ter treinado a rede, procede-se ao seguimento do efector terminal. O objectivo é aproximar o efector terminal do cubo. A primeira acção consiste em utilizar os dados da junta 1 de modo a calcular o plano de trabalho.

c) Escolha correcta do plano de trabalho, através da RNA da junta 1

Na Figura 5.17 apresentam-se dois exemplos, com sucesso, da escolha do valor correcto a colocar na junta 1 de modo a que o manipulador esteja no plano vertical do cubo, sendo de seguida apenas necessário mexer nas juntas 2, 3 e 4 de modo a aproximar o manipulador ao cubo.

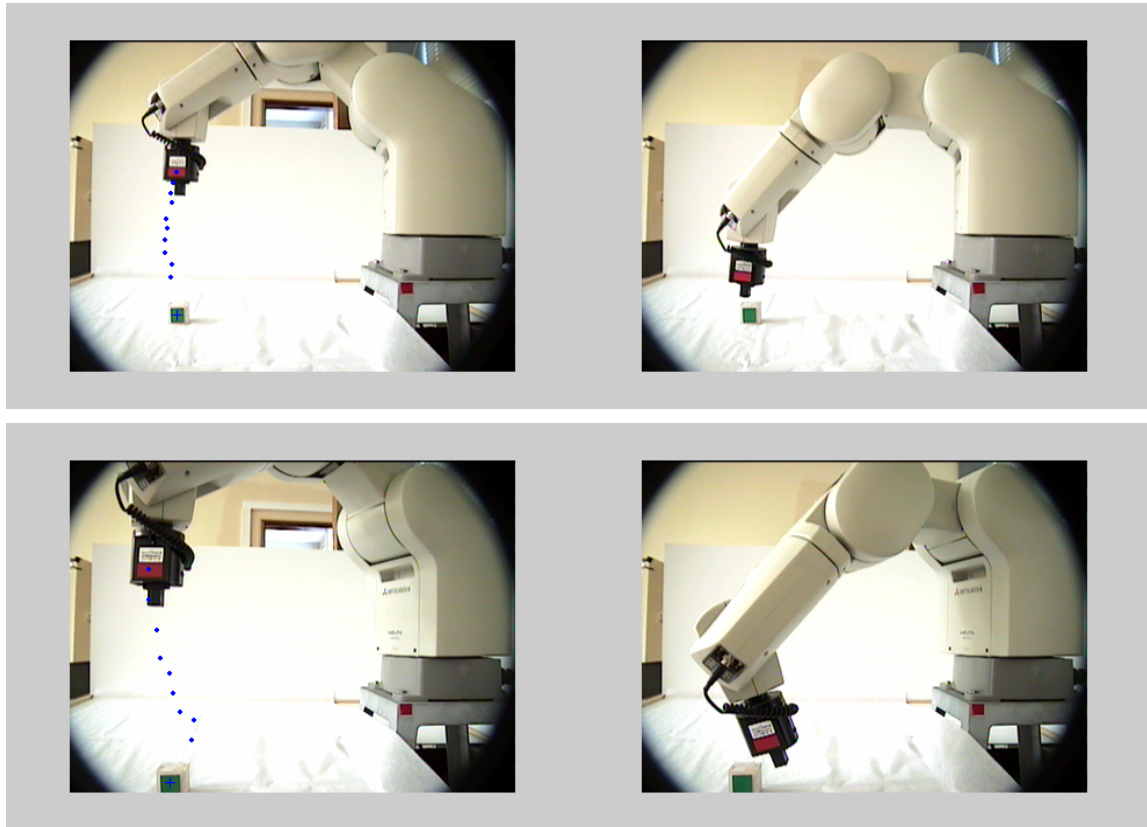


*Figura 5.17 – Escolha do plano de trabalho através da Rede Neuronal para a variável de junta1, para duas colocações diferentes do posicionamento do cubo.*

d) Aproximação do manipulador ao cubo, para duas posições distintas

Uma vez o efector terminal situado no mesmo plano vertical que o objecto, realizam-se testes de aproximação do manipulador ao cubo.

Apresenta-se os resultados para duas posições iniciais distintas.



*Figura 5.18 – Aproximação do manipulador ao cubo, para duas situações distintas. Os pontos a azul representam a trajectória do manipulador ao aproximar-se do cubo.*

Dos exemplos apresentados e dos testes efectuados pode-se concluir que a rede neuronal conseguiu aproximar o Jacobiano completo e os resultados são satisfatórios. Contudo, num trabalho futuro, poder-se-ia melhorar a robustez do sistema de visão e separar o controlo das variáveis de junta 4 e 5, das variáveis de junta 1, 2 e 3. A vantagem desta separação é perfeitamente visível na última imagem da Figura 5.18, onde se verifica que há um problema de alinhamento do manipulador face ao cubo.

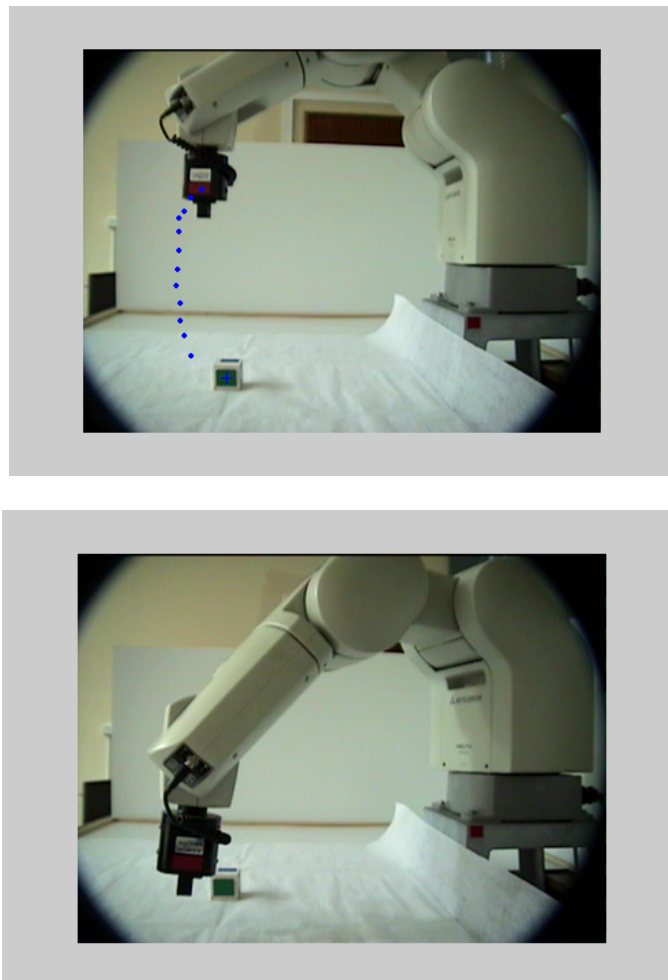
Ao separar-se as variáveis das juntas 4 e 5 das restantes, efectuar-se-ia o posicionamento do manipulador sobre o objecto através das juntas 1, 2 e 3 e com as juntas 4 e 5 tratar-se-ia de fazer o alinhamento do manipulador face ao objecto tendo em vista a sua captura. Ter-se-ia um sistema de controlo para o posicionamento e um sistema de controlo para o alinhamento.

Em alguns casos registou-se uma falha da aproximação do manipulador ao cubo devido ao movimento efectuado segundo o espaço nulo do Jacobiano completo.

Quando a junta 1 é tratada como as restantes juntas, certos movimentos do manipulador, por exemplo os efectuados ao longo do raio de projecção, não produzem nenhuma variação na imagem.

Estas situações ocorrem todas para movimentos no espaço nulo do Jacobiano. Para demonstrar este problema, fez-se um teste em que não houve a separação do espaço de trabalho por zonas.

Na Figura 5.19, o manipulador não consegue aproximar-se do cubo; ao fazer a aproximação ao cubo, o manipulador altera o valor da junta 1, o que faz com que este se desloque na direcção do raio óptico, perdendo-se completamente a noção de profundidade.



*Figura 5.19 – Aproximação falhada ao cubo*



Este resultado vem confirmar a necessidade da separação do controlo da variável da junta 1 das restantes variáveis de junta, visto só estarmos a trabalhar com uma câmara. É através da variável da junta 1 e da colocação de uma marca fixa que é possível recuperar a percepção 3D, perdida pelo facto de só se utilizar uma câmara.

## 6 Conclusões e Trabalho Futuro

O principal objectivo desta Tese foi desenvolver uma metodologia que permite aproximar, um manipulador dotado com sistema de visão monocular, de um objecto, sem a necessidade de efectuar qualquer tipo de calibração da câmara. Não é necessário conhecer os parâmetros intrínsecos da câmara nem efectuar a sua calibração extrínseca. O manipulador é controlado pela informação que retira da imagem e não necessita de perspectiva inversa ou qualquer outra transformação de cinemática. A única restrição é que o objecto a capturar esteja localizado num plano.

O sistema utiliza informações que retira da imagem tendo para isso utilizado a estrutura de arquitectura do tipo “dinâmico vê-e-move”, com controlo baseado na imagem e o terminal em cadeia fechada (ECL).

Optou-se por usar uma só câmara para melhor compreender as variáveis em causa e devido ao alargamento do espaço mensurável de trabalho que uma só câmara tem. Como só se utilizou uma câmara, houve a necessidade de utilizar um método que conseguisse de algum modo aprender a profundidade.

O manipulador utilizado foi o da Mistsubishi, modelo Movemaster E3J, com 5 graus de liberdade. Procedeu-se ao estudo da cinemática directa e inversa, para este manipulador, assim como ao seu Jacobiano. Com estes estudos constatou-se das dificuldades acrescidas quando o manipulador possui 5 graus de liberdade – o facto de não se obterem matrizes quadradas, e ainda as dificuldades resultantes das restrições do cálculo da cinemática inversa e do Jacobiano.

O controlo do manipulador foi realizado no espaço junta. Através dos dados do sistema de visão e à custa de uma rede neuronal conseguiu-se aproximar a função Jacobiano completo.

O bloco da Rede Neuronal Artificial (RNA) simula o Jacobiano completo manipulador-visão, evitando deste modo os processos de calibração do manipulador e da

câmara. Na construção deste bloco recorreu-se às facilidades disponíveis na *Neural Network Toolbox* do programa MatLab da MathWorks. A RNA utilizada tem como entradas a posição actual do manipulador em termos de coordenadas da imagem  $(u, v)$  e a velocidade do vector de imagem  $(du, dv)$ , sendo a saída o diferencial a ser dado a cada uma das juntas  $(d\theta)$ , por forma a que o efector terminal se desloque para a posição pretendida. Os valores de saída do sistema de controlo do manipulador foram aplicados directamente ao controlador do fabricante do manipulador. A RNA é constituída por uma camada de entrada com 4 nós, uma camada escondida com 3 nós e uma camada de saída com um nó.

Antes de se proceder à simulação da arquitectura escolhida, foram feitos estudos sobre a função Jacobiano completo, tendo-se chegado à conclusão da necessidade de separação do espaço de trabalho por zonas.

Após esse estudo procedeu-se à simulação do sistema em cadeia fechada. Para tal criou-se um programa de aquisição de dados para o treino da rede e com o qual se realizaram testes, consistindo estes em provocar perturbações do manipulador em diversos pontos do espaço de trabalho. Obtiveram-se dados satisfatórios, conseguindo-se aproximar o manipulador do cubo.

Quando se passou para a parte experimental com o manipulador real, houve a necessidade de identificar diversos elementos, tais como o efector terminal e o cubo; para simplificar o sistema de visão, foram colocadas fitas coloridas para se proceder à sua identificação. Outro problema surgido foi o de fazer o manipulador deslocar-se no plano correcto, dado ser frequente o sistema terminar, supondo que estava próximo ao cubo, embora tal só acontecesse no plano da imagem e não no plano 3D, devido à perda de informação de profundidade. Assim, separou-se o treino da rede da junta 1, das outras juntas; no treino da rede da junta 1, colocou-se uma marca fixa e efectuaram-se diversos movimentos na junta 1 de forma a conhecer a relação entre a área do objecto fixo face ao objecto.

Com esta alteração já foi possível obter resultados satisfatórios no seguimento do braço manipulador. As falhas, que por vezes se verificaram, tinham a ver com o sistema de visão, quando as condições de iluminação se alteravam.

Em termos gerais pode-se concluir que é possível aproximar um manipulador a um objecto sem que para tal seja necessário conhecer a cinemática do manipulador ou a geometria do objecto.

Em termos de trabalho futuro pretende-se, a curto prazo, melhorar a robustez do sistema de visão face às variações de iluminação, o que permitirá uma melhor identificação de todos os objectos em causa. Surge a necessidade de mudar o plano de trabalho de modo a alargar a área de trabalho útil do manipulador; neste momento o plano de trabalho está numa posição inferior à base do manipulador.

Em termos de investigação futura pretende-se evoluir para um sistema de visão *stereo*, de modo a evitar o problema da profundidade aqui encontrado. Um sistema com duas câmaras também permitirá aumentar a complexidade da aplicação, sendo possível deste modo utilizar a junta de orientação numa aplicação concreta de captura.

Poder-se-ia melhorar ainda o sistema de visão de modo a separar as variáveis da junta 4 e 5, que têm a ver com o alinhamento do manipulador face ao objecto, das variáveis das juntas 1, 2 e 3.

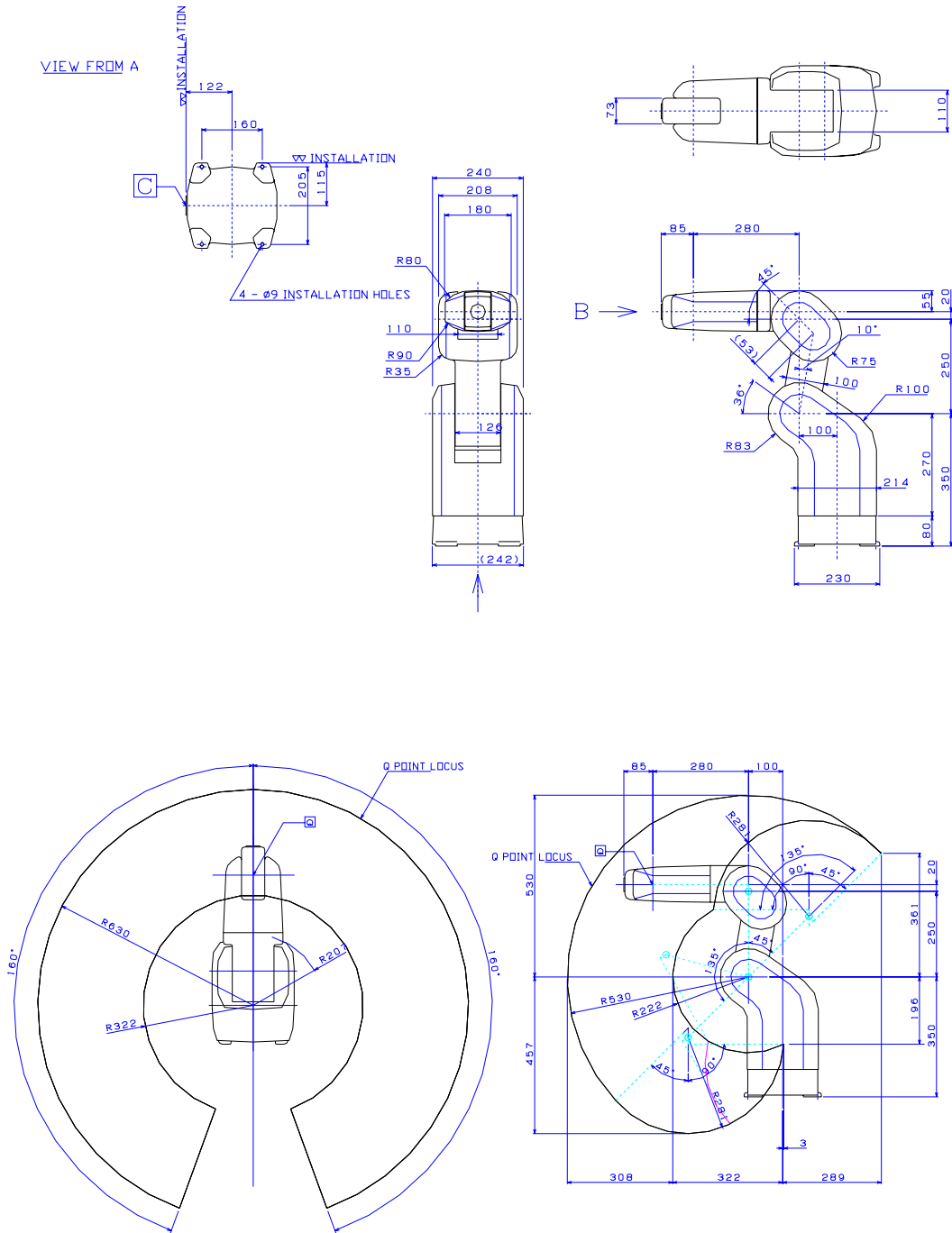
Outra direcção de investigação consistirá em proceder a um estudo comparativo da rede neuronal supervisionada aqui utilizada com a utilização da rede neuronal não supervisionada, Kohonen, utilizada nos trabalhos da equipa de Ritter et al. [14]

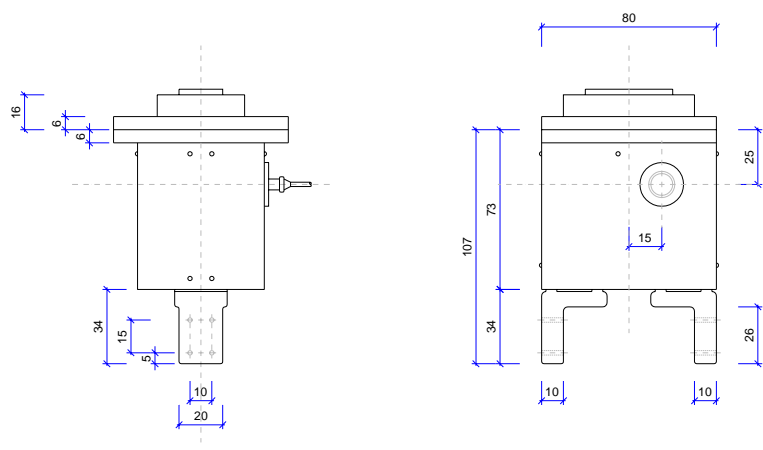
## 7 Referências Bibliográficas

- [1] **Almeida, Luís B.**, “An Introduction to multiplayer perceptrons”, Revista de Engenharia – 92, pág 67 a 87.
- [2] **Batchelor, Bruce G., Paul F. Whelan**, “Intelligent Vision Systems for Industry”, Springer, 1997.
- [3] **Corke, Peter I.**, “A Tutorial on Visual Servo Control”, IEEE - Transactions on Robotics and Automation, vol.12, Outubro 1996.
- [4] **Corke, Peter I.**, “Visual Control of Robots: High-Performance Visual Servoing, Research Studies Press Ltd, 1996.
- [5] **Craig, John J.**, “Introduction to Robotics Mechanics and Control”, Addison-Wesley , 1989.
- [6] **Crane, Carl, Joseph Duffy**, “Kinematic Analysis of Robot Manipulation”, Cambridge University Press, 1998.
- [7] **Demult, H., M. Beale**, “Neural Network Toolbox for use with Matlab”, The Matworks, 1994.
- [8] **Fausett, Laurene**, “Fundamentals of Neural Networks, architectures, Algorithms and applications, Prentice Hall, 1994 .
- [9] **Fernandes, Dinis dos Anjos Guerra**, “Arquitecturas de seguimento visual e captura por um manipulador robótico de objectos em movimento”, Tese de dissertação para obtenção do grau de Mestre, 1997.
- [10] **Fu, K.S., R. C. Gonzalez and C. S. G. Lee**, "Robotics: Control, Sensing, Vision and Intelligence", McGraw-Hill Intern. Ed., Singapore, 1987.
- [11] **Horn, Berthold Klaus Paul**, “Robot Vision”, the MIT Press, 1986.
- [12] **Jordan, M. I., D. E. Rumelhart**, “Forward models: Supervised learning with distal teacher”, Cognitive Science, nº16, 1992.

- [13] **Krose, B. J. A., P. Patrick. Van der Smagt**, “An introduction to Neural Networks”, University of Amsterdam, Fifth Edition, Janeiro 1993.
- [14] **Martinez, Thomas M., Helge J. Ritter, Klaus J. Schulten**, “3D-Neural Net for learning Visiomotor –Coordination of a Robot Arm”, Proceeding of the International Joint Conference on Neural Network, vol 2, pag. 351-356, 1989.
- [15] **Murray, Richard M., Zexiang Li, S. Shankar Satry**, “A Mathematical Introduction to Robotic Manipulation, CRC Press, 1994.
- [16] **Nguyen, Minh-Chinh**, “Stereo-Vision-Based Robot Control Without Calibration, VACETS Techical International Conference (VITC’97), San Jose, Julho 1997.
- [17] **Nguyen, Minh-Chinh, Volker Graefe**, “Self-Learning Vision-Guide Robots for Searching and Grasping Objects, Proceedings of 2000 IEEE, International Conference on Robotics & Automation, San Francisco, CA, Abril 2000.
- [18] **Ritter, Gerard X., Joseph N. Wilson**, “Handbook of Computer Vision Algorithms in Image Algebra”, CRC Press, 1996.
- [19] **Russ, John C.**, “The Image Processing Handbook”, Third Edition, CRC Press, Springer, IEEE Press, 1998.
- [20] **Sciavicco, Lorenzo, Bruno Siciliano**, “Modeling and Control of Robot Manipulators”, Springer, 1996.
- [21] **Siciliano, Bruno**, “Robot Control”, in Perspectives in Control Engineering: Technologies, Tariq Samad Ed., IEEE press, Piscataway, NJ, 2000.
- [22] **Skapura, David M.**, “Bulding Neural Networks”, ACM Press, 1996.
- [23] **Stadler, Wolfram**, “Analytical Robotics and Mechatronics” McGraw-Hill Editions, 1995.
- [24] **Yoshimi, Billibon H., Peter K. Allen**, “Active, Uncalibrated Visual Servoing”, Vol. 4, Pag 156-161, IEEE Robotics & Automation, 1994.
- [25] **Zhuang, Hanqi, Kuanchih Wang, Zvi S. Roth**, “Simultaneous Calibration of a Robot and a Hand-Mounted Camera”, IEEE – Transactions on Robotics and Automation, vol.11, Outubro 1995.
- [26] **Zilouchian, Ali, Mo Jamshidi**, “Intelligent Control Systems Using Soft Computing Methodologies” CRC Press, 2001.

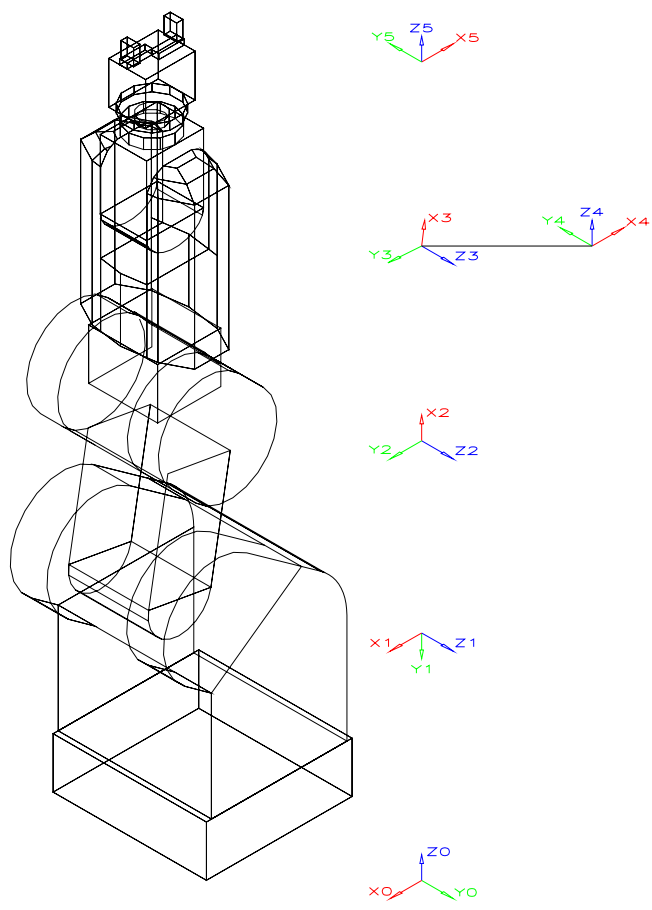
# Anexo A – Manipulador Mitsubishi RV-E3J







## Anexo B - Parâmetros Denavit-Hartenberg



Posição de Repouso  $W=S=E=P=R=0^\circ$

link	alpha	a	theta	d	
1	-90	100	0	350	$\theta_1 = W$
2	0	250	-90	0	$\theta_2 = -90^\circ + S$
3	0	280	-4.1	0	$\theta_3 = -4.1^\circ + E$
4	-90	0	-85.9	0	$\theta_4 = -85.9^\circ + P$
5	0	0	0	208	$\theta_5 = R$

## Anexo C - Cinemática Directa

$${}^0\mathbf{A}_1 = \begin{pmatrix} \cos(t1) & 0 & -\sin(t1) & a1*\cos(t1) \\ \sin(t1) & 0 & \cos(t1) & a1*\sin(t1) \\ 0 & -1 & 0 & d1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$${}^1\mathbf{A}_2 = \begin{pmatrix} \cos(t2) & -\sin(t2) & 0 & a2*\cos(t2) \\ \sin(t2) & \cos(t2) & 0 & a2*\sin(t2) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$${}^2\mathbf{A}_3 = \begin{pmatrix} \cos(t3) & -\sin(t3) & 0 & a3*\cos(t3) \\ \sin(t3) & \cos(t3) & 0 & a3*\sin(t3) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$${}^3\mathbf{A}_4 = \begin{pmatrix} \cos(t4) & 0 & -\sin(t4) & 0 \\ \sin(t4) & 0 & \cos(t4) & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$${}^4\mathbf{A}_5 = \begin{pmatrix} \cos(t5) & -\sin(t5) & 0 & 0 \\ \sin(t5) & \cos(t5) & 0 & 0 \\ 0 & 0 & 1 & d5 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$${}^0\mathbf{T}_3 = \begin{pmatrix} C1C23 & -C1S23 & -S1 & C1(a3C23 + a2C2 + a1) \\ S1C23 & -S1S23 & C1 & S1(a3C23 + a2C2 + a1) \\ -S23 & C23 & 0 & d1 - a3S23 - a2S2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$${}^3\mathbf{T}_5 = \begin{pmatrix} C4C5 & -C4S5 & -S4 & -d5S4 \\ S4C5 & -S4S5 & C4 & d5C4 \\ -S5 & -C5 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$${}^0\mathbf{T}_5 = \begin{pmatrix} C5*C1*C234 + S1*S5 & -S5*C1*C234 + S1*C5 & -C1*S234 & -d5*C1*S234 + a3*C1*C23 + a2*C1*C2 + a1*C1 \\ C5*S1*C234 - C1*S5 & -S5*S1*C234 - C1*C5 & -S1*S234 & -d5*S1*S234 + a3*S1*C23 + a2*S1*C2 + a1*S1 \\ -C5*S234 & S5*S234 & -C234*d1 & d1 - d5*C234 - a3*S23 - a2*S2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

## Anexo D - Cinemática Inversa

$${}^0\mathbf{A}_1^{-1} = \begin{pmatrix} C1 & S1 & 0 & -a1 \\ 0 & 0 & -1 & d1 \\ -S1 & C1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$${}^0\mathbf{A}_1^{-1} * {}^0\mathbf{T}_5 = {}^2\mathbf{T}_5 = {}^1\mathbf{A}_2 * {}^2\mathbf{A}_3 * {}^3\mathbf{A}_4 * {}^4\mathbf{A}_5 =$$

$$\begin{pmatrix} \cos(t1)*nx+\sin(t1)*ny & \cos(t1)*ox+\sin(t1)*oy & \cos(t1)*ax+\sin(t1)*ay & \cos(t1)*px+\sin(t1)*py-a1 \\ -nz & -oz & -az & -pz+d1 \\ -\sin(t1)*nx+\cos(t1)*ny & -\sin(t1)*ox+\cos(t1)*oy & -\sin(t1)*ax+\cos(t1)*ay & -\sin(t1)*px+\cos(t1)*py \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$${}^2\mathbf{T}_5 = {}^1\mathbf{A}_2 * {}^2\mathbf{A}_3 * {}^3\mathbf{A}_4 * {}^4\mathbf{A}_5 =$$

$$\begin{pmatrix} \cos(t2+t3+t4)*\cos(t5) & -\cos(t2+t3+t4)*\sin(t5) & -\sin(t2+t3+t4) & -\sin(t2+t3+t4)*d5+a3*\cos(t2+t3)+a2*\cos(t2) \\ \sin(t2+t3+t4)*\cos(t5) & -\sin(t2+t3+t4)*\sin(t5) & \cos(t2+t3+t4) & \cos(t2+t3+t4)*d5+a3*\sin(t2+t3)+a2*\sin(t2) \\ -\sin(t5) & -\cos(t5) & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Considera-se as seguintes abreviaturas:

$$C1 = \cos(t1); \quad S1 = \sin(t1); \quad C234 = \cos(t2+t3+t4)$$

Cálculo de  $\theta_1$ :

Comparando os elementos (3,4) e (3,3) de ambas as matrizes:

$$-S1ax + C1ay = 0 \quad \Rightarrow \quad \theta_1 = \tan^{-1}\left(\frac{ax}{ay}\right); \quad \theta_1 = \tan^{-1}\left(\frac{-ax}{-ay}\right)$$

$$-S1px + C1py = 0 \quad \Rightarrow \quad \theta_1 = \tan^{-1}\left(\frac{px}{py}\right); \quad \theta_1 = \tan^{-1}\left(\frac{-px}{-py}\right)$$

Cálculo de  $\theta_{234}$ :

Comparando (1,3) e (2,3) de ambas as matrizes:

$$\begin{cases} C1ax + S1ay = -S234 \\ -az = C234 \end{cases}$$

como se conhece  $\theta_1$  vem que  $\theta_{234} = \text{atan2}(c1ax + s1ay, az)$

**Cálculo de  $\theta_3$ :**

Comparando (1,4) e (2,4) de ambas as matrizes:

$$\begin{cases} C1px + S1py - a1 = -d5S234 + a3C23 + a2C2 \\ d1-pz = d5C234 + a3S23 + a2S2 \end{cases}$$

fazendo uma substituição de variáveis:

$$\begin{cases} t1 = a3C23 + a2C2 \\ t2 = a3S23 + a2S2 \end{cases} \quad \text{elevando ambos os termos ao quadrado:}$$

$$C3 = \frac{t_1^2 + t_2^2 - a_3^2 - a_2^2}{2a_2a_3} \quad \text{e portanto } \theta_3 = a \tan 2(-\sqrt{1 - C_3^2}, C_3)$$

e igualmente se pode escrever o 1º sistema como:

$$\begin{cases} C1px + S1py - a1 + d5S234 = t1 \\ d1-pz - d5C234 = t2 \end{cases}$$

**Cálculo de  $\theta_2$ :**

$$\begin{cases} t1 = a3C23 + a2C2 \\ t2 = a3S23 + a2S2 \end{cases} \quad \begin{cases} t1 = a3(C3C2 - S3S2) + a2C2 \\ t2 = a3(S2C3 + S3C2 + a2S2) \end{cases}$$

resolvendo o sistema em ordem a C2 e S2 podemos escrever:

$$\theta_2 = \text{atan2}((a3C3 + a2)t2 - a3S3t1, t1(a3C3+a2) + a3S3t2)$$

**Cálculo de  $\theta_3$ :**

$$\theta_4 = \theta_{234} - \theta_3 - \theta_2$$

**Cálculo de  $\theta_5$ :**

Comparando (3,1) e (3,2)

$$\begin{cases} -S5 = -S1nx + C1ny \\ -C5 = -S1ox + C1oy \end{cases} \quad \theta_5 = \text{atan2}(-S1nx + C1ny, -S1ox + C1oy)$$

## Anexo E - Jacobiano do Manipulador

$$\mathbf{J}_M = \begin{bmatrix} J_{P1} \dots J_{Pn} \\ J_{O1} \dots J_{On} \end{bmatrix} = \begin{bmatrix} JP_{3 \times 5} \\ JO_{3 \times 5} \end{bmatrix}$$

$$\begin{bmatrix} J_{Pi} \\ J_{Oi} \end{bmatrix} = \begin{bmatrix} Z_{i-1} \times (P - P_{i-1}) \\ Z_{i-1} \end{bmatrix} \quad \text{sendo } P_0 = [0 \ 0 \ 0 \ 1]^T \quad \text{e } Z_0 = [0 \ 0 \ 1]^T,$$

$P_{i-1}$  é dado pela 3ª coluna da matriz de Translação

$Z_{i-1}$  é dado por 3ª coluna da matriz de Rotação

Particularizando para o manipulador em causa, vem:

$$\mathbf{J}_{P1} = \begin{pmatrix} d_5 S_1 S_{234} - S_1 a_3 C_{23} - S_1 a_2 C_2 - S_1 a_1 \\ -d_5 C_1 S_{234} + C_1 a_3 C_{23} + C_1 a_2 C_2 + C_1 a_1 \\ 0 \quad d_5 S_{234} - a_3 C_{23} - a_2 C_2 \end{pmatrix} \quad \mathbf{J}_{P2} = \begin{pmatrix} -C_1 d_5 C_{234} - C_1 a_3 S_{23} - C_1 a_2 S_2 \\ -S_1 d_5 C_{234} - S_1 a_3 S_{23} - S_1 a_2 S_2 \end{pmatrix}$$

$$\mathbf{J}_{P3} = \begin{pmatrix} -C_1 d_5 C_{234} - C_1 a_3 S_{23} \\ -S_1 d_5 C_{234} - S_1 a_3 S_{23} \\ d_5 S_{234} - a_3 C_{23} \end{pmatrix} \quad \mathbf{J}_{P4} = \begin{pmatrix} -C_1 d_5 C_{234} \\ -S_1 d_5 C_{234} \\ d_5 S_{234} \end{pmatrix} \quad \mathbf{J}_{P5} = \begin{pmatrix} S_1 S_{234} d_5 C_{234} \\ -C_1 S_{234} d_5 C_{234} \\ 0 \end{pmatrix}$$

$$\mathbf{J}_{O1} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \mathbf{J}_{O2} = \begin{pmatrix} -S_1 \\ C_1 \\ 0 \end{pmatrix} \quad \mathbf{J}_{O3} = \begin{pmatrix} -S_1 \\ C_1 \\ 0 \end{pmatrix} \quad \mathbf{J}_{O4} = \begin{pmatrix} -S_1 \\ C_1 \\ 0 \end{pmatrix} \quad \mathbf{J}_{O5} = \begin{pmatrix} -C_1 S_{234} \\ -S_1 S_{234} \\ 0 \end{pmatrix}$$

nas expressões anteriores utilizou-se uma expressão simplificada para representar os senos e cossenos:

$$S_1 = \text{Sin}(t1); C_1 = \text{Cos}(t1); C_{234} = \text{Cos}(t2+t3+t4); S_{234} = \text{Sin}(t2+t3+t4)$$

## Anexo F - Jacobiano Completo

$$J[1,1]=f_x(S_1f_yS_{234}d_5-S_1f_ya_3C_{23}-S_1f_ya_2C_2-S_1f_ya_1-v_0z)/z/f_y$$

$$J[1,2]=-(f_x^2C_1d_5C_{234}f_y+f_x^2C_1f_ya_3S_{23}+f_x^2C_1f_ya_2S_2+u_0S_{234}d_5f_yf_x-u_0f_yf_xa_3C_{23}-u_0f_yf_xa_2C_2-u_0v_0S_1zf_x-C_1zf_yf_x^2-C_1zf_yu_0^2)/z/f_y/f_x$$

$$J[1,3]=-(f_x^2C_1d_5C_{234}f_y+f_x^2C_1f_ya_3S_{23}+u_0S_{234}d_5f_yf_x-u_0f_yf_xa_3C_{23}-u_0v_0S_1zf_x-C_1zf_yf_x^2-C_1zf_yu_0^2)/z/f_y/f_x$$

$$J[1,4]=-(f_x^2C_1d_5C_{234}f_y+u_0S_{234}d_5f_yf_x-u_0v_0S_1zf_x-C_1zf_yf_x^2-C_1zf_yu_0^2)/z/f_y/f_x$$

$$J[1,5]=S_{234}(f_x^2S_1d_5C_{234}f_y+u_0v_0C_1zf_x-S_1zf_yf_x^2-S_1zf_yu_0^2)/z/f_y/f_x$$

$$J[2,1]=-f_y(C_1f_xS_{234}d_5-C_1f_xa_3C_{23}-C_1f_xa_2C_2-C_1f_xa_1-u_0z)/z/f_x$$

$$J[2,2]=-(f_y^2S_1d_5C_{234}f_x+f_y^2S_1f_xa_3S_{23}+f_y^2S_1f_xa_2S_2+v_0S_{234}d_5f_yf_x-v_0f_yf_xa_3C_{23}-v_0f_yf_xa_2C_2-S_1zf_xf_y^2-S_1zf_xv_0^2-u_0v_0C_1zf_y)/z/f_y/f_x$$

$$J[2,3]=-(f_y^2S_1d_5C_{234}f_x+f_y^2S_1f_xa_3S_{23}+v_0S_{234}d_5f_yf_x-v_0f_yf_xa_3C_{23}-S_1zf_xf_y^2-S_1zf_xv_0^2-u_0v_0C_1zf_y)/z/f_y/f_x$$

$$J[2,4]=-(f_y^2S_1d_5C_{234}f_x+v_0S_{234}d_5f_yf_x-S_1zf_xf_y^2-S_1zf_xv_0^2-u_0v_0C_1zf_y)/z/f_y/f_x$$

$$J[2,5]=-S_{234}(f_y^2C_1d_5C_{234}f_x-C_1zf_xf_y^2-C_1zf_xv_0^2+u_0v_0S_1zf_y)/z/f_y/f_x$$

## Tabela 2 – Características do Movemaster RV-E3J

Graus de liberdade.....	5
Método de tracção.....	Motores AC síncronos de magnetos permanentes
Método de detecção local.....	Codificador absoluto - <i>Encoders</i> ópticos absolutos
Comprimentos do Braço [mm] .....	Deslocamento do Braço 100 Comprimento do Ombro 250 Comprimento do Cotovelo 280 Deslocamento do Pulso 20 Comprimento do Pulso 85
Âmbito de operações..... [Graus] ([graus/s])	Braço W ±160 Ombro S ±180 Cotovelo E 135 Inclinação do Pulso P ±120 Rolagem do Pulso R ±200
Máxima velocidade resultante [mm/s] .	3500
Capacidade de carga [kgf] .....	3
Repetibilidade [mm] .....	0.04
Peso do manipulador [kgf] .....	33
Raio de captura [mm] .....	630
Mão Motorizada .....	Motor DC Força da mão de 0,5 a 7kgf Precisão de repetição 0.03mm

*Tabela 2 - Características do Manipulador Mitsubishi Movemaster RV-E3J*