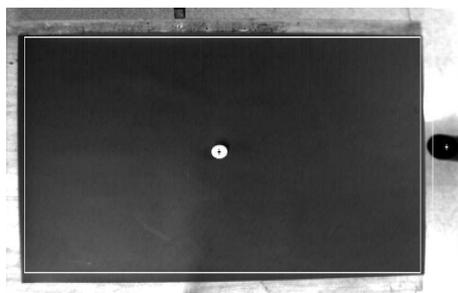




UNIVERSIDADE TÉCNICA DE LISBOA
INSTITUTO SUPERIOR TÉCNICO



ARQUITECTURAS DE SEGUIMENTO VISUAL
E CAPTURA POR UM MANIPULADOR ROBÓTICO
DE OBJECTOS EM MOVIMENTO

DINIS DOS ANJOS GUERRA FERNANDES

(Licenciado)

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

Orientador Científico:

Doutor Pedro Manuel Urbano de Almeida Lima

Constituição do Júri:

Doutor Jorge dos Santos Salvador Marques

Doutor Urbano José Carreira Nunes

Doutor Pedro Manuel Urbano de Almeida Lima

Doutor José Alberto Rosado dos Santos Victor

Lisboa, Dezembro de 1997

Resumo

O seguimento, captura e manipulação de objectos em movimento utilizando sistemas de visão e manipulação são tarefas dinâmicas com grande interesse para diversas aplicações industriais.

Esta Tese tem como objectivo o estudo comparativo de diferentes arquitecturas de controlo de um sistema de seguimento visual e captura de objectos em movimento. Nomeadamente, são implementadas e confrontadas estruturas com o controlo baseado na posição e com o controlo baseado na imagem.

É feita a caracterização dos subsistemas que constituem um sistema deste tipo: (1) um sistema de visão responsável pelo seguimento do objecto e pela computação em tempo real dos parâmetros que caracterizam o seu movimento, (2) um sistema dedicado à predição do movimento do objecto, (3) um sistema de controlo do manipulador adequado às tarefas de intercepção e captura do objecto. Estes subsistemas são abordados em termos genéricos e particularizados relativamente a uma instalação de ensaio, criada no âmbito desta Tese.

Apresentam-se resultados experimentais relativos à utilização da instalação de ensaio na tarefa da captura por um manipulador Puma 560 de bolas de ping-pong que rolam sobre uma mesa.

Palavras Chave: manipulação robótica, visão robótica, seguimento visual, captura baseada em visão, controlo em tempo real, controlo distribuído.

Abstract

Robotic vision and manipulator based servoing, catching and manipulation of moving objects are dynamic tasks of large interest for industrial applications.

The objective of this Thesis is to study and compare different control architectures for visual servoing and catching of moving objects. Notably, position-based and image-based control structures are implemented and confronted.

The characterization of the subsystems composing such a system is made: (1) a vision system responsible for object following and for the real-time computation of the parameters characterizing its motion, (2) a system devoted to the prediction of the object motion, (3) a manipulator control system suitable for the object interception and catching tasks. Those subsystems are tackled in generic terms and particularized for a testbed developed in the framework of this Thesis.

Experimental results are presented concerning the utilization of the testbed for the task of catching, by a Puma 560 manipulator, table-tennis balls rolling on a table.

Keywords: robotic manipulation, robotic vision, visual servoing, visual catching, real-time control, distributed control.

Agradecimentos

Em primeiro lugar devo um agradecimento muito especial ao Prof. Pedro Lima pelo apoio prestado e pela disponibilidade mostrada ao longo da orientação desta Tese de Mestrado.

Desejo também exprimir os meus agradecimentos a todas as outras pessoas que me apoiaram durante este trabalho. Aos meus Pais e à minha irmã, pelo apoio constante e encorajamento durante toda a minha vida. A todos os colegas do LCI, e do ISR em geral, quero agradecer a boa disposição e espírito de grupo criado. Finalmente aos amigos verdadeiros, simplesmente pela amizade.

À Junta Nacional de Investigação Científica, pelo financiamento desta tese, através da bolsa *PRAXIS XXI / BM / 6838 / 95*.

Índice

1	Introdução	1
1.1	Motivação	1
1.2	Perspectiva Histórica	2
1.3	Objectivos	4
1.4	Contribuições Originais	5
1.5	Organização da Tese	6
2	Seguimento Visual e Captura de Objectos Móveis	9
2.1	Sistema de Visão	10
2.1.1	Modelo de formação de imagem	11
2.1.2	Modelo da câmara	12
2.1.3	Calibração da câmara	14
2.1.4	Características da imagem	15
2.1.5	Parâmetros característicos do movimento do objecto	16
2.2	Sistema de Predição	18
2.3	Sistema de Controlo do Manipulador	19
2.3.1	Espaço junta e espaço operacional	20
2.3.2	Controlador PD com compensação gravítica em espaço junta	23
2.3.3	Controlador PD com compensação gravítica em espaço operacional	24
2.4	Arquitecturas de Seguimento Visual	26
3	Descrição da Instalação e Métodos	31
3.1	Sistema de Visão	33
3.1.1	Calibração da câmara	34

3.1.2	Características da imagem	35
3.1.3	Parâmetros característicos do movimento do objecto	38
3.2	Sistema de Predição	38
3.3	Sistema de Controlo do Manipulador	41
3.4	Arquitecturas de Seguimento Visual	43
3.4.1	Controlo baseado na posição	44
3.4.2	Controlo baseado na imagem	46
3.5	Método Experimental	49
4	Resultados	51
4.1	Sistema de Visão	51
4.2	Sistema de Predição	54
4.3	Sistema de Controlo do Manipulador	57
4.4	Arquitecturas de Seguimento Visual	68
5	Conclusões	79
5.1	Trabalho Futuro	81
A	Procedimento de Calibração da Câmara	85

1. Introdução

1.1 Motivação

Actualmente, não obstante o crescimento e evolução tecnológica na área da robótica, verifica-se que a grande maioria dos robôs industriais necessita de ambientes rigorosamente estruturados e adaptados às suas características. Todavia, a utilização de manipuladores robóticos em aplicações onde o ambiente de trabalho ou a localização dos objectos a manipular não são conhecidos com rigor tem um interesse crescente. Tarefas desta natureza exigem um incremento da versatilidade dos manipuladores, o qual é conseguido geralmente através da integração sensorial e, mais concretamente, pela utilização de sistemas de visão robótica.

O seguimento, captura e manipulação de objectos em movimento utilizando sistemas de visão e manipulação começou a despertar nos últimos anos a atenção de alguns investigadores. Estas tarefas dinâmicas constituem a base para variadas aplicações industriais, como por exemplo, o agarrar de objectos em movimento sobre um tapete rolante, a montagem de peças articuladas ou a manipulação a partir de um sistema robótico móvel. Por outro lado, é de notar que se trata de uma área de investigação abrangente, na medida em que requer a fusão de resultados de outras áreas, como a teoria do controlo, a computação em tempo real e o processamento de imagem.

Contudo, constata-se que a captura de objectos em movimento é uma tarefa dinâmica muito melhor desempenhada pelo ser humano, ainda que os manipuladores sejam na sua generalidade mais fortes, rápidos e tenham uma maior repetibilidade. A visão humana tem uma complexidade ainda não totalmente compreendida e as capacidades de coordenação e aprendizagem humanas são muito maiores do que os métodos artificiais desenvolvidos

até ao momento.

De qualquer forma, e em virtude dos recentes avanços tecnológicos, o seguimento visual, captura e manipulação de objectos revela-se como uma linha de investigação bastante promissora. Em particular, o aparecimento recente de sistemas de captura e processamento de imagem em tempo real a preços não proibitivos possibilita a integração eficiente da informação sensorial nas cadeias de controlo dos sistemas robóticos, e o aumento vertiginoso da velocidade de cálculo dos microprocessadores permite o processamento da informação em verdadeiro tempo real.

1.2 Perspectiva Histórica

Desde o trabalho pioneiro de Shirai e Inoue [51], em 1973, no qual foi utilizada retroacção visual para corrigir a posição de um braço robótico, uma grande variedade de trabalhos de investigação têm sido dedicados ao estudo do controlo visual de manipuladores. Uma revisão bastante completa da literatura neste campo encontra-se em [17]. Outras referências importantes são os livros de Corke [18] e Hashimoto [31], e um artigo de Hutchinson, Hager e Corke [40].

No entanto, a utilização de informação visual especificamente para o seguimento e captura de objectos em movimento por braços robóticos foi até ao momento pouco explorada. Ainda assim, os trabalhos de investigação nesta área, apresentados até ao momento, podem ser divididos em dois grupos: aqueles que possuem uma arquitectura de seguimento visual com controlo baseado na posição do effector terminal do manipulador e do(s) objecto(s) a capturar, e aqueles cujo controlo é baseado exclusivamente na imagem (ver secção 2.4).

Os trabalhos enquadrados no primeiro grupo, que a seguir se descrevem, têm a particularidade de requererem uma calibração entre a câmara e o manipulador muito precisa.

Em 1985, Andersson [5][6] usou uma câmara fixa e um braço robótico para capturar bolas de ping-pong que rolavam sobre uma mesa. Nesse trabalho, assumiu-se que a trajectória da bola era perfeitamente rectilínea, simplificando-se assim a predição do ponto de captura em detrimento de uma maior eficácia do sistema. De qualquer forma, o sistema era capaz de apanhar grande parte das bolas atiradas com uma velocidade inferior a 75

cm/s.

Em 1990, Houshangi [38] desenvolveu um sistema dotado de uma câmara estática capaz de capturar objectos animados de um movimento suave sobre uma superfície planar. As poses do manipulador relativas à posição inicial e à posição de captura eram previamente calculadas, o que limitava a flexibilidade do sistema. Por outro lado, assumiu-se que a geometria e a pose do objecto alvo eram conhecidas e que o objecto mantinha uma velocidade aproximadamente constante e sem mudanças de direcção acentuadas. Condicionando experimentalmente esses factores o controlo do manipulador foi simplificado e a predição da posição do objecto ao longo do tempo ficou sujeita a menos erros.

Em 1991, Hove *et al.* [39] propuseram um sistema capaz de apanhar uma bola atirada pelo ar usando um manipulador com quatro graus de liberdade e duas câmaras fixas. A geometria do objecto alvo (uma esfera) possibilitava a sua captura a partir de qualquer ângulo de aproximação e desta forma foram eliminadas quaisquer preocupações ao nível de ângulos de aproximação e alinhamento da garra do manipulador. Cada bola atirada pelo ar tinha uma velocidade típica de 3-6 m/s e o espaço de tempo entre o instante em que saía das mãos do lançador e era interceptada pelo manipulador era de cerca de 0,5 s. Posteriormente, em 1992, Kimura *et al.* [42] efectuaram várias melhorias ao sistema, nomeadamente ao nível do processamento de imagem, por forma a que a informação visual apresentasse uma maior fiabilidade e conseqüentemente o sistema apresentasse uma maior eficácia de captura. Em 1995, Hong *et al.* [34][35] incorporaram um sistema de visão activa (com duas câmaras a cores) no sistema original, o qual apresenta como vantagens um substancial aumento do campo de visão e o facto de requer um ambiente não estruturado, ao contrário do que acontecia com os trabalhos anteriores, uma vez que o objecto é pintado com uma cor pouco comum (fluorescente), e portanto é facilmente detectado através da utilização de técnicas de detecção de cor.

Em 1994, Buttazzo *et al.* [13] utilizaram uma câmara fixa para determinar quando e onde é que um objecto iria cruzar uma linha previamente definida no plano do chão. Esta informação era utilizada pelo sistema de controlo para colocar um cesto, montado no efector terminal do manipulador, sobre o objecto, quando este cruzava a linha pré-definida. A utilização de um cesto como efector terminal eliminou problemas de alinhamento da garra e permitiu alguma tolerância a erros quer de calibração, quer de posicionamento

do manipulador, quer sensoriais. O sistema era capaz de capturar objectos com uma velocidade inferior a 70 cm/s.

Allen *et al.* [3][4] apresentaram um sistema com duas câmaras fixas e um manipulador que agarra um objecto que se move ao longo de uma trajectória oval. O algoritmo de controlo tinha o inconveniente de, uma vez iniciada a trajectória de captura, não permitir posteriores modificações a essa trajectória.

Alguns dos sistemas dotados de uma arquitectura de seguimento visual com controlo baseado na imagem, são descritos de seguida.

Em 1996, Smith *et al.* [53] utilizaram um manipulador com uma câmara montada no braço (*eye-in-hand configuration*) para agarrar objectos (prismas) em movimento. Este sistema tem a particularidade de usar retroacção visual também na fase durante a qual o objecto é agarrado, o que permite um maior número de capturas com sucesso mesmo com a existência de erros associados ao posicionamento do manipulador e à informação visual.

Também em 1996, Ghosh *et al.* [28] serviram-se de uma câmara estática e de um manipulador para agarrar objectos colocados aleatoriamente numa mesa rotativa. É de realçar que, em cada uma das experiências, a localização da câmara podia ser alterada.

Constata-se que, em cada um destes trabalhos foi utilizada uma única arquitectura de controlo, o que torna difícil a comparação directa entre arquitecturas distintas, quer em termos de implementação, quer no que toca ao desempenho do sistema.

1.3 Objectivos

O principal objectivo do trabalho descrito nesta tese é o estudo comparativo de arquitecturas de controlo de um sistema de seguimento visual e captura de objectos em movimento, utilizando uma instalação de ensaio especialmente criada para o efeito.

O cumprimento deste objectivo passa pelo desenvolvimento dos vários subsistemas que integram o sistema de seguimento visual e captura de objectos em movimento propriamente dito:

- um sistema de visão que faça o seguimento de um objecto dinâmico num espaço bidimensional (2D) bem estruturado, e que realize uma computação em tempo real

dos parâmetros que caracterizam o seu movimento;

- um sistema que faça predição do movimento do objecto;
- um sistema de controlo do manipulador adequado às tarefas de intercepção e captura do objecto.

Por outro lado, deste trabalho resulta a criação e desenvolvimento de uma plataforma de seguimento visual que possibilitará futuros trabalhos de investigação nesta área.

Em termos experimentais, os objectivos passam por se conseguir utilizar com sucesso o sistema de seguimento visual e captura de objectos na tarefa que a seguir se descreve. Um indivíduo atira bolas de ping-pong, de modo a rolaem sobre uma mesa, na direcção de um manipulador Puma 560. O sistema de visão faz a aquisição de imagens da bola a rolar sobre a mesa. A partir destas imagens são calculadas estimativas da posição da bola e é feita a predição do ponto onde a bola irá cair - “ponto de captura”. O manipulador é controlado de modo a posicionar o seu effector terminal nesse ponto com a orientação adequada. Presa na garra (*gripper*) do manipulador encontra-se uma pequena “caneca”. A captura tem sucesso quando a bola cai dentro da “caneca”. Esta experiência foi inspirada num artigo de Andersson [5][6], onde é descrita uma aplicação semelhante.

1.4 Contribuições Originais

Em cada uma das arquitecturas de seguimento visual e captura usadas no âmbito do trabalho descrito nesta Tese, o sistema de controlo do manipulador encontra-se implementado em cadeia fechada durante cada intervalo de amostragem da imagem do objecto a capturar. Desta forma, os binários enviados para o manipulador são computados a um ritmo superior ao de amostragem da imagem, por uma cadeia de controlo que utiliza como referência a última predição do ponto de captura. Não são, assim, utilizados geradores de trajectórias, ao contrário do que é comum, os quais não permitem geralmente o replaneamento da trajectória do manipulador durante a sua execução, face a uma nova referência originada pelo sistema de predição.

Outro aspecto original é o facto de a predição do movimento do objecto se basear na soma pesada de um conjunto de funções de base representativas da dinâmica da trajectória

do objecto. A metodologia apresentada pode facilmente ser aplicada a outras situações ou objectos através da escolha de novas funções de base.

A contribuição original mais relevante deste trabalho é a comparação entre a arquitectura de seguimento visual com controlo baseado na posição e a arquitectura cujo controlo é baseado exclusivamente na imagem, usando uma mesma aplicação de ensaio experimental. É assim possível comparar as duas soluções em termos do desempenho e facilidade de implementação.

1.5 Organização da Tese

No Capítulo 2 introduzem-se os principais aspectos que caracterizam um sistema genérico de seguimento visual e captura de objectos em movimento e apresentam-se as várias arquitecturas de seguimento visual que podem servir de base a um sistema desta natureza. Apresenta-se alguma teoria relevante, e introduz-se a terminologia e notação utilizadas. No que concerne ao sistema de visão, são referidas várias alternativas relacionadas com a constituição e configuração de um sistema genérico, e são introduzidas questões como o modelo de formação da imagem e a calibração da câmara. O problema da extracção de características da imagem é relacionado com a extracção de informação relativa aos parâmetros que caracterizam o movimento do objecto. São também definidos alguns conceitos relacionados com a predição da trajectória de objectos em movimento e a sua intercepção e captura por um manipulador robótico. Em relação ao controlo do manipulador, são introduzidas as noções de controlo em espaço junta e controlo em espaço operacional, e são apresentados controladores proporcionais-derivativos (PD) com compensação gravítica. Por fim são descritas as estruturas de controlo olha-e-move dinâmico e servo visual directo que constituem as duas grandes classes de arquitecturas de seguimento visual. Por outro lado, é também feita a distinção entre as arquitecturas com o controlo baseado na posição e com o controlo baseado na imagem, e a separação entre sistemas com o efector terminal em cadeia aberta e fechada, do ponto de vista da imagem observada.

A descrição da instalação experimental que serviu de base para a comparação entre as várias arquitecturas de seguimento visual utilizadas é feita no Capítulo 3. É apresentado

detalhadamente o sistema de visão e quais as características de imagem utilizadas, assim como os métodos usados na sua extracção. Indicam-se também quais os parâmetros característicos do movimento do objecto fornecidos ao sistema de predição. No que toca à predição do movimento do objecto, mostra-se como a trajectória do objecto pode ser modelada por um conjunto de funções de base representativas da sua dinâmica. É apresentada uma metodologia, baseada no princípio dos mínimos quadrados, usada na estimação dos parâmetros que combinam linearmente cada uma dessas funções. A maneira como foram afinados os controladores do manipulador é também ilustrada. É ainda dada uma especial atenção às arquitecturas implementadas: uma com controlo baseado na posição e com o efector terminal em cadeia aberta, e outra com controlo baseado na imagem e com o efector terminal em cadeia fechada. Finalmente descreve-se o método experimental utilizado na comparação entre as arquitecturas de controlo consideradas.

No Capítulo 4 verifica-se qual o desempenho de cada um dos subsistemas que compõem o sistema de seguimento visual e captura de objectos e comparam-se as arquitecturas implementadas, através da apresentação de resultados experimentais.

Finalmente, no Capítulo 5 apresentam-se as conclusões retiradas deste trabalho e estabelecem-se as direcções principais de investigação para trabalho futuro.

2. Seguimento Visual e Captura de Objectos Móveis

Neste capítulo, é feita a caracterização genérica de um sistema de seguimento visual e captura de objectos em movimento e apresentam-se as várias arquitecturas de seguimento visual que podem servir de base a um sistema deste tipo. Abordam-se também alguns tópicos relevantes para o seguimento visual, e introduz-se a terminologia e notação utilizadas.

O termo *Seguimento Visual* (*Visual Servoing*) foi introduzido por Hill e Park [33] em 1979. No que diz respeito à manipulação robótica designa usualmente a utilização de informação visual para controlar a posição e orientação, normalmente designadas por pose, do efector terminal de um braço robô.

Um sistema de seguimento visual e captura de objectos móveis, em termos genéricos, resulta da integração dos seguintes subsistemas:

- um sistema de visão que consiga fazer o seguimento do objecto dinâmico e que faça uma computação em tempo real dos parâmetros que caracterizam o seu movimento;
- um sistema que faça predição do movimento do objecto;
- um sistema de controlo do manipulador adequado às tarefas de intercepção e captura do objecto.

A concepção de um sistema desta natureza envolve a escolha entre várias alternativas relativamente a cada um dos subsistemas que o compõem, cada uma delas com as suas vantagens e inconvenientes. Essas escolhas dependem naturalmente das características da aplicação a que se destina o sistema. Em particular, importa saber qual o tipo de ambiente

onde se desloca o objecto a capturar - se é estruturado ou não, se é bidimensional (2D) ou tridimensional (3D), quais as dimensões do espaço de trabalho - e qual a geometria do objecto a capturar.

As características do ambiente determinam fundamentalmente quais as alternativas mais indicadas no que diz respeito ao sistema de visão: número de câmaras, configuração da(s) câmara(s) e qual a complexidade dos algoritmos de processamento e extracção de características da imagem.

A geometria do objecto desempenha também um papel muito importante. Objectos cilíndricos [38] e prismas [53] requerem um perfeito alinhamento da garra do manipulador. Objectos com dimensões e formatos irregulares requerem mecanismos elaborados de reconhecimento de objectos e um planeamento da captura dos objectos complexo [11][43][54]. Pelo contrário, a utilização de objectos esféricos facilita o reconhecimento dos objectos e o planeamento de trajectórias para alinhamento da garra do manipulador [5][6][34][35][39][42].

2.1 Sistema de Visão

O sistema de visão tem como funções fazer a detecção, o reconhecimento e o seguimento do objecto dinâmico, com a finalidade de extrair em tempo real informação relativa aos parâmetros que caracterizam o seu movimento, os quais serão usados pelo sistema de predição. Além disso, quando o controlo é baseado na imagem (ver secção 2.4) essas mesmas funções dizem respeito não só ao objecto a capturar, mas também ao effector terminal do manipulador.

Em termos físicos, as principais escolhas relativamente aos sistemas de visão residem essencialmente no número de câmaras e na sua configuração.

No que concerne ao número de câmaras, dos sistemas de visão fazem parte, normalmente, uma - sistema monocular - ou duas câmaras - sistema *stereo*. Os sistemas monoculares têm a desvantagem de não poderem fornecer a localização tridimensional de um objecto presente na imagem, a não ser que empreguem algum método de recuperação de profundidade [52], ou que a localização do objecto esteja restringida a um plano conhecido. Por outro lado, os sistemas *stereo* requerem o conhecimento da orientação relativa

entre os sistemas de coordenadas das duas câmaras e a resolução do problema da correspondência entre pontos de duas ou mais imagens, resultantes da projecção de um mesmo ponto 3D.

Os sistemas de visão adoptam tipicamente uma das seguintes configurações da(s) câmara(s): câmara(s) fixa(s) [3][4][5][6][13][38][39][42], câmara montada no efector do manipulador (*eye-in-hand*) [24][41][53][57], ou constituindo um sistema de visão activa com câmara(s) montada(s) num suporte móvel (*pan-tilt*) [34][35] ou num outro manipulador [32].

Nas subsecções seguintes são abordadas questões relacionadas com a modelização e calibração da(s) câmara(s), e com a extracção e utilização de características da imagem para a detecção, reconhecimento e seguimento do objecto em movimento. É também discutida a extracção dos parâmetros que caracterizam o movimento do objecto, usados pelo sistema de predição.

2.1.1 Modelo de formação de imagem

Para controlar um braço robótico usando informação proveniente de um sistema de visão é necessário compreender os aspectos geométricos do processo de formação da imagem. A câmara contém uma lente que forma uma projecção 2D da cena no plano da imagem, onde o sensor está colocado. Esta projecção causa uma perda directa da informação de profundidade, na medida em que cada ponto no plano da imagem corresponde a uma linha no espaço 3D. Por conseguinte, é necessária informação adicional para que se consigam determinar as coordenadas 3D correspondentes ao ponto no plano da imagem. Esta informação pode ser obtida usando duas câmaras (visão *stereo*), múltiplas vistas com uma única câmara, conhecendo a relação geométrica entre diversos pontos do objecto alvo, ou, limitando o espaço de trabalho a um plano, tal como acontece neste trabalho, onde o movimento da bola de ping-pong está confinado ao plano do tampo da mesa, o qual tem uma profundidade conhecida.

Neste trabalho, foi utilizado o modelo de formação de imagem por projecção perspectiva [36], pela sua simplicidade e boa aproximação à realidade. Este modelo encontra-se representado na Figura 2.1.

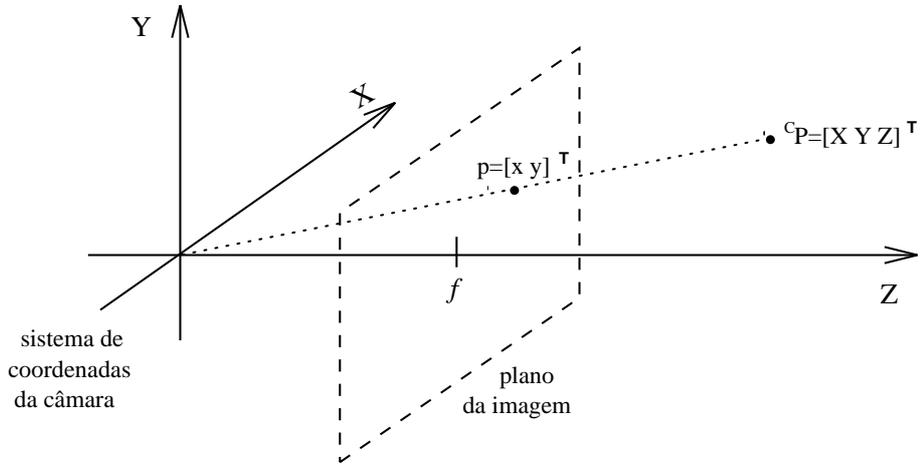


Fig. 2.1: Modelo de formação de imagem por projecção perspectiva.

Desta forma, assumindo que a geometria projectiva da câmara é modelada pela projecção perspectiva, um ponto 3D do mundo ${}^C\mathbf{P} = {}^C[X \ Y \ Z]^T$, expresso no sistema de coordenadas da câmara $\{C\}$, tem como projecção no plano da imagem o ponto $\mathbf{p} = [x \ y]^T$, sendo x e y dados por:

$$\begin{aligned} x &= f \frac{X}{Z}, \\ y &= f \frac{Y}{Z}, \end{aligned} \quad (2.1)$$

onde f é a distância focal da lente.

2.1.2 Modelo da câmara

As equações (2.1) caracterizam a formação de uma imagem através da projecção no plano da imagem de pontos 3D expressos no sistema de coordenadas da câmara $\{C\}$. No entanto, em termos genéricos, é mais conveniente encontrar a projecção no plano da imagem de pontos 3D expressos num sistema de coordenadas do mundo $\{M\}$ não coincidente com o da câmara. Neste trabalho e na maior parte dos sistemas de seguimento visual descritos na literatura, o sistema de coordenadas do mundo escolhido é o sistema de coordenadas da base do manipulador $\{0\}$. Assim, um ponto 3D ${}^C\mathbf{P} = {}^C[X \ Y \ Z]^T$, expresso no sistema de coordenadas da câmara $\{C\}$, relaciona-se com a sua representação ${}^M\mathbf{P} = {}^M[X \ Y \ Z]^T$

no sistema de coordenadas do mundo $\{M\}$, através da seguinte expressão [50]:

$${}^C\mathbf{P} = {}^C\mathbf{R}_M {}^M\mathbf{P} + {}^C\mathbf{o}_M. \quad (2.2)$$

ou, em coordenadas homogéneas [22]:

$${}^C \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} {}^C\mathbf{R}_M & {}^C\mathbf{o}_M \\ 0 & 0 & 0 & 1 \end{bmatrix} {}^M \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (2.3)$$

onde ${}^C\mathbf{R}_M$ representa a matriz de rotação (3×3) entre os dois sistemas de coordenadas, e ${}^C\mathbf{o}_M$ é a origem de $\{M\}$ expressa em $\{C\}$, representando portanto uma translação. Estas equações traduzem uma *transformação de coordenadas* (translação e rotação) entre os dois sistemas de coordenadas. A matriz da equação (2.3) representa os parâmetros extrínsecos da câmara.

Por outro lado, constata-se que nas equações (2.1), os pontos da imagem são expressos em unidades métricas e não no índice dos pixels. Os parâmetros intrínsecos da câmara, que descrevem a dimensão dos pixels da imagem (sob a forma de um escalamento $[k_x \ k_y]$) e a posição $[c_x \ c_y]^T$ do centro da imagem (intersecção do eixo óptico com o plano da imagem), permitem relacionar a representação de um ponto da imagem em pixels $[u \ v]^T$ com a representação do mesmo ponto em unidades métricas $[x \ y]^T$ da seguinte forma:

$$\begin{aligned} u &= k_x x + c_x, \\ v &= k_y y + c_y. \end{aligned} \quad (2.4)$$

As equações (2.1) e (2.4) podem também ser escritas em coordenadas projectivas, obtendo-se respectivamente:

$$\begin{bmatrix} \lambda x \\ \lambda y \\ \lambda \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} {}^C \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (2.5)$$

e

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} k_x & 0 & c_x \\ 0 & k_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda x \\ \lambda y \\ \lambda \end{bmatrix}, \quad (2.6)$$

onde λ é uma constante real.

Juntando as equações anteriores (projectão perspectiva, parâmetros intrínsecos e extrínsecos da câmara) obtém-se o *modelo da câmara* completo que relaciona um ponto ${}^M[X \ Y \ Z]^T$ expresso no sistema de coordenadas do mundo com a sua projectão no plano da imagem $[u \ v]^T$:

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} k_x & 0 & c_x \\ 0 & k_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} c_{\mathbf{R}_M} & & & \\ 0 & 0 & 0 & 1 \\ c_{\mathbf{o}_M} & & & \end{bmatrix} \begin{matrix} {}^M \\ \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \end{matrix}. \quad (2.7)$$

2.1.3 Calibração da câmara

A implementação das equações (2.7), obtidas na secção anterior, requer o conhecimento da distância focal e dos parâmetros intrínsecos e extrínsecos da câmara. Embora estes parâmetros possam ser medidos directamente, é na maior parte dos casos mais prático proceder à sua determinação usando a própria câmara como aparelho de medida.

A expressão (2.7) pode ser escrita na seguinte forma:

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \mathbf{C} \begin{matrix} {}^M \\ \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \end{matrix}, \quad (2.8)$$

onde a matriz

$$\mathbf{C} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \end{bmatrix}, \quad (2.9)$$

é chamada de *matriz de calibração* e encapsula a informação dos parâmetros intrínsecos e extrínsecos da câmara, bem como da projectão perspectiva.

A *calibração da câmara* consiste em estimar os elementos da matriz \mathbf{C} dados vários pontos da imagem $[u \ v]^T$ com coordenadas no mundo ${}^M[X \ Y \ Z]^T$ conhecidas.

Expandindo o produto matricial de (2.8) obtem-se o sistema de equações:

$$\begin{cases} \lambda u = c_{11}X + c_{12}Y + c_{13}Z + c_{14} \\ \lambda v = c_{21}X + c_{22}Y + c_{23}Z + c_{24} \\ \lambda = c_{31}X + c_{32}Y + c_{33}Z + c_{34} \end{cases} \quad (2.10)$$

Por sua vez, eliminando λ do sistema de equações anterior:

$$\begin{cases} u(c_{31}X + c_{32}Y + c_{33}Z + c_{34}) - (c_{11}X + c_{12}Y + c_{13}Z + c_{14}) = 0 \\ v(c_{31}X + c_{32}Y + c_{33}Z + c_{34}) - (c_{21}X + c_{22}Y + c_{23}Z + c_{24}) = 0 \end{cases} \quad (2.11)$$

Por conseguinte, tendo duas equações e doze incógnitas — os coeficientes c_{11}, \dots, c_{34} — são necessários pelo menos seis pontos da imagem $[u \ v]^T$ com coordenadas no mundo ${}^M[X \ Y \ Z]^T$ conhecidas.

Desta forma, o procedimento de calibração consiste em utilizar o conjunto de pontos conhecidos para resolver as equações (2.11) por forma a serem determinados os valores dos coeficientes da matriz de calibração.

2.1.4 Características da imagem

Uma *característica da imagem (image feature)* é qualquer característica estrutural que possa ser extraída da imagem. Tipicamente, corresponde à projecção de uma característica física de um objecto (ou do efector terminal do manipulador) no plano de imagem da câmara. Define-se um *parâmetro característico da imagem* como qualquer quantidade real que possa ser calculada a partir de uma ou mais características de uma imagem [40].

Alguns parâmetros característicos da imagem usados no seguimento visual são, por exemplo, as coordenadas no plano da imagem de pontos presentes na imagem [14][23], a distância entre dois pontos no plano da imagem e a orientação da linha que liga os dois pontos [24][26], a área de uma superfície projectada na imagem [49], a centróide e momentos de maior ordem [36] de uma superfície projectada na imagem [5][6][8][13][34][35][49][57], os parâmetros de linhas no plano da imagem [23], e os parâmetros de uma elipse no plano da imagem [23][28].

A escolha das características da imagem e dos parâmetros característicos da imagem a utilizar, para além de depender da geometria do objecto, baseia-se em critérios relacionados com a capacidade de facilitarem a detecção e reconhecimento do objecto e de fornecerem a informação necessária à derivação dos parâmetros que caracterizam o movimento do objecto.

O problema da selecção das características da imagem a utilizar em aplicações de seguimento visual encontra-se discutido em [24] [40] e [49], onde são apresentadas metodologias de selecção e vários exemplos de aplicação.

Uma vez seleccionado um conjunto de k parâmetros característicos da imagem, é definido um vector de parâmetros característicos da imagem:

$$\mathbf{f} = [f_1 \cdots f_k]^T. \quad (2.12)$$

A título de exemplo, considere-se o caso em que são escolhidos como parâmetros característicos da imagem as coordenadas de um ponto da imagem, correspondente à projecção de um ponto específico do objecto. Assumindo que o modelo de formação de imagem é a projecção perspectiva, o vector de parâmetros característicos da imagem é dado por:

$$\mathbf{f} = [u \ v]^T, \quad (2.13)$$

onde u e v são dados por (2.4).

2.1.5 Parâmetros característicos do movimento do objecto

O movimento de um objecto pode ser completamente descrito pelo conhecimento da sua pose, ou seja, das suas coordenadas de posição e orientação, em todos os instantes do tempo. Existem outras alternativas, como por exemplo, saber as coordenadas de posição e orientação iniciais e conhecer os valores da velocidade e da primeira derivada da orientação em todos os instantes do tempo. Assim, através da integração da velocidade e da primeira derivada da orientação ao longo do tempo podem ser obtidas as coordenadas de posição e orientação.

No entanto, quando se utiliza um sistema de visão, o que se extrai a partir da imagem não são a posição e a orientação tridimensionais do objecto, mas sim a sua projecção

no plano da imagem, como se viu na secção 2.1.1. O mesmo se passa com a velocidade do objecto: a perda de profundidade inerente à projecção perspectiva pode provocar discrepâncias entre a velocidade real do objecto e a sua velocidade no plano da imagem. Por conseguinte, o que se consegue extrair a partir das imagens é um conjunto de quantidades passíveis de descreverem o movimento de um objecto. Estas podem ser, por exemplo, a posição e a orientação do objecto na imagem, um subconjunto destas coordenadas, ou ainda, a velocidade média do objecto na imagem. Tais quantidades podem depois ser usadas directamente pelo sistema de predição, ou ser utilizadas na reconstrução das quantidades tridimensionais correspondentes através de técnicas que envolvem a calibração inversa, o modelo do objecto ou a geometria da aplicação [24]. Em qualquer um dos casos, é obtido um conjunto de quantidades, a que se dá o nome de *parâmetros característicos do movimento do objecto*, que podem estar definidos no espaço da imagem, ou estar definidos no espaço tridimensional do mundo.

A escolha dos parâmetros característicos do movimento do objecto é feita de acordo com a geometria do problema, as capacidades do sistema de visão e os requisitos do sistema de predição, por forma a que a sua utilização em conjunto com restrições ou outra informação relativa à cena permita ao sistema de predição calcular o ponto de captura.

Uma vez seleccionado um conjunto de l parâmetros característicos do movimento do objecto, define-se um *vector de parâmetros característicos do movimento do objecto*:

$$\mathbf{m} = [m_1 \cdots m_l]^T, \quad (2.14)$$

que, no caso dos parâmetros estarem definidos no espaço da imagem, pode coincidir com ou ser um subconjunto do vector de parâmetros de características da imagem.

Por outro lado, constata-se que a utilização de um sistema de visão, ou qualquer outro sensor, tem como consequência, por motivos físicos, que a recolha de informação não seja contínua no tempo, estando dependente do período de amostragem (constante ou variável) do sensor. Desta forma, dada um sequência de n imagens, é obtido um conjunto, discreto no tempo, constituído pelos vectores de parâmetros característicos do movimento do objecto correspondentes a cada uma das imagens:

$$\{\mathbf{m}_{t_1}, \dots, \mathbf{m}_{t_n}\}, \quad (2.15)$$

onde t_i representa o instante de tempo correspondente à imagem i , com $i = 1, \dots, n$. Este conjunto representa a informação que é disponibilizada ao sistema de predição.

Por fim, no que diz respeito à derivação dos parâmetros que caracterizam o movimento de um objecto, esta pode ser feita, a partir de uma sequência de imagens, através de dois métodos distintos:

- A extracção de um conjunto de características bidimensionais das imagens, correspondente a características tridimensionais do objecto na cena. As características extraídas de cada imagem podem ser por exemplo pontos, linhas ou curvas e dão origem aos parâmetros característicos da imagem (ver secção 2.1.4), os quais permitem a computação dos parâmetros característicos do movimento do objecto, ou podem mesmo constituir esses parâmetros.
- A determinação do fluxo óptico (*optical flow*) [37], ou seja, do movimento aparente dos padrões de brilho na imagem. O fluxo óptico usado em conjunto com restrições ou outra informação relativa à cena permite calcular a velocidade relativa entre o(s) objecto(s) da cena e a câmara. Exemplos da utilização do fluxo óptico no seguimento visual são [3] [4] [47].

Em [1] encontra-se uma descrição detalhada de cada uma destas abordagens ao problema da extracção do movimento a partir da análise de sequências de imagens, e são discutidos os méritos e inconvenientes relativos a cada uma delas.

2.2 Sistema de Predição

O sistema de predição tem como finalidade determinar o “ponto de captura”, ou seja, o ponto onde o objecto será interceptado e capturado pelo manipulador, fazendo uma predição do movimento do objecto.

Conforme referido na secção anterior, o sistema de visão fornece ao sistema de predição um conjunto, discreto no tempo, constituído pelos vectores de parâmetros característicos do movimento do objecto correspondentes a cada uma das imagens adquiridas. Esse conjunto de vectores de parâmetros característicos do movimento do objecto utilizado em

conjunto, ou não, com um modelo aproximado da dinâmica do movimento do objecto permite a predição da trajectória do objecto.

Assim, o “ponto de captura” é um ponto pertencente à trajectória prevista para o objecto, já dentro do espaço de trabalho do manipulador, e que é determinado de acordo com critérios de alinhamento da garra do manipulador [53] e de suavidade da trajectória do manipulador [34]. Uma vez determinado o “ponto de captura” pretende-se posicionar o efector do manipulador nesse ponto por forma a que o objecto seja capturado.

Importa notar que a predição do movimento do objecto pode ser efectuada no espaço da imagem ou no espaço tridimensional do mundo, conforme os parâmetros característicos do movimento do objecto estejam definidos num ou noutro espaço. Por conseguinte, o “ponto de captura” predito pode ser obtido em termos da sua posição $\mathbf{f}_{captura} = [u_{captura} \ v_{captura}]^T$ no plano da imagem, ou da sua posição $\mathbf{p}_{captura} = [X_{captura} \ Y_{captura} \ Z_{captura}]^T$ no mundo.

Em qualquer um dos casos, essa informação é convertida numa referência para o sistema de controlo do manipulador, sob a forma de uma pose desejada \mathbf{x}_d , de tal forma que o efector do manipulador seja posicionado nesse “ponto de captura” com uma determinada orientação.

Resta salientar que a predição pode também ser feita em termos temporais, ou seja, é predito não só o “ponto de captura”, mas também o instante do tempo no qual o objecto chegará a esse ponto, o que se revela absolutamente necessário quando se pretende que o manipulador chegue ao ponto ao mesmo tempo que o objecto. Tal acontece, por exemplo, quando se deseja uma trajectória de intercepção do manipulador suave e que acompanhe o movimento natural do objecto [34].

A forma como o preditor é implementado depende naturalmente da aplicação em causa. Na literatura relativa a sistemas de seguimento visual encontram-se vários exemplos de preditores baseados na filtragem de Kalman [34][55], nos filtros $\alpha - \beta$ e $\alpha - \beta - \gamma$ [3][4], e no método dos mínimos quadrados [5][6].

2.3 Sistema de Controlo do Manipulador

Em termos gerais, o controlo de um manipulador consiste na determinação dos binários a aplicar em cada uma das juntas, pelos respectivos actuadores, por forma a que o efec-

tor terminal do manipulador seja posicionado de acordo com as referências fornecidas, e portanto, de modo a que se garanta a execução da tarefa desejada.

2.3.1 Espaço junta e espaço operacional

Uma trajectória a descrever pelo efector terminal do manipulador é normalmente especificada por um conjunto de referências temporais de posição e orientação do efector terminal. A posição pode ser dada por um número mínimo de coordenadas relativas à geometria da estrutura, e a orientação pode ser especificada em termos de uma representação mínima, como por exemplo, as baseadas em ângulos de Euler ou ângulos RPY [22][50], que descreve a orientação do efector terminal relativamente ao sistema de coordenadas da base do manipulador.

Nesse caso, é possível descrever a pose do manipulador com n graus de liberdade por um vector de dimensão m , com $m \leq n$:

$$\mathbf{x} = \begin{bmatrix} \mathbf{p} \\ \phi \end{bmatrix}, \quad (2.16)$$

onde \mathbf{p} e ϕ representam respectivamente a posição e uma representação mínima de orientação da moldura do efector terminal em relação à moldura da base.

O vector \mathbf{x} é definido no espaço no qual a tarefa é especificada, o qual é usualmente designado por *espaço operacional*.

Por outro lado, 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:

$$\theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}, \quad (2.17)$$

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

As equações da cinemática directa do manipulador permitem calcular a posição e orientação do efector terminal em função das variáveis de junta:

$$\mathbf{x} = \mathbf{k}(\theta). \quad (2.18)$$

Por sua vez, o problema da cinemática inversa consiste na determinação dos valores das juntas correspondentes a uma dada pose do efector terminal. A solução deste problema revela-se de uma particular importância, na medida em que permite transformar as referências da tarefa, atribuídas ao efector terminal no espaço operacional, nas referências correspondentes em espaço junta que permitem a execução do movimento desejado. No entanto, a sua resolução pode ser muito complexa, uma vez que as equações a resolver são em geral não-lineares. Assim, nem sempre é possível encontrar uma solução na forma fechada, podendo além disso existir soluções múltiplas [50].

No que concerne a um sistema de seguimento visual e captura de objectos em movimento, em particular, verifica-se que na maioria dos casos são fornecidas ao sistema de controlo do manipulador referências correspondentes à pose desejada \mathbf{x}_d . Dada uma referência \mathbf{x}_d , é possível planear uma trajectória entre a pose actual \mathbf{x} e a pose desejada \mathbf{x}_d , e depois executá-la, ou, usar directamente \mathbf{x}_d como referência para o sistema de controlo do manipulador. A primeira solução tem geralmente o inconveniente de, uma vez iniciada uma trajectória, o gerador de trajectórias não permitir o seu replaneamento, até que o manipulador atinja \mathbf{x}_d , mesmo que o sistema de predição disponibilize entretanto uma nova pose desejada \mathbf{x}_d . Em qualquer dos casos, é desejável recorrer a controladores que assegurem estabilidade assintótica global, assumindo que as cadeias de seguimento dos objectos e de controlo do movimento das juntas apresentam uma relação aceitável entre os respectivos intervalos de amostragem (sendo a cadeia de controlo do movimento das juntas tipicamente pelo menos cinco a seis vezes mais rápida).

Face ao exposto anteriormente, é natural subdividir os esquemas de controlo de manipuladores robóticos em duas áreas gerais: o controlo em *espaço junta* (Figura 2.2) e o controlo em *espaço operacional* (Figura 2.3).

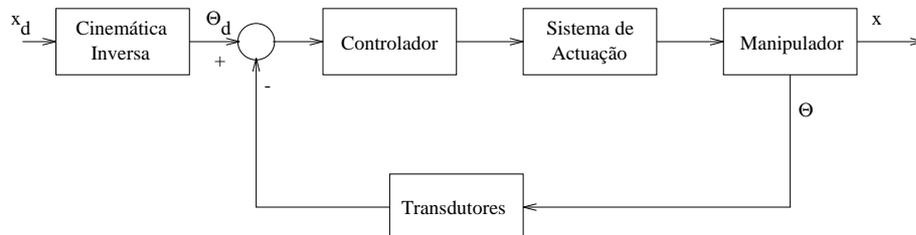


Fig. 2.2: Esquema geral do controlo em espaço junta.

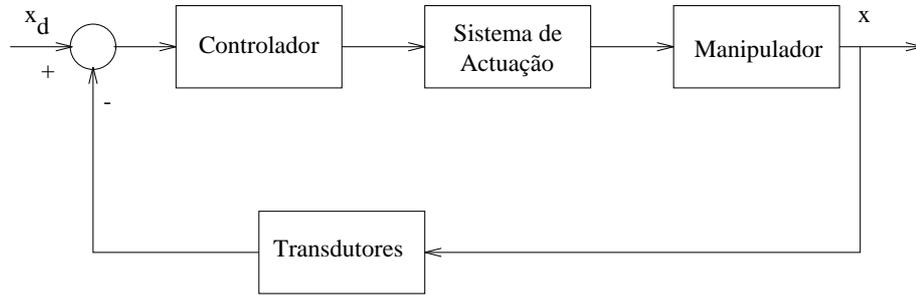


Fig. 2.3: Esquema geral do controlo em espaço operacional.

No controlo em espaço junta as referências são dadas directamente em termos de valores de junta, e por conseguinte, é necessário determinar o vector com os valores desejados das juntas θ_d a partir do vector com a pose desejada \mathbf{x}_d , recorrendo ao cálculo da cinemática inversa. Os esquemas de controlo em espaço junta podem ser distinguidos entre esquemas de controlo *descentralizados*, quando cada uma das juntas é controlada de forma independente das outras, e esquemas de controlo *centralizados*, quando os efeitos de interacção dinâmica entre as juntas são considerados [50]. Na realidade, o manipulador não é constituído por um conjunto de n sistemas desacoplados, mas é sim um sistema multivariável com n entradas (binários aplicados nas juntas) e n saídas (posições das juntas) que interagem entre si através de relações não-lineares. Não obstante, os esquemas descentralizados funcionam suficientemente bem a baixas velocidades e quando o seguimento exacto da trajectória não é crítico.

Em contrapartida, no controlo em espaço operacional o erro é calculado directamente a partir das poses desejada \mathbf{x}_d e actual \mathbf{x} , sendo portanto necessário transformar as variáveis de junta medidas θ nas variáveis em espaço operacional correspondentes. Esta transformação pode ser obtida através da cinemática directa, ou por observação simultânea de \mathbf{x} utilizando um sistema de visão.

Todos os esquemas de controlo em espaço operacional apresentam requisitos computacionais consideráveis, uma vez que têm a necessidade de efectuar uma série de computações na cadeia de controlo onde está implícita a avaliação da cinemática inversa. A presença desta carga computacional elevada aumenta de forma acentuada a duração do período de controlo em malha fechada, o que pode degradar o desempenho do sistema de controlo do manipulador.

Quando se pretende um controlo ponto a ponto onde apenas interessa chegar com exactidão à posição e orientação finais é suficiente um controlador do tipo Proporcional-Derivativo (PD) com compensação gravítica [50], o qual é caracterizado por uma acção linear descentralizada proporcional-derivativa (PD) e uma acção não linear centralizada de compensação dos termos gravíticos da dinâmica do manipulador.

Num sistema de seguimento visual e captura de objectos em movimento é, em geral, mais importante atingir a pose desejada correspondente ao ponto de intercepção e captura do objecto, do que executar essa trajectória com muito rigor, a não ser que haja constrangimentos físicos em relação ao espaço de trabalho. Por conseguinte, vão ser apenas considerados os controladores do tipo PD com compensação gravítica, que serão apresentados nas subsecções seguintes.

2.3.2 Controlador PD com compensação gravítica em espaço junta

Na Figura 2.4 encontra-se representado um controlador em espaço junta do tipo PD com compensação gravítica.

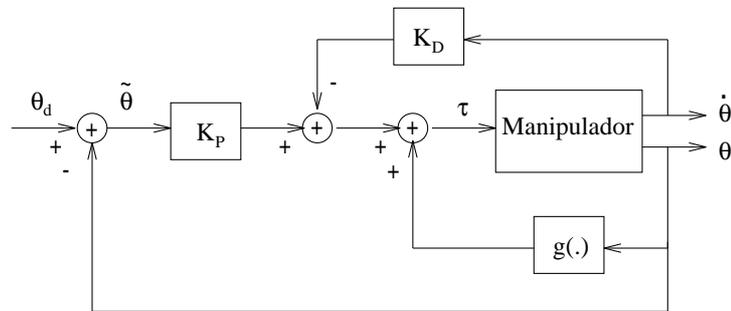


Fig. 2.4: Esquema de blocos do controlador PD com compensação gravítica em espaço junta.

Com este controlador o binário aplicado em cada uma das juntas é dado pela lei de controlo seguinte:

$$\tau = \mathbf{K}_P \tilde{\theta} - \mathbf{K}_D \dot{\theta} + \mathbf{g}(\theta), \quad (2.19)$$

onde

$$\tilde{\theta} = \theta_d - \theta \quad (2.20)$$

representa o erro entre os valores desejados para as variáveis de junta e os valores actuais, respectivamente θ_d e θ .

As características dinâmicas do sistema em malha fechada podem ser moldadas por ajuste das matrizes de ganho \mathbf{K}_P e \mathbf{K}_D , sendo \mathbf{K}_P uma matriz simétrica. A estabilidade assintótica global está assegurada, desde que \mathbf{K}_P e \mathbf{K}_D sejam matrizes definidas positivas [50].

2.3.3 Controlador PD com compensação gravítica em espaço operacional

Sendo o erro em espaço operacional dado por:

$$\tilde{\mathbf{x}} = \mathbf{x}_d - \mathbf{x}, \quad (2.21)$$

onde \mathbf{x}_d e \mathbf{x} representam respectivamente a pose desejada e a pose actual, a escolha da lei de controlo:

$$\tau = \mathbf{J}_A^T(\theta)\mathbf{K}_P\tilde{\mathbf{x}} - \mathbf{J}_A^T(\theta)\mathbf{K}_D\mathbf{J}_A(\theta)\dot{\theta} + \mathbf{g}(\theta) \quad (2.22)$$

garante que o erro tende assintoticamente para zero, sendo \mathbf{K}_P uma matriz simétrica, e desde que \mathbf{K}_P e \mathbf{K}_D sejam matrizes definidas positivas [50].

Em (2.22) $\mathbf{J}_A(\theta)$ representa o *Jacobiano analítico* [50]:

$$\mathbf{J}_A(\theta) = \frac{\partial \mathbf{k}(\theta)}{\partial \theta}, \quad (2.23)$$

que permite estabelecer a relação entre as derivadas no tempo da pose (utilizando uma representação mínima de orientação) e das variáveis de junta:

$$\dot{\mathbf{x}} = \mathbf{J}_A(\theta)\dot{\theta}, \quad (2.24)$$

A equação (2.24) representa a cinemática diferencial do manipulador.

Como pode ser visto na Figura 2.5, o esquema de controlo em espaço operacional revela uma grande analogia com o esquema de controlo em espaço junta da Figura 2.4. Com efeito, a lei de controlo (2.22) traduz-se numa acção não linear de compensação das forças gravíticas em espaço junta e numa acção linear de controlo proporcional-derivativo (PD) em espaço operacional.

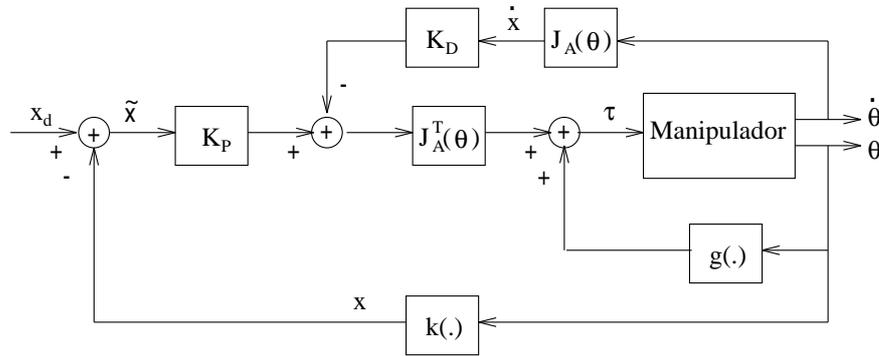


Fig. 2.5: Esquema de blocos do controlador PD com compensação gravítica em espaço operacional.

Para o controlador em espaço operacional não é necessário calcular a cinemática inversa uma vez que o erro é calculado directamente em espaço operacional. Em contrapartida, é necessário proceder-se ao cálculo das funções da cinemática directa $\mathbf{k}(\theta)$ e $\mathbf{J}_A(\theta)$, a fim de serem obtidos os valores de \mathbf{x} e $\dot{\mathbf{x}}$.

É importante notar também que o cálculo do Jacobiano analítico [50] do manipulador, requerido pelo controlador em espaço operacional, acarreta, para além de uma maior carga computacional, uma maior complexidade de controlo, na medida em que a presença de singularidades e/ou redundância do Jacobiano pode repercutir-se num desempenho indesejado do manipulador. Por exemplo, se ocorrer uma singularidade e o erro entrar no espaço nulo do Jacobiano analítico, o manipulador fica preso numa configuração diferente da pretendida. Por oposição, o esquema de controlo em espaço junta é naturalmente transparente a esta situação, uma vez que a questão da redundância do manipulador é resolvida pela cinemática inversa.

Em todo o caso, nas experiências levadas a cabo neste trabalho, este problema não se põe, dado que, como adiante se verá, a localização do efector terminal do manipulador se encontra restringida a uma recta no espaço operacional - “linha de captura” -, ou mais concretamente a uma vizinhança dessa recta que não inclui singularidades do Jacobiano.

2.4 Arquitecturas de Seguimento Visual

Os sistemas de seguimento visual podem ser agrupados em quatro grandes categorias, que correspondem à aplicação de dois critérios principais [40]:

- O tipo de estrutura de controlo usada divide os sistemas em duas classes distintas: *olha-e-move dinâmico* (*dynamic look-and-move*) e *servo visual directo* (*direct visual servo*). Nos sistemas *olha-e-move dinâmico*, a arquitectura de controlo é hierárquica, havendo uma cadeia de controlo exterior, fechada pela informação visual proveniente do sistema de visão, que fornece referências a uma outra cadeia de controlo interna — cadeia de controlo das juntas do manipulador. Pelo contrário, no segundo caso, o controlador do manipulador é inteiramente substituído por um controlador que computa directamente os binários a aplicar nas juntas a partir do erro entre as poses estimadas visualmente e a pose de referência do robot, pelo que é usada unicamente a informação visual na cadeia de controlo.
- Por sua vez a lei de controlo de cada uma das estruturas de controlo acima pode ser *baseada na posição* ou *baseada na imagem*. No *controlo baseado na posição* a lei de controlo reduz o erro entre as poses do manipulador e do objecto, sendo esta última estimada a partir da imagem, usando um modelo geométrico do objecto e o modelo da câmara. No *controlo baseado na imagem*, os valores de controlo são calculados directamente a partir das características da imagem.

Nas Figuras 2.6 a 2.9 encontram-se representações de cada uma destas arquitecturas, nas quais se assumiu a utilização de esquemas de controlo em espaço operacional. Estas figuras foram adaptadas de [40].

Na maior parte das aplicações de seguimento visual, pretende-se posicionar o efector terminal relativamente a um objecto. Assim, importa notar que é também normalmente feita uma distinção entre os sistemas que observam unicamente o objecto alvo — sistemas com o efector terminal em cadeia aberta (EOL - *endpoint open-loop*) — e os que observam o objecto alvo e o efector terminal — sistemas com o efector terminal em cadeia fechada (ECL - *endpoint closed-loop*). Importa notar que, na quase totalidade dos casos, as arquitecturas com a lei de controlo baseada na posição são do tipo EOL, uma vez que

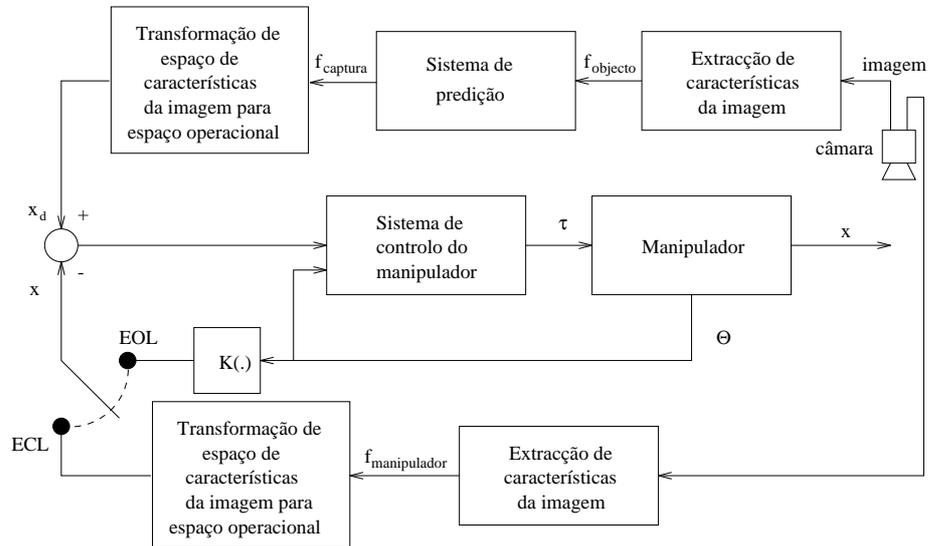


Fig. 2.6: Arquitectura olha-e-move dinâmica com controlo baseado em posição.

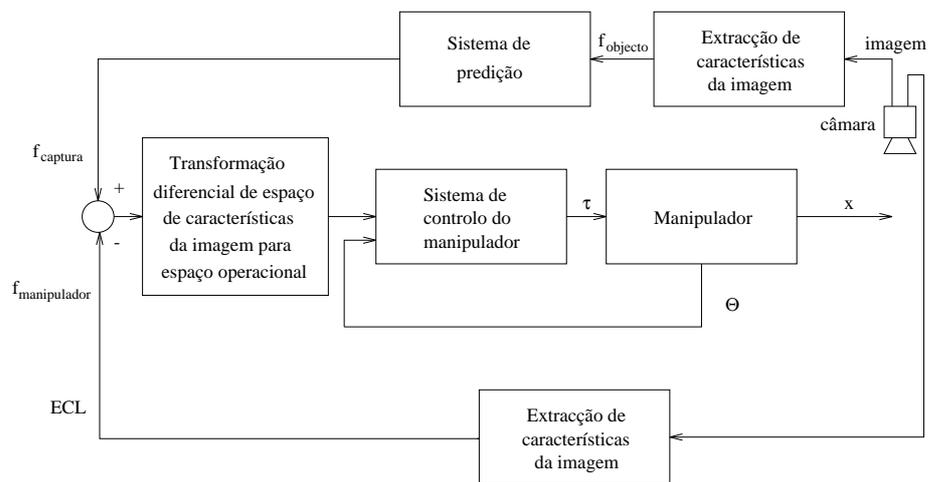


Fig. 2.7: Arquitectura olha-e-move dinâmica com controlo baseado na imagem.

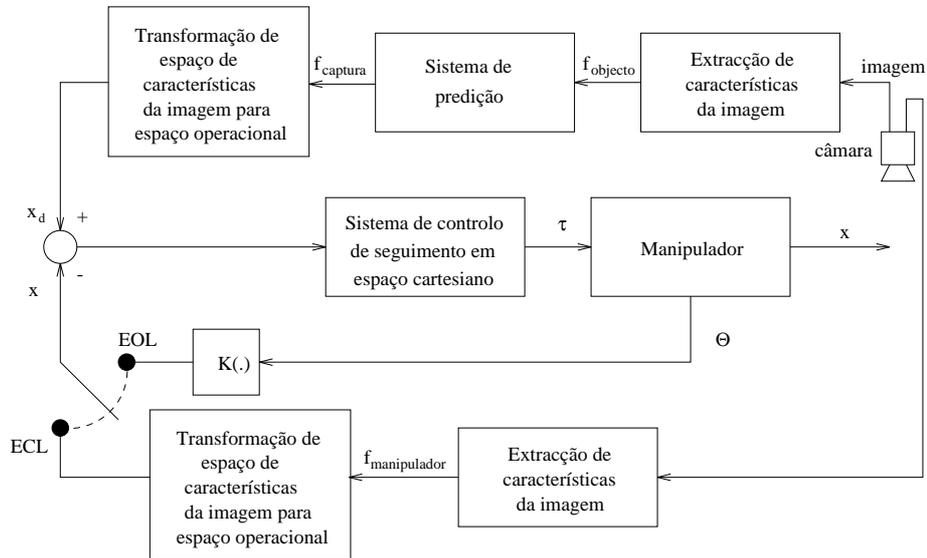


Fig. 2.8: Arquitectura servo visual directo com controlo baseado em posição.

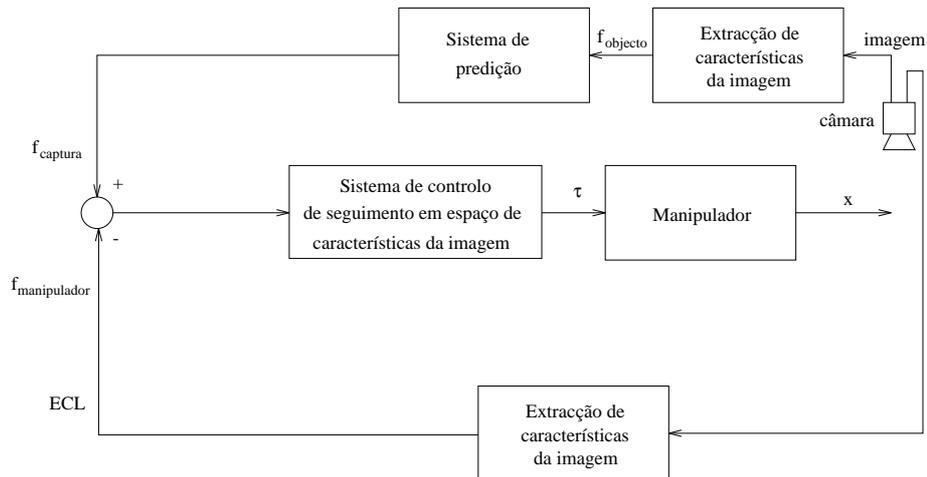


Fig. 2.9: Arquitectura servo visual directo com controlo baseado na imagem.

geralmente as equações da cinemática directa são mais fáceis de calcular e estão sujeitas a menos erros do que a calibração inversa da câmara.

Na quase totalidade dos sistemas descritos na literatura, foi adoptada a estrutura *olha-e-move dinâmico*. Isto, porque as frequências de amostragem dos sistemas de visão não são, regra geral, suficientemente elevadas para permitirem o controlo directo do manipulador. Por outro lado, com esta estrutura, o controlador visual não tem de levar em linha de conta as singularidades do manipulador.

No que diz respeito à lei de controlo, a principal vantagem do controlo baseado em posição é possibilitar a descrição das tarefas em termos de poses cartesianas, tal como é comum na robótica. Contudo, apresenta como desvantagem o facto de os valores de controlo serem computados usando quantidades dependentes dos parâmetros de calibração e da cinemática do sistema, o que torna este tipo de controlo bastante sensível aos erros de calibração da câmara e aos erros na cinemática directa do manipulador. Por outro lado, o controlo baseado na imagem apresenta como vantagens o facto de eliminar os erros inerentes à modelação do sensor e à calibração da câmara, para além de reduzir o tempo computacional e de eliminar a necessidade de interpretação da imagem. Todavia, o projecto do controlador revela-se mais complexo face ao controlo baseado em posição, na medida em que a dinâmica do conjunto composto pelo manipulador e pelo sistema de visão é não-linear e acoplada.

Por fim, note-se que nos sistemas com o efector terminal em cadeia aberta (EOL), a posição do efector em relação ao objecto é apenas determinada indirectamente pela relação existente entre a câmara e o manipulador. Erros nessa relação, devido a uma má calibração da câmara ou conhecimento incorrecto e/ou incompleto da cinemática directa do manipulador, conduzem a erros de posicionamento que não podem ser observados pelo sistema. De qualquer forma, importa salientar que este problema não é tão notório quando a câmara está acoplada ao braço robótico porque, uma vez posicionada a câmara relativamente ao objecto, resta apenas o erro na cinemática câmara/manipulador. Nos sistemas com o efector terminal em cadeia fechada (ECL), pelo contrário, é possível observar e corrigir tais erros de posicionamento, uma vez que o manipulador e o objecto a capturar estão ambos presentes na imagem.

3. Descrição da Instalação e Métodos

No âmbito deste trabalho, foi desenvolvida uma plataforma de seguimento visual e captura de objectos por um manipulador, que permite a implementação e teste das diversas arquitecturas de controlo que se desejavam estudar. Por conseguinte, a plataforma foi dotada de alguma modularidade e flexibilidade, o que permitirá também a sua utilização em posteriores trabalhos de investigação nesta área.

Neste capítulo é feita a descrição dessa instalação experimental. Em particular, é apresentado em detalhe o sistema de seguimento visual e captura de objectos em movimento usado e cada um dos subsistemas que o constituem. Além disso, são fundamentadas as decisões e soluções de compromisso efectuadas no desenvolvimento do sistema utilizado neste trabalho. Por fim, é descrito o método experimental utilizado na obtenção de resultados relativos às arquitecturas de controlo consideradas.

A plataforma é constituída por três subsistemas:

- O sistema de visão, o qual consiste numa câmara Electrim EDC-1000L, a respectiva placa para ligação ao barramento ISA de um PC, e um PC com microprocessador Pentium a 133 MHz e 16 Mb de memória RAM, responsável pelos algoritmos de processamento e de extracção de características da imagem.
- O sistema de predição, cujos algoritmos de predição da trajectória da bola e do “ponto de captura” se encontram também no PC pertencente ao sistema de visão.
- O sistema de manipulação, constituído por um manipulador Puma 560, por um PC com processador Pentium a 133 MHz e 16 Mb de memória RAM, e duas placas

Trident Robotics — TRC-004 e TRC-006 — através das quais é possível, a partir do PC, aceder à retroacção proveniente dos *encoders* e dos potenciómetros do manipulador, e aplicar binários nas suas juntas.

O conjunto destes três sistemas apoia-se numa arquitectura aberta para controlo de manipuladores, desenvolvida nos últimos anos no Instituto de Sistemas e Robótica, pólo do Instituto Superior Técnico (ISR/IST) [44][45]. Com esta arquitectura, a utilização de novas soluções de controlo é possível, ao contrário do que acontecia com a arquitectura de controlo fechada, baseada no controlador MARK-III e na linguagem VAL-II, de que o Puma 560 dispunha originalmente.

O sistema operativo usado nos PCs é o MS-DOS. As comunicações entre os diversos sistemas, mais concretamente entre os dois PC, processam-se através de TCP/IP sobre uma rede Ethernet local, utilizando *sockets*.

O diagrama de blocos da plataforma encontra-se representado na Figura 3.1.

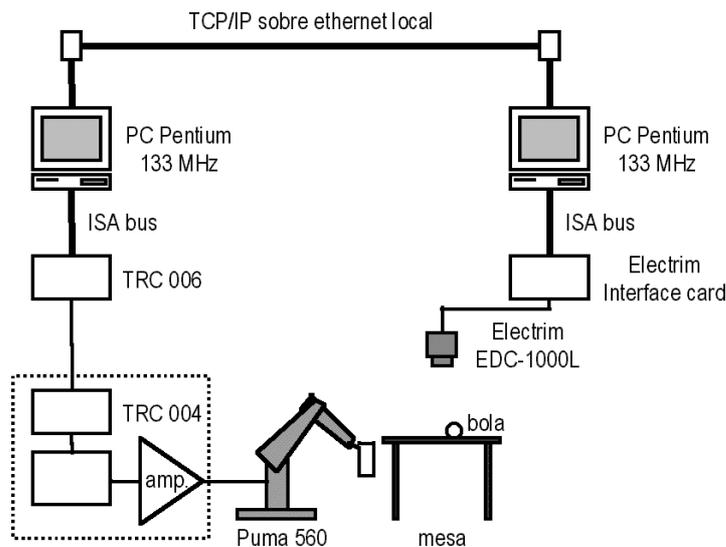


Fig. 3.1: O diagrama de blocos da plataforma de seguimento e captura visual.

A configuração usada na experiência encontra-se esquematizada na Figura 3.2. O tampo da mesa, sobre o qual as bolas se deslocam, tem um comprimento de 1,20 m e uma largura de 0,80 m. A câmara está montada num pórtico, a uma altura de aproximadamente 1.435 m do tampo da mesa. A câmara aponta para a mesa, na vertical, de tal forma que o plano da imagem e o tampo da mesa se encontram paralelos. Deste modo, a

área da bola na imagem é aproximadamente constante, o que facilita a sua detecção (ver secção 3.1.2).

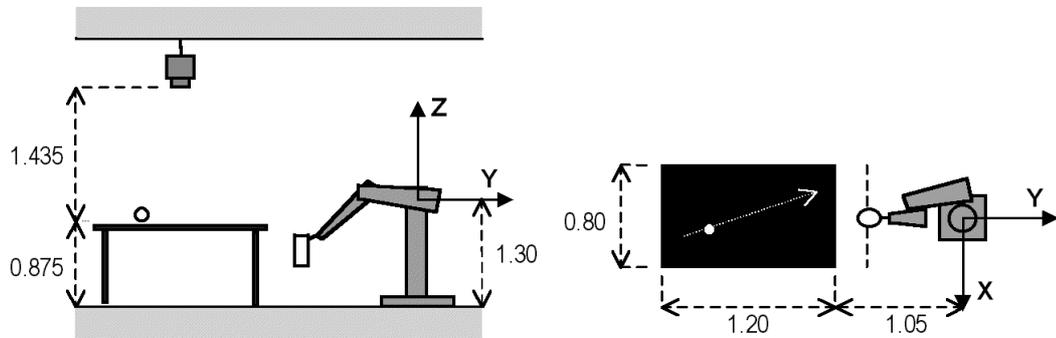


Fig. 3.2: Configuração usada na experiência: a) vista de lado; b) vista de cima.

Uma vez que os objectivos deste trabalho se prendem essencialmente com o estudo de arquitecturas de seguimento e captura visual distintas, utilizou-se um ambiente experimental estruturado, o qual não tem grandes exigências em termos de detecção e reconhecimento de objectos. Desta forma, conseguiu-se que o sistema de visão disponibilizasse a informação necessária à maior velocidade possível e a baixo custo.

Neste trabalho, o objecto alvo é uma bola de ping-pong. Sendo um objecto esférico, questões como o reconhecimento de objectos e o planeamento de trajectórias para alinhamento da garra do manipulador ficam bastante simplificadas, além da própria captura, o que vai de encontro aos objectivos deste trabalho, uma vez que se pretende essencialmente dar ênfase à forma como a escolha da arquitectura de controlo influi nos resultados.

3.1 Sistema de Visão

O sistema de visão contém uma única câmara fixa - modelo Electrim EDC-1000L -, que adquire imagens com uma resolução de 242 linhas por 753 colunas e com 256 tons de cinzento.

O período de amostragem do sistema de visão depende do tempo de exposição usado, o qual, por sua vez, é função da luminosidade ambiental. Em condições normais de luminosidade, e só adquirindo metade das linhas da imagem (linhas ímpares), conseguem-se adquirir, em média, cerca de 9,56 imagens por segundo, o que corresponde a um período

de amostragem médio aproximadamente igual a 104,5 ms.

3.1.1 Calibração da câmara

Neste caso particular, o problema da calibração da câmara, apresentado na secção 2.1.3, encontra-se simplificado, uma vez que a bola se desloca ao longo de um plano - plano da mesa. Este plano é descrito por:

$$Z = Z_{mesa}, \quad (3.1)$$

onde Z_{mesa} é uma constante que representa a altura da mesa face à origem do sistema de coordenadas do mundo - sistema de coordenadas da base do manipulador $\{0\}$.

Dada esta restrição, as equações da calibração da câmara (2.11) simplificam-se para:

$$\begin{cases} u(c_{31}X + c_{32}Y + c_{33}Z_{mesa} + c_{34}) - (c_{11}X + c_{12}Y + c_{13}Z_{mesa} + c_{14}) = 0 \\ v(c_{31}X + c_{32}Y + c_{33}Z_{mesa} + c_{34}) - (c_{21}X + c_{22}Y + c_{23}Z_{mesa} + c_{24}) = 0 \end{cases} \quad (3.2)$$

Agrupando os termos constantes e fazendo:

$$\begin{cases} c'_{13} = c_{13}Z_{mesa} + c_{14} \\ c'_{23} = c_{23}Z_{mesa} + c_{24} \\ c'_{33} = c_{33}Z_{mesa} + c_{34} \end{cases}, \quad (3.3)$$

obtem-se:

$$\begin{cases} c_{11}X + c_{12}Y + c'_{13} - c_{31}uX - c_{32}uY - c'_{33}u = 0 \\ c_{21}X + c_{22}Y + c'_{23} - c_{31}vX - c_{32}vY - c'_{33}v = 0 \end{cases} \quad (3.4)$$

Esta simplificação do problema torna a matriz de calibração numa matriz quadrada de dimensão 3, e portanto invertível, desde que não seja singular:

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \mathbf{C}' \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}, \quad (3.5)$$

onde a matriz

$$\mathbf{C}' = \begin{bmatrix} c_{11} & c_{12} & c'_{13} \\ c_{21} & c_{22} & c'_{23} \\ c_{31} & c_{32} & c'_{33} \end{bmatrix}. \quad (3.6)$$

O número de coeficientes a determinar reduziu-se assim de doze para nove. Contudo, uma vez que o sistema de equações (3.4) é homogéneo, a transformação não é afectada por um escalamento uniforme de todos os coeficientes, pelo que na realidade existem apenas oito graus de liberdade [36]. Assim, pode-se fazer por exemplo $c'_{33} = 1$. Reescrevendo (3.4) na forma matricial:

$$\begin{bmatrix} X & Y & 1 & 0 & 0 & 0 & -uX & -uY \\ 0 & 0 & 0 & X & Y & 1 & -vX & -vY \end{bmatrix} \begin{bmatrix} c_{11} \\ c_{12} \\ c'_{13} \\ c_{21} \\ c_{22} \\ c'_{23} \\ c_{31} \\ c_{32} \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix}, \quad (3.7)$$

Por conseguinte, tendo duas equações e oito incógnitas são necessários quatro pontos da imagem com coordenadas conhecidas no sistema de coordenadas da base do manipulador, tendo em atenção que esses pontos fazem necessariamente parte do plano da mesa.

Mais uma vez, o procedimento de calibração consiste em utilizar o conjunto de pontos conhecidos para resolver as equações (3.7) por forma a obterem-se os valores dos coeficientes da matriz de calibração. O procedimento de calibração da câmara é descrito com detalhe no Apêndice A.

Por fim, uma vez estimados os coeficientes da matriz de calibração \mathbf{C}' , e dado um qualquer ponto $[u \ v]^T$ da imagem, podem ser reconstruídas as coordenadas 3D correspondentes ${}^0[X \ Y \ Z_{\text{mesa}}]^T$ (pertencentes ao plano da mesa) expressas no sistema de coordenadas da base do manipulador, através da inversão daquela matriz:

$$\begin{bmatrix} \lambda X \\ \lambda Y \\ \lambda \end{bmatrix} = \mathbf{C}'^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}. \quad (3.8)$$

3.1.2 Características da imagem

Os parâmetros de características da imagem usados no presente trabalho são as coordenadas no plano da imagem de pontos correspondentes à posição do centro do objecto

alvo (bola de ping-pong) e à posição do efector terminal do manipulador (ou mais concretamente do ponto central da “caneca” presa no efector terminal) na imagem. Mais especificamente, esses parâmetros são dados pela centróide [36] da superfície da bola projectada na imagem, e pela centróide da superfície da “caneca” projectada na imagem.

Em primeiro lugar, a imagem adquirida é convertida numa imagem binária através da comparação com um valor de limiar (*thresholding*), tal como se vê na Figura 3.3.

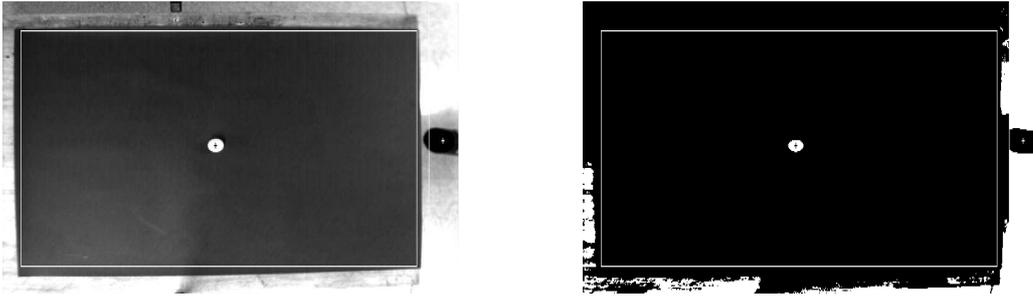


Fig. 3.3: *Thresholding* da imagem: a) imagem original; b) imagem binária.

De seguida, são calculados os momentos de ordem zero e de primeira ordem da imagem binária, que correspondem, respectivamente, à área da bola e à sua posição na imagem - o centro de área, ou centróide.

A área A de um único objecto presente numa imagem binária b é dada por [36]:

$$A = \sum_{i=1}^n \sum_{j=1}^m b_{ij}, \quad (3.9)$$

onde b_{ij} é o valor do pixel da coluna i e linha j , assumindo que a imagem tem n linhas por m colunas. A detecção da bola na imagem é conseguida através da área A . Sendo conhecida a área média \bar{A} da bola nas imagens, considera-se que a bola, e apenas ela, está presente numa determinada imagem com área A , se:

$$|A - \bar{A}| < \epsilon, \quad (3.10)$$

onde $\epsilon > 0$ é um valor determinado empiricamente e corresponde à variação máxima da área da bola para diferentes posições da mesma na imagem.

Por sua vez, a posição do centro de área de um objecto $[u \ v]^T = [\bar{i} \ \bar{j}]^T$ calcula-se da seguinte forma:

$$\bar{i} = \frac{\sum_{i=1}^n \sum_{j=1}^m i b_{ij}}{A},$$

$$\bar{j} = \frac{\sum_{i=1}^n \sum_{j=1}^m j b_{ij}}{A}, \quad (3.11)$$

onde os numeradores são respectivamente o primeiro momento da imagem segundo o eixo i (colunas), e o primeiro momento da imagem segundo o eixo j (linhas).

Na realidade, uma vez que se pretendem obter as coordenadas no plano da imagem dos pontos correspondentes à posição do centro da bola de ping-pong e à posição do ponto central da “caneca” presa no efector terminal na imagem, há a necessidade de calcular duas centróides, pelo que a imagem foi dividida em duas áreas — janelas. Desta forma, o cálculo da centróide [36] da superfície da bola projectada na imagem é efectuado numa janela da imagem, e o cálculo da centróide da superfície da “caneca” projectada na imagem é efectuado noutra janela, tal como se pode constatar a partir da Figura 3.4.

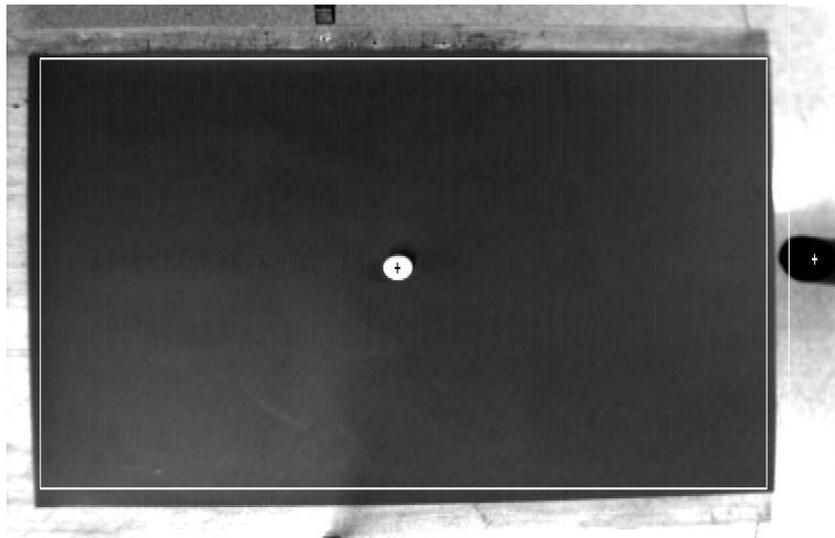


Fig. 3.4: Cálculo das centróides da bola e da “caneca” na imagem.

Sendo assim, quando é usada uma arquitectura de controlo com o efector terminal em cadeia aberta (EOL), o vector de parâmetros característicos da imagem é dado por:

$$\mathbf{f} = [u_{bola} \ v_{bola}]^T, \quad (3.12)$$

onde u_{bola} e v_{bola} correspondem às coordenadas do centro de área da bola e são dados por (3.11).

Por outro lado, quando é usada uma arquitectura de controlo com o efector terminal

em cadeia fechada (ECL), o vector de parâmetros característicos da imagem é dado por:

$$\mathbf{f} = [u_{bola} \ v_{bola} \ u_{canecca} \ v_{canecca}]^T, \quad (3.13)$$

onde u_{bola} e v_{bola} , assim como $u_{canecca}$ e $v_{canecca}$ são dados por (3.11).

3.1.3 Parâmetros característicos do movimento do objecto

Os parâmetros característicos do movimento do objecto utilizados coincidem com os parâmetros característicos da imagem correspondentes à posição da bola na imagem.

Sendo assim, o vector de parâmetros característicos do movimento do objecto é dado por:

$$\mathbf{m} = [u_{bola} \ v_{bola}]^T, \quad (3.14)$$

quando são usadas arquiteturas de controlo com o efector terminal em cadeia aberta (EOL) ou fechada (ECL).

A informação que é disponibilizada ao sistema de predição, é dada pelo conjunto, discreto no tempo, constituído pelos vectores de parâmetros característicos do movimento do objecto correspondentes a cada uma das n imagens:

$$\{[u_{bola} \ v_{bola}]_1^T, \dots, [u_{bola} \ v_{bola}]_n^T\}. \quad (3.15)$$

3.2 Sistema de Predição

A partir das amostras (3.15) da trajectória seguida pelo objecto, o sistema de predição estima a trajectória contínua completa do objecto em movimento de forma a predizer o “ponto de captura”.

A trajectória da bola no plano da imagem é descrita pela equação:

$$v_{bola}(i) = r(u_{bola}(i)) + e(i), \quad (3.16)$$

onde r representa a função que modela o movimento da bola, e e é um resíduo que representa a dinâmica não modelada. A função r consiste na combinação linear de m funções de base que modelam a dinâmica do objecto:

$$r(u_{bola}(i)) = k_1 r_1(u_{bola}(i)) + \dots + k_m r_m(u_{bola}(i)). \quad (3.17)$$

A escolha das funções de base r_j revela-se de uma importância fulcral, na medida em que quanto melhor representarem a dinâmica do movimento tanto menor será o resíduo ϵ e portanto tanto melhor será a predição. A expressão (3.17) pode ser também escrita na seguinte forma compacta:

$$r(u_{bola}(i)) = \mathbf{K}^T \varphi(u_{bola}(i)), \quad (3.18)$$

onde

$$\mathbf{K} = \begin{bmatrix} k_1 \\ \vdots \\ k_m \end{bmatrix}, \quad (3.19)$$

representa o vector dos parâmetros e

$$\varphi(u_{bola}(i)) = \begin{bmatrix} r_1(u_{bola}(i)) \\ \vdots \\ r_m(u_{bola}(i)) \end{bmatrix}, \quad (3.20)$$

representa o vector de dados.

A trajectória da bola seria, no caso ideal, rectilínea, mas verifica-se que, por vezes, o seu comportamento apresenta um carácter curvilíneo, devido a diversos factores:

- a mesa não está perfeitamente nivelada;
- a mesa encontra-se coberta por uma tela preta que apresenta algumas sobrelevações;
- por vezes, a bola é lançada de tal forma que saltita durante a fase inicial;
- no acto do lançamento é muitas vezes aplicada uma força centrífuga na bola, que faz com que esta rode sobre si própria.

Desta forma, as funções de base escolhidas para modelar o movimento do objecto são as seguintes:

$$\varphi(u_{bola}(i)) = \begin{bmatrix} 1 \\ u_{bola}(i) \\ u_{bola}^2(i) \end{bmatrix}, \quad (3.21)$$

onde a função linear traduz a componente rectilínea da trajectória e a função quadrática a vertente curvilínea.

O problema que se considera consiste na determinação de estimativas $\hat{k}_1, \hat{k}_2, \hat{k}_3$ dos parâmetros k_1, k_2, k_3 , a partir do conjunto, discreto no tempo, constituído pelos vectores de parâmetros característicos do movimento do objecto correspondentes a cada uma das imagens adquiridas, dado pela expressão (3.15).

Neste trabalho, essas estimativas foram determinadas utilizando o *método dos mínimos quadrados* [10], segundo o qual as estimativas são determinadas por forma a que o quadrado das diferenças entre os valores efectivamente observados e os calculados seja mínima.

Desta forma, tendo n imagens, e por conseguinte n posições consecutivas da bola na imagem, de acordo com o princípio dos mínimos quadrados, a estimativa $\hat{\mathbf{K}}$ é construída de modo a minimizar a função:

$$J(\hat{\mathbf{K}}) = \frac{1}{2} \sum_{i=1}^n [v_{bola}(i) - \hat{v}_{bola}(i)]^2, \quad (3.22)$$

com:

$$\hat{v}_{bola}(i) = \hat{\mathbf{K}}^T \varphi(u_{bola}(i)). \quad (3.23)$$

O valor mínimo de J é obtido para $\hat{\mathbf{K}}$ tal que:

$$\mathbf{\Phi}^T \mathbf{\Phi} \hat{\mathbf{K}} = \mathbf{\Phi}^T \mathbf{V}, \quad (3.24)$$

onde

$$\mathbf{\Phi} = \begin{bmatrix} \varphi^T(u_{bola}(1)) \\ \vdots \\ \varphi^T(u_{bola}(n)) \end{bmatrix}, \quad (3.25)$$

e

$$\mathbf{V} = \begin{bmatrix} v_{bola}(1) \\ \vdots \\ v_{bola}(n) \end{bmatrix}, \quad (3.26)$$

Se a matriz $\mathbf{\Phi}^T \mathbf{\Phi}$ for não singular, o mínimo é único, e é dado por:

$$\hat{\mathbf{K}} = (\mathbf{\Phi}^T \mathbf{\Phi})^{-1} \mathbf{\Phi}^T \mathbf{V}. \quad (3.27)$$

A estimativa obtida a partir da equação (3.27) é uma estimativa não recursiva, uma vez que é obtida a partir de um conjunto de n observações. Assim, à medida que são adquiridas

novas imagens e obtidas novas posições da bola na imagem, todo o processo tem de ser repetido. Este aumento consecutivo do número de dados reflete-se num aumento da matriz Φ e do vector \mathbf{V} , o que exige mais memória e envolve uma maior carga computacional na construção das estimativas. No presente sistema, como o número de amostras provenientes do sistema de visão é baixo este problema não é significativo, mas, se assim não fosse, poderia ser utilizada uma versão recursiva da estimação por mínimos quadrados [10], na qual a estimativa é actualizada usando apenas cada novo dado e a estimativa anterior. Por outro lado, o número de amostras obtido não permite a convergência da versão recursiva do método, levando a resultados inferiores aos obtidos pela versão não recursiva.

No que concerne à determinação do “ponto de captura”, definiu-se que a captura seria sempre executada ao longo de uma linha recta paralela ao lado da mesa onde as bolas caem — “linha de captura”. Desta forma, o problema de captura original em 2-D foi reduzido a um problema em 1-D. Assim, o “ponto de captura” é dado pela intersecção da trajectória da bola predita com a “linha de captura”, e é obtido em termos das suas coordenadas no plano da imagem:

$$\mathbf{f}_{captura} = [u_{captura} \ v_{captura}]^T. \quad (3.28)$$

Por fim, resta notar que a predição do instante de tempo no qual o objecto chega ao “ponto de captura” não é feita, uma vez que se pretende apenas posicionar o manipulador nesse ponto e ficar à espera do objecto.

3.3 Sistema de Controlo do Manipulador

O sistema de controlo do manipulador encontra-se implementado de forma a que de cada vez que é recebida uma nova referência relativa ao ponto de intercepção e captura do objecto, o sinal de controlo seguinte seja calculado de acordo com a lei de controlo escolhida e com o erro entre a nova referência e a configuração actual do manipulador, como é vulgar num sistema de controlo em cadeia fechada. Não são, assim, utilizados geradores de trajectórias, ao contrário do que é comum, na medida em que não permitem geralmente o replaneamento da trajectória do manipulador enquanto esta está a ser executada.

Os controladores implementados têm a vantagem de permitir correcções, a qualquer

momento, assim que se obtêm novas referências. Contudo, esta característica implica um grande cuidado no dimensionamento do próprio controlador, para que o movimento do manipulador não saia fora de certos limites, de modo a que não entre em conflito com possíveis restrições no espaço de trabalho, uma vez que o trajecto efectuado entre referências consecutivas não é planeado.

No que diz respeito à parametrização dos controladores utilizados, os termos gravíticos, as cinemáticas directa e inversa e o jacobiano do Puma 560 são os descritos em [27]. Os ganhos dos controladores PD foram ajustados a partir de simulações realizadas em Matlab. O modelo dinâmico do manipulador utilizado nas simulações faz parte das funções Matlab descritas em [21] e inclui a limitação dos valores de binários enviados para as juntas, através de uma não linearidade do tipo saturação. No entanto, existem limitações de velocidade e aceleração máxima para as juntas que não foram simuladas e que obrigam a alterações dos ganhos nas experiências com o sistema real, nomeadamente para os controladores em espaço operacional.

Os ganhos obtidos através das simulações foram:

- Controlador PD com compensação gravítica em espaço junta

$$K_p = [1000 \ 2000 \ 500 \ 60 \ 50 \ 60]^T$$

$$K_D = [110 \ 220 \ 55 \ 7 \ 6 \ 7]^T.$$

- Controlador PD com compensação gravítica em espaço operacional

$$K_p = [10000 \ 10000 \ 10000 \ 10000 \ 10000 \ 10000]^T$$

$$K_D = [760 \ 760 \ 760 \ 760 \ 760 \ 760]^T.$$

Os ganhos utilizados nos controladores cujos resultados são apresentados no Capítulo 4, foram os seguintes:

- Controlador PD com compensação gravítica em espaço junta

$$K_p = [1000 \ 2000 \ 500 \ 60 \ 50 \ 60]^T$$

$$K_D = [110 \ 220 \ 55 \ 7 \ 6 \ 7]^T.$$

- Controlador PD com compensação gravítica em espaço operacional

$$K_p = [1000 \ 1000 \ 1000 \ 200 \ 200 \ 200]^T$$

$$K_D = [170 \ 170 \ 170 \ 34 \ 34 \ 34]^T.$$

De forma a tornar mais eficiente a implementação do algoritmo de controlo em MS-DOS (sistema operativo uni-tarefa) foi utilizado um *scheduler* primitivo, que não permite a retoma de processos que haviam sido interrompidos. Este *scheduler*, inicialmente desenvolvido pela equipa da Trident Robotics, implica a definição de um período de amostragem base, do qual os períodos de amostragem de cada um dos processos a executar serão submúltiplos. Assim, cada vez que uma interrupção é desencadeada pelo pulso do temporizador correspondente ao início de um novo período de amostragem base, os processos a executar durante esse período são lançados sequencialmente para execução. É importante notar que a situação mais desfavorável a considerar é a correspondente aos períodos de amostragem em que todos os processos devem ser executados, caso em que o tempo total de execução não deve exceder o período de amostragem base.

Na implementação actual existem três processos:

- Controlador, executado com uma frequência de amostragem de 400 Hz;
- Leitura de referências originadas pelo sistema de predição, através de *sockets*, executado com uma frequência de amostragem de 40 Hz;
- Memorização dos sinais obtidos durante uma experiência, executado com uma frequência de amostragem de 400 Hz.

3.4 Arquitecturas de Seguimento Visual

Neste trabalho, assim como na quase totalidade dos sistemas descritos na literatura, foi adoptada a estrutura *olha-e-move dinâmico*. Isto, porque a frequência de amostragem do sistema de visão não é suficientemente elevada por forma a permitir o controlo directo do manipulador. Usando como base esta estrutura, foram implementadas duas arquitecturas de seguimento visual distintas:

- controlo baseado na posição e o efector terminal em cadeia aberta (EOL);
- controlo baseado na imagem e o efector terminal em cadeia fechada (ECL).

3.4.1 Controlo baseado na posição

Espaço Junta

No diagrama da Figura 3.5 encontra-se representada a arquitectura do tipo *olha-e-move dinâmico* com controlo baseado na posição, e com o efector terminal em cadeia aberta, que tem como base um controlador em espaço junta do tipo PD com compensação gravítica, estudado na secção 2.3.2, o qual é caracterizado por uma acção linear proporcional-derivativa (PD) e uma acção não linear de compensação dos termos gravíticos da dinâmica do manipulador.

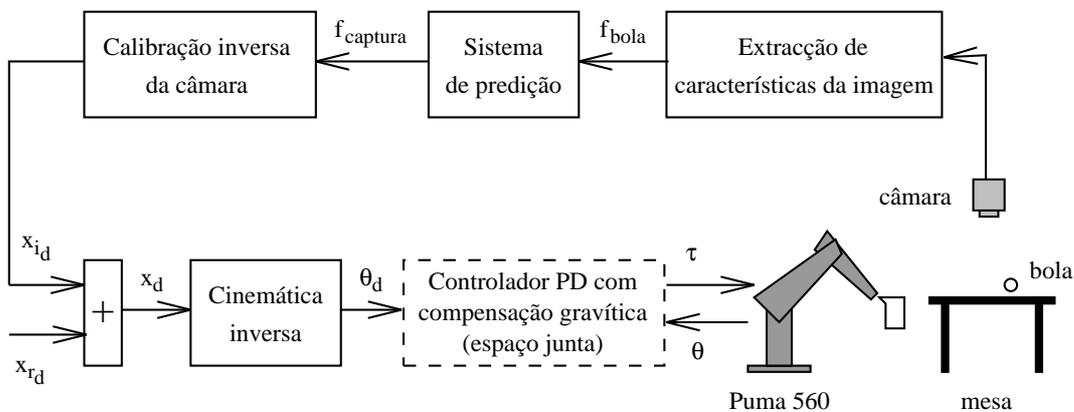


Fig. 3.5: Arquitectura *olha-e-move dinâmico* com controlo baseado na posição, em espaço junta.

Convém notar que, nesta aplicação, como a captura do objecto se processa sempre ao longo da “linha de captura”, e como se pretende que ao longo dessa recta o efector do manipulador mantenha a sua orientação e a sua coordenada Z constantes, a pose desejada x_d apresenta sempre os mesmos valores para a coordenada Z e para as coordenadas de orientação.

Defina-se então o vector com as coordenadas desejadas X_d e Y_d :

$$\mathbf{x}_{id} = \begin{bmatrix} X_d \\ Y_d \end{bmatrix}, \quad (3.29)$$

o qual é obtido a partir do “ponto de captura” $\mathbf{f}_{captura} = [u_{captura} \ v_{captura}]^T$ predito pelo sistema de predição, através da equação da calibração inversa (equação (3.8)).

Por outro lado, defina-se também o vector \mathbf{x}_{rd} com as coordenadas de orientação e a coordenada Z desejadas:

$$\mathbf{x}_{rd} = \begin{bmatrix} Z_d \\ \alpha_d \\ \beta_d \\ \gamma_d \end{bmatrix}, \quad (3.30)$$

onde α_d , β_d e γ_d são os ângulos de Euler ZYX desejados para a orientação.

Este vector é constante, uma vez que se pretende que ao longo da “linha de captura” o efector do manipulador mantenha a sua orientação e a sua coordenada Z constantes.

A pose desejada para o manipulador \mathbf{x}_d é obtida então a partir dos vectores \mathbf{x}_{id} e \mathbf{x}_{rd} :

$$\mathbf{x}_d = \begin{bmatrix} \mathbf{x}_{id} \\ \mathbf{x}_{rd} \end{bmatrix}. \quad (3.31)$$

A cinemática inversa do manipulador é por fim utilizada para o cálculo do vector θ_d das variáveis de junta desejadas, a partir da pose desejada \mathbf{x}_d .

Espaço Operacional

A arquitectura do tipo *olha-e-move dinâmico* com controlo baseado na posição, e com o efector terminal em cadeia aberta, tendo por base um controlador em espaço operacional do tipo PD com compensação gravítica, descrito na secção 2.3.3, encontra-se representada na Figura 3.6.

Esta arquitectura revela uma grande analogia com a anterior, mas verifica-se que com este controlador não é necessário calcular a cinemática inversa uma vez que o erro é calculado directamente em espaço operacional. Em contrapartida, é necessário proceder-se ao cálculo das funções da cinemática directa $\mathbf{k}(\theta)$ e $\mathbf{J}_A(\theta)$, a fim de se obterem os valores de \mathbf{x} e $\dot{\mathbf{x}}$.

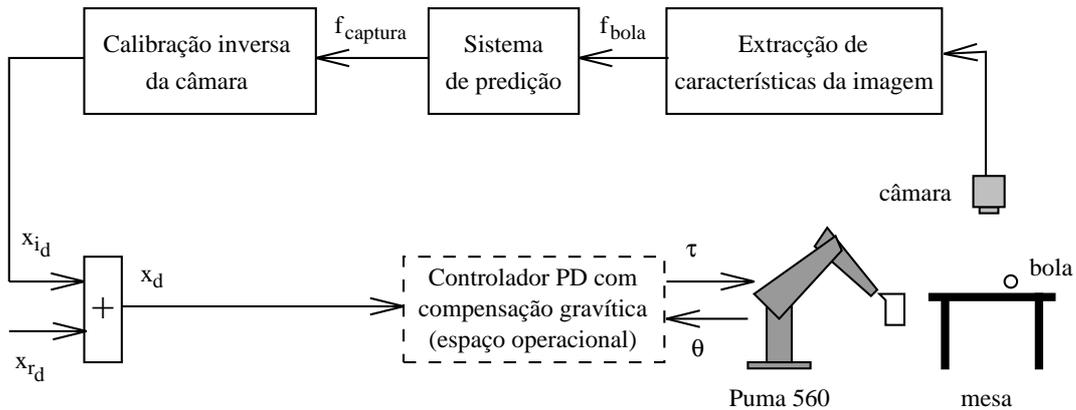


Fig. 3.6: Arquitectura *olha-e-move dinâmico* com controlo baseado na posição, em espaço operacional.

3.4.2 Controlo baseado na imagem

A arquitectura *olha-e-move dinâmico* com controlo baseado na imagem implementada encontra-se representada no diagrama de blocos da Figura 3.7.

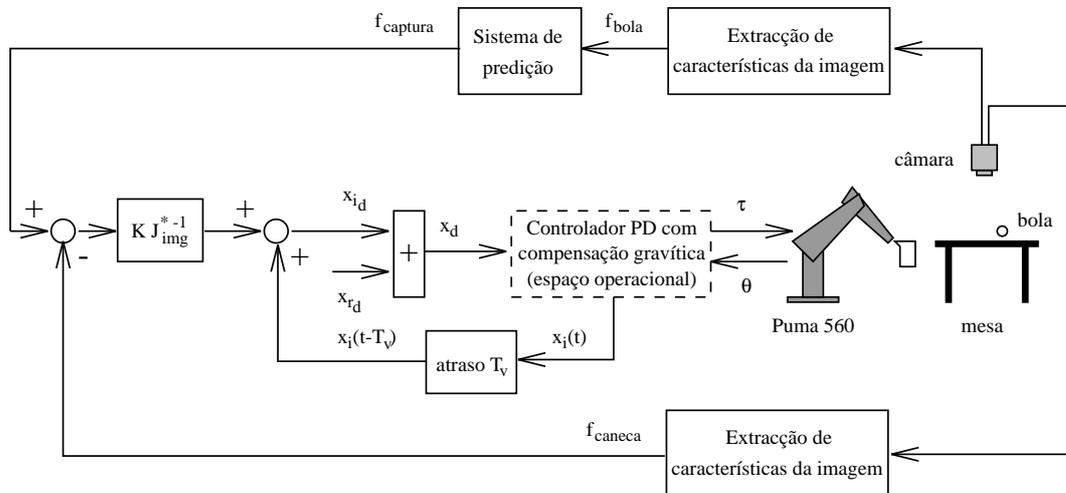


Fig. 3.7: Arquitectura *olha-e-move dinâmico* com controlo baseado na imagem (ECL).

Tendo presente o esquema genérico da arquitectura de seguimento visual *olha-e-move dinâmico* baseada na imagem apresentada na secção 2.4 verifica-se que a cadeia de controlo externa tem o erro de controlo $\Delta \mathbf{f}$ definido no espaço de parâmetros característicos da imagem, da seguinte forma:

$$\Delta \mathbf{f} = \mathbf{f}_{captura} - \mathbf{f}_{caneca}, \quad (3.32)$$

onde $\mathbf{f}_{captura} = [u_{captura} \ v_{captura}]^T$ representa a posição no plano da imagem do “ponto de captura” predito pelo sistema de predição, e $\mathbf{f}_{caneca} = [u_{caneca} \ v_{caneca}]^T$ representa as coordenadas no plano da imagem da posição da “caneca” presa ao efector terminal do Puma 560.

Contudo, ainda que o erro $\Delta \mathbf{f}$ seja definido no espaço de parâmetros característicos da imagem, as referências do sistema de controlo do manipulador são dadas em termos de coordenadas em espaço junta ou em espaço operacional. A relação entre os parâmetros característicos da imagem e as coordenadas em espaço operacional correspondentes é dada pelo *Jacobiano da imagem* [40]. Por sua vez, a multiplicação do Jacobiano da imagem pelo Jacobiano do manipulador permite encontrar a relação com as coordenadas em espaço junta, o que é equivalente ao que é feito no controlo em espaço operacional. Por conseguinte, efectuou-se o controlo do manipulador em espaço operacional.

O Jacobiano da imagem \mathbf{J}_{img} permite descrever a forma como os parâmetros característicos da imagem se alteram em função de mudanças na pose do manipulador da seguinte maneira:

$$d\mathbf{f} = \mathbf{J}_{img} d\mathbf{x}, \quad (3.33)$$

sendo o Jacobiano da imagem dado por:

$$\mathbf{J}_{img} = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right]. \quad (3.34)$$

Nesta aplicação em particular, como a captura do objecto se processa sempre ao longo da “linha de captura”, e como se pretende que ao longo dessa recta o efector do manipulador mantenha a sua orientação e a sua coordenada Z constantes, pode-se admitir que $dZ = d\alpha = d\beta = d\gamma \approx 0$. Sendo assim, deseja-se apenas determinar qual a variação nas coordenadas X e Y do efector do manipulador a partir da variação nos parâmetros característicos da imagem $\Delta \mathbf{f} = [\Delta u \ \Delta v]^T$.

Desta forma, a expressão (3.33) simplifica-se da seguinte maneira:

$$d\mathbf{f} = \mathbf{J}_{img}^* d\mathbf{x}_i, \quad (3.35)$$

com:

$$d\mathbf{x}_i = \begin{bmatrix} dX \\ dY \end{bmatrix}, \quad (3.36)$$

e

$$\mathbf{J}_{img}^* = \begin{bmatrix} \frac{\partial u(X,Y)}{\partial X} & \frac{\partial u(X,Y)}{\partial Y} \\ \frac{\partial v(X,Y)}{\partial X} & \frac{\partial v(X,Y)}{\partial Y} \end{bmatrix}, \quad (3.37)$$

onde as expressões $u(X, Y)$ e $v(X, Y)$ são dadas pelo modelo da câmara (expressão (2.7)) com $Z = Z_{mesa}$.

Na cadeia de controlo externa foi assim utilizada a seguinte lei de controlo:

$$\Delta \mathbf{x}_i = \mathbf{K}_{img} \mathbf{J}_{img}^{*-1} \Delta \mathbf{f}, \quad (3.38)$$

onde \mathbf{K}_{img} é uma matriz de ganhos constantes que permite a afinação do Jacobiano face a erros nos parâmetros da câmara, e

$$\Delta \mathbf{x}_i = \begin{bmatrix} \Delta X \\ \Delta Y \end{bmatrix}, \quad (3.39)$$

representa a variação desejada nas coordenadas X e Y do efector do manipulador.

O vector \mathbf{x}_{id} com as coordenadas desejadas X_d e Y_d é agora obtido através da expressão seguinte:

$$\mathbf{x}_{id}(t) = \mathbf{x}_i(t - T_v) + \Delta \mathbf{x}_i(t - T_v), \quad (3.40)$$

onde $\mathbf{x}_i(t - T_v)$ representa a posição do efector terminal do manipulador no instante de tempo no qual a imagem foi adquirida. T_v é o período médio de amostragem do sistema de visão.

Mais uma vez, defina-se o vector \mathbf{x}_{rd} , constante, com as coordenadas de orientação e a coordenada Z desejadas:

$$\mathbf{x}_{rd} = \begin{bmatrix} Z_d \\ \alpha_d \\ \beta_d \\ \gamma_d \end{bmatrix}. \quad (3.41)$$

A pose desejada para o efector terminal do manipulador \mathbf{x}_d é dada mais uma vez por:

$$\mathbf{x}_d = \begin{bmatrix} \mathbf{x}_{id} \\ \mathbf{x}_{rd} \end{bmatrix}. \quad (3.42)$$

O controlo do manipulador é efectuado numa cadeia de controlo interna por um controlador PD com compensação gravítica em espaço operacional, estudado na secção 2.3.3.

A necessidade desta cadeia de controlo interna, deve-se não só ao facto de a frequência de amostragem do sistema de visão ser demasiado baixa para permitir um controlo satisfatório do manipulador, mas também porque a partir da imagem só se consegue extrair informação relativa a X e Y (assumindo que Z se mantém constante).

3.5 Método Experimental

As bolas de ping-pong são atiradas, por um individuo, na direcção do manipulador, rolando desde o extremo da mesa mais afastado até ao extremo dentro do seu espaço de trabalho. A cada 104,5 ms uma nova imagem da mesa com a bola a rolar é adquirida pelo sistema de visão, e é calculada uma nova estimativa do “ponto de captura” predito pelo sistema de predição. Paralelamente, o sistema de controlo do manipulador utiliza essa informação para posicionar o efector terminal do Puma 560, de acordo com a arquitectura usada. Presa na garra do manipulador encontra-se uma pequena “caneca” com um diâmetro de 8 cm. A captura tem sucesso quando a bola, cujo diâmetro é de 4 cm, cai dentro da “caneca”.

Optou-se por testar cada um dos subsistemas separadamente, obtendo-se depois os resultados para cada uma das arquitecturas completas. Assim:

- O subsistema de visão foi testado quanto à sua capacidade de seguimento da bola e taxa de amostragem máxima possível (ver 4.1);
- O subsistema de predição foi testado quanto à sua fiabilidade na predição do “ponto de captura” e suavidade da trajectória enviada para o subsistema de controlo do manipulador (ver 4.2);
- Os três subsistemas de controlo do manipulador, ajustados a cada uma das três arquitecturas estudadas, foram testados para respostas a um escalão sobre a recta de intercepção da bola, com uma amplitude correspondente ao pior caso, isto é, aquele em que a predição inicial aponta para uma excursão maior do efector terminal (ver 4.3).

A apresentação dos resultados para as arquitecturas completas reflecte o comportamento do sistema em cadeia fechada, caso em que a trajectória enviada ao subsistema de

controle do manipulador consiste numa sequência de escalões de amplitude variável e a uma frequência aproximadamente constante (9.56 imagens por segundo).

4. Resultados

Neste capítulo são apresentados e comentados os resultados relativos a cada um dos subsistemas que compõem o sistema de seguimento visual e captura de objectos em movimento. São também analisados resultados correspondentes a capturas com sucesso para cada uma das arquitecturas estudadas.

4.1 Sistema de Visão

Nas Figuras 4.1, 4.2 e 4.3 encontram-se representados os resultados do sistema de visão para três lançamentos da bola.

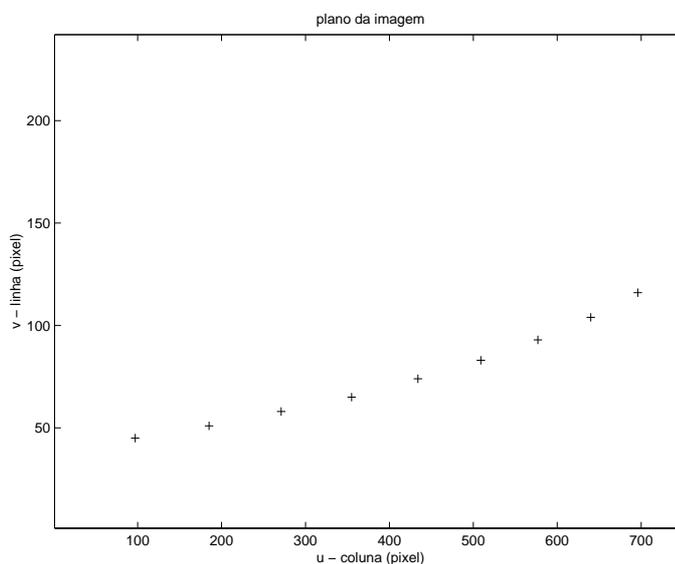


Fig. 4.1: Resultados do lançamento nº 1. Posição da bola no plano da imagem.

A partir de uma sequência de trinta lançamentos da bola, constatou-se que, por

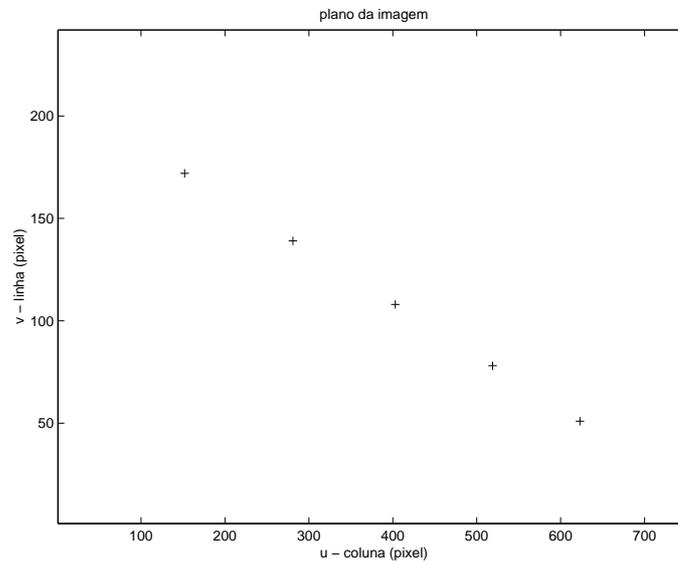


Fig. 4.2: Resultados do lançamento nº 2. Posição da bola no plano da imagem.

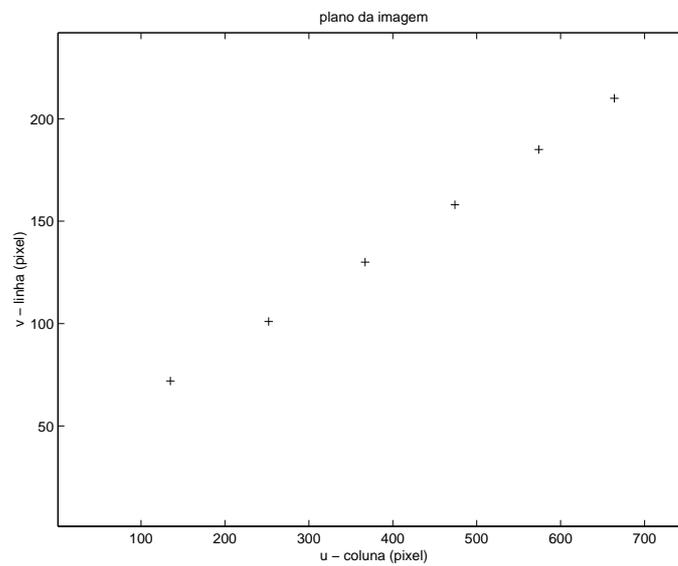


Fig. 4.3: Resultados do lançamento nº 3. Posição da bola no plano da imagem.

lançamento, o número de imagens adquiridas da bola a rolar varia geralmente entre 6 e 12 imagens, sendo o número de amostras médio aproximadamente igual a 9,4 imagens. Por conseguinte, o deslocamento médio entre duas amostras do centro de área da bola segundo o eixo u (colunas), que corresponde à direcção principal do deslocamento, é de cerca de 72 *pixels* por imagem. Por outro lado, sabendo que o comprimento da superfície da mesa aproveitado é igual a 100 cm, o deslocamento da bola na direcção principal do movimento (eixo u) é em média igual a 10,5 cm, entre duas imagens. Por conseguinte, atendendo a que o período de amostragem médio do sistema de visão é 104,5 ms, a velocidade da bola é em média cerca de 100 cm/s.

No que concerne à fiabilidade da localização da bola através do processo de cálculo do centro de área da imagem, note-se em primeiro lugar que a luz não é uniforme no local onde a bola se move, o que implica que essa mesma luz não é reflectida uniformemente pela superfície da bola. Desta forma, o centro de área da imagem pode não coincidir com o centro de área da bola tal como se pode observar na Figura 4.4-(b).

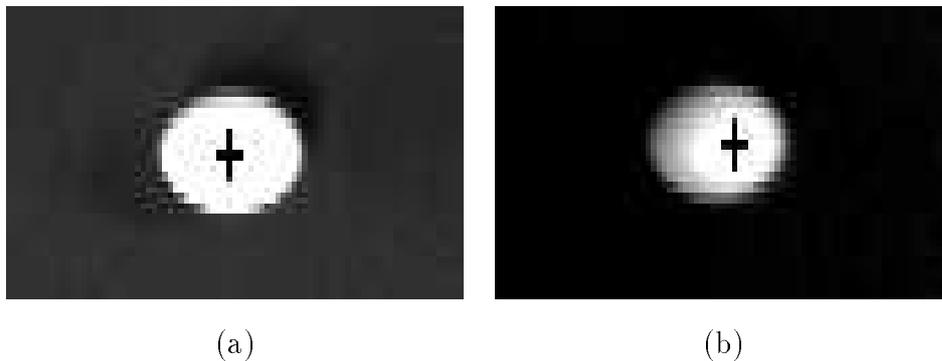


Fig. 4.4: Dependendo das condições de iluminação e dos parâmetros do sistema, o centro de área da imagem pode não coincidir com o centro de área da bola: (a) centro de área da imagem em condições normais; (b) centro de área da imagem não coincidente com o centro da bola

Por este motivo, o programa utilizado permite o ajuste do tempo de exposição utilizado na aquisição da imagem e do valor de limiar usado na transformação da imagem numa imagem binária. O ajuste destes parâmetros e da abertura da câmara é fundamental para que se obtenha uma imagem da bola isenta de sombras. O acerto destes parâmetros é bastante simples e por conseguinte, em condições normais, o centro de área da imagem

coincide com o centro de área da bola, tal como se pode ver na Figura 4.4-(a).

Em condições normais, o erro na localização da bola coincide com o erro mínimo: ± 1 *pixel*, o que corresponde a um erro de 3,5 mm segundo a direcção vertical, e de 1,5 mm segundo a direcção horizontal. Se não houver o cuidado de se ajustar os parâmetros acima referidos, o que deve ser efectuado de cada vez que houver alterações significativas na luminosidade ambiental, esse erro pode atingir, no pior caso, os 2 cm — raio da bola —, se bem que não ultrapasse usualmente o valor de 1 cm.

4.2 Sistema de Predição

Na Figura 4.5 encontra-se representada uma sequência de trajectórias preditas, calculadas pelo sistema de predição, à medida que são disponibilizadas novas posições da bola no plano da imagem pelo sistema de visão. Estes resultados do sistema de predição são relativos ao primeiro lançamento da bola apresentado na secção anterior.

O lançamento foi efectuado propositadamente por forma a aplicar uma força centrífuga na bola, fazendo com que esta rodasse sobre si própria e descrevesse uma trajectória um pouco curvilínea. Na realidade, a maior parte das trajectórias têm um carácter marcadamente rectilíneo, à semelhança do que se verifica com os lançamentos apresentados nas Figuras 4.2 e 4.3. Contudo, nem sempre a trajectória é rectilínea, devido quer a irregularidades na mesa, quer à forma como é lançada a bola (saltitando, rodando sobre si própria).

A variação da coordenada v do “ponto de captura” predito encontra-se na Figura 4.6. Verifica-se que com este método, a evolução da predição do “ponto de captura” aproxima-se mais rapidamente do valor final, do que usando por exemplo um método de predição da trajectória baseado no ajuste de uma recta às duas últimas posições da bola, o que começou por ser feito numa fase inicial deste trabalho.

Na Tabela 4.1 encontra-se a evolução dos valores dos parâmetros estimados $\hat{\mathbf{K}}$, correspondentes a cada uma das trajectórias representadas na Figura 4.5. Nesta tabela encontra-se também o erro quadrático $J(\hat{\mathbf{K}})$ associado a cada estimativa. Note-se que este erro foi obtido utilizando-se todas as nove posições da bola conhecidas, e não apenas as posições conhecidas na altura da estimação. A primeira trajectória estimada é uma

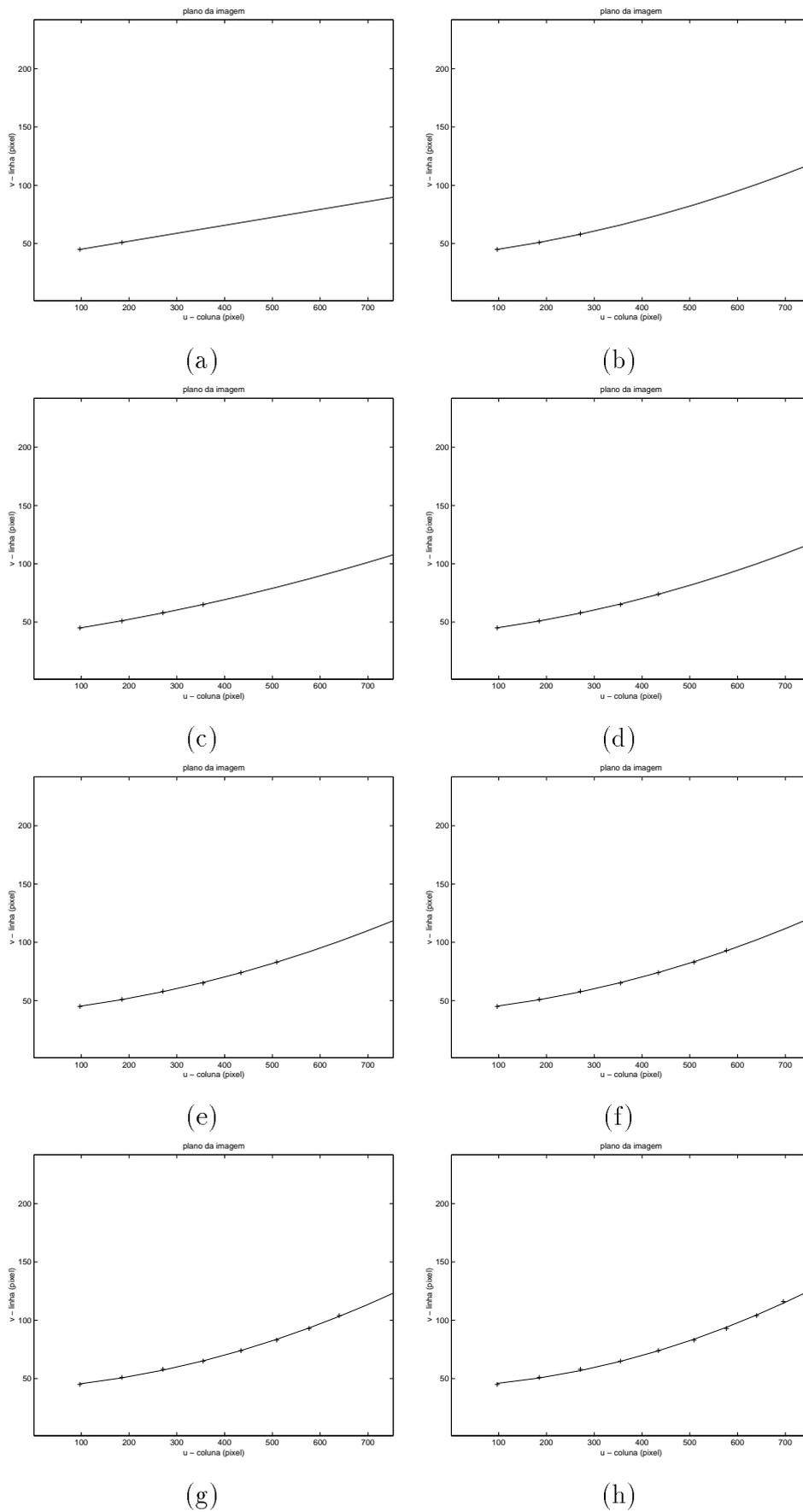


Fig. 4.5: Resultados do lançamento nº 1. Predição da trajetória da bola no plano da imagem.

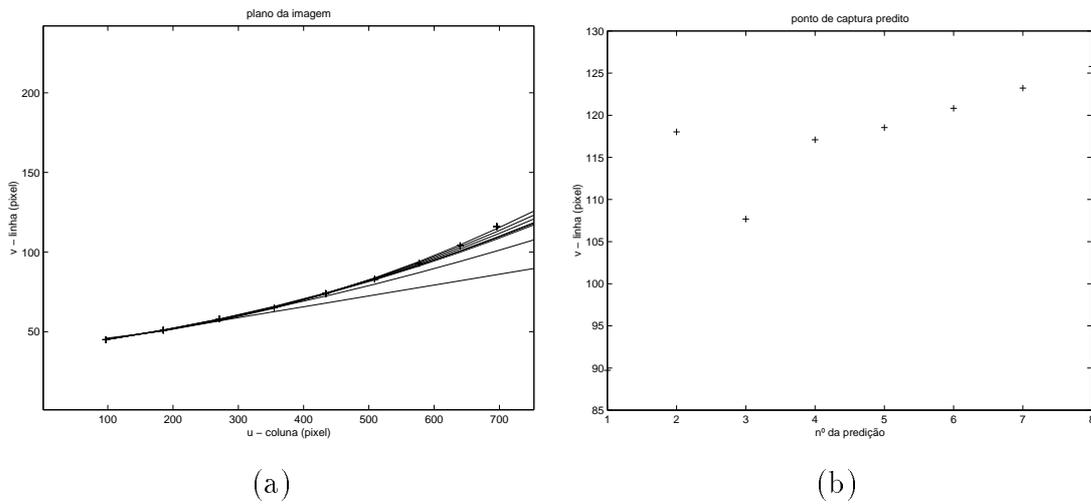


Fig. 4.6: Resultados do lançamento nº 1. (a) Variação da predição da trajectória da bola no plano da imagem. (b) Variação da coordenada v do ponto de captura predito

recta, uma vez que é determinada unicamente a partir das duas primeiras localizações da bola. Por conseguinte, verifica-se que o parâmetro $\hat{k}_3 = 0$.

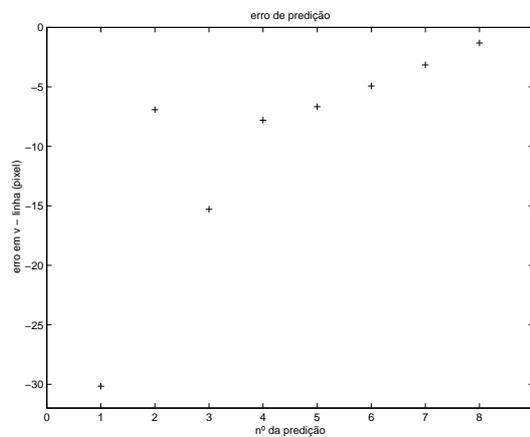


Fig. 4.7: Erro de predição em relação à coordenada v (linha) da última posição da bola conhecida.

Na Figura 4.7 encontra-se representado o erro de predição em relação à coordenada v (linha) da última posição da bola conhecida. Como pode ser observado, o erro correspondente à última trajectória predita é de apenas 1 pixel, o que se traduz em cerca de 3,5 mm. No entanto, as predições imediatamente anteriores apresentam erros de 7, 5 e 3 pixel consecutivamente, que correspondem a distâncias de aproximadamente 2,45 cm, 1,75 cm

e 1,05 cm. Estes erros já são significativos, uma vez que são da ordem de grandeza da margem de segurança de 2 cm dada pela largura da “caneca” utilizada. Em todo o caso, o erro de predição só se torna notório quando o número de posições conhecidas da bola é pequeno, ou seja, quando a bola tem uma velocidade elevada. Mas de qualquer forma, estes resultados sugerem a implementação de eventuais melhorias no sistema de predição.

Nº da predição	$\hat{\mathbf{K}}$	$J(\hat{\mathbf{K}})$
1	$[38.38640.06820]^T$	883.6929
2	$[39.74910.04680.0001]^T$	30.0714
3	$[38.93350.05770.00004]^T$	189.0689
4	$[40.07750.04420.0001]^T$	40.4025
5	$[40.33770.04150.0001]^T$	28.1835
6	$[40.92000.03590.0001]^T$	14.3500
7	$[41.76670.02830.0001]^T$	5.9963
8	$[42.99720.01810.0001]^T$	3.1027

Tabela 4.1: Evolução dos valores dos parâmetros estimados $\hat{\mathbf{K}}$ e do erro quadrático associado $J(\hat{\mathbf{K}})$.

4.3 Sistema de Controlo do Manipulador

O teste ao sistema de controlo do manipulador foi efectuado através da aplicação de escalões com uma amplitude de 0,32 m segundo a direcção X do sistema de coordenadas da base do manipulador, e portanto no espaço operacional, correspondentes a uma diferença elevada entre referências sucessivas enviadas pelo sistema de predição.

Importa salientar que depois das primeiras experiências se decidiu implementar um interpolador de caminho de modo a suavizar as trajectórias correspondentes às referências dadas. Esse interpolador consiste num filtro passa-baixo com uma frequência de corte variável, em função da amplitude da diferença entre a posição do efector do manipulador e a referência dada. O seu comportamento é similar ao de uma rampa, mas com a

vantagem de produzir caminhos mais suaves no início e no fim.

Nas Figuras 4.8 e 4.9 encontra-se a resposta ao escalão em X do controlador PD com compensação gravítica em espaço junta a utilizar na arquitectura de seguimento visual e captura de objectos com controlo baseado na posição, em termos de variáveis de junta. Sendo este controlador em espaço junta, o escalão foi convertido através da cinemática inversa. Na Figura 4.10 encontra-se a resposta em termos da posição e orientação do efector terminal do manipulador. Nas várias figuras, as linhas a tracejado representam os valores correspondentes à referência dada, e as linhas contínuas representam a resposta do manipulador a essa referência.

A resposta do controlador PD com compensação gravítica em espaço operacional a utilizar na arquitectura de seguimento visual e captura de objectos com controlo baseado na posição, encontra-se nas Figuras 4.11 e 4.12, em termos das posições angulares das juntas. Na Figura 4.13 encontra-se a resposta em termos da posição e orientação do efector terminal do manipulador.

Comparando o desempenho dos dois controladores, verifica-se que o controlador em espaço junta apresenta um tempo de estabelecimento menor, o que não é de estranhar, uma vez que os binários aplicados são mais elevados. A este facto não será alheio o facto de os ganhos utilizados no controlador em espaço operacional serem inferiores aos que foram ajustados através das simulações, os quais correspondiam a um desempenho idêntico ao do controlador em espaço junta (ver a secção 3.3).

O erro em regime estacionário é também menor quando se utiliza o controlador em espaço junta.

Constata-se também que com o controlador em espaço operacional, os binários aplicados nas juntas apresentam maiores oscilações, do que quando se utiliza o controlador em espaço junta. Consequentemente, a resposta do controlador em espaço operacional apresenta um comportamento mais nervoso do que a do controlador em espaço junta.

Nas Figuras 4.14, 4.15 e 4.16 encontra-se a resposta ao escalão do controlador PD com compensação gravítica em espaço operacional a utilizar na arquitectura de seguimento visual e captura de objectos com controlo baseado na imagem. Neste caso o escalão fornecido consistiu num Δf correspondente ao escalão com a amplitude de 0,32 m considerado atrás.

Relativamente ao desempenho deste controlador, verifica-se que apresenta uma resposta ainda mais nervosa do que a do controlador em espaço operacional utilizado na arquitectura com controlo baseado na posição. Tal deve-se ao facto de o período de amostragem do sistema de visão não ser fixo, razão pela qual o seu valor não coincide algumas vezes com o valor médio T_v (ver Secção 3.4.2) utilizado no controlo, o que leva a oscilações nos sinais de controlo e consequentes vibrações no efector terminal do manipulador e sobre-elevações na sua trajectória.

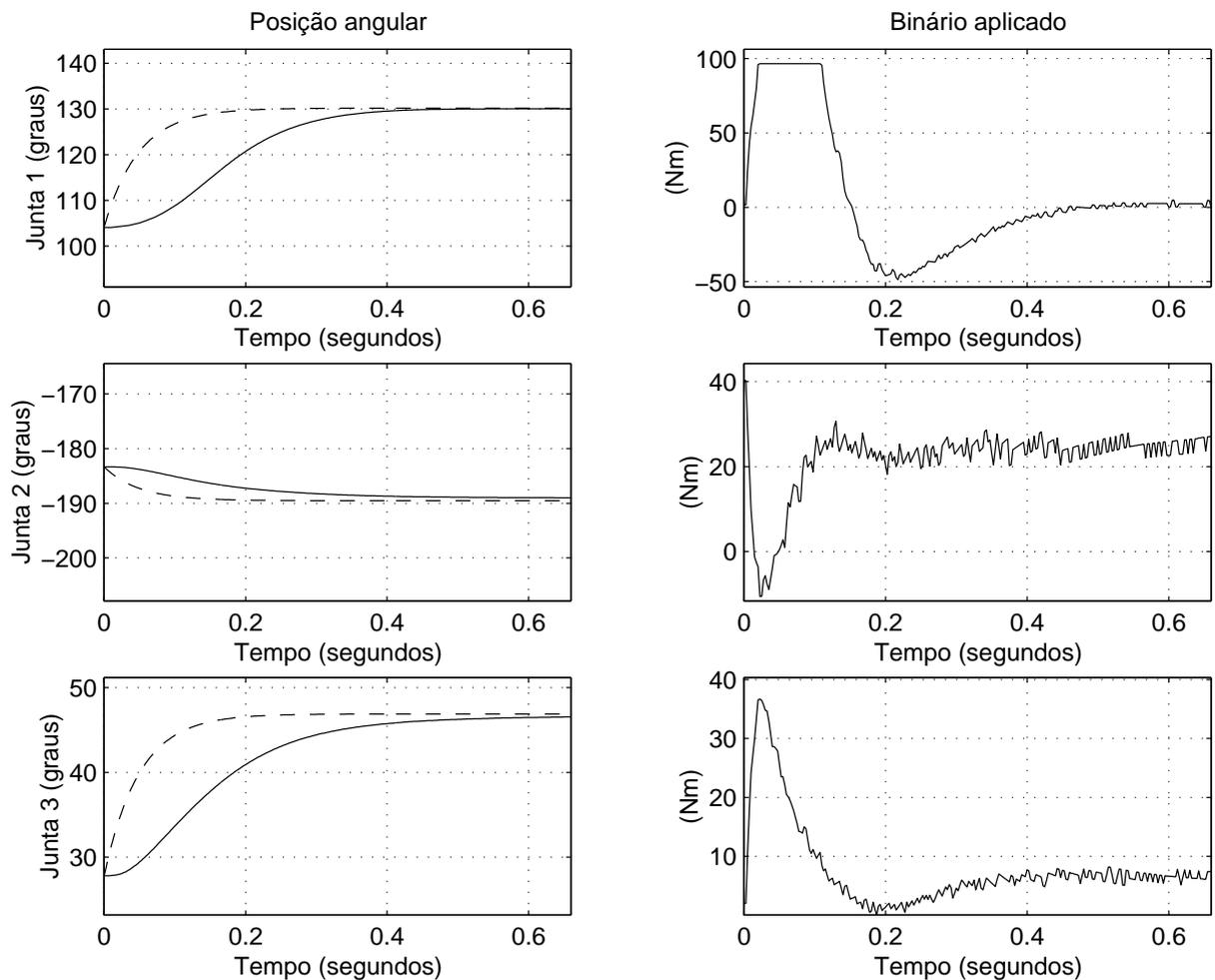


Fig. 4.8: Resposta a um escalão em X do controlador PD com compensação gravítica em espaço junta a utilizar na arquitectura de seguimento visual e captura de objectos com controlo baseado na posição. Posição angular das juntas e binários aplicados nas três primeiras juntas.

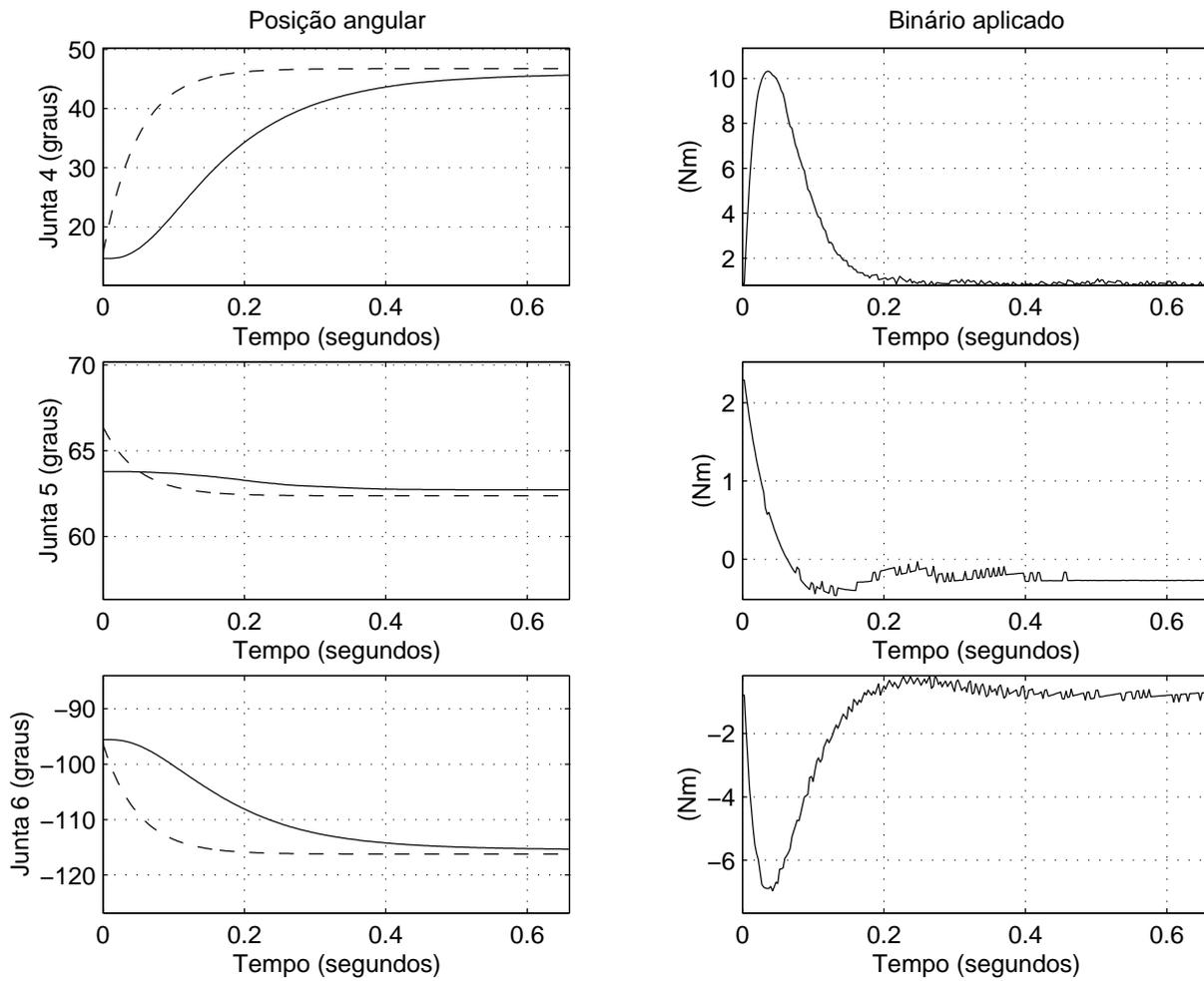


Fig. 4.9: Resposta a um escalão em X do controlador PD com compensação gravítica em espaço junta a utilizar na arquitectura de seguimento visual e captura de objectos com controlo baseado na posição. Posição angular das juntas e binários aplicados nas três últimas juntas.

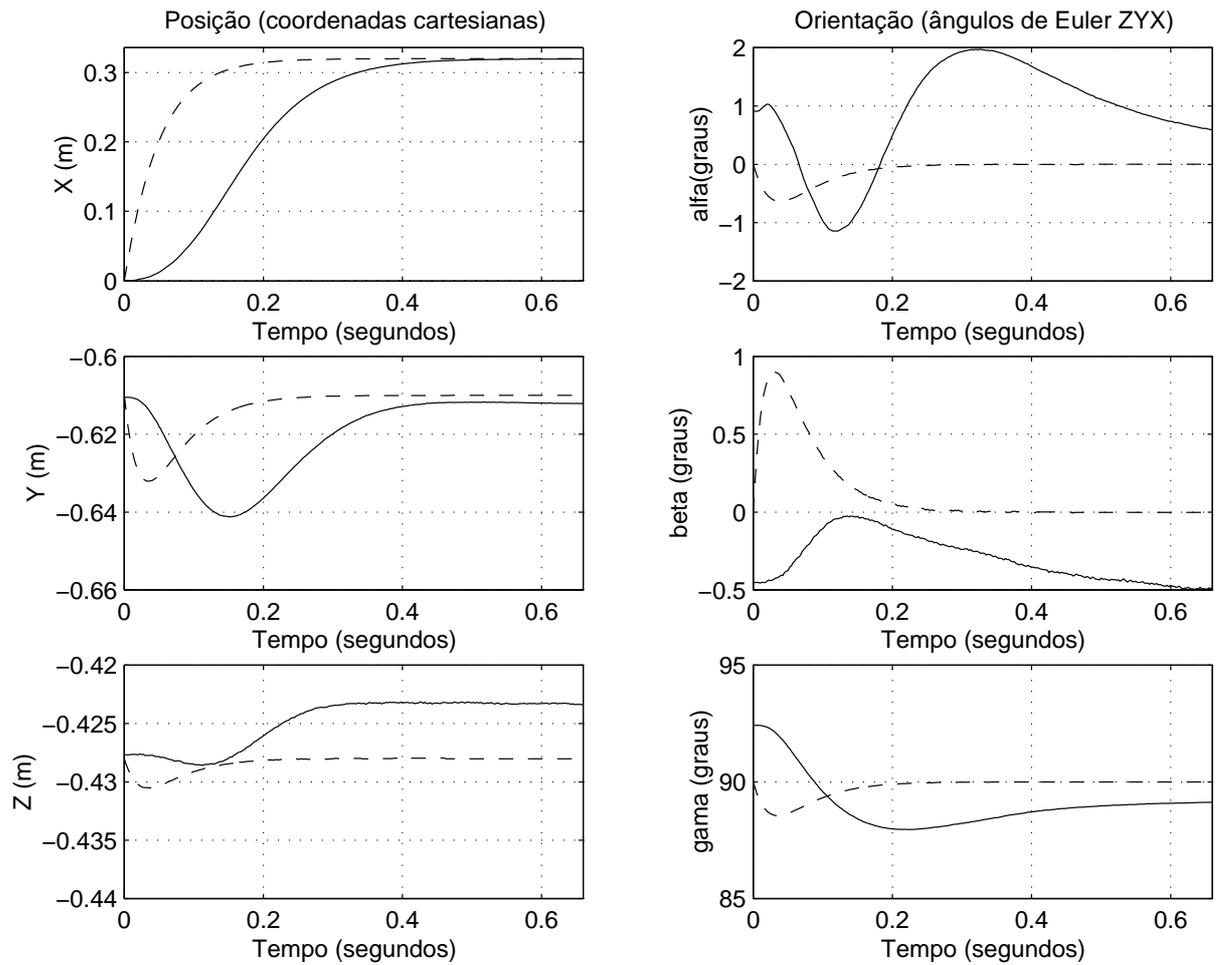


Fig. 4.10: Resposta a um escalão em X do controlador PD com compensação gravítica em espaço junta a utilizar na arquitectura de seguimento visual e captura de objectos com controlo baseado na posição. Posição e orientação do efector terminal do manipulador.

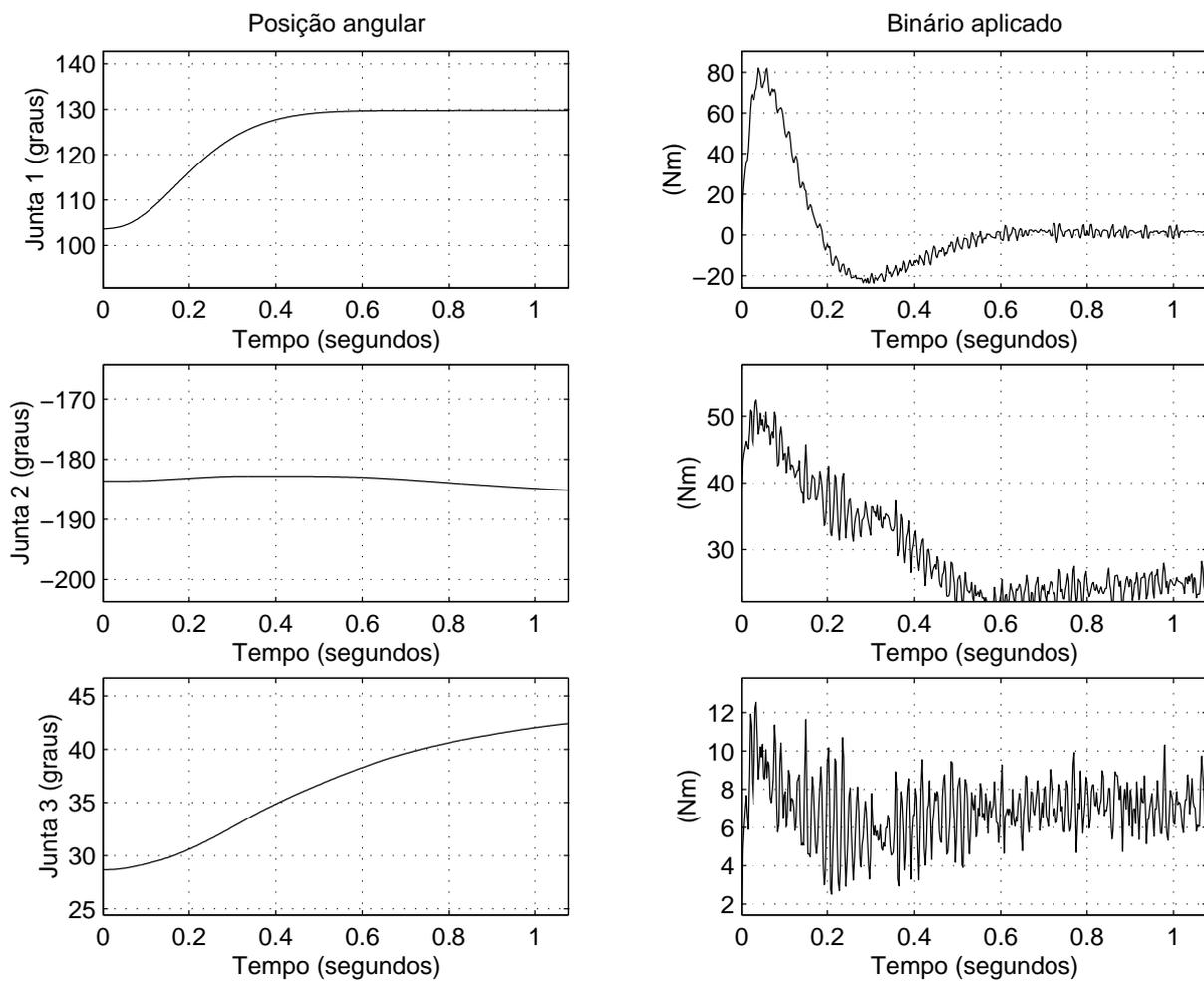


Fig. 4.11: Resposta a um escalão em X do controlador PD com compensação gravítica em espaço operacional a utilizar na arquitectura de seguimento visual e captura de objectos com controlo baseado na posição. Posição angular das juntas e binários aplicados nas três primeiras juntas.

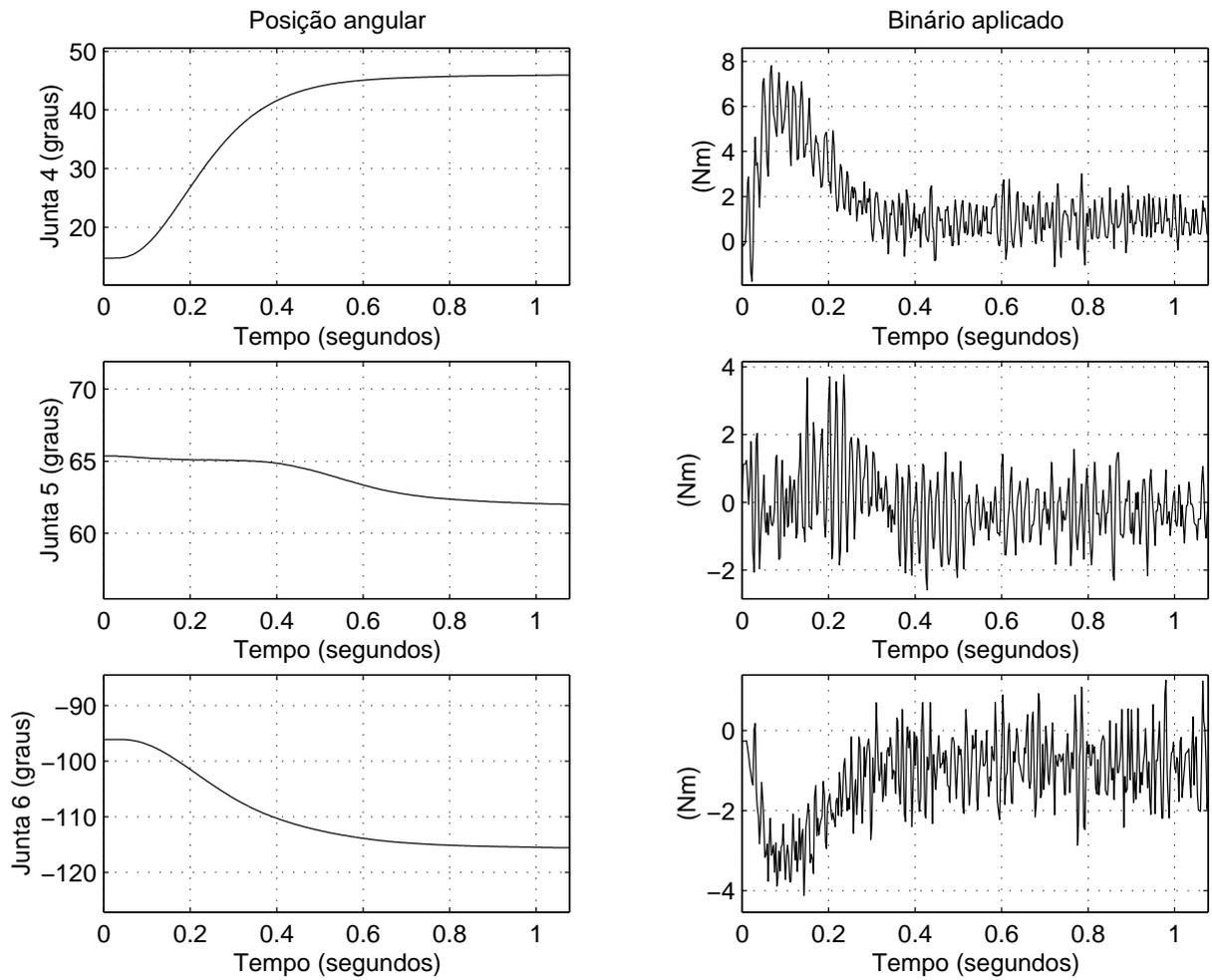


Fig. 4.12: Resposta a um escalão em X do controlador PD com compensação gravítica em espaço operacional a utilizar na arquitectura de seguimento visual e captura de objectos com controlo baseado na posição. Posição angular das juntas e binários aplicados nas três últimas juntas.

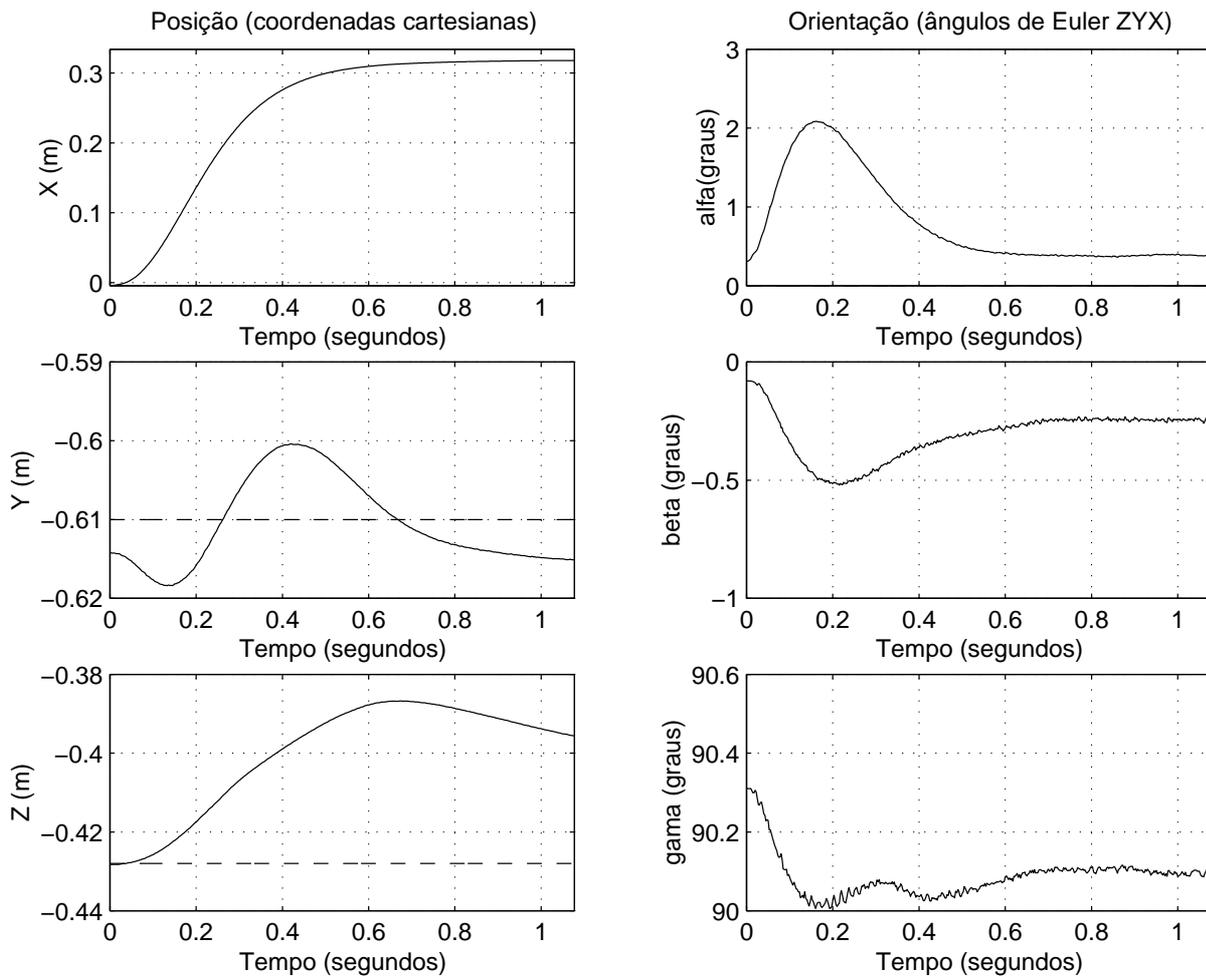


Fig. 4.13: Resposta a um escalão em X do controlador PD com compensação gravítica em espaço operacional a utilizar na arquitectura de seguimento visual e captura de objectos com controlo baseado na posição. Posição e orientação do efector terminal do manipulador.

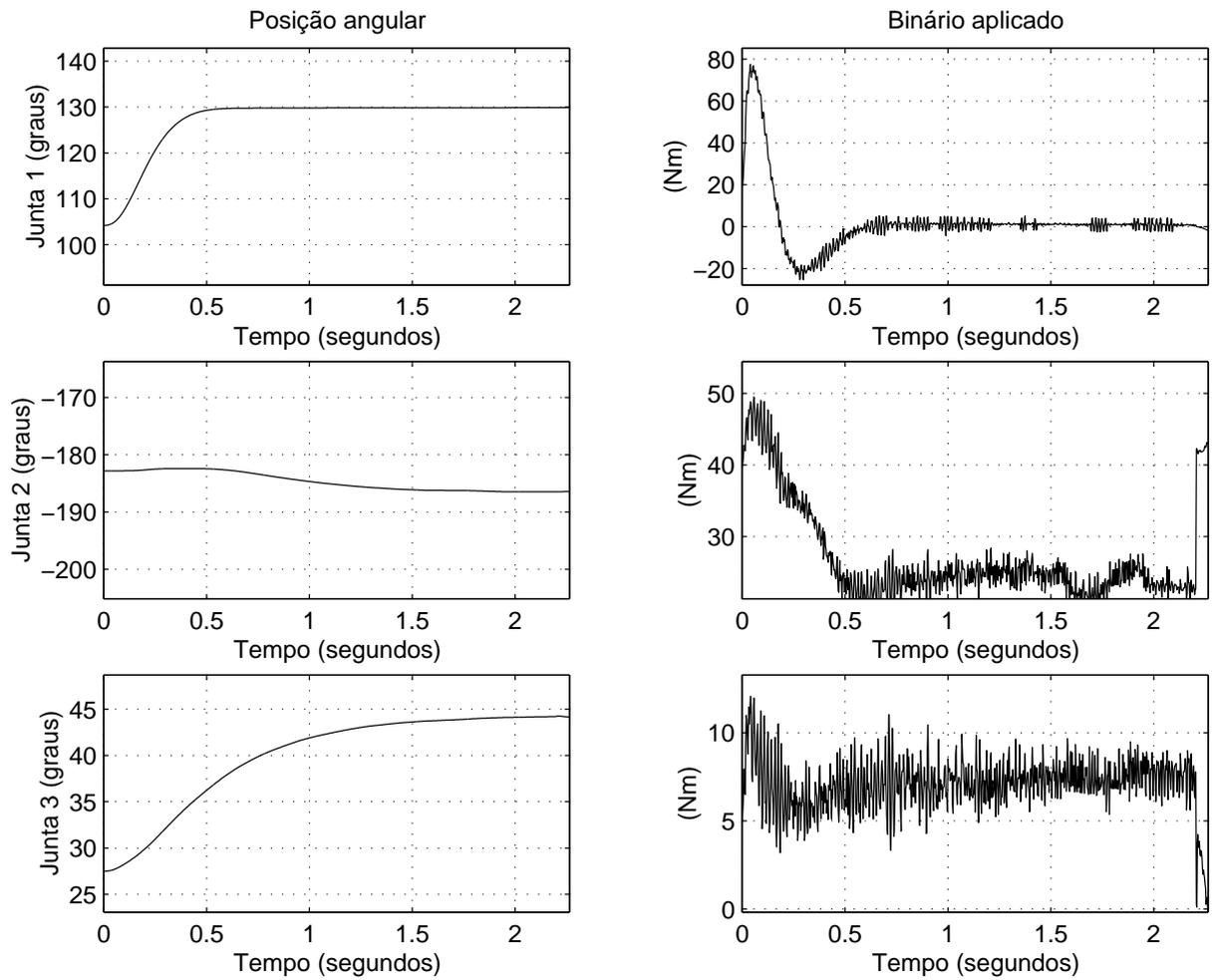


Fig. 4.14: Resposta a um escalão em X do controlador PD com compensação gravítica em espaço operacional a utilizar na arquitectura de seguimento visual e captura de objectos com controlo baseado na imagem. Posição angular das juntas e binários aplicados nas três primeiras juntas.

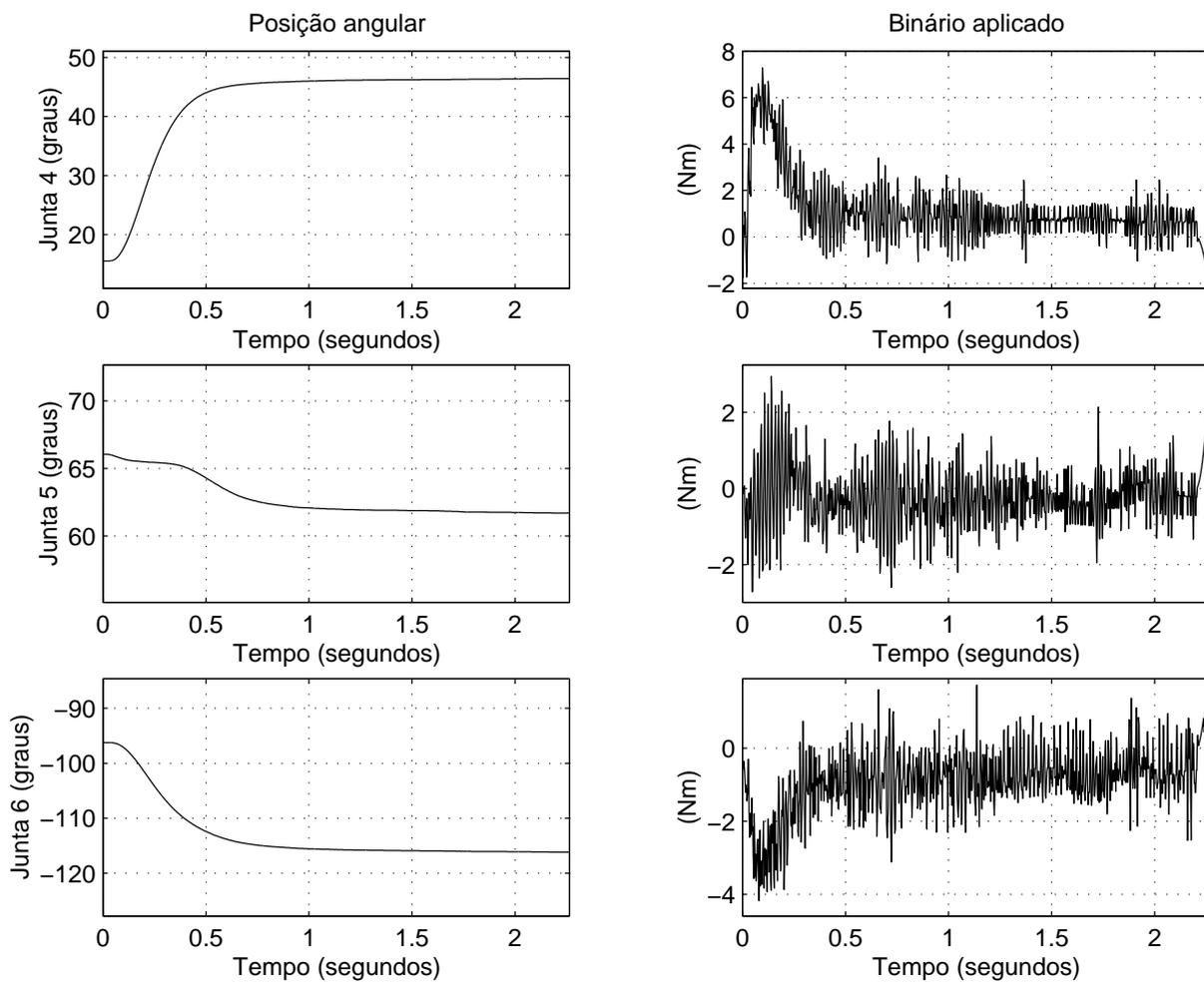


Fig. 4.15: Resposta a um escalão em X do controlador PD com compensação gravítica em espaço operacional a utilizar na arquitectura de seguimento visual e captura de objectos com controlo baseado na imagem. Posição angular das juntas e binários aplicados nas três últimas juntas.

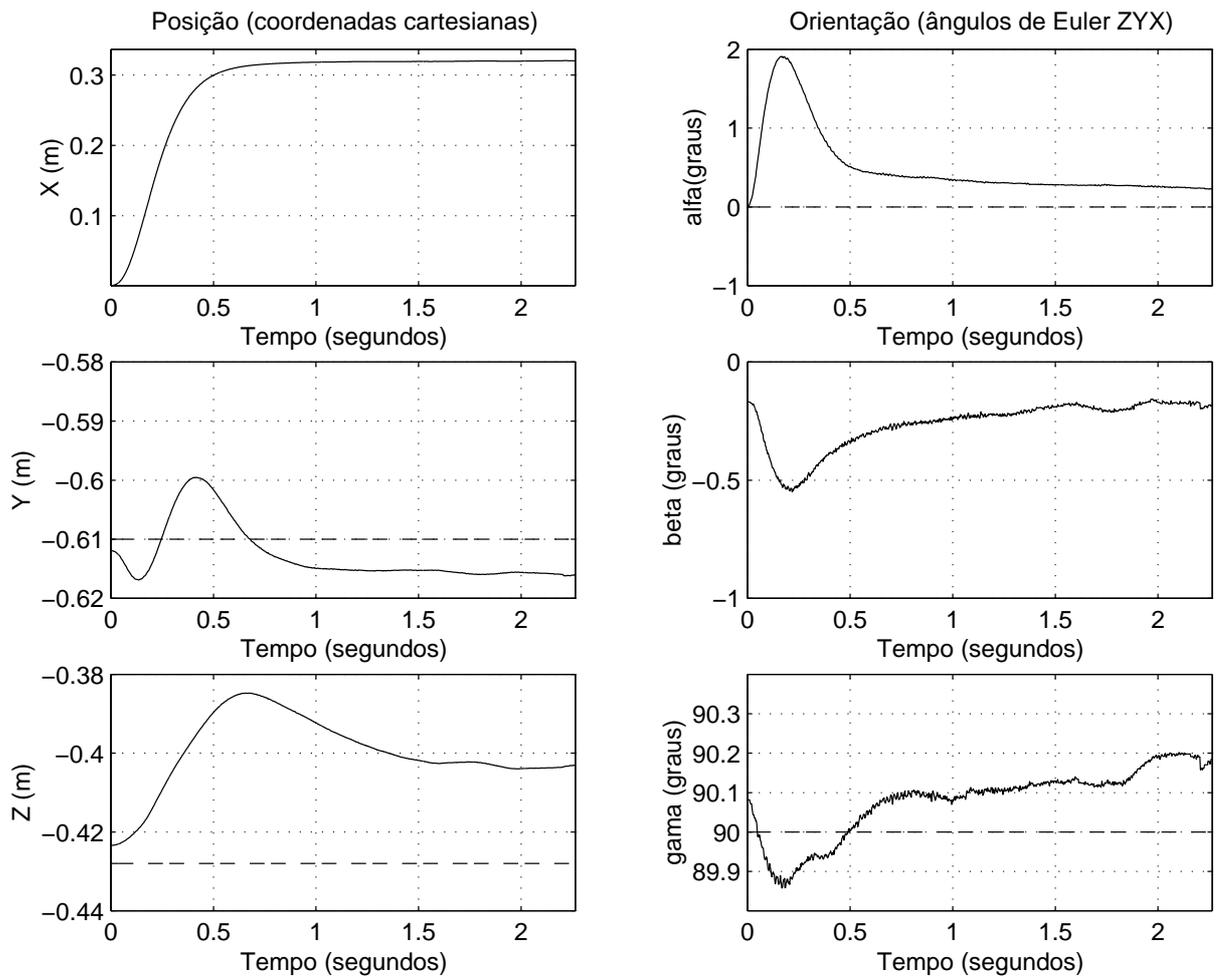


Fig. 4.16: Resposta a um escalão em X do controlador PD com compensação gravítica em espaço operacional a utilizar na arquitectura de seguimento visual e captura de objectos com controlo baseado na imagem. Posição e orientação do efector terminal do manipulador.

4.4 Arquitecturas de Seguimento Visual

Nesta secção são apresentados resultados correspondentes a capturas com sucesso para cada uma das arquitecturas estudadas.

Nas Figuras 4.17 e 4.18 encontram-se resultados, em termos de ângulos de junta, relativos à arquitectura de seguimento visual e captura de objectos com controlo baseado na posição, em espaço junta. Na Figura 4.19 encontra-se a resposta em termos da posição e orientação do efector terminal do manipulador.

Por sua vez, os resultados correspondentes à arquitectura com controlo baseado na posição, em espaço operacional, encontram-se nas Figuras 4.20 e 4.21, em termos das posições angulares das juntas. Os resultados em termos da posição e orientação do efector terminal do manipulador encontram-se na Figura 4.22.

Nas Figuras 4.23, 4.24 e 4.25 encontram-se resultados da arquitectura de seguimento visual e captura de objectos com controlo baseado na imagem.

Comparando o desempenho das várias arquitecturas, verifica-se que a arquitectura com controlo baseado na posição (EOL), em espaço junta, apresenta o melhor comportamento. A resposta do manipulador é a mais rápida e não apresenta sobrelevações, o que se traduz numa percentagem de capturas com sucesso na ordem dos 90%.

A arquitectura com controlo baseado na posição, em espaço operacional, apresenta uma percentagem de sucessos na ordem dos 82%. O facto de o controlador em espaço operacional apresentar um desempenho inferior ao do controlador em espaço junta, como foi visto na secção anterior, justifica esta diferença na taxa de sucessos. A resposta do controlador em espaço operacional é mais lenta e apresenta sobrelevações, uma vez que os ganhos de controlo utilizados são inferiores aos valores ajustados na simulação e que correspondiam ao melhor desempenho, similar ao do controlador em espaço junta.

Em contrapartida, a arquitectura com controlo baseado na imagem e o efector terminal em cadeia fechada (ECL) revela uma percentagem de capturas com sucesso de apenas 73%. O incremento no tempo de processamento da imagem, que introduz atrasos significativos na cadeia de controlo, e a diferença entre o período de amostragem do sistema de visão e o valor médio T_v , provocam oscilações nos sinais de controlo e consequentes vibrações no efector terminal do manipulador e sobrelevações na sua trajectória. Como pode ser

observado na Figura 4.25 numa vizinhança do instante de tempo $t = 0.5s$, muitas vezes a “caneca” vibra enquanto se processa a tentativa de captura, razão pela qual esta é a arquitectura que apresenta pior percentagem de capturas com sucesso.

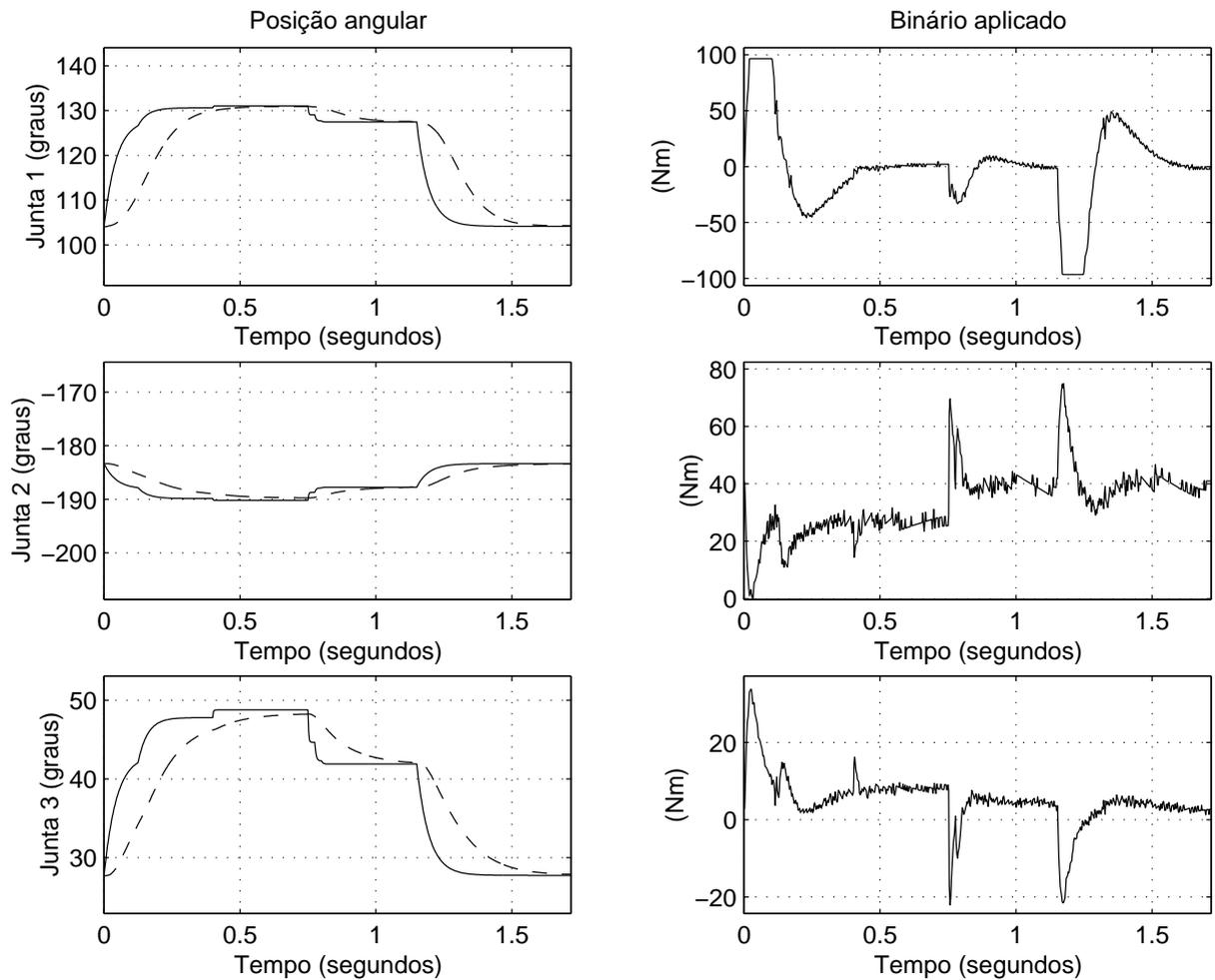


Fig. 4.17: Resultados da arquitectura de seguimento visual e captura de objectos com controlo baseado na posição, em espaço junta. Posição angular das juntas e binários aplicados nas três primeiras juntas.

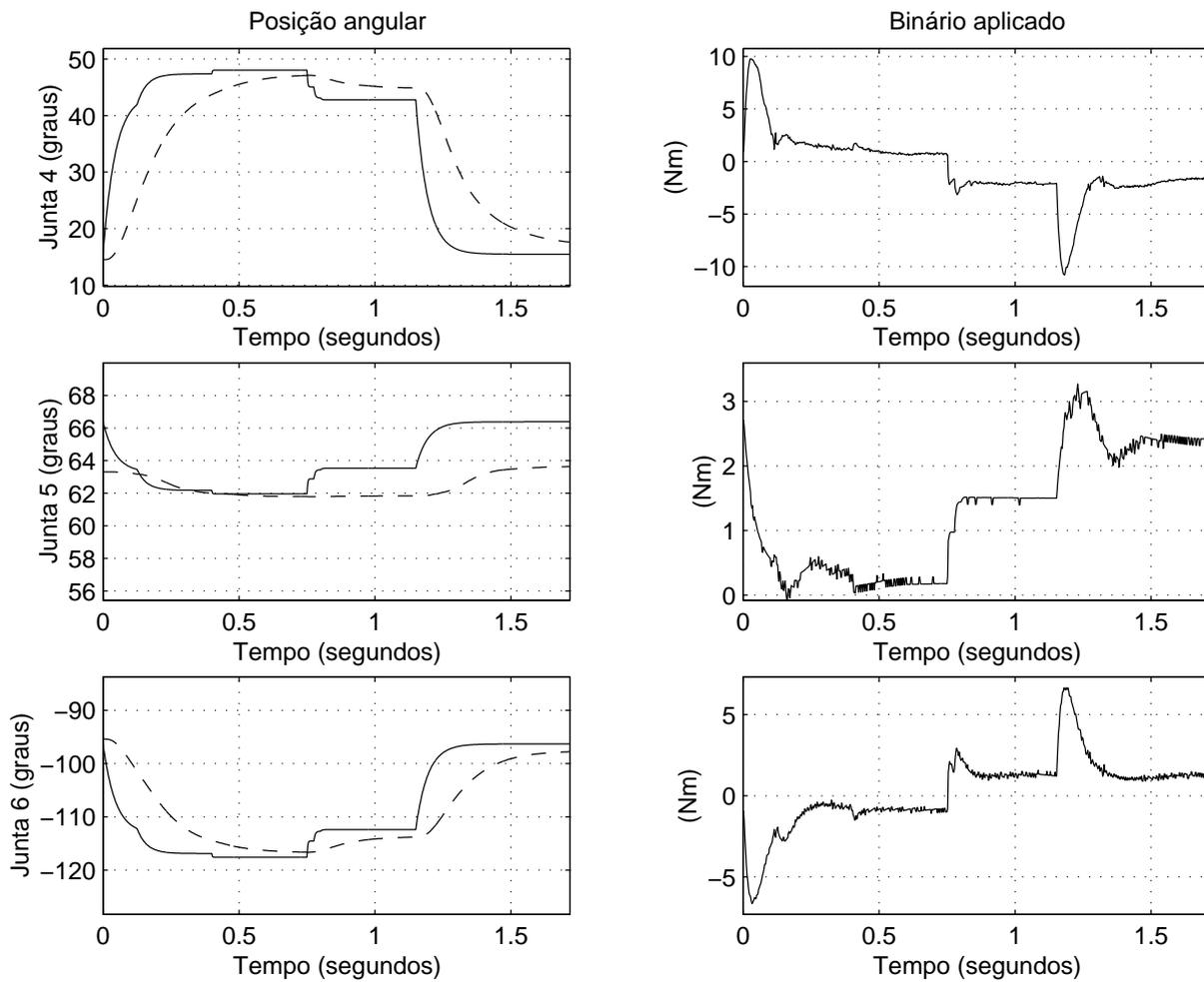


Fig. 4.18: Resultados da arquitectura de seguimento visual e captura de objectos com controlo baseado na posição, em espaço junta. Posição angular das juntas e binários aplicados nas três últimas juntas.

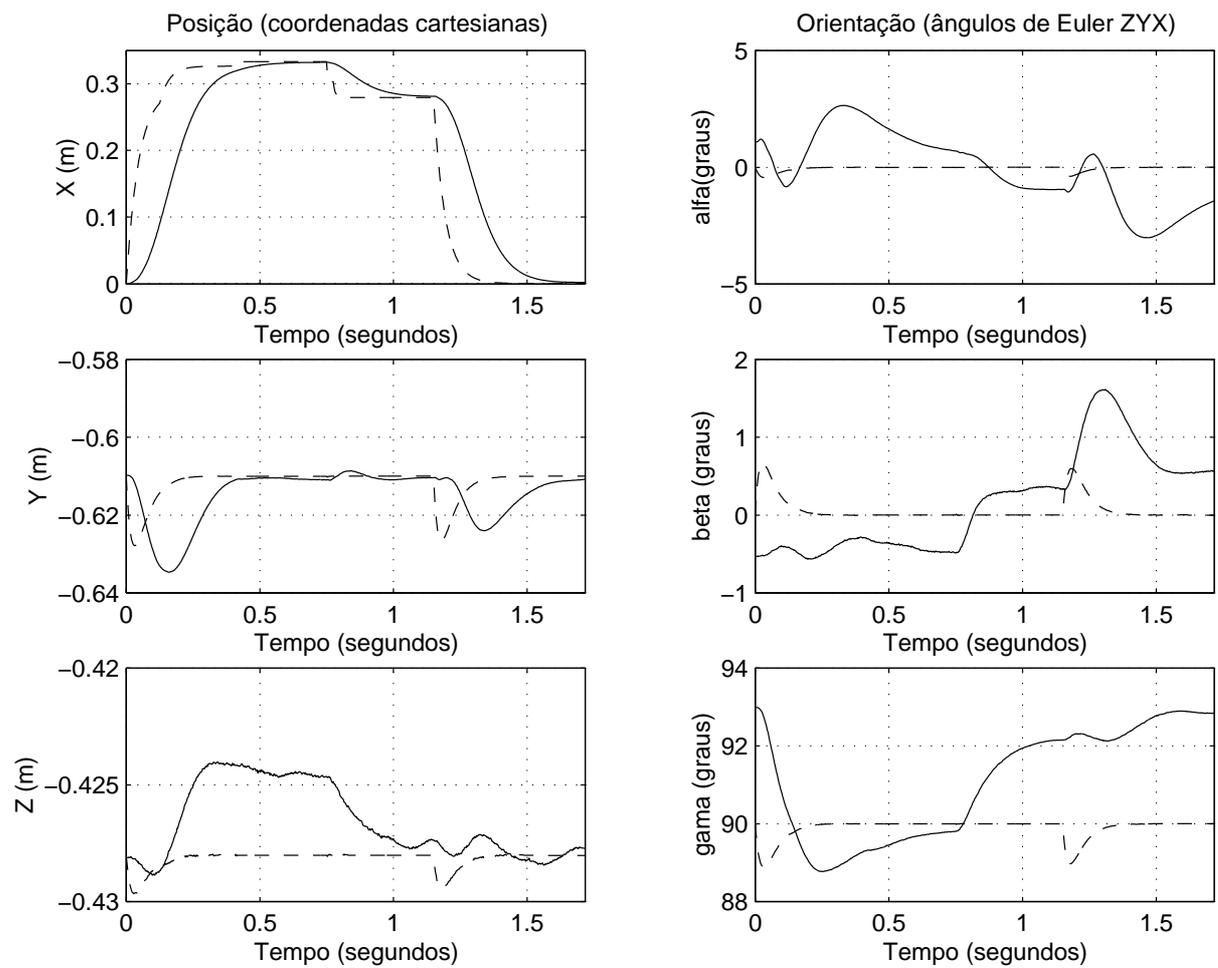


Fig. 4.19: Resultados da arquitectura de seguimento visual e captura de objectos com controlo baseado na posição, em espaço junta. Posição e orientação do effector terminal do manipulador.

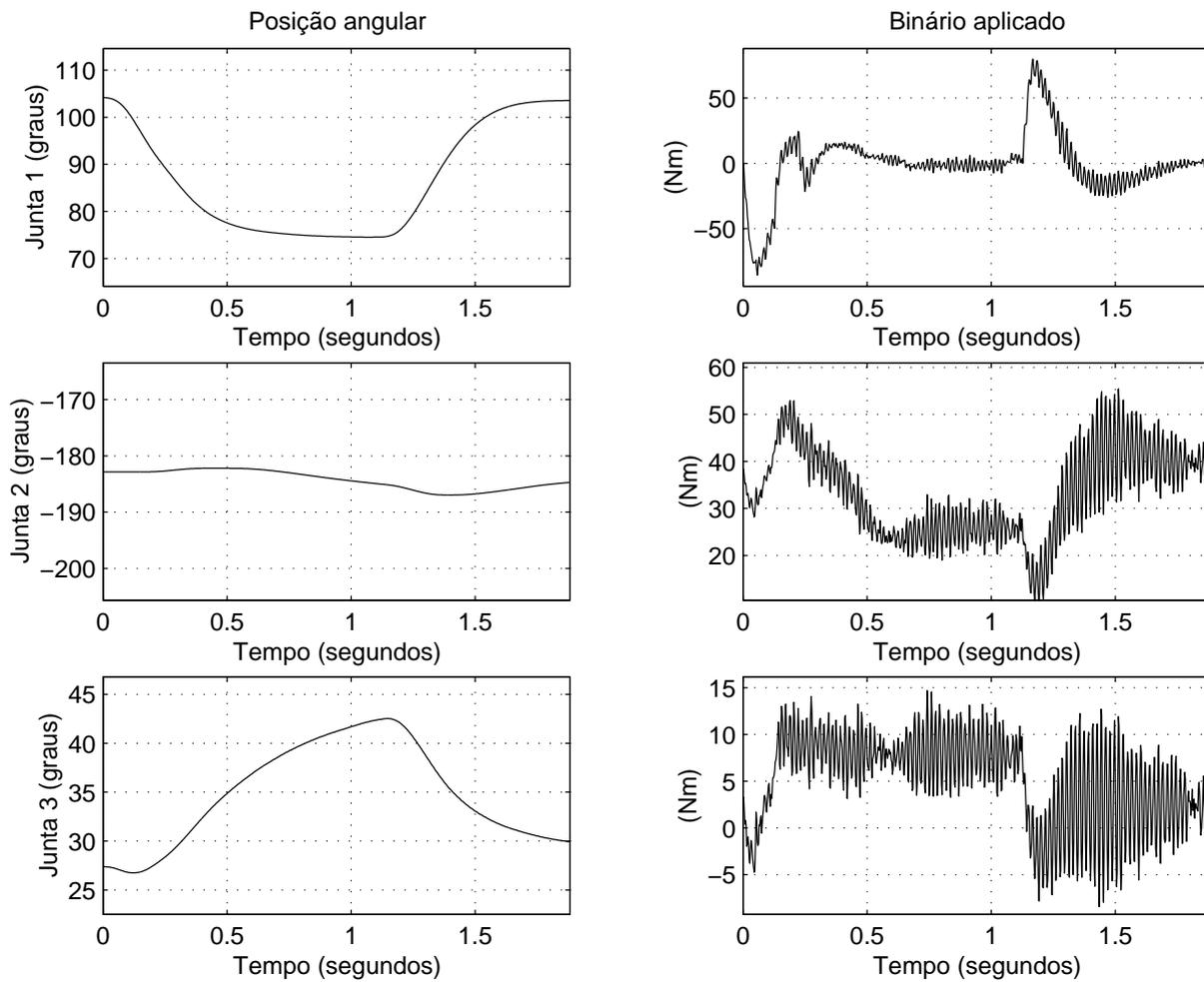


Fig. 4.20: Resultados da arquitectura de seguimento visual e captura de objectos com controlo baseado na posição, em espaço operacional. Posição angular das juntas e binários aplicados nas três primeiras juntas.

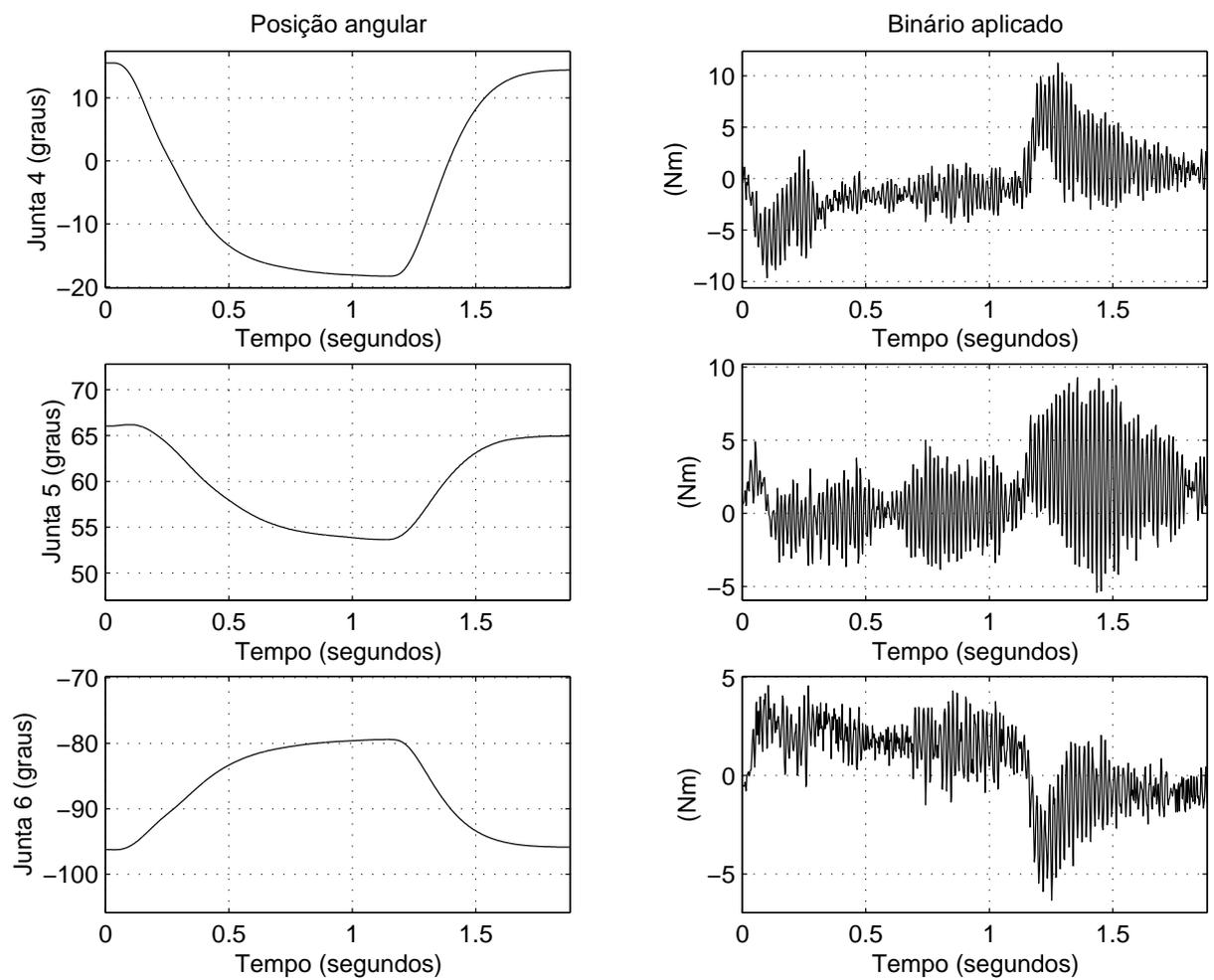


Fig. 4.21: Resultados da arquitectura de seguimento visual e captura de objectos com controlo baseado na posição, em espaço operacional. Posição angular das juntas e binários aplicados nas três últimas juntas.

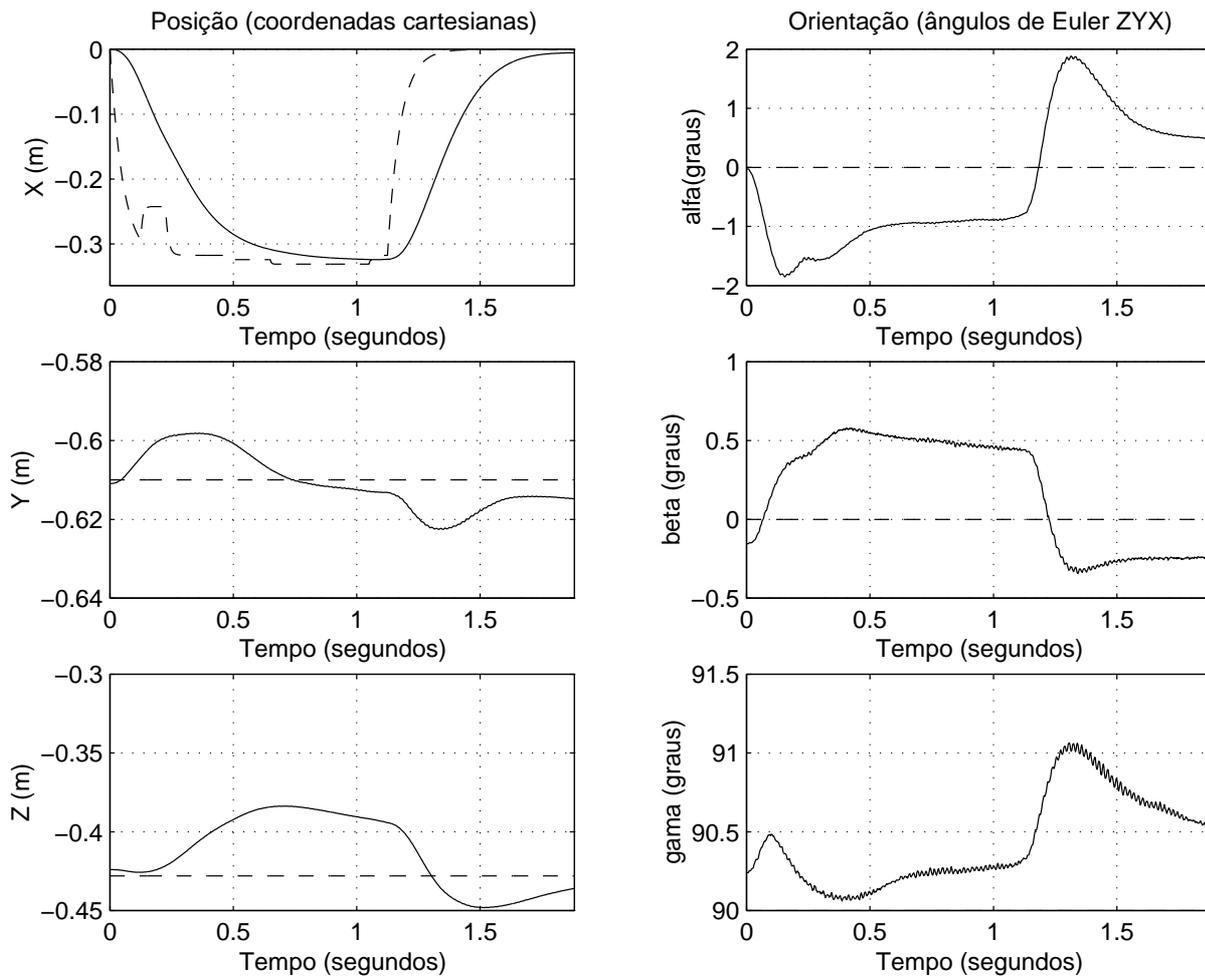


Fig. 4.22: Resultados da arquitectura de seguimento visual e captura de objectos com controlo baseado na posição, em espaço operacional. Posição e orientação do efector terminal do manipulador.

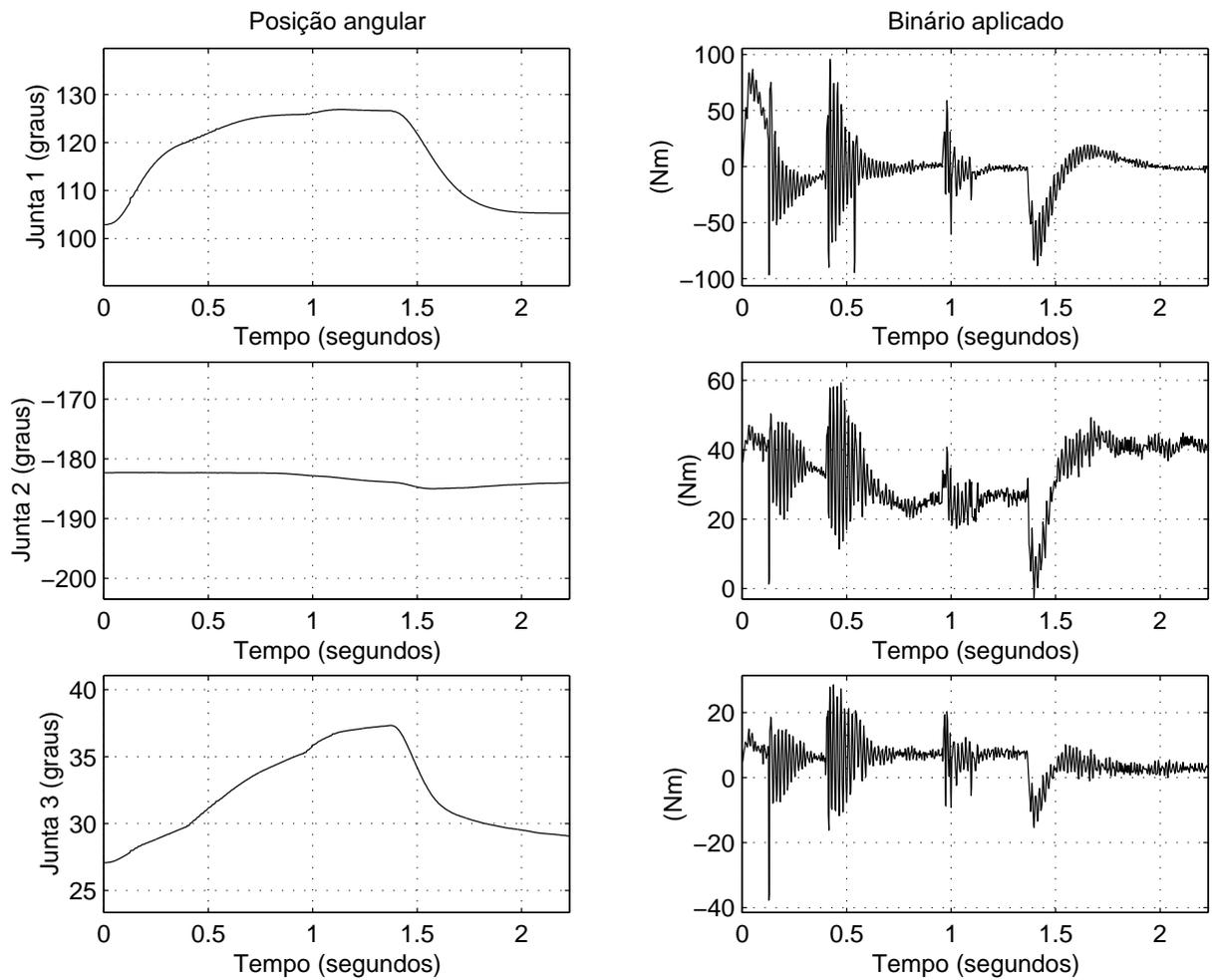


Fig. 4.23: Resultados da arquitectura de seguimento visual e captura de objectos com controlo baseado na imagem. Posição angular das juntas e binários aplicados nas três primeiras juntas.

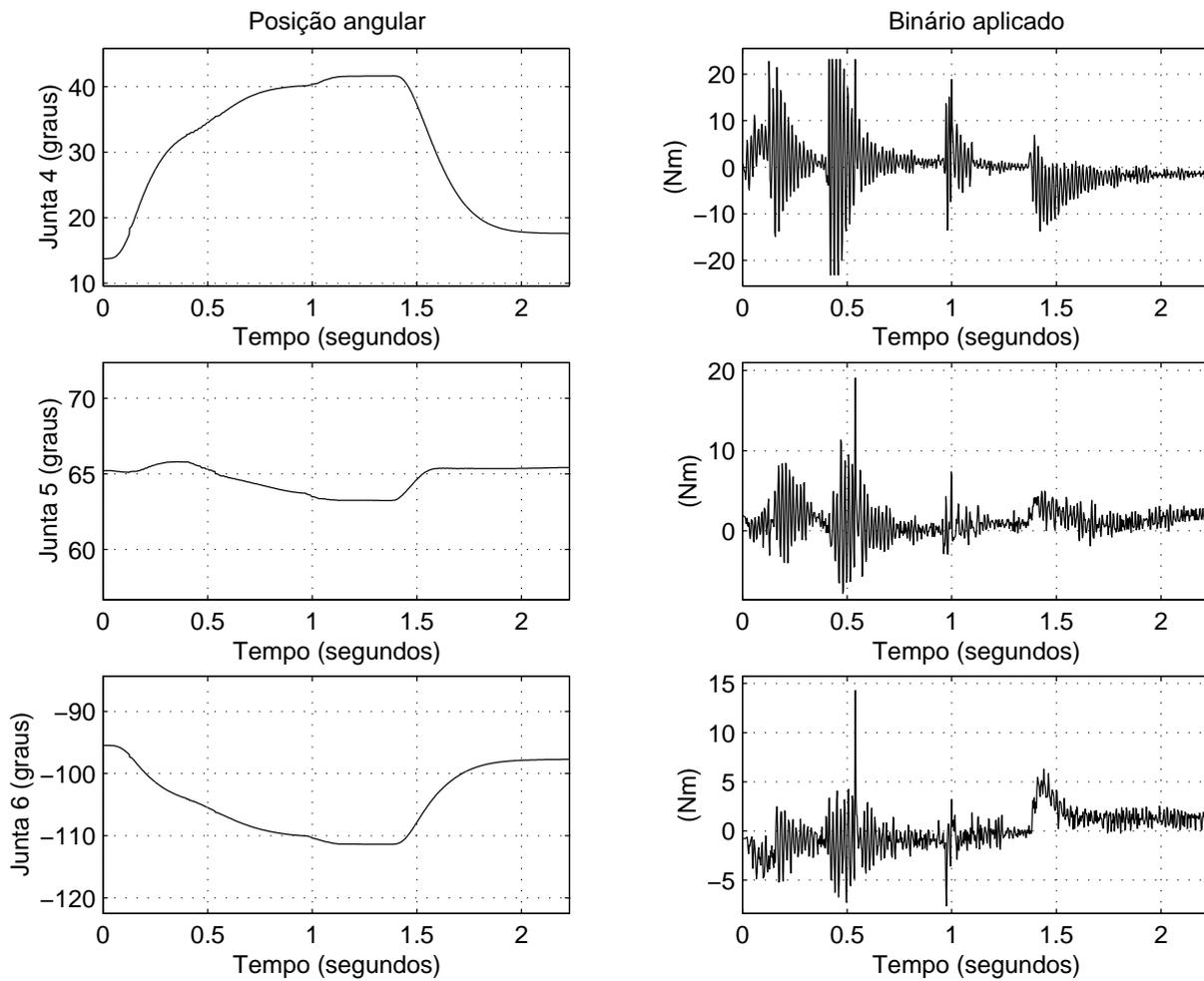


Fig. 4.24: Resultados da arquitectura de seguimento visual e captura de objectos com controlo baseado na imagem. Posição angular das juntas e binários aplicados nas três últimas juntas.

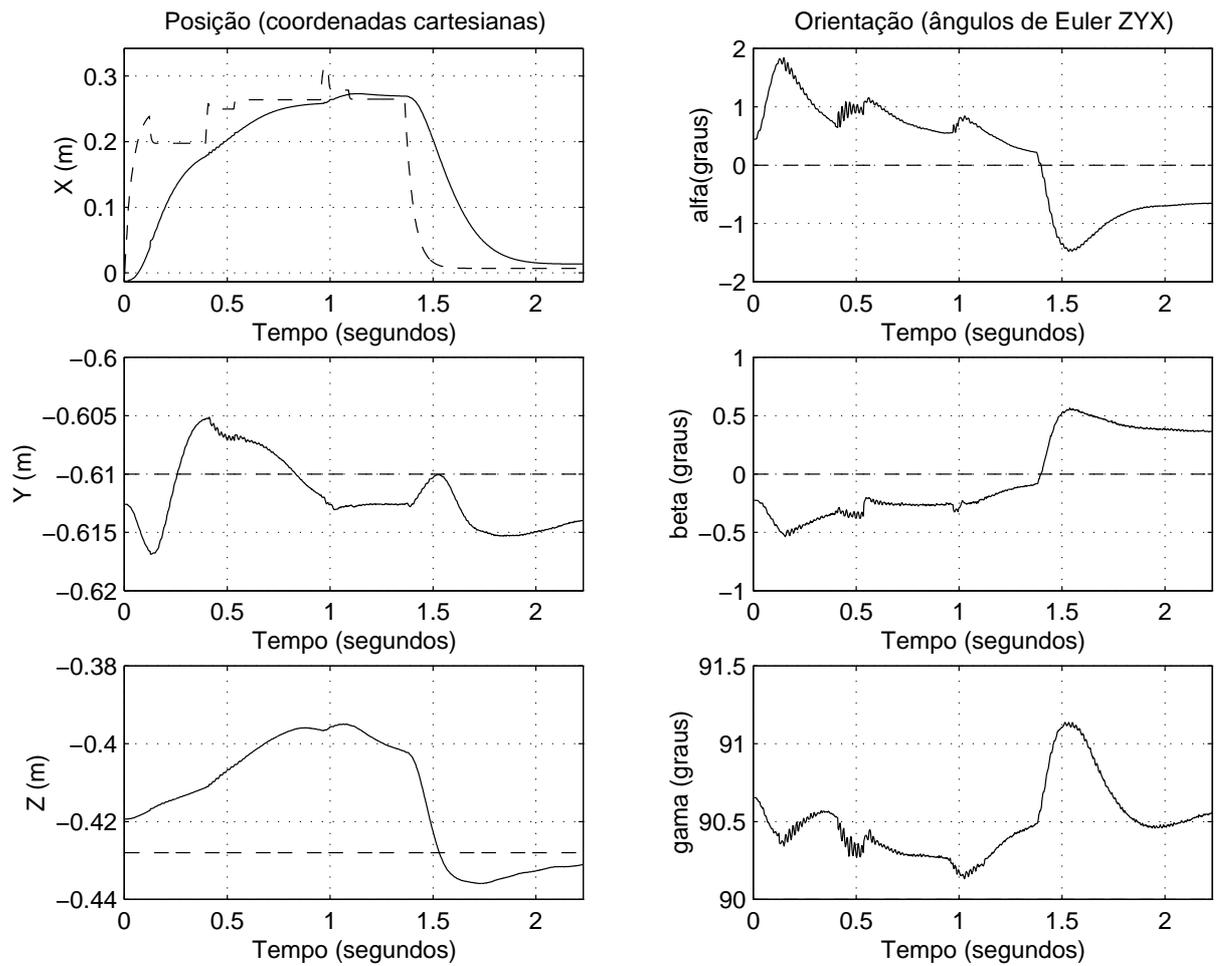


Fig. 4.25: Resultados da arquitectura de seguimento visual e captura de objectos com controlo baseado na imagem. Posição e orientação do efector terminal do manipulador.

5. Conclusões

Neste trabalho foi feito um estudo comparativo de diferentes arquitecturas de seguimento e captura de objectos em movimento por um manipulador robótico. Concretamente, foram implementadas duas arquitecturas distintas usando como base uma estrutura do tipo *olha-e-move dinâmico*:

- controlo baseado na posição e o efector terminal em cadeia aberta (EOL);
- controlo baseado na imagem e o efector terminal em cadeia fechada (ECL).

Note-se que em relação à arquitectura com controlo baseado na posição foram implementadas duas soluções distintas, uma delas servindo-se de um controlador em espaço junta do tipo PD com compensação gravítica, e a outra também com um controlador PD com compensação gravítica, mas em espaço operacional. No que toca à arquitectura com controlo baseado na imagem, foi usado apenas o controlador PD com compensação gravítica em espaço operacional.

Procedeu-se também à caracterização, quer em termos genéricos, quer relativamente a uma instalação de ensaio criada no âmbito desta Tese, de cada um dos subsistemas que constituem um sistema de seguimento e captura de objectos em movimento: o sistema de visão, o sistema de predição e o sistema de controlo do manipulador.

Em termos experimentais, foram apresentados resultados relativos à utilização da instalação de ensaio na tarefa da captura por um manipulador Puma 560 de bolas de ping-pong que rolam sobre uma mesa.

Constata-se que a arquitectura de seguimento visual e captura baseada na posição que utiliza o controlador PD com compensação gravítica em espaço junta, apresenta o melhor comportamento. A resposta do manipulador é a mais rápida e não apresenta sobrelevações, o que se traduz numa percentagem de capturas com sucesso na ordem dos

90%. Por seu turno, a arquitectura baseada na posição com o controlador do manipulador em espaço operacional apresenta uma percentagem de sucessos na ordem dos 82%. Tal é justificado pelo facto de o controlador em espaço operacional apresentar um desempenho inferior ao do controlador em espaço junta. A resposta é mais lenta e apresenta sobreelevações, o que é justificado pelo facto de os ganhos de controlo utilizados serem inferiores aos valores ajustados na simulação e que correspondiam ao melhor desempenho, similar ao do controlador em espaço junta. Note-se que esta limitação imposta nos ganhos de controlo revelou-se necessária devido à relativamente baixa frequência do sistema de controlo do Puma 560.

Em contrapartida, a arquitectura com controlo baseado na imagem e o efector terminal em cadeia fechada (ECL) revela uma percentagem de capturas com sucesso de apenas 73%. Embora o controlo baseado na imagem elimine os erros de calibração da câmara, e o esquema ECL reduza os erros de posicionamento do manipulador, verifica-se que na realidade a utilização desta arquitectura não se traduz numa melhoria face à utilização da arquitectura com controlo baseado na posição, pelo facto de os erros de calibração e da cinemática do sistema em causa não serem muito significativos e como tal serem facilmente compensados pela largura da “caneca” utilizada. Além disso, devido ao incremento no tempo de processamento da imagem, são introduzidos atrasos significativos na cadeia de controlo, e por outro lado, uma vez que o período de amostragem do sistema de visão não é fixo, o seu valor não coincide muitas vezes com o valor médio T_v (ver Secção 3.4.2) utilizado no controlo, o que leva a oscilações nos sinais de controlo e consequentes vibrações no efector terminal do manipulador e sobreelevações na sua trajectória. Em consequência, esta é a arquitectura que apresenta pior percentagem de capturas com sucesso.

Nas várias experiências apresentadas, constatou-se que a qualidade dos resultados era em grande parte influenciada pelas limitações do sistema de controlo do manipulador e do sistema de visão. Convém assim notar que o desempenho de um sistema de seguimento e captura de objectos em movimento depende não só da arquitectura utilizada, mas também das características e limitações dos subsistemas que o constituem, as quais são impostas pela tecnologia utilizada para os implementar.

Em termos genéricos, a escolha entre uma arquitectura baseada na posição ou baseada na imagem é um problema importante e que depende fundamentalmente dos requisitos e características da aplicação.

Uma arquitectura de seguimento visual e captura com controlo baseado na imagem apresenta uma estrutura de controlo com maior complexidade do que uma arquitectura com o controlo baseado na posição. Além do mais, requer a computação em tempo real do Jacobiano da imagem, o que representa uma dificuldade acrescida, uma vez que esta quantidade depende da distância da câmara ao objecto alvo, que pode ser difícil de calcular, em especial num sistema monocular. Ainda que nalguns casos se possa usar um Jacobiano da imagem constante, mas que apenas será válido para uma determinada região do espaço, é normalmente necessário estimar a distância do objecto à câmara para obter uma estimativa do Jacobiano da imagem, ou estimar directamente o Jacobiano da imagem, o que em qualquer um dos casos introduz uma complexidade significativa e requer uma carga computacional adicional.

Assim, a sua utilização em detrimento do uso de uma arquitectura com o controlo baseado na posição só se justificará nos casos em que a necessidade de suprimir os erros de calibração da câmara, ou a própria calibração, por ser muitas vezes um processo moroso, compensa as desvantagens acima referidas.

Por seu turno, as arquitecturas com o efector terminal do manipulador em cadeia fechada (ECL) são as únicas que garantem imunidade face aos erros na cinemática do manipulador, reduzindo os erros de posicionamento do manipulador. Infelizmente, apresentam um problema adicional de aplicabilidade prática e que consiste na necessidade de se garantir que o efector terminal do manipulador não saia da imagem. O que implica um grande cuidado no dimensionamento dos ganhos dos controladores das juntas, ou na geração de trajectórias do manipulador, de modo a que o efector terminal do manipulador não saia do campo de visão da câmara.

5.1 Trabalho Futuro

Em termos de trabalho futuro, pretende-se, a curto prazo, substituir o actual sistema de visão por outro mais rápido, muito provavelmente utilizando um *frame grabber* para

bus PCI, o que permitirá tornar a predição mais precisa pelo maior número de pontos disponíveis. Além disso, o novo sistema deverá permitir a aquisição das imagens com um período de amostragem fixo, o que possibilitará a predição em termos temporais do movimento do objecto, podendo, por conseguinte, ser implementadas outras formas de interceptação e captura do objecto mais elaboradas. Nomeadamente, em vez da captura se processar através do posicionamento do efector terminal do manipulador no “ponto de captura”, permanecendo aí à espera do objecto, poderia ser implementada uma solução na qual fosse calculada uma trajectória de interceptação para o efector terminal de tal forma que este atingisse o “ponto de captura” no mesmo instante que o objecto, e em vez de parar nesse ponto, seguisse a trajectória do objecto de forma a prolongar o movimento natural deste, tal como no trabalho de Hong *et al.* [34][35]. O uso desta técnica traduz-se numa captura mais suave e mais imune a problemas relativos ao impacto do objecto com o efector terminal.

No que concerne às direcções de investigação para um trabalho futuro, uma delas passa necessariamente por um aumento da complexidade da aplicação. Numa primeira fase, teria interesse a evolução da actual captura da bola ao longo de uma linha, a “linha de captura”, para a captura dentro de um plano (por exemplo, o plano da mesa), o que constitui uma verdadeira captura em 2D. Posteriormente, poder-se-ia evoluir para a captura 3D em qualquer ponto do espaço de trabalho do manipulador, tal como a captura de bolas atiradas pelo ar.

Outra direcção de investigação é a da implementação de um mecanismo de reconhecimento de objectos mais elaborado que permitisse a detecção e localização da bola num ambiente menos estruturado, ou mesmo não estruturado. Seria também de explorar a aplicação desse mecanismo no reconhecimento de objectos mais complexos do que uma bola de ping-pong.

Embora o sistema de predição implementado tenha obtido bons resultados, vê-se com interesse a utilização de outras funções de base, assim como de outras soluções de predição, como a filtragem de Kalman (conhecendo-se a dinâmica do objecto) ou filtros $\alpha - \beta - \gamma$, em especial quando se pretender efectuar capturas no espaço 3D.

Relativamente ao controlo do manipulador, a utilização de outros sistemas de controlo centralizado, como por exemplo os que utilizam a dinâmica inversa do manipulador,

será também de considerar, em particular quando se pretender gerar trajectórias de interceptação para o manipulador, de forma a que a captura se processe de forma suave com o manipulador em movimento e a acompanhar a trajectória do objecto, o que exige um bom seguimento de trajectórias. Em todo o caso, no que toca ao sistema de controlo do manipulador, dever-se-à proceder, em primeiro lugar, a um aumento da frequência do controlador que permita um melhor desempenho do esquema de controlo em espaço operacional e de esquemas que utilizem a dinâmica inversa do manipulador.

Teria também grande interesse o estudo formal da cadeia de controlo completa em termos de estabilidade e robustez, bem assim como a sistematização de métodos de projecto dos controladores para seguimento e captura de objectos em movimento.

A. Procedimento de Calibração da Câmara

O procedimento de calibração da câmara consiste, em primeiro lugar, no envio do efector terminal do manipulador para quatro pontos com coordenadas conhecidas no sistema de coordenadas da base do manipulador ${}^0[X'_i \ Y'_i \ Z_{mesa}]^T$, com $i = 1, \dots, 4$, tendo em atenção que esses pontos deverão fazer necessariamente parte do plano da mesa. Para posicionar o efector terminal nesses pontos é utilizada uma Interface Gráfica do Puma 560. Na garra do manipulador é presa uma barra de material plástico, cuja finalidade é a de aumentar o comprimento total do manipulador por forma a que o seu ponto mais extremo seja visível na imagem. A orientação do efector terminal é escolhida por forma a que a barra de material plástico fique paralela ao eixo Y do sistema de coordenadas da base do manipulador.

São então adquiridas quatro imagens, correspondentes a cada um dos pontos. Na Figura A.1 encontram-se as imagens usadas numa calibração da câmara, nas quais é visível a barra de material plástico que foi presa na garra do manipulador. As coordenadas da ponta da barra de material plástico visível nas imagens, ${}^0[X_i \ Y_i \ Z_{mesa}]^T$, com $i = 1, \dots, 4$, relacionam-se com as coordenadas do efector terminal do manipulador da seguinte forma:

$$\begin{aligned} X_i &= X'_i, \\ Y_i &= Y'_i + d, \end{aligned} \tag{A.1}$$

onde d representa a distância entre o ponto efector terminal e a ponta da barra de material plástico. Na realidade, nas imagens c) e d) da Figura A.1 em vez de se considerar a ponta da barra, considerou-se um ponto correspondente a cerca de metade do comprimento da barra e que se encontra marcado a negro.

Por fim é utilizada uma rotina em Matlab, que dadas as quatro imagens, pede ao utilizador para marcar, em cada uma delas, o ponto da imagem $[u_i v_i]^T$ correspondente a ${}^0[X_i Y_i Z_{mesa}]^T$, com $i = 1, \dots, 4$. Os coeficientes da matriz de calibração C' são estimados usando a equação (3.7) da seguinte forma:

$$\begin{bmatrix} X_1 & Y_1 & 1 & 0 & 0 & 0 & -u_1 X_1 & -u_1 Y_1 \\ 0 & 0 & 0 & X_1 & Y_1 & 1 & -v_1 X_1 & -v_1 Y_1 \\ X_2 & Y_2 & 1 & 0 & 0 & 0 & -u_2 X_2 & -u_2 Y_2 \\ 0 & 0 & 0 & X_2 & Y_2 & 1 & -v_2 X_2 & -v_2 Y_2 \\ X_3 & Y_3 & 1 & 0 & 0 & 0 & -u_3 X_3 & -u_3 Y_3 \\ 0 & 0 & 0 & X_3 & Y_3 & 1 & -v_3 X_3 & -v_3 Y_3 \\ X_4 & Y_4 & 1 & 0 & 0 & 0 & -u_4 X_4 & -u_4 Y_4 \\ 0 & 0 & 0 & X_4 & Y_4 & 1 & -v_4 X_4 & -v_4 Y_4 \end{bmatrix} \begin{bmatrix} c_{11} \\ c_{12} \\ c'_{13} \\ c_{21} \\ c_{22} \\ c'_{23} \\ c_{31} \\ c_{32} \end{bmatrix} = \begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \\ u_4 \\ v_4 \end{bmatrix}. \quad (\text{A.2})$$

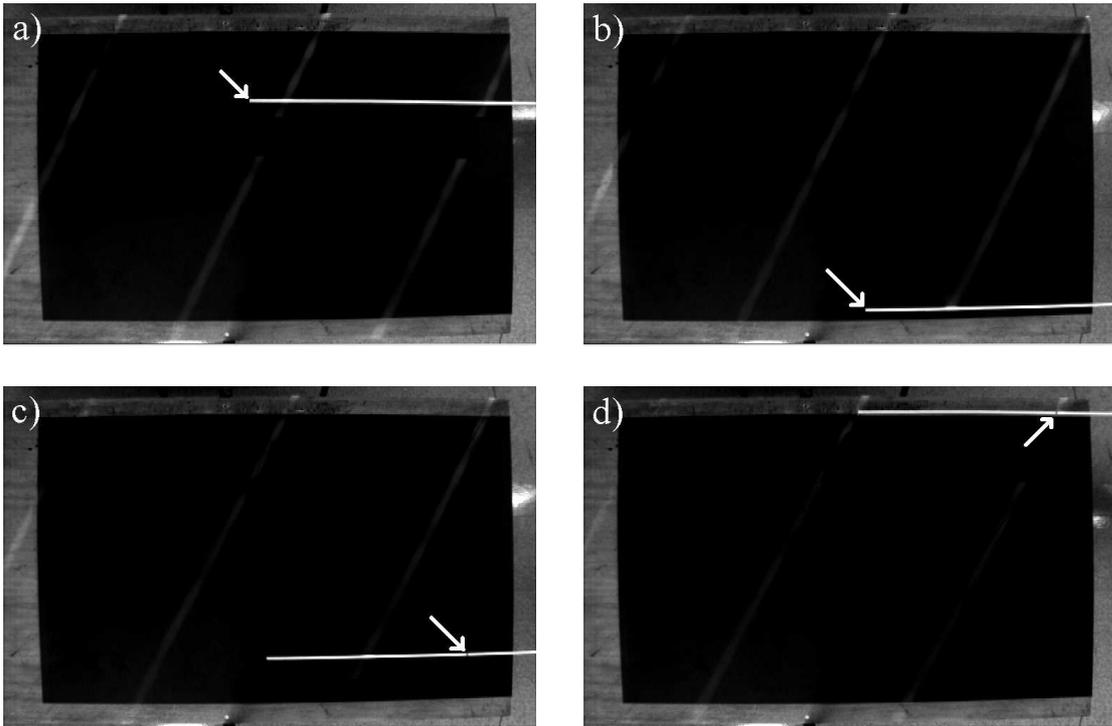


Fig. A.1: Imagens usadas na calibração da câmara. As setas indicam o ponto da barra utilizado na calibração.

Referências

- [1] J. Aggarwal and N. Nandhakumar. On the computation of motion from sequences of images. *Proceedings of the IEEE*, 76(8), August 1988.
- [2] P. Allen, A. Timcenko, B. Yoshimi, and P. Michelman. Hand-eye coordination for robotic tracking and grasping. In K. Hashimoto, editor, *Visual Servoing*, volume 7 of *Robotics and Automated Systems*. World Scientific, 1993.
- [3] P. Allen, B. Yoshimi, and A. Timchenko. Real time visual servoing. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 851–856, April 1991.
- [4] P. Allen, B. Yoshimi, A. Timchenko, and P. Michelman. Trajectory filtering and prediction for automated tracking and grasping of a moving object. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 1850–1856, May 1992.
- [5] R. Andersson. Real-time gray-scale video processing using a moment-generating chip. *IEEE Journal of Robotics and Automation*, RA-1(2), June 1985.
- [6] R. Andersson. Real time intelligent visual control of a robot. In *Proc. of IEEE Workshop on Intelligent Control*, August 1985.
- [7] R. Andersson. Aggressive trajectory generator for a robot ping-pong player. *IEEE Control Systems Magazine*, February 1989.
- [8] R. Andersson. Dynamic sensing in a ping-pong playing robot. *IEEE Transactions on Robotics and Automation*, 5(6), December 1989.

- [9] B. Armstrong, O. Khatib, and J. Burdick. The explicit dynamic model and inertial parameters of the Puma 560 arm. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 510–518, 1986.
- [10] K. Åström and B. Wittenmark. *Computer Controlled Systems*. Prentice-Hall, 1984.
- [11] A. Bendiksen and G. Hager. A vision-based grasping system for unfamiliar planar objects. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 2844–2849, May 1994.
- [12] Z. Bien, W. Jang, and J. Park. Characterization and use of feature-jacobian matrix for visual servoing. In K. Hashimoto, editor, *Visual Servoing*, volume 7 of *Robotics and Automated Systems*. World Scientific, 1993.
- [13] G. Buttazzo, B. Allotta, and F. Fanizza. Mousebuster: a robot for real-time catching. *IEEE Control Systems Magazine*, 14(1), February 1994.
- [14] A. Castaño and S. Hutchinson. Visual compliance: Task-directed visual servo control. *IEEE Transactions on Robotics and Automation*, 10(3), June 1994.
- [15] F. Chaumette, P. Rives, and B. Espiau. Classification and realization of the different vision-based tasks. In K. Hashimoto, editor, *Visual Servoing*, volume 7 of *Robotics and Automated Systems*. World Scientific, 1993.
- [16] P. Corke. Video-rate robot visual servoing. In K. Hashimoto, editor, *Visual Servoing*, volume 7 of *Robotics and Automated Systems*. World Scientific, 1993.
- [17] P. Corke. Visual control of robot manipulators - a review. In K. Hashimoto, editor, *Visual Servoing*, volume 7 of *Robotics and Automated Systems*. World Scientific, 1993.
- [18] P. Corke. *Visual Control of Robots: High Performance Visual Servoing*. Research Studies Press (John Wiley), 1996.
- [19] P. Corke and B. Armstrong-Hélouvry. A search for consensus among model parameters reported for the PUMA 560 robot. In *Proc. of the IEEE International Conference on Robotics and Automation*, May 1994.

- [20] P. Corke and M. Good. Dynamic effects in visual closed-loop systems. *IEEE Transactions on Robotics and Automation*, 12(5), October 1996.
- [21] P.I. Corke. A robotics toolbox for MATLAB. *IEEE Robotics and Automation Magazine*, 3(1):24–32, March 1996.
- [22] J. Craig. *Introduction to Robotics: Mechanics and Control*. Addison-Wesley, 1989.
- [23] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Transactions on Robotics and Automation*, 8(3), June 1992.
- [24] J. Feddema, C. Lee, and O. Mitchell. Weighted selection of image features for resolved rate visual feedback control. *IEEE Transactions on Robotics and Automation*, 7(1), February 1991.
- [25] J. Feddema, C. Lee, and O. Mitchell. Feature-based visual servoing of robotic systems. In K. Hashimoto, editor, *Visual Servoing*, volume 7 of *Robotics and Automated Systems*. World Scientific, 1993.
- [26] J. Feddema and O. Mitchell. Vision-guided servoing with feature-based trajectory generation. *IEEE Transactions on Robotics and Automation*, 5(5), October 1989.
- [27] K. Fu, R. Gonzalez, and C. Lee. *Robotics — Control, Sensing, Vision and Intelligence*. McGraw-Hill, 1987.
- [28] B. Ghosh, T. Tarn, N. Xi, Z. Yu, and D. Xiau. Calibration free visually controlled manipulation of parts in a robotic manufacturing workcell. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 3197–3202, April 1996.
- [29] E. Grosso, G. Metta, A. Oddera, and G. Sandini. Robust visual servoing in 3-D reaching tasks. *IEEE Transactions on Robotics and Automation*, 12(5), October 1996.
- [30] G. Hager. A modular system for robust positioning using feedback from stereo vision. *IEEE Transactions on Robotics and Automation*, 13(4), August 1997.

- [31] K. Hashimoto, editor. *Visual Servoing*, volume 7 of *Robotics and Automated Systems*. World Scientific, 1993.
- [32] K. Hashimoto, T. Kimoto, T. Ebine, and H. Kimura. Manipulator control with image-based visual servo. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 2267–2272, April 1991.
- [33] J. Hill and W. Park. Real time control of a manipulator with a mobile camera. In *Proc. 9th ISIR*, pages 233–246, Washington, D.C., March 1979.
- [34] W. Hong. Robotic catching and manipulation using active vision. Master's thesis, MIT, Cambridge, MA, August 1995.
- [35] W. Hong and J. Slotine. Experiments in hand-eye coordination using active vision. In *Experimental Robotics IV, Springer-Verlag, Proc. of ISER 95*, Stanford, CA, July 1995.
- [36] B. Horn. *Robot Vision*. MIT Press, McGraw-Hill, 1986.
- [37] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1983.
- [38] N. Houshangi. Control of a robotic manipulator to grasp a moving target using vision. In *Proc. of IEEE International Conference on Robotics and Automation*, pages 604–609, May 1990.
- [39] B. Hove and J. Slotine. Experiments in robotic catching. In *Proc. of the 1991 American Control Conference*, pages 380–385, Boston, MA, June 1991.
- [40] S. Hutchinson, G. Hager, and P. Corke. A tutorial on visual servo control. *IEEE Transactions on Robotics and Automation*, 12(5), October 1996.
- [41] W. Jang and Z. Bien. Feature-based visual servoing of an eye-in-hand robot with improved tracking performance. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 2254–2260, April 1991.

- [42] H. Kimura, N. Mukai, and J. Slotine. Adaptive visual tracking and gaussian network algorithms for robotic catching. In *Advances in Robust and Nonlinear Control Systems, Winter Annual Meeting of the ASME*, volume DSC-43, Anaheim, CA, November 1992.
- [43] Y. Li and M. Lee. Applying vision guidance in robotic food handling. *IEEE Robotics and Automation Magazine*, 3(1):4–12, March 1996.
- [44] P. Lima, N. Martins, and D. Fernandes. Arquitectura aberta de controlo do puma 560. Technical report, Instituto de Sistemas e Robótica, 1997.
- [45] N. Moreira, P. Alvito, and P. Lima. First steps towards an open control architecture for a puma 560. In *Proc. of the Portuguese Conference on Automatic Control, CONTROLO 96*, Porto - PORTUGAL, September 1996.
- [46] B. Nelson, N. Papanikolopoulos, and P. Khosla. Robotic visual servoing and robotic assembly tasks. *IEEE Robotics and Automation Magazine*, June 1996.
- [47] N. Papanikolopoulos, P. Khosla, and T. Kanade. Vision and control techniques for robotic visual tracking. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 857–864, April 1991.
- [48] A. Rizzi and D. Koditschek. An active visual estimator for dexterous manipulation. *IEEE Transactions on Robotics and Automation*, 12(5), October 1996.
- [49] A. Sanderson, L. Weiss, and C. Neuman. Dynamic sensor-based control of robots with visual feedback. *IEEE Transactions on Robotics and Automation*, RA-3:404–417, October 1987.
- [50] L. Sciavicco and B. Siciliano. *Modeling and Control of Robot Manipulators*. McGraw-Hill, 1996.
- [51] Y. Shirai and H. Inoue. Guiding a robot by visual feedback in assembly tasks. *Pattern Recognition*, 5:99–108, 1973.

- [52] C. Smith and N. Papanikolopoulos. Computation of shape through controlled active exploration. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 2516–2521, May 1994.
- [53] C. Smith and N. Papanikolopoulos. Vision-guided robotic grasping: Issues and experiments. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 3203–3208, April 1996.
- [54] M. Taylor, A. Blake, and A. Cox. Visually guided grasping in 3D. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 761–766, May 1994.
- [55] W. Wilson. Visual servo control of robots using Kalman filter estimates of robot pose relative to work-pieces. In K. Hashimoto, editor, *Visual Servoing*, volume 7 of *Robotics and Automated Systems*. World Scientific, 1993.
- [56] W. Wilson, C. Hulls, and G. Bell. Relative end-effector control using cartesian position based visual servoing. *IEEE Transactions on Robotics and Automation*, 12(5), October 1996.
- [57] B. Yoshimi and P. Allen. Active, uncalibrated visual servoing. In *Proc. of the IEEE International Conference on Robotics and Automation*, pages 156–161, May 1994.