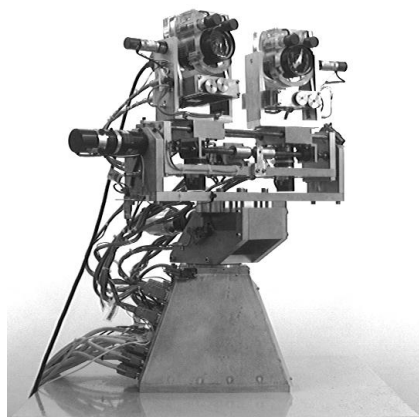


SOLUÇÕES PARA CONTROLO VISUAL DE  
MOVIMENTO: APLICAÇÃO EM SEGUIMENTO  
COM VISÃO ACTIVA



PROVAS DE APTIDÃO PEDAGÓGICA E CAPACIDADE CIENTÍFICA  
SUBMETIDAS AO DEPARTAMENTO DE ENGENHARIA ELECTROTÉCNICA  
FACULDADE DE CIÊNCIAS E TECNOLOGIA DA UNIVERSIDADE DE COIMBRA

João Pedro de Almeida Barreto  
Maio 2000



# Conteúdo

<b>1</b>	<b>O Sistema de Seguimento Visual MDOF</b>	<b>5</b>
1.1	Seguimento Binocular . . . . .	7
1.1.1	Seguimento com o Olho Ciclópico . . . . .	9
1.1.2	Seguimento em Vergência . . . . .	10
1.2	Arquitectura do Sistema de Seguimento da MDOF. Diagrama de Blocos . . . . .	11
1.2.1	Seguimento Monocular . . . . .	13
1.2.2	Seguimento em Vergência . . . . .	14
<b>2</b>	<b>Trajectórias de Referência do Alvo e Avaliação do Desempenho do Sistema</b>	<b>17</b>
2.1	Trajectórias de Referência . . . . .	18
2.1.1	Gerar com Rigor os Sinais de Perturbação . . . . .	20
2.1.2	Equações das Trajectórias de Referência. Seguimento Monocular . . . . .	22
2.1.3	Equações das Trajectórias de Referência. Controlo de Vergência . . . . .	28
2.2	Avaliação do Desempenho do Sistema . . . . .	31
2.2.1	Seguimento Monocular . . . . .	31
<b>3</b>	<b>Optimização do desempenho do processamento visual</b>	<b>37</b>
3.1	Estimação da Posição do Alvo na Imagem. . . . .	38
3.2	Estimação da Velocidade do Alvo na Imagem. . . . .	41
3.2.1	Método dos resíduos . . . . .	42
3.2.2	Método da máscara variável . . . . .	44
3.2.3	Resultados Experimentais . . . . .	44
3.3	Filtragem de Kalman . . . . .	47
3.4	Optimização do Controlo da Vergência . . . . .	53

3.4.1	Projecto do Filtro de Kalman . . . . .	55
3.4.2	Avaliação do Desempenho . . . . .	60
<b>4</b>	<b>Arquitectura do Sistema e Controlo de Fixação</b>	<b>65</b>
4.1	Controlo de Baixo Nível dos Servo-Motores . . . . .	67
4.2	Interpolação para Compensação do Atraso Visual . . . . .	72
4.3	Controlador de Fixação . . . . .	76
4.4	Controlo de Fixação Usando um Controlador DMC . . . . .	78
4.4.1	O controlador preditivo de Matriz Dinâmica (DMC) . . . . .	79
4.4.2	O sinal de referência $w(n)$ . . . . .	80
4.4.3	Desempenho do controlador DMC . . . . .	83
<b>5</b>	<b>Conclusões</b>	<b>85</b>

# Capítulo 1

## O Sistema de Seguimento Visual MDOF

O sistema MDOF, exibido na Fig 1.1, é um sistema mecânico binocular de elevado desempenho com um total de 16 graus de liberdade independentes. O desenho desta cabeça robótica tem uma inspiração antropomórfica. Cada olho tem dois graus de liberdade rotacionais: azimute, movimento direita-esquerda também designado por "eye pan", e elevação, movimento cima-baixo também conhecido por "eye tilt". Os dois olhos estão montados sobre uma plataforma com três graus de liberdade: azimute ou "neck pan", inclinação ou "neck swing" e elevação ou "neck tilt". Os graus de liberdade da plataforma constituem o módulo do pescoço. O centro óptico das lentes de cada câmara pode ser alinhado com o eixo de rotação do azimute do olho. Este grau de liberdade de ajuste tem como objectivo assegurar rotações perfeitas nos movimentos oculares. A distância intra-ocular pode também ser variada usando o grau de liberdade da "baseline". As câmaras do sistema MDOF estão equipadas com lentes motorizadas. Os seis graus de liberdade que faltam referir são os graus ópticos dos olhos: focagem, zoom e abertura. Para mais detalhes sobre a mecânica da cabeça MDOF consultar [1] [2].

No sistema MDOF o movimento é gerado por motores DC com caixas de engrenagem "harmonic drive". Cada motor está equipado com um codificador óptico que permite obter informação da sua posição/velocidade. O controlo dos motores é feito usando placas DCX-AT100 da "Precision Micro-control" [3] que permitem controlar as juntas em posição e/ou em velocidade. A posição de cada junta é controlada localmente por uma malha fechada com um filtro PID digital. A frequência de amostragem do servo-controlador local

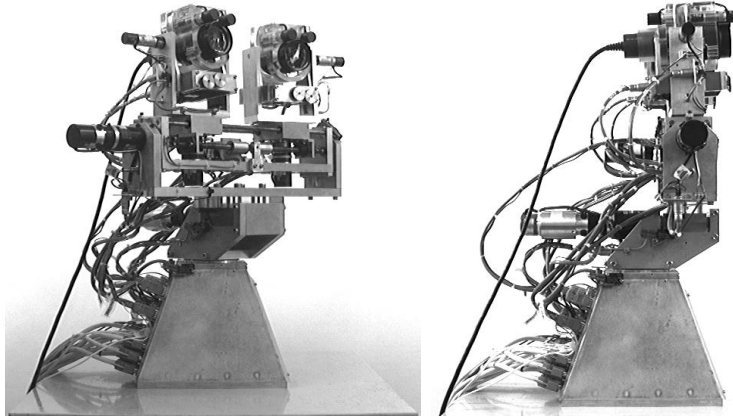


Figura 1.1: O sistema de visão activo ISR-MDOF.

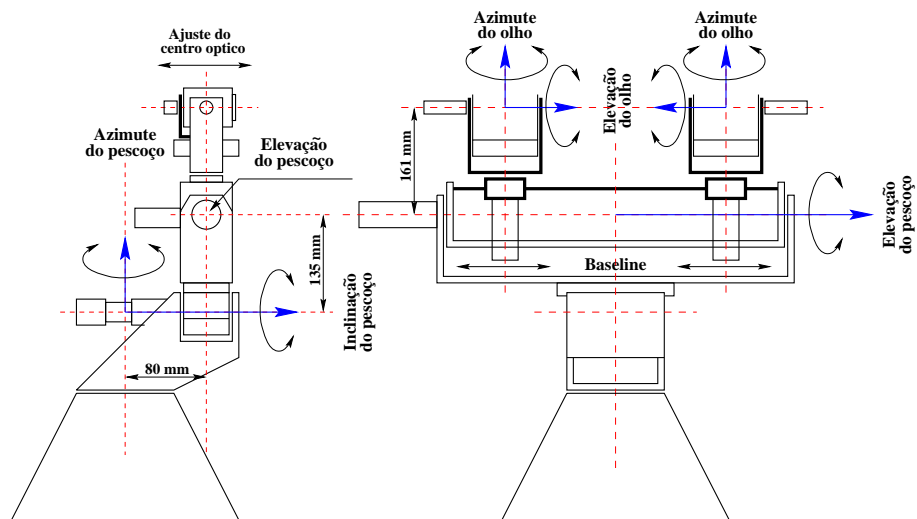


Figura 1.2: Geometria e graus de liberdade do sistema MDOF.

	Resolução	Velocidade Máxima	Aceleração Máxima
Azimute do Olho	0.0036°	$\sim 360^\circ/s$	$\sim 15000^\circ/s^2$
Elevação do Olho	0.0031°	$\sim 330^\circ/s$	$\sim 10000^\circ/s^2$
Azimute do Pescoço	0.0036°	$\sim 180^\circ/s$	$\sim 3000^\circ/s^2$
Elevação do Pescoço	0.0036°	$\sim 180^\circ/s$	$\sim 5000^\circ/s^2$

Figura 1.3: Desempenho e resolução dos graus de liberdade da MDOF utilizados.

é de 1KHz. A realimentação é obtida através dos codificadores ópticos. O controlo em velocidade é conseguido usando um gerador de perfil para gerar a referência da malha de posição. Este gerador de perfil integra o comando de velocidade fornecido pelo processo utilizador de alto nível. A comunicação com os servo-controladores locais é síncrona com frequência 166Hz. O controlo dos actuadores será analisado detalhadamente mais à frente.

Dos 16 graus de liberdade do sistema MDOF só interessam para o presente trabalho o azimute e elevação dos olhos e do pescoço. A tabela 1.3 mostra um quadro resumo com o desempenho e resolução de cada um destes graus de liberdade.

## 1.1 Seguimento Binocular

O sistema MDOF foi utilizado para implementar seguimento visual de alvos em movimento. O seguimento de um alvo, ou "smooth pursuit", consiste em acompanhar visualmente um objecto em movimento de forma a que a posição de projecção nas retinas do sistema se mantenha constante. Para isso, sempre que o alvo induz movimento na imagem, o sistema redirecciona a sua atenção de forma a compensar o movimento retinal e manter a projecção do objecto na zona foveal (centro da imagem). O comportamento de seguimento visual foi implementado em tempo-real depois de uma cuidadosa modelação cinemática do sistema mecânico e escolha adequada dos algoritmos de processamento visual [4] [5] [6] [7].

A cabeça MDOF utiliza o azimute e a elevação do pescoço e o azimute dos olhos para fazer o seguimento binocular do alvo. Considere que  $\alpha_p$  e  $\alpha_t$  são, respectivamente, as posições angulares do azimute e da elevação do pescoço. O azimute dos olhos é designado por vergência, sendo  $\beta_l$  o ângulo

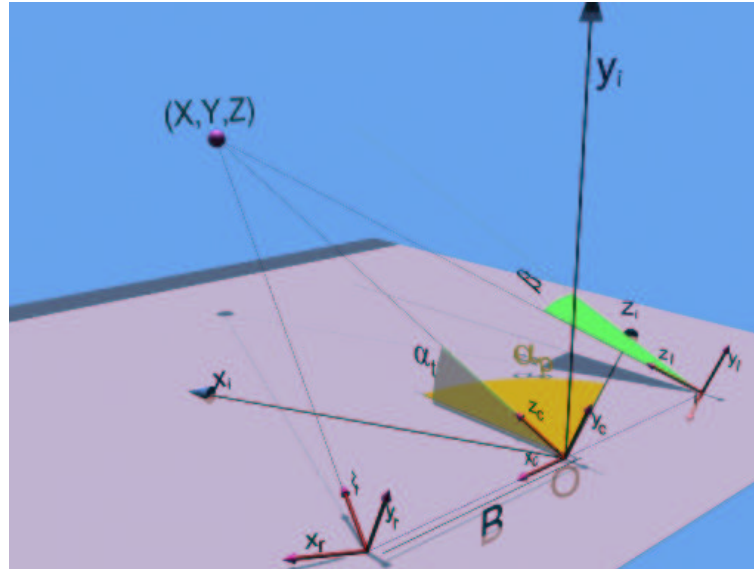


Figura 1.4: Sistema de Seguimento Binocular

de vergência do olho esquerdo e  $\beta_r$  o ângulo de vergência do olho direito. A MDOF tem uma geometria de fixação fronto-simétrica, o que significa que os ângulos de vergência esquerdo e direito tomam valores simétricos ( $\beta_l = -\beta_r = \beta$ ). A Fig. 1.4 esquematiza esta geometria. O seguimento binocular é feito controlando os três graus de liberdade independentes:  $\alpha_p$ ,  $\alpha_t$  e  $\beta$ .

Imagine uma terceira câmara colocada entre os olhos esquerdo e direito. Esta câmara, posicionada no meio da linha interocular, é designada por olho ciclópico. Numa geometria de fixação fronto-simétrica os graus de liberdade do pescoço são utilizados para fazer convergir o eixo óptico do olho ciclópico no eixo de projecção do ponto de seguimento. Desta forma o alvo é projectado na zona foveal da retina ciclópica. O azimute das câmaras, designado por ângulo de vergência, é utilizado para ajustar a posição dos olhos de forma a que o alvo seja projectado no centro de ambas as imagens (esquerda e direita). Considerando que o seguimento do alvo é feito mantendo os ângulos de vergência simétricos,  $\beta$  só depende do movimento do alvo ao longo do eixo óptico do olho ciclópico (eixo  $Z$ ), i.e. só depende da profundidade a que se encontra o objecto no referencial ciclópico. Assim, podemos dividir o problema do seguimento binocular na MDOF em dois sub-problemas: controlo do olho ciclópico e controlo do ângulo de vergência.



### 1.1.1 Seguimento com o Olho Ciclópico

O azimute e a elevação do módulo do pescoço são controlados de forma a alinhar o eixo óptico do olho ciclópico com o alvo. Este seguimento com o olho ciclópico é também designado como seguimento em versão [1].

Considere uma câmara com dois graus de liberdade: azimute e elevação. O objectivo de um algoritmo de seguimento monocular é o de controlar o movimento da câmara de forma a que a projecção do alvo na imagem se mantenha estacionária, normalmente no centro da retina. Isto é conseguido fazendo convergir dinamicamente o eixo óptico da câmara no eixo de projecção do objecto em movimento. Note que a questão do seguimento em versão é similar ao problema de seguimento monocular com a câmara ciclópica. Os graus de liberdade do pescoço são controlados de forma a alinhar o eixo ciclópico do  $Z$  com o alvo móvel. Assim o problema do seguimento em versão resume-se ao seguimento monocular desde que se consiga transferir para a retina ciclópica a informação visual extraída nas câmaras esquerda e direita.

$$\begin{bmatrix} x_c \\ y_c \end{bmatrix} = \begin{bmatrix} \frac{x_l + x_r}{2} \\ \frac{y_l + y_r}{2} \end{bmatrix} \quad (1.1)$$

Considere que  $(x_l, y_l)^t$  e  $(x_r, y_r)^t$  são, respectivamente, a posição de projecção do alvo na imagem esquerda e direita. Imagine uma câmara colocada no meio da linha interocular (olho ciclópico).  $(x_c, y_c)^t$  corresponde à posição de projecção na retina ciclópica. A equação 1.1 permite estimar a posição do alvo na imagem ciclópica conhecendo a posição na imagem esquerda e direita. A relação da equação 1.1 é válida quando o ângulo de vergência é nulo ( $\beta = 0$ ). Para a maior parte das condições de operação do sistema, a distância interocular (22cm) é consideravelmente inferior à distância do alvo à cabeça MDOF (tipicamente superior a 2m). Assim o ângulo de vergência toma valores muito pequenos sendo aceitável usar o resultado da equação 1.1 para estimar a posição de projecção no olho ciclópico [1].

$$\begin{bmatrix} \dot{x}_c \\ \dot{y}_c \end{bmatrix} = \begin{bmatrix} \frac{\dot{x}_l + \dot{x}_r}{2} \\ \frac{\dot{y}_l + \dot{y}_r}{2} \end{bmatrix} \quad (1.2)$$

O sistema MDOF utiliza simultaneamente informação de posição e de velocidade para fazer o seguimento de alvos em movimento. A informação de velocidade permite compensar a latência do processamento visual mel-

horando a estabilidade e suavidade do sistema enquanto a informação de posição compensa os erros devido ao escorregamento do alvo nas retinas [1]. A equação 1.2 permite estimar a velocidade na retina ciclópica  $(\dot{x}_c, \dot{y}_c)^t$  conhecendo a velocidade induzida pelo alvo na imagem esquerda  $(\dot{x}_l, \dot{y}_l)^t$  e na imagem direita  $(\dot{x}_r, \dot{y}_r)^t$ . Mais uma vez esta aproximação é válida se o ângulo de vergência for pequeno ( $\beta \approx 0$ ) e se o alvo estiver a ser projectado aproximadamente no centro de ambas as imagens ( $(x_l, y_l)^t \approx (0, 0)^t$  e  $(x_r, y_r)^t \approx (0, 0)^t$ ).

### 1.1.2 Seguimento em Vergência

No seguimento binocular do sistema MDOF o movimento de azimute dos olhos é designado por movimento de vergência. Chama-se ponto de vergência no espaço ao ponto de intersecção do eixo óptico da câmara esquerda com o eixo óptico da câmara direita. Numa geometria de fixação fronto-simétrica os ângulos de azimute dos olhos são simétricos e o ponto de vergência está sempre sobre o eixo  $Z$  do olho ciclópico. O objectivo do controlo de vergência é fazer coincidir o ponto de vergência com o alvo. Isto acontece sempre que o ângulo de vergência  $\beta$  verifica a equação  $\tan(\beta) = \frac{B}{z}$ , onde  $z$  é a coordenada em  $Z$  do alvo no sistema de coordenadas cartesianas do olho ciclópico (ver Fig. 1.4).

$$\Delta\psi = \arctan\left(\frac{x_l - x_r}{2f}\right) \quad (1.3)$$

O azimute dos olhos é controlado de forma a fazer convergir o ponto de vergência na posição do alvo sobre o eixo óptico do olho ciclópico. Considere  $\Delta\phi$  o erro angular em posição da vergência. Isto significa que corrigindo este erro o sistema fica vergido no alvo ou, matematicamente,  $\tan(\Delta\phi + \beta) = \frac{B}{z}$ . A equação 1.3 permite estimar o erro de posição da vergência em função da disparidade entre a imagem esquerda e direita. A relação 1.3, onde  $f$  é a distância focal das câmaras, verifica-se sempre que o ângulo de vergência é nulo ou sempre que o alvo está sobre o eixo dos  $Z$  da câmara ciclópica ( $(x_c, y_c)^t = (0, 0)^t$ ).

$$\dot{\psi} = \frac{\dot{x}_l - \dot{x}_r}{2f} \quad (1.4)$$

Considere que os graus de liberdade do pescoço estão a ser controlados de tal forma que o seguimento feito pela câmara ciclópica é perfeito. Isto implica que  $(x_c, y_c)^t = (0, 0)^t$  e  $(\dot{x}_c, \dot{y}_c)^t = (0, 0)^t$ . Assuma ainda que o

erro em posição da vergência é nulo ( $\Delta\phi = 0$ ) o que significa que  $x_l = x_r = 0$ . A velocidade angular do azimute dos olhos  $\dot{\phi}$  que mantém o sistema vergido é dada por 1.4. Note que para esta geometria a velocidade angular de seguimento para a vergência é função da componente horizontal da disparidade entre o fluxo óptico das duas retinas [7].

## 1.2 Arquitectura do Sistema de Seguimento da MDOF. Diagrama de Blocos

Com o foi referido anteriormente o seguimento em versão é similar ao seguimento monocular com a câmara ciclópica imaginária. As equações 1.1 e 1.2 permitem transferir a informação visual das imagens esquerda e direita para a retina ciclópica. Assim, o problema de seguimento binocular resume-se ao problema de seguimento monocular mais o seguimento em vergência. Este trabalho concentra-se nestes dois últimos problemas.

O seguimento activo de alvos é um problema de controlo de movimento utilizando visão. Neste caso o sistema a ser controlado é a própria plataforma que suporta as câmaras. Em muitas situações o controlo de movimento usando visão é analisado como um sistema de controlo tradicional. A visão funciona como o sensor que fornece a informação de realimentação na malha de controlo [8] [9]. Um sistema de seguimento activo pode ser modelado como um servomecanismo onde a entrada de referência é a posição/velocidade do alvo no espaço, e a saída é a posição/velocidade dos graus de liberdade da plataforma. Considere o seguimento monocular. De acordo com o paradigma servo-mecânico a entrada de referência é dada pela orientação da recta de projecção do alvo na imagem, enquanto a saída corresponde à orientação do eixo óptico da câmara. O objectivo do algoritmo de seguimento é alinhar no espaço 3D o eixo óptico com o objecto ou, usando a terminologia do controlo, fazer convergir a saída do sistema para a entrada.

Corke em [10] e [11] apresenta um modelo alternativo para um sistema de posicionamento com realimentação visual. Nesse modelo as saídas do sistema estão associadas à informação visual extraída das imagens. O sistema de seguimento visual é interpretado como um regulador cujo objectivo é manter a posição do alvo nas imagens estacionária. O movimento do objecto no espaço 3D é encarado como uma perturbação com dinâmica desconhecida. Se a perturbação afecta a posição e/ou velocidade na imagem o sistema tem

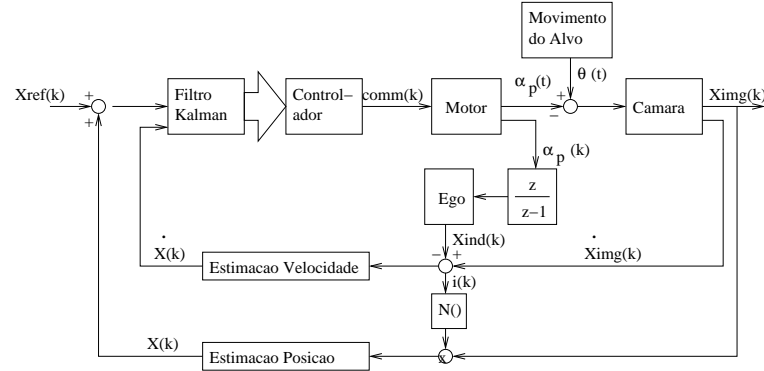


Figura 1.5: Diagrama de blocos de um dos graus de liberdade no sistema de seguimento monocular (azimute).  $\theta$  e  $\alpha_p$  são as posições angulares do alvo e do motor.  $x_{img}(k)$  e  $\dot{x}_{img}(k)$  correspondem à velocidade e posição do objecto na imagem.  $\dot{x}_{ind}(k)$  é a velocidade auto-induzida.  $x(k)$  é a medida de posição e  $\dot{x}(k)$  a medida da velocidade induzida pelo alvo na imagem.  $comm(k)$  é o comando enviado para o motor.

de reagir de forma a compensar esse efeito. O desempenho do sistema é medido pela sua capacidade de rejeição da perturbação.

O paradigma de regulação na modelação do sistema melhora a compreensão do processo de seguimento. O facto de associar as entradas e saídas do sistema com a informação visual relega para um plano secundário os aspectos mecânicos e cinemáticos do sistema. Considere ainda a situação do seguimento monocular. De acordo com o primeiro modelo o objectivo do processo de seguimento é fazer convergir o eixo óptico da câmara (saída do sistema) na recta de projecção do alvo (entrada de referência do sistema). Deste modo a imagem do alvo é mantida no centro da imagem. No caso do modelo de regulação a entrada de referência é a posição desejada do alvo na imagem, que poderá ser o centro ou não. O sistema é controlado de forma a compensar o movimento do objecto (perturbação) e manter a posição de projecção estacionária. O modelo de regulação é absolutamente flexível quanto à posição desejada na imagem. Por estas razões este trabalho usa o paradigma de regulação para modelar o sistema de seguimento MDOF.

### 1.2.1 Seguimento Monocular

Considere uma câmara com dois graus de liberdade rotativos: azimute e elevação. Assuma  $\alpha_p$  como sendo o ângulo de azimute e  $\alpha_t$  como sendo a posição angular da elevação. Ambos os eixos de rotação passam pelo centro óptico da lente de forma à câmara descrever movimentos de rotação puros. A Fig. 1.5 esquematiza a malha de controlo do azimute. Assuma dois graus de liberdade, azimute e elevação, similares e não acoplados. Esta aproximação é válida sempre que a posição do alvo esteja próxima do centro da imagem e o ângulo de elevação  $\alpha_t$  não seja muito elevado.

No regulador da Fig 1.5 o movimento do objecto actua como uma perturbação que tem de ser compensada. O objectivo do algoritmo de seguimento é controlar o movimento da câmara de modo a que a projecção do alvo em movimento se mantenha estacionária na imagem. O controlo é feito usando informação de posição e velocidade. O processamento visual destina-se a determinar a posição do alvo na imagem ( $x$ ) e a velocidade induzida na retina pelo movimento do objecto no espaço ( $\dot{x}$ ).

O movimento na imagem depende do movimento do objecto no espaço e do próprio movimento da câmara. Esta última componente de movimento é designada por movimento auto-induzido ou "egomotion". Para controlar o sistema pretende-se obter uma estimativa da velocidade do objecto que está relacionada com a velocidade que este induz na imagem. Assim a componente devida ao "egomotion" tem de ser compensada. Considere que a câmara tem movimento de rotação puro e que não existe movimento na cena. Duas imagens adquiridas em instantes de tempo sucessivos são relacionadas por uma homografia [12]. A homografia é facilmente determinada conhecendo a rotação. O algoritmo de compensação de "egomotion" usa as leituras dos codificadores dos motores para determinar a homografia e compensar nas imagens o movimento auto-induzido [13] [1].

Considere duas imagens consecutivas após a compensação do movimento auto-induzido. A segmentação do alvo é obtida subtraindo uma imagem da outra. A imagem de diferença contém os pontos da cena onde ocorreu movimento independente. Assuma que todos os pontos do objecto se movem com a mesma velocidade no plano imagem. A velocidade induzida pelo alvo na retina é estimada calculando o fluxo óptico entre duas imagens consecutivas. Note que o fluxo óptico só é determinado nos pontos em que ocorreu movimento independente, ou seja nos pontos que aparecem na imagem de

diferença.

$$\begin{cases} N(i(k)) = 1 \Leftarrow i(k) \neq 0 \\ N(i(k)) = 0 \Leftarrow i(k) = 0 \end{cases} \quad (1.5)$$

A posição do alvo é estimada como a localização média do conjunto de pontos com fluxo óptico não nulo em duas imagens consecutivas. Usando este método o valor estimado é nulo sempre que o objecto está parado. Neste caso não há movimento induzido na imagem sendo o fluxo óptico depois da compensação do "egomotion" zero ( $\dot{x} = 0$ ). A função  $N()$  definida na equação 1.5 modela esta situação.

$$comm(k) = \dot{x}(k) + G.x(k) \quad (1.6)$$

O motor é controlado em velocidade. O comando  $comm(k)$  que é enviado é dado pela equação 1.6. A estimativa de posição  $x(k)$ , multiplicada por um ganho  $G$ , é adicionada ao valor de velocidade  $\dot{x}(k)$ . O controlador inclui um filtro de Kalman com o objectivo de limitar os efeitos dos erros na medição de posição e velocidade na imagem e assegurar um comportamento suave do sistema [14] [15] [16].

### 1.2.2 Seguimento em Vergência

No seguimento binocular o controlo de vergência ajusta o azimute das duas câmaras de forma ao alvo ser projectado no centro de ambas as retinas. 1.3 e 1.4 são as equações de controlo da vergência usando disparidade. A Fig. 1.6 mostra a malha de controlo onde a disparidade horizontal entre as retinas é regulada controlando o ângulo de vergência  $\beta$ . A velocidade horizontal do alvo na retina esquerda e direita são medidas depois de compensar o movimento auto-induzido. Determina-se a diferença entre elas (disparidade). Um filtro de Kalman é utilizado de forma a garantir um controlo suave. O comando em velocidade é obtido de acordo com a equação 1.4.

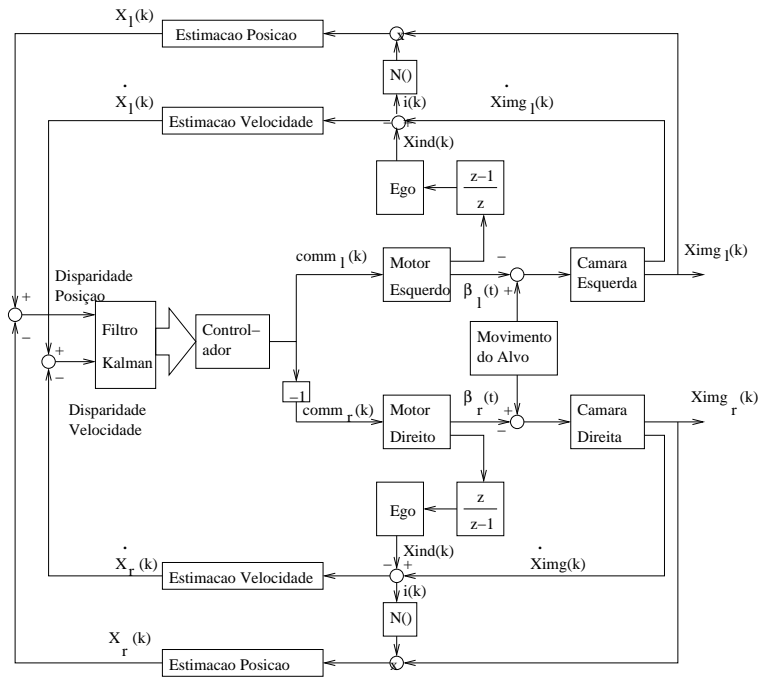


Figura 1.6: Diagrama de blocos do controlo de vergência simétrica.  $\beta_L$  e  $\beta_R$  são as posições angulares do olho esquerdo e direito. O processamento visual é idêntico ao do seguimento monocular (ver Fig. 1.5). A velocidade e posição do alvo é estimada em ambas as imagens. A disparidade em posição e velocidade é utilizada para controlar os motores. Note que o comando enviado para os olhos é simétrico de forma aos ângulos de vergência serem simétricos.





## Capítulo 2

# Trajectórias de Referência do Alvo e Avaliação do Desempenho do Sistema

A robustez e o desempenho de um sistema de controlo de movimento usando visão depende não só dos algoritmos de processamento visual das imagens, mas também de aspectos relacionados com a teoria do controlo e a arquitectura do sistema [17] [18] [?]. O estudo da dinâmica do sistema é fundamental para se obter comportamentos robustos de alto desempenho. Assuntos como estabilidade, resposta transitória e atraso não podem ser negligenciados [10] [19] [20] [21].

Hill e Park nos fins dos anos 70 foram dos primeiros a abordar problemas de instabilidade em sistemas em malha fechada com visão [22]. Esta instabilidade foi atribuída à latência em malha aberta devida ao atraso do processamento visual. Em [10], Corke refere que é frequente o relato de sistemas de visão ou com um comportamento ligeiramente oscilante, ou com uma resposta lenta e com muito atraso. A primeira situação revela um sistema em malha-fechada próximo do limite de estabilidade enquanto a segunda é muitas vezes motivada por um mau desenho e/ou sintonização do controlador. Isto demonstra a importância dos aspectos dinâmicos na optimização do desempenho de sistemas de controlo que usam visão.

Para implementar comportamentos de seguimento robustos e com elevado desempenho tem de se ter em atenção o processamento visual, a estratégia de controlo e a arquitectura do próprio sistema. Ferramentas para caracterizar o sistema de seguimento são fundamentais para a optimização do seu

desempenho e aumento da robustez. Estas ferramentas devem avaliar simultaneamente os aspectos visuais e de controlo. Só assim se consegue ter uma perspectiva global duma determinada aproximação ao problema. Por exemplo, considere o seguimento visual num ambiente variável e com elevado grau de incerteza, de tal forma que se torna difícil ou impossível a modelação. Operar em tal ambiente exige algoritmos de visão mais complexos e/ou algoritmos de controlo mais robustos. Decisões e escolhas adequadas quanto à arquitectura e algoritmos só podem ser feitas se todos os aspectos do sistema puderem ser caracterizados conjuntamente de forma a otimizar o comportamento global do sistema. A teoria do controlo tem um conjunto de ferramentas que permitem uma caracterização global do desempenho de sistemas controlados com visão [23]. Este capítulo começa por estabelecer uma série de métricas e conceitos para fazer a avaliação do desempenho dum sistema visual [24] [25].

A parte final do capítulo é dedicada à caracterização do sistema de seguimento MDOF. Vários aspectos são discutidos e estudados com vista a uma optimização do desempenho [24] [26].

## **2.1 Trajectórias de Referência**

Muitos trabalhos abordam os sistemas de posicionamento controlados por visão como servo-sistemas tradicionais que utilizam um sensor visual na malha de realimentação [8, 9]. De acordo com esta perspectiva o sistema de seguimento binocular MDOF pode ser estudado como um servo-mecanismo. As entradas de referência são relacionadas com os vectores de posição e velocidade do alvo no espaço, e as saídas com a posição e velocidade das juntas da plataforma. O objectivo do seguimento monocular é alinhar o eixo óptico com o raio de projecção do alvo. De acordo com uma perspectiva servo-mecânica a orientação do raio projectivo é a entrada de referência, e a orientação do eixo óptico a saída do sistema. No caso particular do seguimento binocular o objectivo é fazer coincidir a ponto de vergência com a posição do alvo no espaço. Neste caso a entrada de referência e a saída do sistema são, respectivamente, as coordenadas 3D do alvo e do ponto de vergência.

O sistema MDOF pretende seguir um objecto que se movimenta no espaço com uma dinâmica completamente desconhecida. Considere o caso particular do seguimento monocular. Sem restrições adicionais é teoricamente impossível determinar o movimento 3D do alvo a partir da informação visual.

O processo de seguimento só pretende manter a posição de projecção do alvo na imagem estacionária, o que não implica a reconstrução do movimento do objecto no espaço. O objectivo do seguimento visual está claramente ao nível do sinal de imagem, fazendo pouco sentido assumir como entrada/saída do sistema os parâmetros de movimento 3D.

Considere ainda a entrada do sistema como a posição do alvo no espaço. O processo de formação de imagem é descrito por um modelo de projecção central que estabelece uma relação não-linear entre o movimento no espaço e na imagem. Esta não-linearidade aparece explicitamente na malha de realimentação do servo-mecanismo. Imagine que pretende identificar o sistema e avaliar o seu desempenho. No caso do sistema ser linear a caracterização pode ser feita aplicando sinais de teste sendo os mais frequentes o degrau, a rampa, a parábola e a sinusóide. Assumindo que a referência do sistema é a posição do alvo no espaço, temos uma entrada em rampa sempre que o objecto tiver movimento rectilíneo uniforme. No entanto, devido à não-linearidade introduzida pela transformação projectiva, isto não significa que o alvo se movimenta com velocidade constante na imagem. Pode até acontecer, no caso da velocidade 3D ter a direcção do eixo de projecção, que a posição de projecção do objecto na imagem se mantenha constante. De um modo geral, um movimento uniforme no espaço 3D não define uma rampa na imagem e, conseqüentemente, não constitui um estímulo adequado para fazer a identificação dos algoritmos de processamento visual e a avaliação do desempenho global do sistema.

O estudo do sistema de seguimento como um servo-mecanismo é uma opção discutível. Por um lado os objectivos do seguimento visual estão ao nível do plano imagem. O movimento da plataforma de forma ao eixo óptico convergir para o raio de projecção do alvo, no caso do seguimento monocular, e do ponto de vergência convergir para a posição do objecto, no caso do seguimento binocular, são um meio para atingir um fim. O objectivo não é recuperar os parâmetros de movimento do alvo mas sim manter a sua posição na imagem estacionária. Por outro lado a relação entre o movimento no espaço e na imagem é não-linear. Os sinais de teste para identificação de sistemas lineares não são um estímulo adequado para caracterizar o desempenho do sistema de seguimento se a entrada de referência for definida no espaço 3D .

Como já foi referido no capítulo anterior este trabalho assume uma paradigma de regulação para modelar o seguimento monocular (Fig. 1.5) e o controlo

de vergência (Fig. 1.6). Em ambos os casos a entrada de referência do sistema é a posição desejada de projecção do alvo na imagem (normalmente o centro). A saída é dada pela posição de projecção nas imagens adquiridas e o movimento do alvo no espaço actua como uma perturbação. O sistema é controlado de forma a compensar essa perturbação e manter a projecção do objecto na imagem estacionária.

Os sinais de teste para fins de identificação e caracterização do sistema devem ser definidos no plano imagem e não no espaço 3D. Por um lado é a opção mais coerente com o paradigma de regulação que considera as entradas e saídas grandezas medidas na imagem. Por outro permite lidar com as não-linearidades introduzidas pela transformação projectiva. Um movimento rectilíneo uniforme no espaço 3D não corresponde a uma perturbação em rampa no plano imagem. No caso extremo da velocidade estar alinhada com o eixo de projecção o sistema nem sequer reage pois não é induzido movimento na retina. O estímulo pretendido é obtido quando o alvo descreve uma trajectória que projecta um movimento de velocidade constante na imagem. Só desta forma se obtém um sinal em rampa que permite caracterizar os algoritmos de processamento visual e avaliar o desempenho do sistema. O que foi dito para a rampa é também aplicável a uma perturbação em degrau, parábola e sinusóide.

O estabelecimentos dos sinais de teste para identificação e avaliação do desempenho de sistemas controlados por visão levanta dois problemas fundamentais:

- Gerar com rigor os sinais de perturbação do sistema
- Determinar trajectórias de movimento do alvo que definam sinais de perturbação em degrau, em rampa, em parábola e em sinusóide no plano imagem.

### **2.1.1 Gerar com Rigor os Sinais de Perturbação**

Para caracterizar a capacidade do sistema em compensar a perturbação induzida pelo movimento do alvo são utilizados sinais de teste específicos. Em vez de usar alvos reais para perturbar o sistema optou-se por utilizar um alvo virtual. A projecção do objecto é gerada na imagem em cada instante de aquisição sendo função da posição do alvo e da geometria do sistema naquele momento. A vantagem desta aproximação é que torna possível gerar com

rigor trajectórias de movimento do alvo no espaço 3D que correspondem a funções matemáticas pré-estabelecidas. O conhecimento da perturbação em relação à qual se está a medir o desempenho do sistema é fundamental para uma caracterização rigorosa.

As imagens obtidas são usadas como entradas de perturbação do sistema de visão activo. Dado um movimento do objecto pré-definido, as imagens capturadas dependem, não só da posição do alvo no espaço, mas também da orientação da câmara. Durante a operação de seguimento a geometria do sistema varia ao longo do tempo. Isto obriga a que as imagens sejam geradas em cada instante de aquisição de forma a entrar em linha de conta com a geometria específica do sistema naquele momento. Deste modo é necessário conhecer a posição do alvo no espaço e a orientação da câmara num mesmo referencial inercial. A posição do objecto é calculada usando um modelo de movimento específico definido por uma expressão matemática adequada. A orientação da câmara no instante de aquisição é obtida usando as leituras dos codificadores ópticos dos motores e a cinemática inversa da plataforma de suporte. Por razões de simplicidade a origem do sistema de coordenadas inercial é colocada no centro óptico para o caso do seguimento monocular, e na origem do referencial ciclópico para o caso do seguimento binocular e controlo de vergência.

$$\begin{cases} x_{img} = f \frac{X_c}{Z_c} \\ y_{img} = f \frac{Y_c}{Z_c} \end{cases} \quad (2.1)$$

Em resumo, o movimento desejado do objecto no espaço é descrito rigorosamente por equações matemáticas. As coordenadas de movimento do alvo são convertidas em coordenadas cartesianas inerciais através da transformação adequada. As coordenadas de posição do corpo em movimento são convertidas para coordenadas da câmara. Esta transformação depende da geometria da plataforma que é obtida lendo os codificadores dos motores. Conhecendo a orientação da câmara relativamente ao alvo, a imagem é gerada assumindo projecção em perspectiva. A equação 2.1 calcula a posição do alvo na imagem  $(x_{img}, y_{img})$ , em que  $(X_c, Y_c, Z_c)$  corresponde às coordenadas do objecto no referencial da câmara e  $f$  à distância focal. Estes cálculos são feitos para cada instante de aquisição de imagem (de 40ms em 40ms).

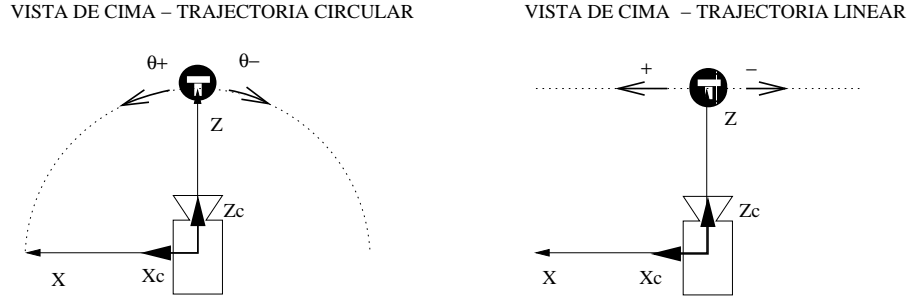


Figura 2.1: Esquerda: Posição do alvo em coordenadas esféricas  $(\rho, \theta, \phi)$ . Trajectória circular de movimento ( $\theta$  varia,  $\phi = 0$  e  $\rho$  é constante). Direita: Posição do alvo em coordenadas cartesianas. Trajectória rectilínea de movimento ( $X$  varia e  $Y = Z = 0$ )

### 2.1.2 Equações das Trajectórias de Referência. Seguimento Monocular

Na secção anterior foi apresentada a técnica para gerar perturbações no sistema de forma a caracterizar o seu desempenho. Os sinais de teste utilizados em teoria do controlo linear são o degrau, a rampa, a parábola e a sinusóide. Já foi discutido que estes sinais devem ser definidos no plano imagem. Esta secção deduz as equações de movimento do alvo no espaço, que induzem na imagem os sinais desejados. Estes sinais de teste permitem avaliar simultaneamente os aspectos de controlo e de visão num sistema de seguimento visual.

#### Sinais de Teste Definidos em Função dos Actuadores

$$\begin{cases} X = \rho \cos(\phi) \sin(\theta) \\ Y = \rho \sin(\phi) \\ Z = \rho \cos(\phi) \cos(\theta) \end{cases} \quad (2.2)$$

Assuma que a posição do alvo no referencial inercial é dada em coordenadas esféricas  $(\rho, \theta, \phi)$  em que  $\rho$  é o raio ou profundidade,  $\theta$  é o ângulo de azimute e  $\phi$  é o ângulo de elevação. A equação 2.2 mostra a relação entre as coordenadas esféricas e cartesianas. Considere ainda o actuador como sendo um motor rotativo de tal forma que a câmara descreve um movimento de rotação puro em torno do eixo do  $Y$  (azimute).

$$\theta(t) = \begin{cases} Const \Leftarrow t > 0 \\ 0 \Leftarrow t = 0 \end{cases} \quad (2.3)$$

$$\theta(t) = \omega.t \quad (2.4)$$

$$\theta(t) = \frac{\gamma}{2}.t^2 \quad (2.5)$$

$$\theta(t) = A \sin\left(\frac{2\pi}{T}.t\right) \quad (2.6)$$

Considere a perturbação na perspectiva do actuador/motor. A posição do alvo é descrita em coordenadas esféricas com  $\rho$  constante,  $\phi$  nulo e  $\theta$  a variar no tempo (ver trajectória da Fig. 2.1). As equações de movimento 2.3, 2.4, 2.5 e 2.6 definem, respectivamente, uma perturbação em degrau, rampa, parábola e sinusóide. Considere, por exemplo, que o movimento do alvo é descrito pela equação 2.4. O motor tem que rodar com uma velocidade angular constante  $\omega$  de forma a seguir o alvo.

$$X(t) = \begin{cases} Const \Leftarrow t > 0 \\ 0 \Leftarrow t = 0 \end{cases} \quad (2.7)$$

$$X(t) = v.t \quad (2.8)$$

$$X(t) = \frac{a}{2}.t^2 \quad (2.9)$$

$$X(t) = A \sin\left(\frac{2\pi}{T}.t\right) \quad (2.10)$$

Assuma a posição do alvo dada em coordenadas cartesianas  $(X, Y, Z)$  e que o actuador movimenta a câmara linearmente ao longo do eixo dos  $X$ . As equações 2.7 a 2.10 definem os sinais de teste (degrau, rampa, parábola e sinusóide) para um actuador linear. Note que a coordenada em  $Z$  é constante.

### Sinais de Teste Definidos em Função da Imagem. Câmara Estática

A Fig. 1.5 mostra o diagrama de blocos do sistema de seguimento monocular. De forma a relacionar as saídas com a informação medida nas imagens, os sinais de teste que perturbam o sistema devem ser gerados no plano imagem. Assim um degrau em posição corresponde a uma mudança abrupta na posição de projecção do alvo na imagem. Uma rampa/parábola ocorre quando o movimento 3D do objecto induz um movimento com velocidade/aceleração constante no plano imagem. Uma perturbação sinusoidal é gerada sempre

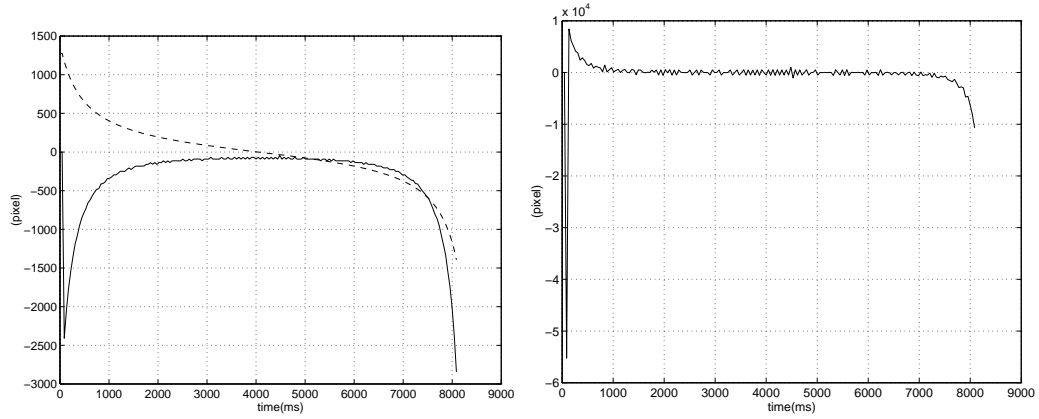


Figura 2.2: Projecção de movimento numa câmara estática quando o alvo percorre a trajectória circular da Fig. 2.1 com velocidade angular constante  $\omega = 5\text{rad/s}$ . Esquerda: Posição do alvo na imagem (linha tracejada [- -]), velocidade do alvo na imagem em  $\text{pixel/imagem}$  (linha a cheio [-]). Direita: Aceleração do alvo na imagem em  $\text{pixel/imagem}^2$  (linha a cheio [-])

que a posição e velocidade do alvo na imagem são descritas por funções sinusoidais com diferença de fase de  $90^\circ$ .

$$x_{img} = f \cdot \tan(\theta) \quad (2.11)$$

$$\frac{dx_{img}}{dt} = f \cdot \frac{d\theta}{dt} \cdot \frac{1}{\cos^2(\theta)} \quad (2.12)$$

$$\frac{d^2x_{img}}{dt^2} = f \cdot \frac{d^2\theta}{dt^2} \cdot \frac{1}{\cos^2(\theta)} + 2 \cdot f \cdot \left(\frac{d\theta}{dt}\right)^2 \cdot \frac{\tan(\theta)}{\cos^2(\theta)} \quad (2.13)$$

Assuma a câmara estática de tal forma que o referencial da câmara coincide com o referencial inercial. O movimento do alvo é descrito em coordenadas esféricas. As equações 2.11 a 2.13 são obtidas a partir da expressão de projecção 2.1 onde  $(X_c, Y_c, Z_c)$  é definido em função de  $(\rho, \theta, \phi)$  de acordo com 2.2. A equação 2.11 relaciona a posição angular do alvo no espaço  $\theta$  com a posição horizontal de projecção na imagem  $x_{img}$ . As equações 2.12 e 2.13 estabelecem, respectivamente, a relação de velocidade e aceleração.

Considere o movimento do objecto no espaço em que  $\rho$  é constante,  $\phi = 0$  e  $\theta$  varia no tempo segundo 2.4 com  $\omega = 5\text{rad/s}$ . A Fig. 2.2 mostra a posição, velocidade e aceleração do alvo na imagem. Note que, apesar da aceleração angular ser nula ( $\gamma = 0\text{rad/s}^2$ ), existe uma aceleração residual na imagem



correspondente ao segundo termo da equação 2.13. No caso da câmara ser estática, um movimento do objecto com velocidade angular constante não induz uma velocidade constante no plano imagem. Isto é explicado pela distorção provocada pela projecção em perspectiva. A distorção na projecção do movimento é tanto maior quanto maior for o  $\theta$ . Conclui-se assim que as equações 2.4 a 2.6 não geram as perturbações desejadas na imagem quando a câmara está parada.

$$x_{img} = f \cdot \frac{X}{Z} \quad (2.14)$$

$$\frac{dx_{img}}{dt} = \frac{f}{Z} \cdot \frac{dX}{dt} - \frac{f \cdot X}{Z^2} \cdot \frac{dZ}{dt} \quad (2.15)$$

$$\frac{d^2x_{img}}{dt^2} = \frac{f}{Z} \cdot \frac{d^2X}{dt^2} - \frac{f \cdot X}{Z^2} \cdot \frac{d^2Z}{dt^2} - \frac{2 \cdot f}{Z^2} \cdot \frac{dX}{dt} \cdot \frac{dZ}{dt} - \frac{2 \cdot f \cdot X}{Z^3} \cdot \left(\frac{dZ}{dt}\right)^2 \quad (2.16)$$

O movimento do objecto é agora descrito em coordenadas cartesianas. As equações 2.14, 2.15 e 2.16, deduzidas a partir de 2.1, determinam a posição, velocidade e aceleração horizontal na imagem. Assuma  $Z$  constante e  $X$  a variar de acordo com 2.8. O objecto move-se com velocidade constante ao longo duma trajectória rectilínea paralela ao plano imagem. Este movimento do alvo no espaço 3D induz uma velocidade constante na imagem (ver equação 2.15). O movimento linear descrito por 2.8 origina uma perturbação em rampa na imagem. O mesmo se pode concluir para os movimento descritos por 2.9 e 2.10 e as perturbações em parábola e sinusóide.

### Sinais de Teste Definidos em Função da Imagem. Câmara Móvel

O sistema binocular MDOF utiliza actuadores/motores rotativos. Na situação monocular (seguimento com a elevação e azimute de um dos olhos) a câmara move-se ao longo de uma trajectória circular. Observe a Fig. 2.3. A rotação em azimute é uma rotação de ângulo  $\alpha_p$  em torno do eixo do  $Y$ . O alvo a move-se numa trajectória circular de raio constante  $\rho$  em que a posição angular é dada por  $\theta$ . Assuma um modelo de projecção em perspectiva para a formação da imagem 2.1.

$$x_{img} = f \cdot \tan(\theta - \alpha_p) \quad (2.17)$$

$$\frac{dx_{img}}{dt} = f \cdot \frac{d\theta}{dt} \cdot \frac{1}{\cos^2(\theta - \alpha_p)} - f \cdot \frac{d\alpha_p}{dt} \cdot \frac{1}{\cos^2(\theta - \alpha_p)} \quad (2.18)$$

VISTA DE CIMA – TRAJECTORIA CIRCULAR

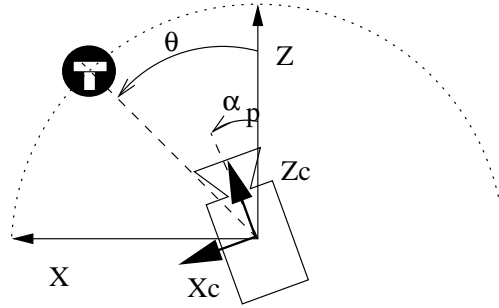


Figura 2.3: Alvo a movimentar-se ao longo de uma trajectória plana circular, seguido pela câmara com movimento rotativo.  $\theta$  é a posição angular do alvo e  $\alpha_p$  a posição angular da câmara.

$$\frac{d^2 x_{img}}{dt^2} = f \cdot \left( \frac{d^2 \theta}{dt^2} - \frac{d^2 \alpha_p}{dt^2} \right) \cdot \frac{1}{\cos^2(\theta - \alpha_p)} + 2f \cdot \left( \frac{d\theta}{dt} - \frac{d\alpha_p}{dt} \right)^2 \cdot \frac{\tan(\theta - \alpha_p)}{\cos^2(\theta - \alpha_p)} \quad (2.19)$$

As equações 2.17 a 2.19 são deduzidas das equações 2.11 a 2.13 assumindo agora a câmara móvel. A posição do alvo na imagem  $x_{img}$  depende simultaneamente da posição angular da câmara  $\alpha_p$  e da posição angular do alvo  $\theta$  ( 2.17). A velocidade na imagem é dada por 2.18 obtida derivando 2.17 em ordem ao tempo. Note que  $\alpha_p$  e  $\theta$  são ambas funções do tempo. Derivando 2.18 tem-se 2.19 que calcula o valor da aceleração na imagem.

Estas equações mostram que o movimento na imagem é causado pelo movimento do alvo no espaço e o movimento da câmara (movimento auto-induzido ou "egomotion"). Numa situação de seguimento perfeito o primeiro é compensado pelo segundo de forma ao movimento na imagem ser nulo. Caso contrário a projecção do alvo desloca-se no plano imagem como resultado dos erros de seguimento. O objectivo do algoritmo de seguimento activo é controlar o movimento da câmara de forma ao "egomotion" compensar o movimento induzido pelo alvo na imagem, que actua como uma perturbação do sistema.

$$\omega_{img} = f \cdot \frac{d\theta}{dt} \cdot \frac{1}{\cos^2(\theta - \alpha_p)} \quad (2.20)$$

$$\gamma_{img.t} = f \cdot \frac{d\theta}{dt} \cdot \frac{1}{\cos^2(\theta - \alpha_p)} \quad (2.21)$$

$$A_{img} \frac{2\pi}{T} \cos\left(\frac{2\pi}{T} \cdot t\right) = f \cdot \frac{d\theta}{dt} \cdot \frac{1}{\cos^2(\theta - \alpha_p)} \quad (2.22)$$

Considere os dois termos do membro direito da equação de velocidade 2.18. O primeiro termo corresponde à velocidade induzida pelo alvo na imagem, enquanto o segundo termo calcula a velocidade auto-induzida pela rotação da câmara. Imagine que se pretende gerar uma perturbação em rampa. Para isso tem de se determinar a trajectória do alvo  $\theta(t)$  que induz na imagem um movimento de velocidade constante  $\omega_{img}$ . A trajectória pretendida é calculada resolvendo a equação diferencial 2.20. Esta equação é deduzida fazendo o primeiro termo de 2.18, que corresponde à componente de velocidade na imagem induzida pelo alvo, igual à velocidade constante  $\omega_{img}$ . As trajectórias de referência que originam uma perturbação em parábola e em sinusóide são calculadas resolvendo as equações 2.21 e 2.22 deduzidas a partir de 2.20. Note que em qualquer um dos casos a trajectória de referência  $\theta(t)$  depende da reacção do sistema à perturbação  $\alpha_p(t)$ . Isto é explicado pelo facto da imagem adquirida ser função simultânea do movimento do alvo no espaço e do movimento da câmara. Assim, e regressando ao exemplo da rampa, para induzir uma velocidade constante na imagem  $\omega_{img}$ , a velocidade angular do alvo no espaço deve ser calculada em cada instante de aquisição como função do erro de seguimento  $\theta - \alpha_p$ .

Assuma agora que o sistema está a fazer seguimento perfeito do objecto. Neste caso o erro de seguimento é nulo e  $\theta(t) - \alpha_p(t) = 0$ . As soluções das equações diferenciais 2.20 a 2.22 são, respectivamente, as funções 2.4 a 2.6, onde  $\omega = \omega_{img}/f$ ,  $\gamma = \gamma_{img}/f$  e  $A = A_{img}/f$ . Estas são as trajectórias de referência que vão ser utilizadas para a caracterização do sistema. Voltando ao exemplo da perturbação em rampa, é verdade que a trajectória descrita por 2.4 só induz uma velocidade constante no plano imagem se o erro de seguimento for nulo. No entanto, e atendendo que durante a operação de seguimento  $\alpha_p$  converge para  $\theta$ , as variações na velocidade induzida devido aos erros de seguimento são tendencialmente pequenas. Por outro lado a perturbação gerada é independente da reacção do sistema permitindo avaliar a capacidade dos algoritmos em recuperar os erros de seguimento.

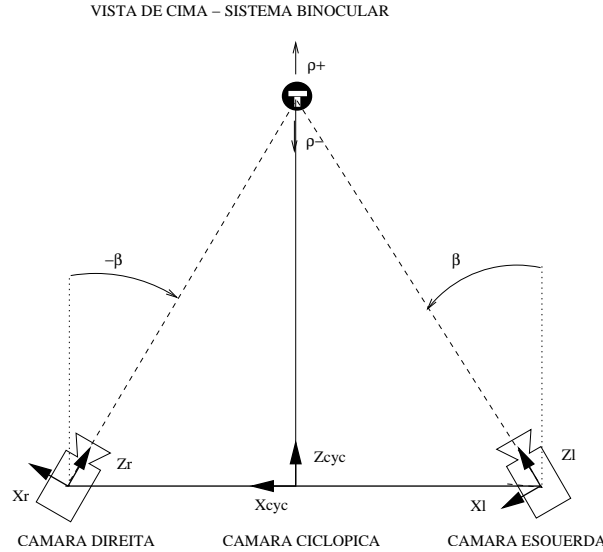


Figura 2.4: Sistema binocular visto de cima. A distância entre as câmaras é  $2B$  e a vergência é simétrica ( $\beta_L = -\beta_R = \beta$ ). O alvo move-se ao longo do eixo ciclópico sendo  $\rho(t)$  a respectiva função de posição

### 2.1.3 Equações das Trajectórias de Referência. Controlo de Vergência

Esta secção estuda as trajectórias de movimento do alvo no espaço que geram nas imagens os sinais de teste para avaliar e caracterizar o algoritmo de controlo de vergência. A Fig. 2.4 mostra que distância entre as câmaras ("baseline") é  $2B$  e que a vergência é simétrica ( $\beta_l = -\beta_r = \beta$ ). Assuma que o azimute e a elevação do pescoço asseguram o alinhamento do eixo óptico do olho ciclópico com o alvo. Assim a coordenada em  $Z$  da posição do objecto no referencial ciclópico é dada por  $\rho$ . A vergência tem como objectivo controlar o azimute dos olhos de forma a projectar o alvo no centro das retinas. Para isso deve fazer convergir o eixo óptico de cada olho no raio projecção do objecto na imagem, ou, por outras palavras, fazer coincidir o ponto de vergência com a posição do alvo no espaço. Isto acontece sempre que  $\beta = \arctan(\frac{B}{\rho})$ .

Já foi referido que o controlo de vergência utiliza a disparidade entre as retinas. A diferença de posição e velocidade do alvo entre a imagem esquerda e direita são os estímulos do sistema. Desta forma os sinais de teste têm de estar relacionados com a disparidade nas retinas. Considere uma perturbação

em rampa/parábola. Esta perturbação ocorre sempre que o movimento do alvo no espaço induz uma diferença constante de velocidade/aceleração entre a imagem esquerda e direita. Nesta perspectiva a perturbação em degrau corresponde a uma variação súbita da disparidade em posição. Quando a disparidade retinal varia de forma sinusoidal temos uma perturbação em sinusóide.

$$\Delta x_{img} = 2f \cdot \frac{-\rho \cdot \sin(\beta) + B \cdot \cos(\beta)}{\rho \cdot \cos(\beta) + B \cdot \sin(\beta)} \quad (2.23)$$

Assuma as condições da Fig. 2.4. A equação 2.23 permite calcular a disparidade de posição  $\Delta x_{img} = x_l - x_r$ , conhecendo a posição do alvo sobre o eixo ciclópico  $\rho$  e o ângulo de vergência  $\beta$ . A sua dedução é trivial e por isso omitida. Basta determinar as coordenadas do alvo no referencial de cada uma das câmaras e calcular a posição na imagem assumindo o modelo de projecção em perspectiva (ver equação 2.1). Note que no caso de  $\tan(\beta) = \frac{B}{\rho}$  a disparidade  $\Delta x_{img}$  é nula e o seguimento é perfeito.

$$\frac{d\Delta x_{img}}{dt} = -\frac{2fB}{(\rho \cos(\beta) + B \sin(\beta))^2} \dot{\rho} - \frac{2f(\rho^2 + B^2)}{(\rho \cos(\beta) + B \sin(\beta))^2} \dot{\beta} \quad (2.24)$$

Derivando 2.23 em ordem ao tempo obtém-se a equação 2.24 onde a disparidade entre as velocidades das duas imagens é expressa em função da velocidade do alvo ao longo do eixo óptico  $\dot{\rho}$  e a da velocidade das câmaras  $\dot{\beta}$ . O primeiro termo de 2.24 refere-se à velocidade induzida na imagem pelo movimento do alvo, o segundo termo corresponde à componente auto-induzida ou "egomotion". O objectivo do controlo de vergência é movimentar as câmaras de forma a cancelar o movimento induzido e manter a posição e velocidade da projecção estacionária.

$$\frac{d\Delta x_{img}}{dt} = -\frac{2fB}{(\rho^2 + B^2)} \dot{\rho} - 2f \dot{\beta} \quad (2.25)$$

Assumindo que o ponto de vergência coincide com o alvo vem que  $\beta = \arctan(\frac{B}{\rho})$ . Substituindo no denominador de 2.24 e simplificando obtém-se a expressão 2.25. Note que a equação 1.4 pode ser deduzida a partir de 2.25. Considere  $\dot{x}_l$  e  $\dot{x}_r$  a velocidade na imagem esquerda e direita medidas após o cancelamento do "egomotion".  $\dot{x}_l - \dot{x}_r$  corresponde à disparidade da velocidade induzida pelo movimento do alvo (primeiro termo de 2.25).

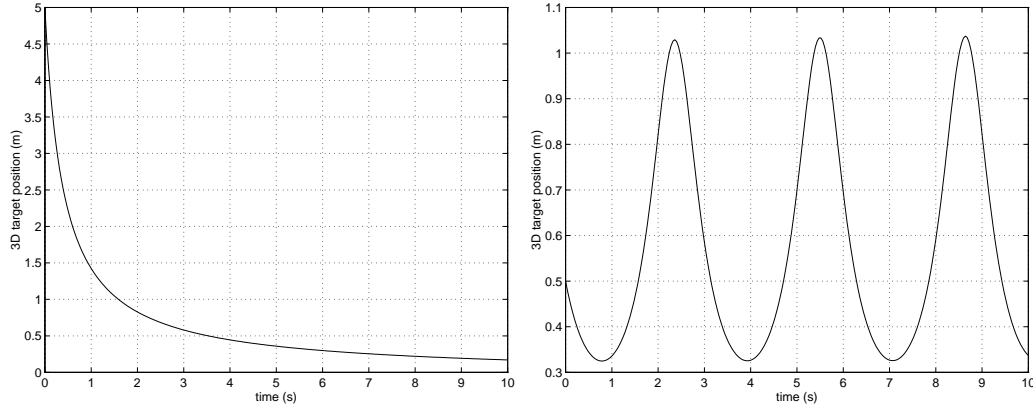


Figura 2.5: Posição do alvo sobre o eixo ciclópico ao longo do tempo (função  $\rho(t)$ ). Esquerda: Perturbação em rampa. Movimento do alvo para gerar uma disparidade em velocidade constante de  $1\text{pixel/imagem}$  ( $\rho(0) = 5(\text{m})$ ). Direita: Perturbação em sinusóide. Movimento do alvo para gerar uma variação sinusoidal de disparidade entre imagens ( $A = 2(\text{pixel})$ ,  $\omega = 2(\text{rad/s})$  e  $\rho(0) = 1(\text{m})$ )

Pretende-se obter o valor de  $\dot{\beta}$  que garanta  $\frac{d\Delta x_{img}}{dt} = 0$ . Substituindo em 2.25 e resolvendo em ordem a  $\dot{\beta}$  obtém o resultado 1.4 onde  $\psi$  é a velocidade angular da vergência para o seguimento perfeito.

$$2fB \cdot \frac{d\rho}{dt} + v_{img} \cdot \rho^2 = -v_{img} \cdot B^2 \tag{2.26}$$

$$\frac{2fB}{\rho^2 + B^2} \cdot \frac{d^2\rho}{dt^2} - \rho \cdot \frac{4fB}{(\rho^2 + B^2)^2} \cdot \left(\frac{d\rho}{dt}\right)^2 = -a_{img} \tag{2.27}$$

$$2fB \cdot \frac{d\rho}{dt} + A_{img} \frac{2\pi}{T} \cos\left(\frac{2\pi}{T}t\right) \cdot \rho^2 = -A_{img} \frac{2\pi}{T} \cos\left(\frac{2\pi}{T}t\right) \cdot B^2 \tag{2.28}$$

Pretende-se uma perturbação do sistema em rampa. Para isso é necessário determinar o movimento do alvo ao longo do eixo ciclópico  $\rho(t)$  que induz uma disparidade constante da velocidade.  $\rho(t)$  é calculado resolvendo a equação diferencial 2.26, onde  $v_{img}$  é a disparidade de velocidade constante entre as imagens. A expressão 2.26 é deduzida a partir do primeiro termo de 2.25. Para uma perturbação em parábola resolve-se 2.27, em que  $a_{img}$  é a disparidade de aceleração nas imagens. A equação 2.28 permite determinar a trajectória de referência correspondente a uma perturbação sinusoidal.

A Fig. 2.5 mostra as trajectórias de referência  $\rho(t)$  que geram uma per-

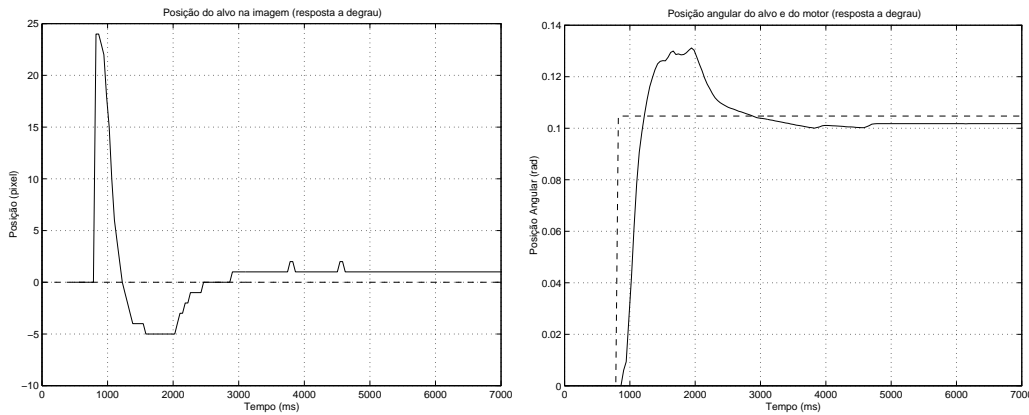


Figura 2.6: Resposta a degrau. Esquerda: Regulação na imagem. Posição de projecção do alvo na imagem (linha a cheio [-]) e centro da imagem (linha tracejada [- -]). Direita: Resposta servo-mecânica. Posição angular do alvo  $\theta$  (linha tracejada [- -]) e do motor  $\alpha_p$  (linha a cheio [-])

turbação em rampa e em sinusóide. A perturbação em rampa foi calculada resolvendo numericamente a equação diferencial 2.26 com  $v_{img} = 25\text{pixel/s}$  e  $\rho(0) = 5m$ . Repare que para o alvo induzir nas imagens uma disparidade em velocidade constante, a sua velocidade 3D ao longo do eixo óptico aumenta com a profundidade a que ele se encontra. A trajectória da perturbação sinusoidal foi obtida resolvendo numericamente a equação diferencial 2.28.

## 2.2 Avaliação do Desempenho do Sistema

Esta secção avalia a capacidade do sistema em compensar as perturbações induzidas pelo movimento do alvo. São estudadas as respostas a degrau, rampa, parábola e sinusóide do seguimento monocular e da vergência.

### 2.2.1 Seguimento Monocular

O sistema de seguimento monocular tem dois graus de liberdade: azimute e elevação. O estudo apresentado refere-se ao grau de liberdade do azimute.

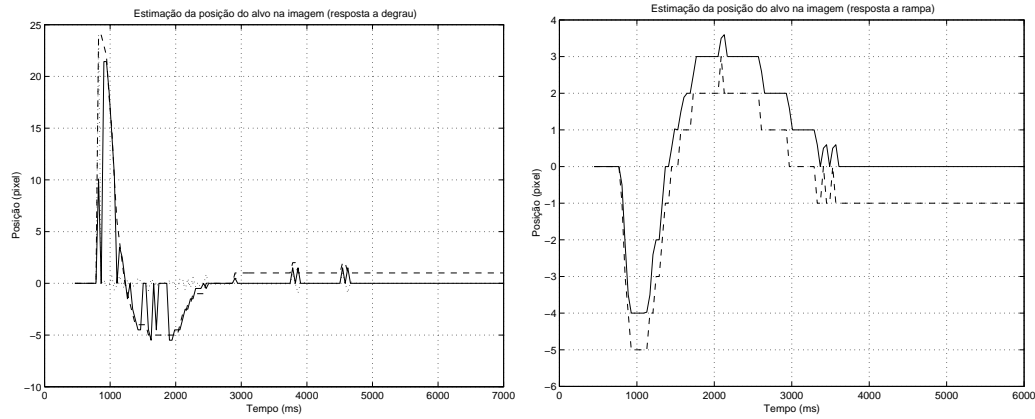


Figura 2.7: Estimação da posição do alvo na imagem. Esquerda: Resposta a degrau. Posição estimada (linha a cheio [-]), posição real (linha tracejada [- -]) e velocidade residual na imagem devido a mau cancelamento (linha pontuada [...]). Direita: Resposta a rampa. Posição estimada (linha a cheio [-]) e posição real (linha tracejada [- -])

### Resposta a Degrau

O sistema é perturbado com uma entrada em degrau o que significa que o alvo muda subitamente de posição. A Fig. 2.6(E) mostra a posição de projecção do alvo na imagem ao longo do tempo. O sistema reage à perturbação com um "overshoot" de cerca de 20% e um erro em regime final de 1 "pixel". A Fig. 2.6(D) exhibe a resposta servo-mecânica do sistema mostrando a posição angular do motor e do alvo. Trata-se tipicamente de uma resposta a degrau de segunda ordem de um sistema de tipo 0. Note que os resultados exibidos nas duas figuras são perfeitamente coerentes.

Realizaram-se experiências com perturbações em degrau de diferentes amplitudes. Observou-se que para amplitudes pequenas a regulação é perfeita e a câmara é reorientada de forma ao alvo ser projectado no centro da imagem. Neste caso o erro de regulação é 0 e o sistema tem um comportamento de tipo 1. O facto do tipo de resposta do sistema depender da amplitude da entrada em degrau é um comportamento não-linear. A Fig. 1.5 mostra a malha de controlo do seguimento monocular. O processo de formação de imagem e os algoritmos de visão para a medição de posição e velocidade claramente introduzem não-linearidades que explicam os comportamentos observados.

A Fig. 2.7(E) mostra os resultados de estimação da posição do alvo na im-



agem. A posição do objecto é estimada como a localização média dos pontos com fluxo óptico não nulo. O fluxo é calculado entre duas imagens consecutivas após a compensação do movimento auto-induzido ("egomotion"). Note que o ponto estimado é o centro de movimento e não a posição do alvo. Estes dois pontos são próximos sempre que o deslocamento do objecto entre duas imagens é pequeno. Tal não é o caso das imagens adquiridas antes e depois da transição do degrau. O alvo passa da posição 0 para a posição 24 e a posição medida é 12 "pixeis", ou seja o ponto médio do deslocamento. Muitos dos erros de estimação observados na Fig. 2.7 explicam-se pelo facto do algoritmo utilizado aproximar a posição do alvo pelo centro de movimento.

Considere o alvo parado no espaço. A velocidade induzida pelo objecto é nula e qualquer deslocamento detectado na imagem é devido ao movimento da câmara. Isto significa que a velocidade total na imagem  $\dot{x}_{img}$  é igual à velocidade induzida  $\dot{x}_{ind}$ , e portanto  $\dot{x}_{img} - \dot{x}_{ind} = 0$  (ver Fig. 1.5). O fluxo óptico ( $\dot{x}$ ) é nulo em todos os pontos da imagem e a estimativa de posição do alvo  $x$  é zero. De acordo com este raciocínio a posição do objecto só deveria ser estimada no instante de transição da entrada em degrau. A Fig. 2.7 mostra que tal não acontece devido a erros na compensação do movimento auto-induzido. O "egomotion" só poder ser exactamente determinado quando a câmara descreve rotações puras. A amostragem e ligeiros desalinhamentos entre o centro de projecção e o eixo de rotação da câmara introduzem pequenos erros. Estes erros são suficientes para fazer com que a diferença  $\dot{x}_{img} - \dot{x}_{ind}$  seja não nula mesmo quando o alvo está parado. Sempre que tal acontece a posição do alvo é estimada (ver Fig. 2.7). Note no entanto que o algoritmo não é capaz de medir erros em regime final (câmara e alvo parados), e obviamente que o sistema não pode compensar um erro que não é capaz de detectar.

### Resposta a Rampa

A Fig. 2.8 mostra a resposta do sistema para uma perturbação em rampa. A velocidade angular do alvo no espaço é de  $-10^\circ/s$ , o que induz uma velocidade na imagem de  $-1.5 \text{ pixel/imagem}$ . O alvo afasta-se 5 "pixeis" do centro da imagem até que o sistema reaja e comece a compensar a perturbação. A inércia na reacção é explicada pela latência do processamento visual e pela acção do filtro de Kalman (ver Fig. 1.5). A medição de posição e velocidade na imagem obtida por processamento visual é ruidosa. O filtro de Kalman

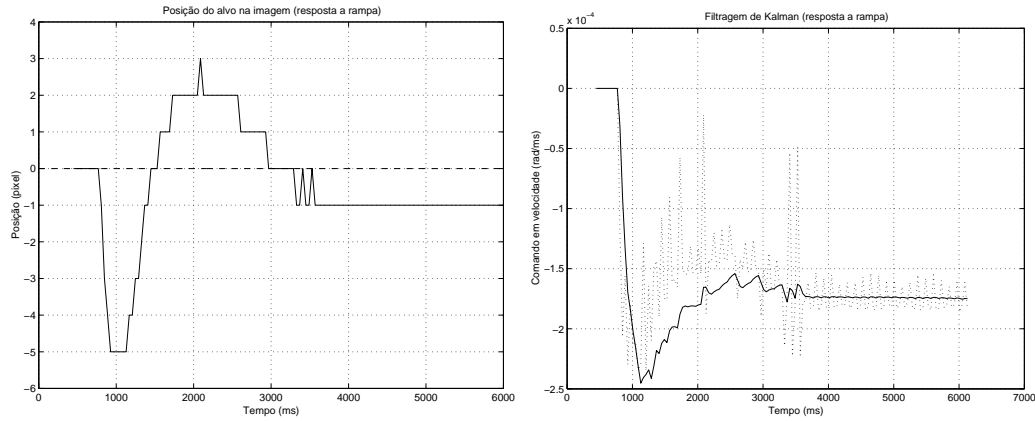


Figura 2.8: Resposta a rampa. Esquerda: Regulação na imagem. Posição de projecção do alvo na imagem (linha a cheio [-]) e centro da imagem (linha tracejada [- -]). Direita: Acção do filtro de Kalman. Sinal de entrada (linha pontuada [..]) e sinal de saída (linha a cheio [-]).

destina-se a limitar o efeito dos erros de medida, permitindo um comportamento de seguimento suave e sem oscilações. A Fig. 2.8(D) mostra a acção do filtro de Kalman na suavização do comando de velocidade calculado pela equação 1.6. O erro em regime final é de  $-1$  "pixels" sendo resposta a rampa de um sistema de tipo 1. Mais uma vez as insuficiências na estimação de posição fazem com que o sistema não seja capaz de anular o erro. Para perturbações em rampa o erro de estimação é aproximadamente metade da velocidade induzida pelo alvo na imagem. Assim, e atendendo a que a velocidade induzida é  $-1.5 \text{ pixels/imagem}$ , o erro na medida de posição é cerca de  $-0.75$  "pixels" (ver Fig. 2.7).

O sistema de seguimento só é capaz de medir correctamente velocidades na imagem até  $2 \text{ pixel/imagem}$ . Esta limitação está relacionada com a dimensão das máscaras utilizadas no cálculo do gradiente da função de brilho. O algoritmo de medição de velocidade é discutido detalhadamente no capítulo seguinte. Para todo o efeito este facto condiciona severamente a resposta do sistema a uma perturbação em rampa. Sempre que o alvo induz velocidades acima de  $2 \text{ pixel/imagem}$  o fluxo óptico é subestimado. Esta limitação é parcialmente compensada pela componente de posição no comando do motor (ver equação 1.6). As experiências feitas permitem concluir que o sistema é capaz de seguir movimentos com velocidade constante até  $20^\circ/s$ .

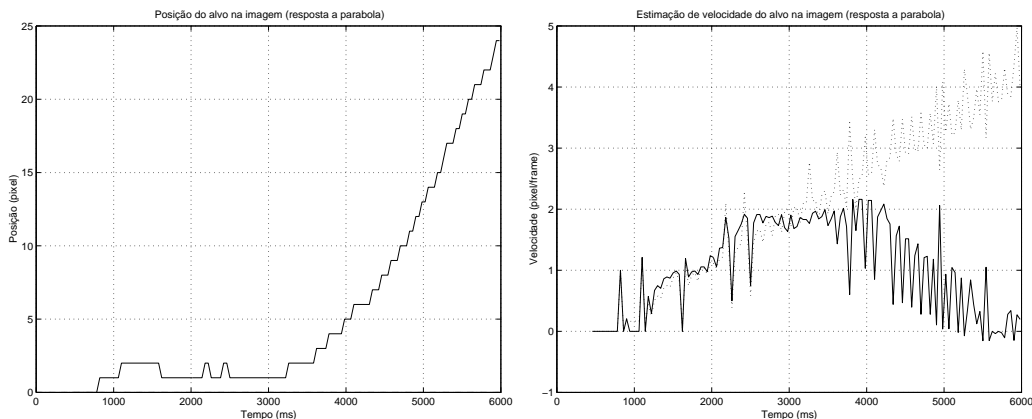


Figura 2.9: Resposta a parábola. Esquerda: Regulação na imagem. Posição de projecção do alvo na imagem (linha a cheio [-]) e centro da imagem (linha tracejada [- -]). Direita: Estimação da velocidade induzida pelo alvo na imagem. Velocidade estimada (linha a cheio [-]) e velocidade real (linha ponteadada [..]).

### Resposta a Parábola

O alvo movimenta-se com uma aceleração angular constante de  $5^\circ/s^2$  e uma velocidade inicial de  $1^\circ/s$ . A Fig. 2.9(E) mostra a regulação na imagem que indica a capacidade do sistema em compensar esta perturbação. A Fig. 2.9(D) permite visualizar os resultados da estimação da velocidade induzida na imagem. Como se pode ver o sistema só é capaz de medir velocidades até  $2 \text{ pixel/imagem}$ . Quando a velocidade sobe acima deste valor a subestimação do fluxo óptico limita o desempenho global da regulação. O sistema é incapaz de seguir o objecto de forma a compensar a velocidade induzida na imagem e, conseqüentemente, a posição do alvo na imagem afasta-se do centro.

### Resposta a Sinusóide.

A Fig. 2.10 mostra vários aspectos da resposta a uma perturbação sinusoidal de amplitude  $10^\circ$  e velocidade angular  $2 \text{ rad/s}$ . A resposta servo-mecânica do motor apresenta um considerável atraso de fase. Este atraso é motivado pelas latências na malha de controlo e pelo efeito suavizador do filtro de Kalman. Velocidades acima dos  $2 \text{ pixel/imagem}$  são incorrectamente medidas. Note que o filtro de Kalman minora os efeitos da subestimação do fluxo.

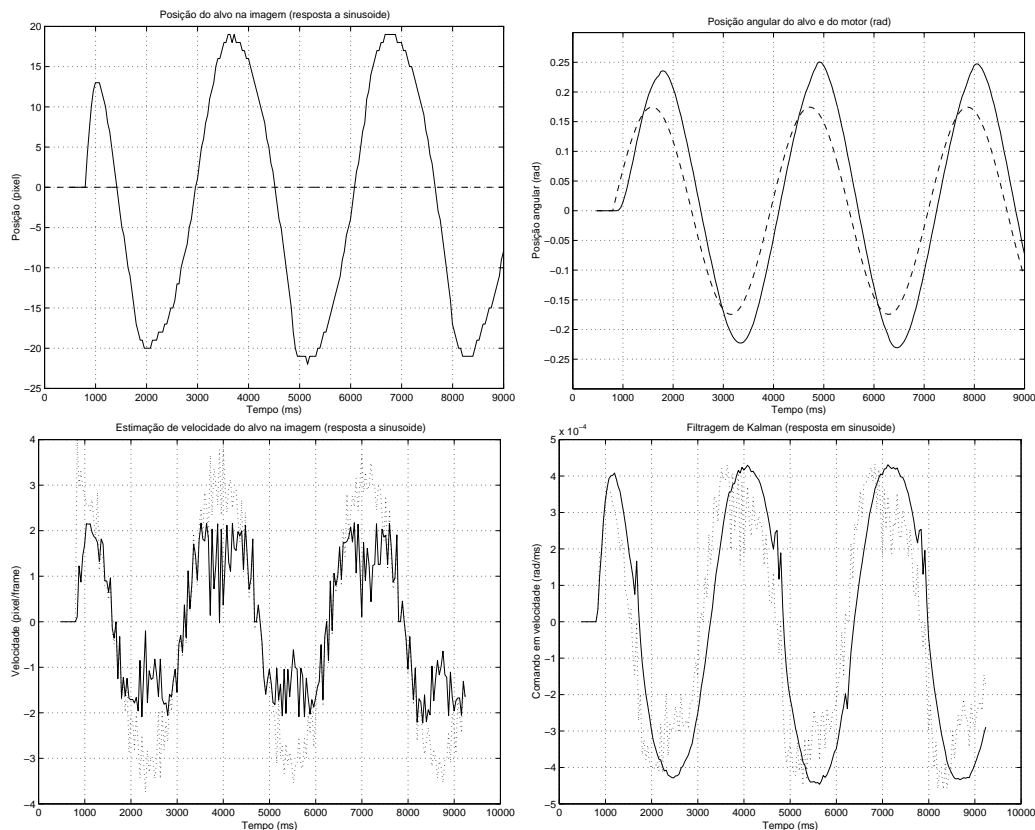


Figura 2.10: Resposta a sinusóide. Em cima à esquerda: Regulação na imagem. Posição de projecção do alvo na imagem (linha a cheio [-]) e centro da imagem (linha tracejada [- -]). Em cima à direita: Resposta servo-mecânica. Posição angular do alvo  $\theta$  (linha tracejada [- -]) e do motor  $\alpha_p$  (linha a cheio [-]). Em baixo à esquerda: Estimação da velocidade induzida pelo alvo na imagem. Velocidade estimada (linha a cheio [-]) e velocidade real (linha pontuada [..]). Em baixo à direita: Acção do filtro de Kalman. Sinal de entrada (linha pontuada [..]) e sinal de saída (linha a cheio [-]).

## Capítulo 3

# Optimização do desempenho do processamento visual

O capítulo anterior estabeleceu um conjunto de técnicas para caracterizar o sistema de visão activo. A avaliação feita permitiu identificar diversos aspectos que limitam o desempenho global do seguimento.

Verificou-se que o sistema estima o centro de movimento em vez da posição do alvo na imagem. Por outro lado a estimação de posição só é feita quando a velocidade induzida pelo alvo na imagem é diferente de zero. Concluiu-se ainda que a estimação da posição do alvo na imagem é fundamental para garantir um erro em regime final nulo. Observou-se que os problemas na medição da posição deterioram o desempenho global da regulação. O algoritmo visual de estimação de posição tem de ser melhorado de forma a otimizar a capacidade de seguimento do sistema.

Outra limitação importante é o facto do sistema só ser capaz de medir velocidades até 2pixel/imagem. Isto conduz a maus desempenhos no seguimento de alvos cujo movimento induz velocidades na imagem acima deste valor. O aumento da gama de velocidades mensuráveis pelo sistema irá permitir acréscimos do desempenho global.

Este capítulo discute melhorias no processamento visual de forma a superar os problemas detectados. A estimação de posição e velocidade na imagem tem sido objecto de estudo por vários autores [27, 28, 29, 30, 31, 32, 33, 34]. No entanto é importante não esquecer que as soluções implementadas devem ser eficientes do ponto de vista computacional. Técnicas de estimação muito rigorosa tendem a ser complexas e computacionalmente pesadas. Os atrasos introduzidos na realimentação visual podem ter consequências desastrosas no



Figura 3.1: Sequência de imagens capturadas durante o seguimento. A linha de cima mostra as imagens originais, a segunda linha as imagens depois da compensação do movimento auto-induzido, a terceira linha o conjunto de pontos na imagem onde ocorreu movimento independente e a quarta linha os pontos considerados para estimar a posição do alvo na imagem

desempenho conduzindo inclusive a comportamentos instáveis [21, 20, 35]. A escolha de algoritmos que equilibrem bem o rigor, a robustez e a eficiência computacional é fundamental para se obter um sistema de seguimento em tempo real com elevado desempenho [36].

### 3.1 Estimação da Posição do Alvo na Imagem.

Observe a sequência de imagens da Fig. 3.1 capturadas em instantes de tempo sucessivos durante uma acção de seguimento. Note que ao fundo aparecem partes da ombreira de uma porta e de uma prateleira. Apesar de estes elementos da cena serem estáticos a sua posição de projecção na imagem altera-se ao longo do tempo. Este movimento nas imagens, conhecido por movimento auto-induzido, é provocado pela alteração da geometria do sistema. O movimento das câmaras entre duas aquisições sucessivas pode ser conhecido através das leituras dos codificadores ópticos dos motores.

No caso de seguimento monocular a câmara descreve rotações puras. As velocidades auto-induzidas na imagem podem ser calculadas sem conhecer a profundidade dos pontos da cena (distância ao centro de projecção), existindo uma relação homográfica entre as imagens adquiridas antes e depois da rotação [12]. Esta relação homográfica deixa de se verificar quando movi-



Figura 3.2: Estimação da posição do alvo na imagem. Localização média dos pontos com movimento independente (x). Localização média dos pontos com movimento independente e onde o gradiente espacial da função de brilho na segunda imagem é diferente de zero (o)

mento da câmara tem componente de translação. No seguimento binocular os movimentos do pescoço introduzem uma componente de translação. Os estudos feitos em [37] demonstram que a velocidade induzida pela translação é proporcional ao quociente entre metade da distância interocular (braço  $B$  na Fig. 1.4) e a distância da cena ao centro de projecção do olho ciclópico (profundidade). Atendendo a que normalmente a primeira é muito inferior à segunda o movimento auto-induzido devido à translação pode ser negligenciado na maior parte dos casos.

As câmaras do sistema MDOF têm um campo de visão de  $12^\circ$ . Para campos de visão estreitos a velocidade auto-induzida é aproximadamente igual em todos os pontos da imagem [37]. Considere duas imagens adquiridas em instantes de tempo sucessivos. A velocidade auto-induzida é calculada usando as leituras dos codificadores. Uma nova imagem é determinada translacionando a primeira imagem segundo o vector de velocidade calculado. O resultado, obtido por interpolação bilinear, pode ser observado na segunda linha da Fig. 3.1. A terceira sequência da Fig. 3.1 é determinada subtraindo a imagem translacionada à imagem adquirida no instante de tempo seguinte. A imagem de diferença contém os pontos onde ocorreu movimento independente.

O estudo feito no capítulo anterior detectou limitações na estimação de posição e velocidade que interferem com o desempenho global do sistema. A secção seguinte discute em detalhe a estimação de velocidade. Para estimar a velocidade o sistema determina o gradiente da função de brilho  $\nabla I = (I_x, I_y, I_t)$  nos pontos da imagem com movimento independente. A imagem de diferença da Fig. 3.1 é utilizada como máscara neste cálculo. A posição do alvo é calculada como a localização média do conjunto de pontos com fluxo óptico não nulo entre duas imagens adquiridas em instantes de tempo

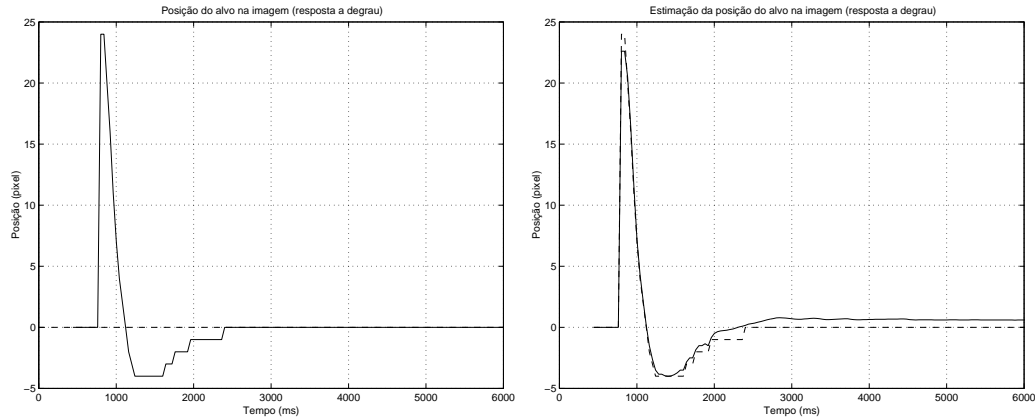


Figura 3.3: Resposta a degrau. Esquerda: Regulação na imagem. Posição de projecção do alvo na imagem (linha a cheio [-]) e centro da imagem (linha tracejada [- -]). Direita: Estimação da posição do alvo na imagem. Posição estimada (linha a cheio [-]), posição real (linha tracejada [- -]).

sucessivos (pontos de movimento independente). Na realidade o sistema calcula o centro de movimento em vez da posição do alvo. Esta aproximação é grosseira quando o corpo induz velocidades elevadas na imagem. Considere o subconjunto dos pontos que, para além de fluxo não nulo, têm derivadas parciais do brilho em  $X$  e  $Y$  não nulas na imagem adquirida por último. Estes pontos são mostrados na quarta sequência da Fig. 3.1. A localização média destes pontos aproxima o centro de massa do alvo na imagem adquirida. Note que esta modificação do algoritmo não envolve qualquer computação adicional dado que o gradiente do brilho tem que ser calculado para estimar a velocidade. As melhorias na estimação de posição podem ser observadas na sequência de imagens da Fig. 3.2.

$$x(k) = x(k - 1) + T_i \cdot \dot{x}_{ind}(k) \quad (3.1)$$

No capítulo anterior referiu-se ainda que a posição só era estimada quando o alvo induzia movimento na imagem. Se a velocidade na imagem, após compensação do movimento auto-induzido, é nula então é porque o alvo está parado. Assim a nova posição na imagem há de ser igual à estimativa anterior compensada pelo integral da velocidade induzida pelo movimento da câmara (equação 3.1 onde  $T_i$  é o período de amostragem (40ms)).

Considere a perturbação em degrau utilizada na caracterização do capítulo



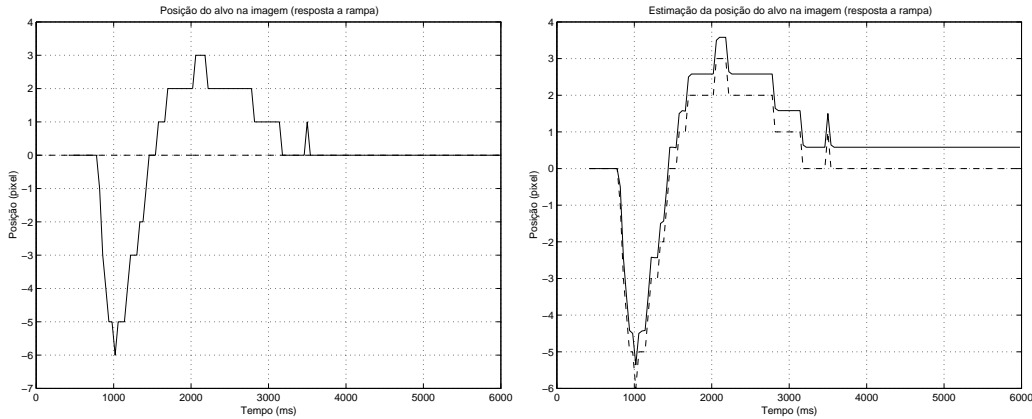


Figura 3.4: Resposta a rampa. Esquerda: Regulação na imagem. Posição de projecção do alvo na imagem (linha a cheio [-]) e centro da imagem (linha tracejada [- -]). Direita: Estimação da posição do alvo na imagem. Posição estimada (linha a cheio [-]), posição real (linha tracejada [- -]).

anterior. Comparando os resultados da Fig. 2.5(E) e Fig. 3.3(E) verifica-se que as melhorias no algoritmo de estimação de posição aumentaram consideravelmente o desempenho de regulação do sistema.

Para a perturbação em rampa as melhorias também são assinaláveis. Note que agora o erro em regime final é nulo.

## 3.2 Estimação da Velocidade do Alvo na Imagem.

Assuma que existe um único corpo em movimento no campo de visão do sistema. A imagem de diferenças, calculada após a compensação do movimento auto-induzido, mostra os "pixels" onde ocorre movimento independente. Atendendo a que só existe uma fonte de movimento, todos estes pontos fazem parte da projecção do alvo na imagem. Assuma ainda que todos os "pixels" pertencentes ao alvo se movem com a mesma velocidade  $(u, v)$  cujo valor o sistema pretende estimar.

$$I_x \cdot u + I_y \cdot v + I_t = 0 \quad (3.2)$$

Considere uma sequência de imagens bidimensionais adquirida continu-

amente ao longo do tempo. A cada instante de tempo  $t$  corresponde uma imagem cujos pontos são referenciados pelas coordenadas  $(x, y)$ .  $I(x, y, t)$  é a função escalar de brilho e  $\nabla I = (I_x, I_y, I_t)$  o respectivo gradiente. Se  $I(x, y, t)$  é uma função contínua e se o brilho de cada ponto é invariante ao longo do tempo, então a equação 3.2 verifica-se para todos os pontos onde ocorre movimento independente. O vector velocidade  $(u, v)$  pode ser calculado aplicando o método dos mínimos quadrados[38].

A restrição do fluxo expressa pela equação 3.2 é verdadeira se a função  $I(x, y, t)$  for contínua no espaço e houver conservação do brilho ao longo do tempo. No entanto a função de brilho é amostrada pelo sistema no tempo e no espaço. As derivadas parciais  $(I_x, I_y, I_t)$  são calculadas a partir da função amostrada. Fenómenos de "aliasing" podem comprometer uma correcta estimação do vector de velocidade  $(u, v)$ .

Assuma que o alvo está a movimentar-se muito devagar. Neste caso é necessário uma elevada resolução espacial de forma a calcular correctamente as derivadas espaciais  $I_x$  e  $I_y$ . A estimação de baixas velocidades na imagem implica uma elevada resolução espacial, caso contrário o movimento não é perceptível pelo sistema. Um movimento de alvo de alta velocidade corresponde a elevadas frequências no tempo.  $I_t$  tem de ser calculado com um período de amostragem reduzido. No entanto a frequência de amostragem temporal está limitada a 25Hz. Uma solução para estimar velocidades elevadas na imagem é reduzir a resolução espacial. A redução da resolução tem como inconveniente a perda das altas frequências espaciais o que torna os pequenos movimentos do alvo indetectáveis.

Pretende-se um sistema de seguimento de elevado desempenho capaz de seguir alvos em movimento quer lento quer rápido. No primeiro caso o movimento só é perceptível em imagens de alta resolução. No segundo caso o gradiente da função de brilho deve ser calculado em imagens de baixa resolução. Esta secção apresenta dois métodos para aumentar a gama de velocidades que o sistema é capaz de estimar correctamente. Em ambos os casos o algoritmo de estimação de velocidade é multi-escalar [39, 40].

### 3.2.1 Método dos resíduos

O sistema de seguimento adquire imagens com resolução  $64 \times 64$  e utiliza uma máscara  $2 \times 2$  para calcular as derivadas espaciais da função de brilho  $I_x$  e  $I_y$  [41]. O limite máximo da gama de velocidades estimáveis é dado pela

dimensão desta máscara. Isto explica a incapacidade do sistema em medir velocidades acima dos 2pixel/imagem.

As imagens sofrem uma sub-amostragem de forma a obter uma resolução  $32 \times 32$ . Utilizando a mesma máscara  $2 \times 2$  para calcular as derivadas espaciais a velocidade máxima que o sistema é capaz de estimar passa a ser 4pixel/imagem. Isto porque uma velocidade de 2 pixel/imagem a uma resolução de  $32 \times 32$  corresponde a uma velocidade de 4 pixel/imagem na resolução  $64 \times 64$ . Assim, reduzindo a resolução das imagens, o sistema é capaz de estimar deslocamentos mais elevados do alvo utilizando o mesmo algoritmo de cálculo de fluxo. Note que as imagens de baixa resolução são obtidas sub-amostrando as imagens originais depois de uma filtragem passa-baixo.

Este novo algoritmo começa por construir uma pirâmide de imagens com diferentes resoluções. A resolução de cada nível é metade da resolução no nível inferior. Na implementação feita a pirâmide tem três níveis em que as resoluções respectivas são  $64 \times 64$ ,  $32 \times 32$  e  $16 \times 16$ . O fluxo óptico é calculado simultaneamente nos três níveis usando uma máscara  $2 \times 2$  na determinação das derivadas espaciais. Os valores obtidos são, respectivamente,  $\dot{x}_{64 \times 64}$ ,  $\dot{x}_{32 \times 32}$  e  $\dot{x}_{16 \times 16}$ . Atendendo às considerações feitas anteriormente  $\dot{x}_{64 \times 64}$  estima correctamente velocidades até 2pixel/imagem. Velocidades entre 2 e 4pixel/imagem e 4 e 8pixel/imagem são dadas respectivamente por  $\dot{x}_{32 \times 32}$  e  $\dot{x}_{16 \times 16}$ . A questão está em decidir em cada iteração qual das estimativas  $\dot{x}_{64 \times 64}$ ,  $\dot{x}_{32 \times 32}$  e  $\dot{x}_{16 \times 16}$  está mais próxima da velocidade real induzida pelo alvo na imagem.

$$\sum_{i=1}^N (I_x^i \cdot u + I_y^i \cdot v + I_t^i)^2 = 0 \quad (3.3)$$

Assuma que existem  $N$  pontos na imagem onde ocorreu movimento independente. A velocidade  $(u, v)$  é estimada calculando o vector que minimiza o erro quadrático da equação 3.3. Normalmente o ajuste não é perfeito e a equação 3.2 não se verifica exactamente. Neste caso, para a estimativa  $(u, v)$ , cada ponto da imagem tem um resíduo associado. O valor médio destes resíduo é um indicador do desempenho do estimador. O algoritmo escolhe como estimativa da velocidade o valor obtido no nível com menor resíduo médio.

### 3.2.2 Método da máscara variável

De forma semelhante ao que acontece no método dos resíduos, este algoritmo começa por construir uma pirâmide de imagens com três níveis de resolução. O fluxo óptico é calculado no nível mais elevado ( $16 \times 16$ ) usando uma máscara  $2 \times 2$ . O resultado desta operação controla o tamanho da máscara utilizada para estimar as derivadas espaciais no nível imediatamente inferior. Assim,  $\dot{x}_{32 \times 32}$  é determinado usando uma máscara  $m \times n$  com  $m, n \in \{2, 3, 4\}$ , em que  $n$  depende de  $\dot{x}_{16 \times 16}$  e  $m$  depende de  $\dot{y}_{16 \times 16}$  (a componente de velocidade segundo  $Y$ ). O controlo de dimensão da máscara é baseado em intervalos entre valores de limiar definidos experimentalmente. Considere os limiares  $L_1$  e  $L_2$  em que  $L_1 \leq L_2$ .  $n$  é 2,3 ou 4 conforme  $|\dot{x}_{16 \times 16}| \leq L_1$ ,  $L_1 \leq |\dot{x}_{16 \times 16}| \leq L_2$  ou  $|\dot{x}_{16 \times 16}| \geq L_2$ .

O processo descrito em cima repete-se para os dois últimos níveis da pirâmide.  $\dot{x}_{16 \times 16}$  e  $\dot{x}_{32 \times 32}$  são utilizados para definir o tamanho da máscara para o cálculo de  $\dot{x}_{64 \times 64}$ . A dimensão da máscara é agora  $m \times n$  com  $m, n \in \{2, 3, 4, 5, 6, 7, 8\}$ . A lei de controlo continua a basear-se nos intervalos entre valores de limiar. Note que, enquanto no primeiro método o algoritmo define o nível de resolução a que é feita a estimativa, neste método a velocidade é sempre estimada sobre a imagem de maior resolução.

### 3.2.3 Resultados Experimentais

Para ambos os algoritmos o limite máximo teórico da gama de velocidades mensuráveis aumenta de 2pixel/frame para 8pixel/frame. Considere a perturbação em parábola utilizada na caracterização do capítulo anterior. A Fig. 3.5 mostra a resposta do sistema fazendo a estimação com o método dos resíduos (em cima) e com o método da máscara variável (em baixo). Compare com os resultados da Fig. 2.9. A velocidade é correctamente estimada pelos dois métodos e os gráficos da regulação na imagem mostram uma compensação da perturbação quase perfeita.

A Fig. 3.6 mostra a resposta do sistema a uma perturbação em sinusóide. Compare estes resultados com os obtidos na Fig. 2.10. Ambos os métodos estimam correctamente a velocidade induzida na imagem. No entanto a melhoria na estimação de velocidade não é acompanhada por uma melhoria efectiva da capacidade de regulação na imagem. Isto é explicado pela acção do filtro de Kalman. Note na Fig. 2.10 que a filtragem de Kalman tende a compensar os erros na estimação da velocidade. O modelo de movimento

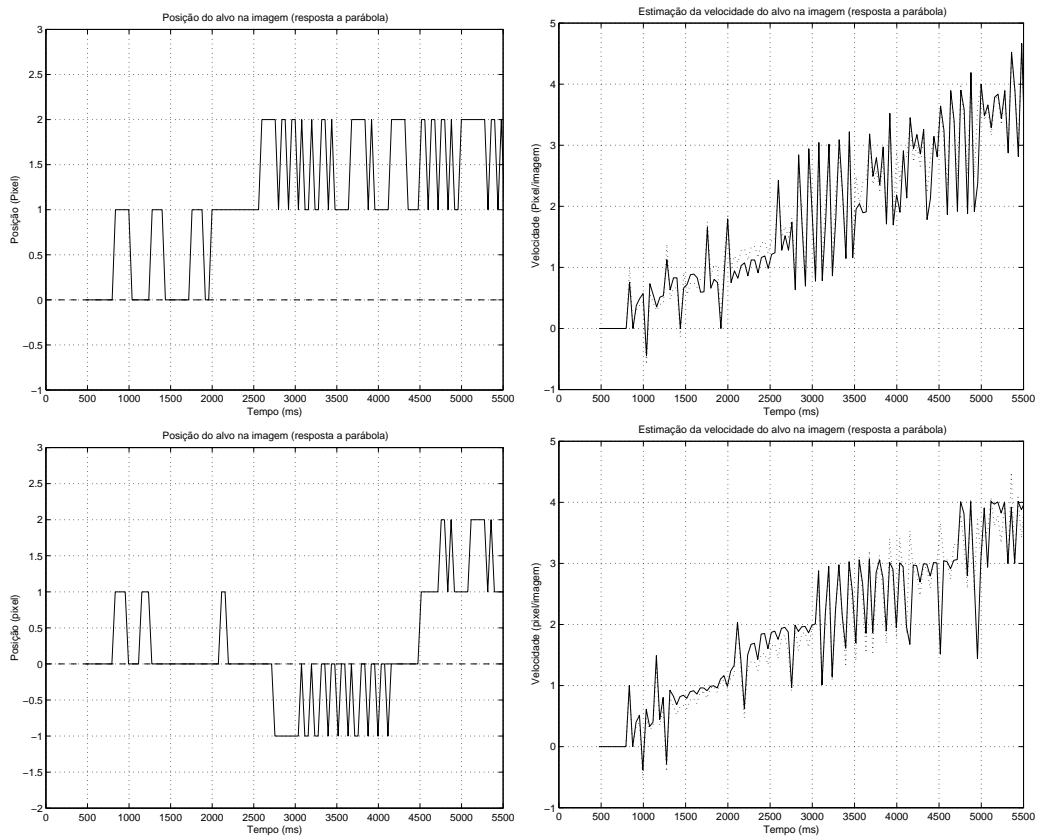


Figura 3.5: Resposta a parábola. Em cima à esquerda: Regulação na imagem usando o método dos resíduos. Posição de projecção do alvo na imagem (linha a cheio [-]) e centro da imagem (linha tracejada [- -]). Em cima à direita: Estimação da velocidade induzida pelo alvo na imagem usando o método dos resíduos. Velocidade estimada (linha a cheio [-]) e velocidade real (linha pontuada [..]). Em baixo à esquerda: Regulação na imagem usando o método da máscara variável. Posição de projecção do alvo na imagem (linha a cheio [-]) e centro da imagem (linha tracejada [- -]). Em baixo à direita: Estimação da velocidade induzida pelo alvo na imagem usando o método da máscara variável. Velocidade estimada (linha a cheio [-]) e velocidade real (linha pontuada [..]).

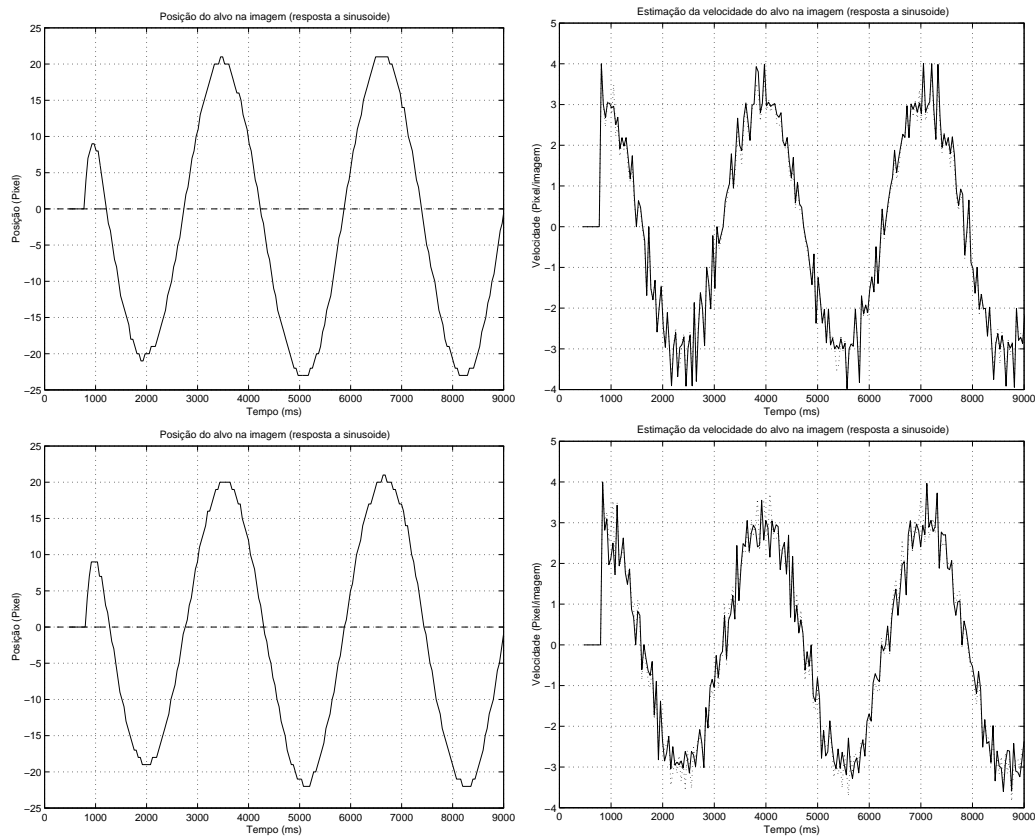


Figura 3.6: Resposta a parábola. Em cima à esquerda: Regulação na imagem usando o método dos resíduos. Posição de projecção do alvo na imagem (linha a cheio [-]) e centro da imagem (linha tracejada [- -]). Em cima à direita: Estimação da velocidade induzida pelo alvo na imagem usando o método dos resíduos. Velocidade estimada (linha a cheio [-]) e velocidade real (linha ponteadada [..]). Em baixo à esquerda: Regulação na imagem usando o método da máscara variável. Posição de projecção do alvo na imagem (linha a cheio [-]) e centro da imagem (linha tracejada [- -]). Em baixo à direita: Estimação da velocidade induzida pelo alvo na imagem usando o método da máscara variável. Velocidade estimada (linha a cheio [-]) e velocidade real (linha ponteadada [..]).

assumido compensa parcialmente as consequências do súbito colapso do estimador. As melhorias do desempenho da medição de velocidade não se fazem notar devido a este efeito suavizador. Obviamente que a inércia imposta pelo filtro de Kalman é mais notória para perturbação em sinusóide, onde há constantes mudanças de aceleração, do que para uma entrada em parábola. A secção seguinte discute a sintonização do filtro de Kalman.

A gama de velocidades mensuráveis pode ser aumentada utilizando pirâmides com um maior número de níveis. No método da máscara variável a escolha dos valores de limiar é crítica para um bom desempenho de estimação. No entanto este algoritmo tem a vantagem de desacoplar a estimação de velocidade em  $X$  e  $Y$ . Assuma que a velocidade em  $X$  (horizontal) é muito elevada enquanto em  $Y$  (vertical) é muito baixa. O método dos resíduos utiliza sempre uma máscara quadrada para determinar as derivadas espaciais  $I_x$  e  $I_y$ . Desta forma não é possível obter-se uma correcta estimação da velocidade em ambas as direcções. Calculando a velocidade numa imagem de baixa resolução a componente vertical é mal estimada, utilizando uma imagem de alta resolução a componente horizontal é determinada erradamente. O método da máscara variável lida com esta situação utilizando uma máscara rectangular  $m \times n$  com  $m$  pequeno e  $n$  elevado. Por esta razão o método da máscara variável é preferível ao método dos resíduos.

### 3.3 Filtragem de Kalman

A filtragem de Kalman [42] é utilizada para estimar, usando a informação visual, alguns parâmetros do movimento 3D do alvo. Considere o movimento em azimuth no seguimento monocular. Imagine que pretende estimar o movimento de rotação do alvo em torno do eixo dos  $Y$  de forma a controlar a correspondente rotação da câmara. De acordo com a Fig. 2.3 a posição angular do alvo em relação ao referencial inercial é  $\theta$ . Assuma que o objecto tem um movimento de rotação uniformemente acelerado e que a velocidade e aceleração angulares são, respectivamente,  $\omega$  e  $\gamma$  ( $\omega = \dot{\theta}$  e  $\gamma = \ddot{\theta}$ ).

$$x = f \tan(\Delta\theta) \approx f\Delta\theta \quad (3.4)$$

$$\dot{x} = f \cdot \omega \frac{1}{\cos^2(\Delta\theta)} \approx f\omega \quad (3.5)$$

Considere ainda que  $x$  e  $\dot{x}$  são a posição e a velocidade medidas ao longo

do eixo dos  $X$  da imagem (ver Fig. 1.5). As equações 3.4 e 3.5 estabelecem a relação entre as grandezas medidas na imagem e o movimento angular do alvo. Este resultado é deduzido a partir das equações 2.18 e 2.19. A linearização é feita por expansão em série de Taylor em torno de  $x = 0$  (centro da imagem). Note que enquanto a velocidade medida após cancelamento de "egomotion" está directamente relacionada com a velocidade angular do alvo no referencial inercial, a posição na imagem está associada à posição angular do objecto no referencial da câmara.  $\Delta\theta$  representa o erro angular de seguimento ( $\Delta\theta = \theta - \alpha_p$ ).

$$\theta(n) = \theta(n-1) + \omega(n-1).T_i + \gamma(n-1).\frac{T_i^2}{2} \quad (3.6)$$

$$\Delta\theta(n) = \theta(n-1) + \omega(n-1).T_i + \gamma(n-1).\frac{T_i^2}{2} - \alpha_p(n) \quad (3.7)$$

$$\Delta\theta(n) = \Delta\theta(n-1) + \omega(n-1).T_i + \gamma(n-1).\frac{T_i^2}{2} - (\alpha_p(n) - \alpha_p(n-1)) \quad (3.8)$$

O filtro modela o comportamento do alvo através de um movimento angular uniformemente acelerado entre os instantes de aquisição de imagem. Esta aproximação é aceitável atendendo à frequência de amostragem de 25Hz (40ms entre cada imagem). Assim a posição angular no momento actual  $\theta(n)$ , é calculada pela equação 3.6, conhecendo a posição, velocidade e aceleração da iteração anterior.  $T_i$  representa o tempo entre duas iterações sucessivas (período de amostragem). Subtraindo a ambos os membros da equação a posição actual do motor  $\alpha_p(n)$ , lida pelo codificador óptico, obtém-se a expressão do erro angular de seguimento 3.7. A equação 3.8 é deduzida de 3.7 fazendo  $\theta(n-1) = \Delta\theta(n-1) + \alpha_p(n-1)$ .

$$\mathbf{u}(n) = \begin{bmatrix} 1 & T_i & \frac{T_i^2}{2} \\ 0 & 1 & T_i \\ 0 & 0 & 1 \end{bmatrix} \mathbf{u}(n-1) - \begin{bmatrix} \Delta\alpha_p \\ 0 \\ 0 \end{bmatrix} + \mathbf{q}(n-1) \quad (3.9)$$

$$\mathbf{v}(n) = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \end{bmatrix} \mathbf{u}(n) + \mathbf{r}(n) \quad (3.10)$$

A estrutura do filtro de Kalman utilizado para o seguimento monocular é descrita por 3.9 e 3.10. Os parâmetros a ser estimados pelo filtro são: o erro angular de seguimento  $\Delta\theta$ , a velocidade angular do alvo  $\omega$  e a aceleração



angular  $\gamma$ . O vector de estado é  $\mathbf{u} = (\Delta\theta, \omega, \gamma)^t$ . A equação 3.9 é a equação do processo assumindo um modelo de movimento uniformemente acelerado. A primeira linha da matriz de transição de estados é deduzida a partir de 3.8 em que  $\Delta\alpha_p = \alpha_p(n) - \alpha_p(n-1)$ . 3.10 é a equação de medida que estabelece a relação entre o vector de estado  $\mathbf{u}$  e as grandezas medidas na imagem  $\mathbf{v} = (x, \dot{x})^t$ . Esta equação é derivada a partir de 3.4 e 3.5. Os vectores  $\mathbf{q}$  e  $\mathbf{r}$  são termos de ruído branco.

$$\mathbf{Q} = \begin{bmatrix} \sigma_p^2 & 0 & 0 \\ 0 & \sigma_v^2 & 0 \\ 0 & 0 & \sigma_a^2 \end{bmatrix} \quad (3.11)$$

$$\mathbf{R} = \begin{bmatrix} \sigma_{pi}^2 & 0 \\ 0 & \sigma_{vi}^2 \end{bmatrix} \quad (3.12)$$

$\mathbf{Q}$  e  $\mathbf{R}$  são as matrizes de correlação de  $\mathbf{q}$  e  $\mathbf{r}$ . Assuma que os erros do modelo e das medidas são não correlacionados e que  $\mathbf{Q}$  e  $\mathbf{R}$  são matrizes diagonais. Observe as equações 3.11 e 3.12.  $\sigma_p$ ,  $\sigma_v$  e  $\sigma_a$  são, respectivamente, o desvio padrão do erro do modelo considerado para a posição, velocidade e aceleração angular do alvo. Já  $\sigma_{pi}$  e  $\sigma_{vi}$  representam o desvio padrão do erro de medida da posição e velocidade na imagem. Considere ainda que o valor inicial do vector de estado  $\mathbf{u}$  é  $\mathbf{u}_0 = (0, 0, 0)^t$  (alvo parado e alinhado com o eixo óptico). A sintonização do Kalman consiste na escolha adequada dos valores de desvio padrão  $\sigma_p$ ,  $\sigma_v$ ,  $\sigma_a$ ,  $\sigma_{pi}$  e  $\sigma_{vi}$ .

Os valores atribuídos a  $\sigma_{pi}$  e  $\sigma_{vi}$  indicam o grau de confiança nas medições feitas na imagem. Se os valores destes desvios padrões forem baixos o filtro assume que as estimativas da posição e velocidade têm uma elevada probabilidade de estar correctas. No caso da medição das grandezas na imagem serem muito ruidosas  $\sigma_{pi}$  e  $\sigma_{vi}$  devem tomar valores elevados. Da mesma forma  $\sigma_p$ ,  $\sigma_v$  e  $\sigma_a$  traduzem o grau de confiança no modelo de movimento. Imagine que se pretende seguir um alvo com um movimento suave sem grandes mudanças de aceleração. Este tipo de movimentação é adequadamente descrito por 3.9 e  $\sigma_p$ ,  $\sigma_v$  e  $\sigma_a$  devem ser sintonizados com valores baixos. No caso do alvo apresentar mudanças de aceleração súbitas e frequentes o modelo de aceleração constante deixa de ser adequado o que se deve reflectir num aumento do desvio padrão da aceleração. O filtro tem um efeito suavizador compensando nalguns casos a má prestação do processamento visual. A desvantagem desta sintonização é a inércia que introduz no sistema. Se a confiança dada às

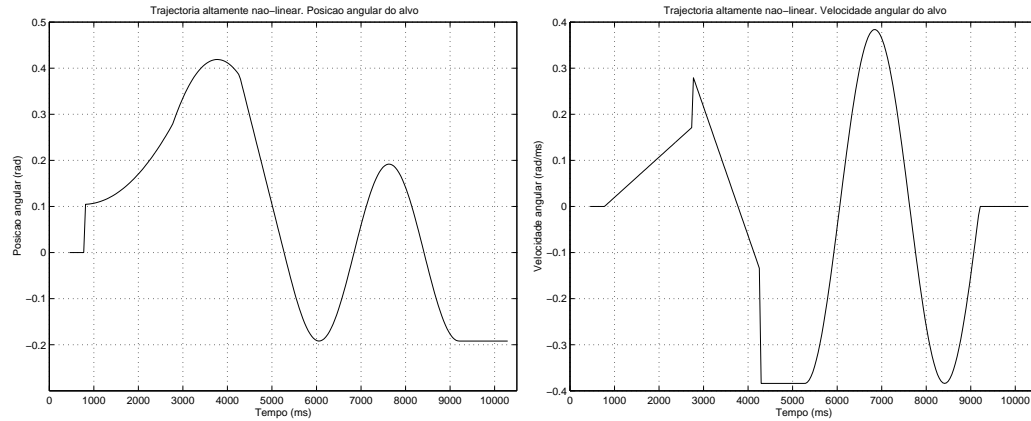


Figura 3.7: Trajectória de teste altamente não-linear. A fim de se testar a reacção do sistema a perturbações não-lineares juntou-se ao conjunto dos sinais de teste a trajectória da figura. O gráfico da esquerda mostra a variação da posição angular do alvo e o da direita a evolução da velocidade angular correspondente.

medições na imagem é baixa diminui-se a capacidade do sistema em reagir a mudanças bruscas no movimento do alvo. As melhorias no processamento visual devem ser acompanhadas por uma sintonização menos conservativa do filtro. Desta forma a capitaliza-se o acréscimo de capacidade de estimação visual num aumento efectivo de desempenho global do sistema.

A Fig. 3.8 compara as estimativas dos parâmetros angulares de movimento para diferentes sintonizações do filtro. Os dados foram obtidos assumindo uma trajectória de referência altamente não linear (ver Fig. 3.7). As modificações feitas nos algoritmos de processamento visual permitiram uma melhoria na estimação da posição e velocidade do alvo na imagem. Este decréscimo nos erros de medida deve ser acompanhado por uma nova sintonização do filtro. A sintonização utilizada até ao momento é referenciada por  $S1$ . A sintonização  $S2$  foi estabelecida tendo em conta o aumento de confiança nas medições de posição e velocidade. Note que a nova sintonização é obtida diminuindo o desvio padrão dos erros de medição ( $\sigma_{pi}, \sigma_{vi}$ ) e aumentando o desvio padrão dos erros do modelo ( $\sigma_p, \sigma_v, \sigma_a$ ). A sintonização  $S1$  é mais conservativa e o efeito suavizador mais pronunciado. A sintonização  $S2$  apresenta uma convergência mais rápida para os parâmetros reais do movimento, o que teoricamente se traduz numa maior capacidade de reacção do sistema. É preciso referir que uma confiança excessiva nos resultados do processamen-

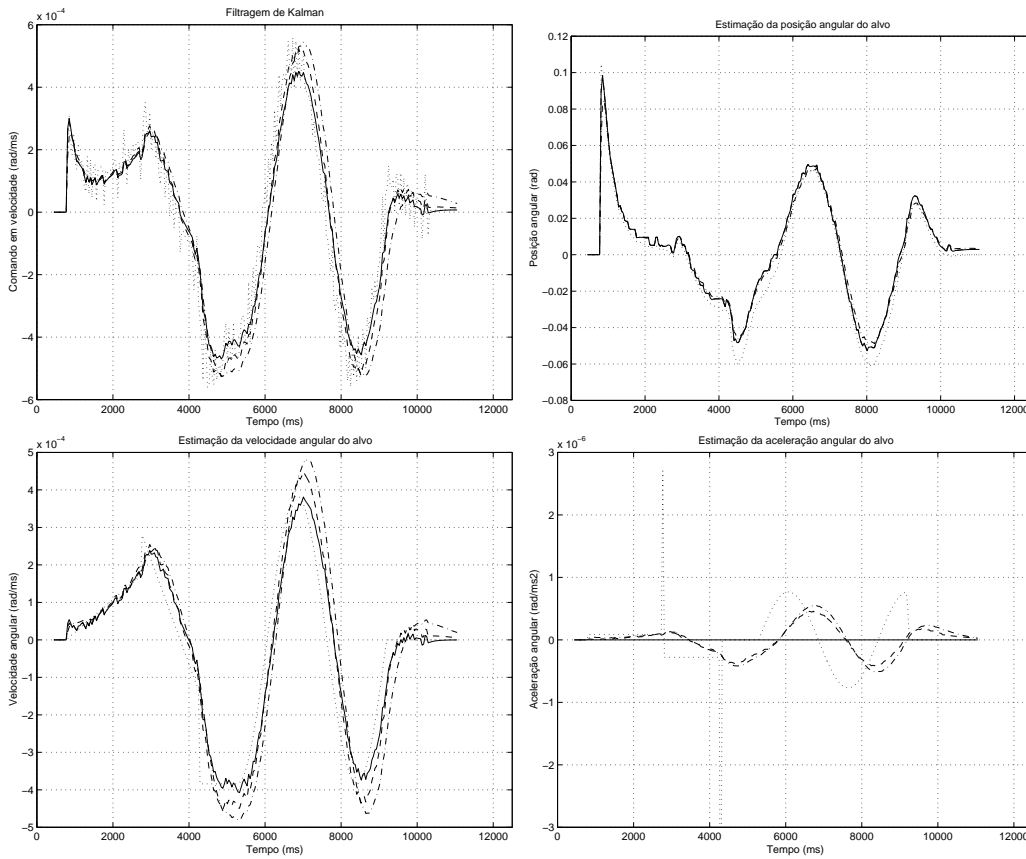


Figura 3.8: Comparação das diferentes sintonizações do filtro de Kalman. Sintonização inicial **S1** ( $\sigma_p = 0.5^\circ$ ,  $\sigma_v = 0.5^\circ/ms$ ,  $\sigma_a = 1.5^\circ/ms^2$ ,  $\sigma_{pi} = 2pixel$ ,  $\sigma_{vi} = 2pixel/imagem$ ). Nova sintonização **S2** ( $\sigma_p = 1^\circ$ ,  $\sigma_v = 1.5^\circ/ms$ ,  $\sigma_a = 2^\circ/ms^2$ ,  $\sigma_{pi} = 1.5pixel$ ,  $\sigma_{vi} = 1.5pixel/imagem$ ). Sintonização assumindo um modelo de velocidade constante **S3** ( $\sigma_p = 1^\circ$ ,  $\sigma_v = 2^\circ/ms$ ,  $\sigma_a = 0^\circ/ms^2$ ,  $\sigma_{pi} = 1.5pixel$ ,  $\sigma_{vi} = 1.5pixel/imagem$ ). Os dados referentes a **S1** aparecem como linha traço-ponto-traço [-.-]. Os dados referentes a **S2** aparecem em linha tracejada [- -]. A sintonização **S3** aparece em linha a cheio [-]. Em cima à esquerda: Acção do filtro de Kalman. Sinal de entrada (linha pontuada [..]) e saídas para as diferentes sintonizações. Em cima à direita: Estimção do erro angular de seguimento. Erro angular real (linha pontuada [..]) e estimativa das diferentes sintonizações. Em baixo à esquerda: Estimção da velocidade angular do alvo. Velocidade angular real (linha pontuada [..]) e estimativa das diferentes sintonizações. Em baixo à direita: Estimção da aceleração angular do alvo. Aceleração angular real (linha pontuada [..]) e estimativa das diferentes sintonizações.

to visual conduzem a comportamentos de seguimento muito oscilatórios. A procura duma maior capacidade de reacção torna o sistema mais vulnerável em relação aos erros de estimação visual. Uma boa sintonização do filtro passa sempre por estabelecer um compromisso entre reacção e suavidade.

Considere um sinal com largura de banda  $BW$ . A sua derivada de primeira ordem é um sinal com uma largura de banda  $BW_d^1$  maior que  $BW$ . Por sua vez a derivada de segunda ordem é um sinal em que a respectiva largura de banda  $BW_d^2$  é superior a  $BW_d^1$ , e assim sucessivamente. Decorre do teorema da amostragem de Nyquist que são necessárias elevadas frequências de amostragem da informação visual no espaço e no tempo de forma a ser possível o cálculo correcto das derivadas de segunda ordem. Obviamente que a quantificação destas frequências depende da natureza do sinal e do movimento na imagem. De uma forma geral a tecnologia corrente de aquisição de imagem proporciona frequências de amostragem insuficientes. Assim, em termos práticos, a medição directa da aceleração na imagem tende a ser de tal forma ruidosa que os resultados obtidos são de pouca ou nenhuma utilidade. A filtragem de Kalman tenta inferir a aceleração angular do alvo a partir de informação de posição e velocidade na imagem. Observando a Fig. 3.8 verifica-se que a eficiência da estimação da aceleração é substancialmente inferior à da velocidade e posição angular. Tal facto é natural face ao explanado quanto à dificuldade na medição de aceleração a partir de informação visual. A estimativa da aceleração varia de forma suave convergindo lentamente para o valor real. A escolha duma sintonização em que  $\sigma_a$  tome valores mais elevados permitiria uma convergência mais rápida. No entanto a filtragem de ruído diminuiria passando a existir componentes de alta frequência que conduziriam a comportamentos oscilatórios durante o processo de seguimento.

A estimação de aceleração a partir de informação visual é um processo complexo e com muitos erros. Muitos autores preferiram assumir modelos de movimento uniforme em vez de uniformemente acelerado [20] de forma a evitarem a estimação da aceleração. O primeiro modelo é um caso particular do segundo em que a aceleração é nula. Fazendo  $\sigma_a = 0$  e atendendo a que o vector de estado inicial  $u_0$  é nulo a equação de processo 3.9 descreve movimento uniforme e o filtro de Kalman passa a assumir um modelo de velocidade constante. O resultado desta sintonização  $S3$  pode ser observado na Fig. 3.8. Note que em ambas as sintonizações  $S1$  e  $S2$  existe um "overshoot" na estimação da velocidade sempre que a aceleração do movimento

muda de sinal. As deficiências na medição da aceleração fazem com que o sistema demore algum tempo a aperceber-se que o movimento passou de acelerado/desacelerado a desacelerado/acelerado. Os resultados obtidos com a sintonização  $S3$  são muito melhores para esta situação particular.

A Fig. 3.9 compara o desempenho das sintonizações  $S1$ ,  $S2$  e  $S3$  na compensação de diferentes entradas de perturbação. Os dois primeiros gráficos referem-se a uma perturbação em parábola. A sintonização  $S2$  é a que apresenta melhores resultados de regulação. O filtro é bem sucedido na estimação da aceleração constante compensando os erros de medição na imagem. A diferença na regulação é particularmente notória quando a velocidade sobe acima da gama mensurável pelo algoritmo multi-escalar descrito na secção anterior. Os gráficos da segunda linha da Fig. 3.9 mostram a regulação para uma perturbação em sinusóide. O desempenho da sintonização  $S2$  é muito superior ao da sintonização  $S1$ . A comparação das sintonizações  $S2$  e  $S3$  é mais complexa. A sintonização  $S2$  apresenta melhores resultados quando a velocidade varia de forma aproximadamente linear. No entanto, quando o alvo sofre mudanças de sinal na aceleração a sintonização  $S3$  apresenta uma menor inércia de reacção e, conseqüentemente, melhores desempenhos. Os dois últimos gráficos exibem a resposta do sistema para uma trajectória altamente não linear que confirmam as observações feitas até aqui.

Desta discussão pode-se concluir que a qualidade do processamento visual está intimamente associada com a sintonização do filtro de Kalman. As melhorias nas medições de posição e velocidade na imagem devem ser acompanhadas por um nova sintonização de forma a otimizar o desempenho global do sistema. Por outro lado a utilização de um modelo de movimento uniforme ou uniformemente acelerado depende muito do tipo de movimento a seguir. Se a velocidade variar de uma forma aproximadamente linear o filtro de Kalman é capaz de estimar razoavelmente a aceleração. Neste caso o desempenho do sistema é melhor se for assumido movimento uniformemente acelerado. Caso contrário o modelo de movimento uniforme permite obter melhores resultados.

### 3.4 Otimização do Controlo da Vergência

O seguimento binocular do sistema MDOF utiliza os dois graus de liberdade do pescoço (azimute e elevação) para alinhar o alvo com o eixo óptico da câmara ciclópica. O algoritmo de vergência tem como objectivo controlar

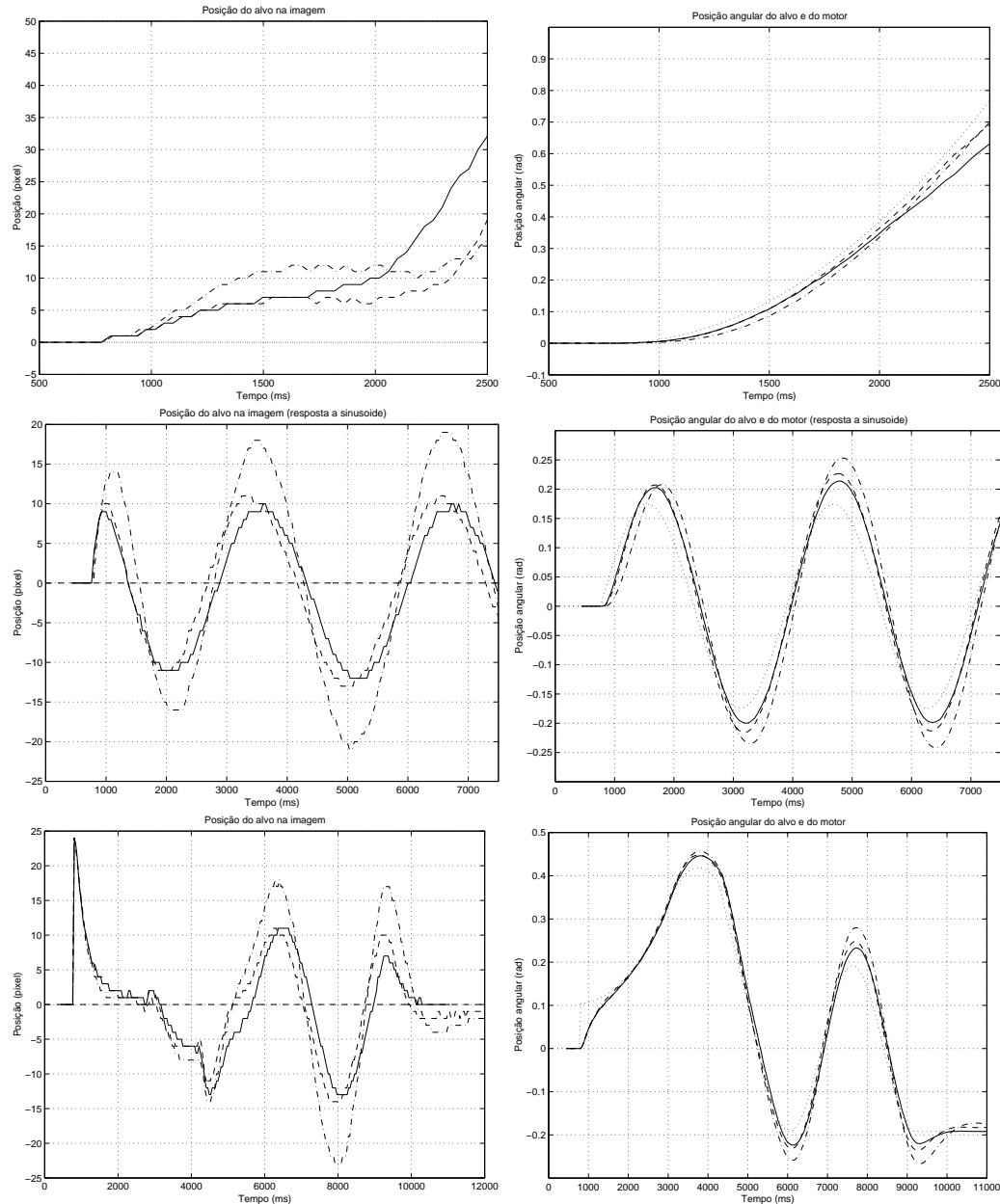


Figura 3.9: Avaliação do desempenho das diferentes sintonizações do filtro de Kalman. Os dados referentes a **S1** aparecem como linha traço-ponto-traço [-.-]. Os dados referentes a **S2** aparecem em linha tracejada [- -]. A sintonização **S3** aparece em linha a cheio [-]. Os resultados da linha de cima foram obtidos usando como sinal de teste uma perturbação em parábola com aceleração  $30^\circ/s^2$ . A segunda linha refere-se à perturbação sinusoidal. Os dados exibidos na última linha foram obtidos para uma perturbação composta por uma parábola, rampa e sinusóide. A primeira coluna mostra os resultados de regulação na imagem. O centro da imagem é assinalado pela linha tracejada [- -]) e as outras linhas dão a posição do alvo na imagem para as diferentes sintonizações. A segunda coluna mostra os resultados da regulação servo-mecânica. A posição real do alvo aparece como a linha pontuada [..] e as outras linhas mostram a posição angular do motor para as diferentes sintonizações.

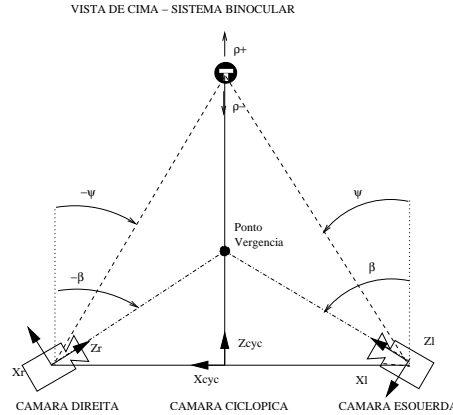


Figura 3.10: Sistema binocular visto de cima. A distância entre as câmaras é  $2B$ . Os ângulos de azimute do olho esquerdo e direito são respectivamente  $\beta_L$  e  $\beta_R$ . A vergência é simétrica e por isso  $\beta_L = -\beta_R = \beta$  em que  $\beta$  é designado por ângulo de vergência. O alvo move-se ao longo do eixo ciclópico sendo  $\rho(t)$  a respectiva função de posição.  $\psi$  e  $\dot{\psi}$  representam a posição e velocidade angular para um seguimento perfeito em vergência.

o azimute dos olhos de forma ao objecto ser projectado no centro das duas retinas. A Fig. 1.6 esquematiza a estrutura do controlo em vergência. Este controlo é feito através da disparidade em posição e velocidade entre as duas imagens (equações 1.3 e 1.4). Assim o objectivo do processamento visual é estimar a posição e velocidade horizontal do alvo nas duas retinas. Note que o algoritmo de visão é similar ao utilizado para o seguimento monocular e que as optimizações discutidas no inicio do capítulo são igualmente aplicáveis. Esta secção aborda a questão do filtro de Kalman a utilizar e faz a avaliação de desempenho do controlo da vergência apresentando a resposta do sistema às perturbações deduzidas no capítulo anterior.

### 3.4.1 Projecto do Filtro de Kalman

A Fig. 3.10 esquematiza o processo de seguimento da vergência. As duas câmaras estão separadas por uma distância  $2B$ .  $\beta_L$  e  $\beta_R$  são, respectivamente, a posição angular em azimute da câmara esquerda e direita. A vergência é simétrica o que significa que  $\beta_L = -\beta_R = \beta$  em que  $\beta$  é designado por ângulo de vergência. Assuma que o alvo se movimentava ao longo do eixo ciclópico e que  $\rho(t)$  é a respectiva função de posição.  $\psi$  é o ângulo de vergência ideal.

O sistema diz-se vergido, ou seja o ponto de vergência coincide com o alvo, sempre que  $\beta = \psi$ .  $\dot{\psi}$  representa a velocidade angular ideal de vergência. O filtro de Kalman é utilizado para estimar a posição e velocidade de vergência a partir da informação extraída das imagens.

### Equação de medida do Filtro de Kalman

$$\begin{cases} \Delta\psi = \arctan(\frac{\Delta x}{2f}) \approx \frac{\Delta x}{2f} \\ \dot{\psi} = \frac{\Delta \dot{x}}{2f} \end{cases} \quad (3.13)$$

Considere que  $\Delta\psi$  é o erro em posição angular da vergência. Note que  $\Delta\psi = \psi - \beta$  em que  $\psi$  e  $\beta$  são respectivamente o ângulo de vergência ideal e actual. Sempre que o erro  $\Delta\psi$  é nulo, ou seja  $\beta = \psi$ , o sistema diz-se vergido. A equação 1.3 estabelece uma relação entre a diferença das posições de projecção do alvo nas duas retinas (disparidade em posição) e  $\Delta\psi$ . Por sua vez a equação 1.4 relaciona a velocidade ideal de vergência  $\dot{\psi}$  com a disparidade da velocidade induzida pelo alvo nas imagens. Estas duas relações são linearizadas por expansão em série de Taylor e reescritas em 3.13.  $\Delta x$  é a disparidade de posição na imagem ( $\Delta x = x_l - x_r$ ).  $\Delta \dot{x}$  é a disparidade em velocidade, ou seja é a diferença entre a velocidade estimada na imagem esquerda e direita após cancelamento de "egomotion" ( $\Delta \dot{x} = \dot{x}_l - \dot{x}_r$ ).

$$\mathbf{v}(\mathbf{n}) = 2f \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{u}(\mathbf{n}) + \mathbf{r}(\mathbf{n}) \quad (3.14)$$

A equação de medida 3.14, directamente deduzida de 3.13, estabelece a relação entre o vector de estado  $\mathbf{u}$  e o vector de medida  $\mathbf{v}$ .  $\mathbf{u} = (\Delta\psi, \dot{\psi})^t$ ,  $\mathbf{v} = (\Delta x, \Delta \dot{x})^t$  e  $\mathbf{r}$  é um termo de ruído branco com média zero.

### Equação de processo do Filtro de Kalman

$$\begin{cases} \rho(n) = \rho(n-1) + \dot{\rho}(n-1)T_i \\ \dot{\rho}(n) = \dot{\rho}(n-1) \end{cases} \quad (3.15)$$

O filtro de Kalman usado no algoritmo de seguimento monocular considera que o alvo tem um movimento uniformemente acelerado em torno do eixo de rotação da câmara. Para o controlo de vergência vamos assumir que o movimento do alvo ao longo do eixo ciclópico entre os instantes de aquisição



de imagem   descrito por um modelo de velocidade constante. Assim, a equa o 3.15 permite estimar os par metros actuais de movimento do alvo conhecendo a sua posi o e velocidade na itera o anterior.  $T_i$  representa o per odo de amostragem que   tipicamente  $40ms$

$$\psi(n) = \arctan\left(\frac{B}{\rho(n)}\right) \approx \frac{B}{\rho(n)} \quad (3.16)$$

$$\dot{\psi}(n) \approx -\frac{B}{\rho^2(n)}\dot{\rho}(n) \quad (3.17)$$

Considere  $\psi(n)$  como sendo o  ngulo de verg ncia ideal no instante  $nT_i$ . Se  $\beta(n) = \psi(n)$  o ponto de verg ncia coincide com o objecto e o alvo   projectado no centro das duas retinas. A equa o 3.16 estabelece a rela o entre a posi o do alvo sobre o eixo cicl pico  $\rho(n)$  e o  ngulo  $\psi(n)$  atendendo a que  $\rho(n) \gg B$ . A equa o 3.17 relaciona a velocidade angular ideal  $\dot{\psi}(n)$  com a velocidade linear do alvo  $\dot{\rho}(n)$ . Este resultado   obtido derivando 3.16 em ordem ao tempo.

$$\psi(n) \approx \frac{B}{\rho(n-1) + \dot{\rho}(n-1)T_i} \quad (3.18)$$

$$\dot{\psi}(n) \approx -\frac{B}{(\rho(n-1) + \dot{\rho}(n-1)T_i)^2}\dot{\rho}(n-1) \quad (3.19)$$

As equa es 3.18 e 3.19 s o deduzidas a partir de 3.16 e 3.17.  $\rho(n)$  e  $\dot{\rho}(n)$  s o escritos em fun o de  $\rho(n-1)$ ,  $\dot{\rho}(n-1)$  e  $T_i$  de acordo com o modelo de movimento rectil neo uniforme da equa o 3.15.

$$\psi(n) \approx \frac{\psi^2(n-1)}{\psi(n-1) - \dot{\psi}(n-1)T_i} \quad (3.20)$$

$$\dot{\psi}(n) \approx \frac{\psi^2(n-1)}{(\psi(n-1) - \dot{\psi}(n-1)T_i)^2}\dot{\psi}(n-1) \quad (3.21)$$

De 3.16 e 3.17 vem que  $\rho(n) \approx \frac{B}{\psi(n)}$  e que  $\dot{\rho}(n) \approx -\frac{B}{\psi^2(n)}\dot{\psi}(n)$ . Substituindo em 3.18 e 3.19 para o instante  $n-1$  obt m-se 3.20 e 3.21. Estas equa es permitem estimar a posi o e velocidade ideal de verg ncia conhecendo os valores desses mesmos par metros angulares no instante de aquisi o anterior.

$$\begin{cases} \Delta\psi(n) \approx \frac{(\Delta\psi(n-1)+\beta(n-1))^2}{\Delta\psi(n-1)+\beta(n-1)-\dot{\psi}(n-1)T_i} - \beta(n) & +q_0(n-1) \\ \dot{\psi}(n) \approx \frac{(\Delta\psi(n-1)+\beta(n-1))^2}{(\Delta\psi(n-1)+\beta(n-1)-\dot{\psi}(n-1)T_i)^2} \dot{\psi}(n-1) & +q_1(n-1) \end{cases} \quad (3.22)$$

O vector de estado  $\mathbf{u}$  é composto pelo erro angular de posição  $\Delta\psi$  e pela velocidade  $\dot{\psi}$ . Já vimos que  $\Delta\psi = \psi - \beta$  o que implica que  $\psi(n) = \Delta\psi(n) + \beta(n)$  e que  $\psi(n-1) = \Delta\psi(n-1) + \beta(n-1)$ . Fazendo a substituição em 3.20 e 3.21 obtém-se a equação de transição de estados 3.22. 3.22 é a equação de processo do filtro de Kalman assumindo que o alvo descreve um movimento rectilíneo uniforme ao longo do eixo ciclópico. O ângulo de vergência  $\beta$  é lido em cada instante usando os codificadores ópticos dos motores.  $q_0$  e  $q_1$  são termos de ruído branco.

$$\Psi(\mathbf{n} - 1) = \begin{bmatrix} \frac{(\Delta\psi(n-1)+\beta(n-1))(\Delta\psi(n-1)+\beta(n-1)-2\dot{\psi}(n-1)T_i)}{(\Delta\psi(n-1)+\beta(n-1)-\dot{\psi}(n-1)T_i)^2} \\ -2 \frac{(\Delta\psi(n-1)+\beta(n-1))\dot{\psi}^2(n-1)T_i}{(\Delta\psi(n-1)+\beta(n-1)-\dot{\psi}(n-1)T_i)^3} \\ - \frac{(\Delta\psi(n-1)+\beta(n-1))^2 T_i}{(\Delta\psi(n-1)+\beta(n-1)-\dot{\psi}(n-1)T_i)^2} \\ \frac{(\Delta\psi(n-1)+\beta(n-1))^2 (\Delta\psi(n-1)+\beta(n-1)+\dot{\psi}(n-1)T_i)}{(\Delta\psi(n-1)+\beta(n-1)-\dot{\psi}(n-1)T_i)^3} \end{bmatrix} \quad (3.23)$$

A equação dinâmica 3.22 é não-linear. O Filtro de Kalman Estendido (FKE) é um algoritmo de estimação não-óptimo para sistemas não-lineares. Este filtro utiliza a expansão em série de Taylor das equações não-lineares para calcular as covariâncias [43]. Neste caso particular a equação de transição de estados é linearizada no último estado estimado através de expansão em série de Taylor de primeira ordem (filtro de primeira ordem). A matriz Jacobiana obtida é apresentada em 3.23.

### As matrizes de covariância

$$\begin{aligned} m_p(n) = E[\hat{\rho}(n)] &= \rho(n-1) + \dot{\rho}(n-1)T_i \\ &= \frac{B}{\Delta\psi(n-1)+\beta(n-1)} - \frac{B\dot{\psi}(n-1)T_i}{(\Delta\psi(n-1)+\beta(n-1))^2} \end{aligned} \quad (3.24)$$

$$\begin{aligned}
m_v(n) = E[\hat{\rho}(n)] &= \dot{\rho}(n-1) \\
&= -\frac{B\dot{\psi}(n-1)}{(\Delta\psi(n-1)+\beta(n-1))^2}
\end{aligned} \tag{3.25}$$

A equação 3.15 utiliza o modelo de velocidade constante para estimar a posição e velocidade do alvo sobre o eixo ciclópico. As estimativas obtidas em cada instante de aquisição têm um certo grau de incerteza o que implica uma caracterização estatística. Considere que  $\hat{\rho}(n)$  e  $\hat{\dot{\rho}}(n)$ , as estimativas de posição e velocidade, são variáveis aleatórias com média  $m_p(n)$  e  $m_v(n)$ . A média estatística da estimativa de posição num certo instante de tempo é o próprio valor estimado. A equação 3.24 permite calcular o valor de  $m_p(n)$  em função da posição e velocidade do alvo no instante de tempo  $n-1$ . Esta equação pode também ser escrita em função dos componentes do vector de estado  $\Delta\psi(n)$  e  $\dot{\psi}(n)$  (equação 3.16 e 3.17). A equação 3.25 calcula a média estatística da estimativa da velocidade.

Assuma ainda que  $\sigma_p$  e  $\sigma_v$  representam o desvio padrão de  $\hat{\rho}(n)$  e  $\hat{\dot{\rho}}(n)$ . O desvio padrão é constante ao longo do tempo e as duas variáveis aleatórias são não correlacionadas.

$$\Delta\hat{\psi}(n) = \frac{B}{\hat{\rho}(n)} \approx \frac{B}{m_p(n)} - \frac{B}{m_p^2(n)}(\hat{\rho}(n) - m_p(n)) \tag{3.26}$$

$$E[\Delta\hat{\psi}(n)] = \frac{B}{m_p(n)} - \beta(n) \tag{3.27}$$

$$E[(\Delta\hat{\psi}(n) - E[\Delta\hat{\psi}(n)])^2] = \frac{B^2}{m_p^4(n)}\sigma_p^2 \tag{3.28}$$

O filtro de Kalman tem por objectivo estimar o erro em posição e a velocidade da vergência. A estimativa do erro em posição da vergência  $\Delta\hat{\psi}(n)$  é função da estimativa de posição  $\hat{\rho}(n)$ . Uma função de uma variável aleatória é também uma variável aleatória. A equação 3.26 mostra essa dependência funcional. É feita uma expansão em série de Taylor em torno da média de  $\hat{\rho}(n)$  (ver equação 3.24) [44]. A média e a variância de  $\Delta\hat{\psi}(n)$  são calculadas pelas equações 3.27 e 3.28 deduzidas a partir da expansão em série de Taylor de 3.26.

$$\begin{aligned}
\hat{\dot{\psi}}(n) &= -\frac{B}{\hat{\rho}^2(n)}\hat{\dot{\rho}}(n) \\
&\approx -\frac{B}{m_p^2(n)}m_v(n) + \frac{2B}{m_p^3(n)}m_v(n)(\hat{\rho}(n) - m_p(n)) - \frac{B}{m_p^2(n)}(\hat{\dot{\rho}}(n) - m_v(n))
\end{aligned} \tag{3.29}$$

$$E[\hat{\psi}(n)] = -\frac{B}{m_p^2(n)}m_v(n) \quad (3.30)$$

$$E[(\hat{\psi}(n) - E[\hat{\psi}(n)])^2] = \frac{4B^2}{m_p^6(n)}m_v^2(n)\sigma_p^2 + \frac{B^2}{m_p^4(n)}m_v^2(n)\sigma_v^2 \quad (3.31)$$

A estimativa da velocidade ideal de vergência  $\hat{\psi}(n)$  é função de  $\hat{\rho}(n)$  e  $\hat{\dot{\rho}}(n)$  (equação 3.29). As equações 3.30 e 3.31 calculam a média e variância de  $\hat{\psi}(n)$  assumindo que  $\hat{\rho}(n)$  e  $\hat{\dot{\rho}}(n)$  são não correlacionadas.

$$E[(\Delta\hat{\psi}(n) - E[\Delta\hat{\psi}(n)])(\hat{\psi}(n) - E[\hat{\psi}(n)])] = -\frac{2B^2m_v\sigma_p^2}{m_p^5} \quad (3.32)$$

A equação 3.32 calcula o valor da covariância entre  $\Delta\hat{\psi}(n)$  e  $\hat{\psi}(n)$ . Este resultado é facilmente deduzido a partir da expansão em série de Taylor da equação 3.26 e 3.29.

$$\mathbf{Q}_{\text{verg}}(n) = \frac{B^2}{m_p^4(n)} \begin{bmatrix} \sigma_p^2 & -\frac{2m_v(n)\sigma_p^2}{m_p(n)} \\ -\frac{2m_v(n)\sigma_p^2}{m_p(n)} & \frac{4m_v^2(n)}{m_p^2(n)}\sigma_p^2 + \sigma_v^2 \end{bmatrix} \quad (3.33)$$

A matriz de covariância do modelo é dada pela equação 3.33. A matriz  $\mathbf{Q}_{\text{verg}}(n)$  é directamente deduzida dos resultados de 3.28, 3.31 e 3.32.

$$\mathbf{R}_{\text{verg}} = \begin{bmatrix} \sigma_{pi}^2 & 0 \\ 0 & \sigma_{vi}^2 \end{bmatrix} \quad (3.34)$$

$\mathbf{R}_{\text{verg}}$  é a matriz de covariância das medidas que se assume constante ao longo do tempo. A sintonização do filtro de Kalman é feita escolhendo valores adequados para os parâmetros  $\sigma_p$ ,  $\sigma_v$ ,  $\sigma_{pi}$  e  $\sigma_{vi}$ .

### 3.4.2 Avaliação do Desempenho

A Fig. 3.11 permite observar o desempenho do filtro de Kalman apresentado na secção anterior. O alvo movimenta-se ao longo do eixo ciclópico variando a sua distância ao sistema entre 0.5 e 10.5 metros. A função de posição  $\rho(t)$  é sinusoidal com amplitude 10m e velocidade angular 1.5rad/s. O gráfico em baixo à direita mostra a posição do objecto sobre o eixo ciclópico ao longo do tempo (linha pontuada). Nesse mesmo gráfico pode observar a variação da matriz de ganho do Kalman. Note que à medida que o objecto se afasta

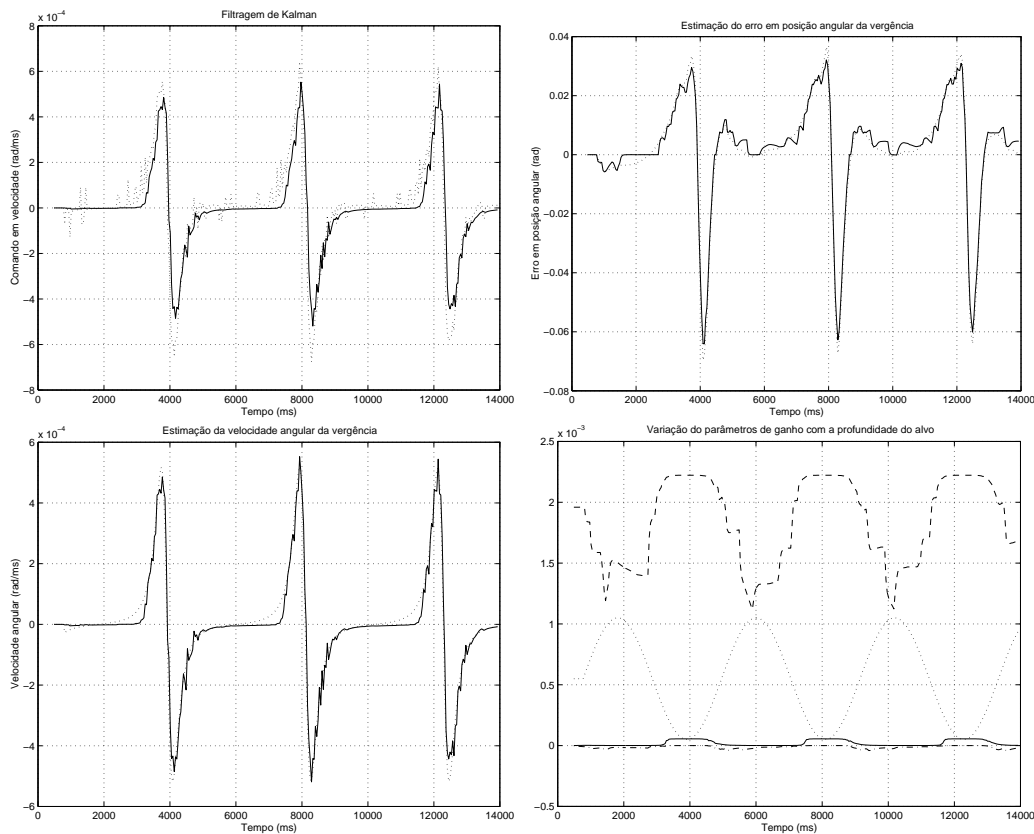


Figura 3.11: Filtragem de Kalman no controlo da vergência (desempenho para uma variação sinusoidal da profundidade do alvo (amplitude 5 metros, velocidade angular 1.5 rad/s, ponto de partida 5.5 metros)). Em cima à esquerda: Acção do filtro de Kalman. Sinal de entrada (linha pontuada [..]) e sinal de saída (linha a cheio[-]). Em cima à direita: Estimação do erro em posição angular da vergência. Erro angular real (linha pontuada [..]) e estimativa do erro angular (linha a cheio [-]). Em baixo à esquerda: Estimação da velocidade angular da vergência. Velocidade angular real (linha pontuada [..]) e estimativa da velocidade angular (linha a cheio [-]). Em baixo à direita: Variação da matriz de ganho do filtro de Kalman com a profundidade do alvo. A posição do alvo sobre o eixo ciclópico em dezenas de quilómetros (linha pontuada[.]). O escalamento em dezenas de quilómetros é feito por razões gráficas (a profundidade do alvo varia sinusoidalmente entre 0.5 e 10.5 metros). A matriz de ganho  $\mathbf{K}$  é uma matriz  $2 \times 2$  com estrutura  $\mathbf{K} = [k_{11}k_{12}; 0k_{22}]$ . O parâmetro  $k_{11}$  aparece com linha tracejada [- -],  $k_{21}$  com linha pontuada/tracejada [-.] e  $k_{22}$  com linha a cheio [-].

do sistema os valores da matriz de ganho diminuem. Isto é explicado pelo facto de que um mesmo movimento do alvo no espaço induz diferentes deslocamentos na imagem dependendo da distância a que o objecto se encontra do sistema. Imagine que o objecto se está a movimentar sobre o eixo ciclópico com uma velocidade constante positiva. A velocidade induzida na imagem diminui à medida que o alvo se afasta. A partir de certo ponto a resolução espacial não é suficiente para detectar o deslocamento no plano imagem e o processamento visual torna-se incapaz de medir a variação de profundidade. Assim, à medida que o alvo se afasta, o filtro de Kalman aumenta a confiança no modelo de movimento em detrimento dos resultados do processamento visual que lhe chegam.

Considere que a posição do alvo sobre o eixo ciclópico muda subitamente. Isto representa um perturbação em degrau para o sistema de controlo de vergência. A Fig 3.12 mostra o desempenho do sistema para entradas em degrau. Nos gráficos de cima o alvo, inicialmente a 0.8 metros, aproxima-se subitamente da cabeça MDOF ficando a 0.5 metros. O sistema consegue compensar completamente a perturbação fazendo coincidir o ponto de vergência com a nova posição do objecto. Os dois gráficos de baixo mostram o desempenho do sistema quando o alvo se afasta subitamente passando a profundidade de 1 metro para 10 metros. A resposta do sistema apresenta um erro em regime final. O alvo encontra-se demasiado afastado e a resolução espacial das imagens não é suficiente para estimar com rigor a profundidade.

No capítulo anterior definiu-se uma perturbação em rampa como o movimento ao longo do eixo ciclópico que induz uma disparidade de velocidade constante nas retinas (equação 2.27). A Fig. 3.13 mostra a resposta do sistema para uma perturbação em rampa que induz uma disparidade constante de -2 pixel/imagem.

A Fig. 3.14 mostra o desempenho do controlo em vergência para uma perturbação em sinusóide definida de acordo com os critérios do capítulo anterior. Note a latência de resposta quando ocorre uma inversão no sentido de movimento e a profundidade do alvo é elevada. Isto é explicado pelo facto do movimento do objecto não estar de acordo com o modelo de movimento rectilíneo uniforme assumido.

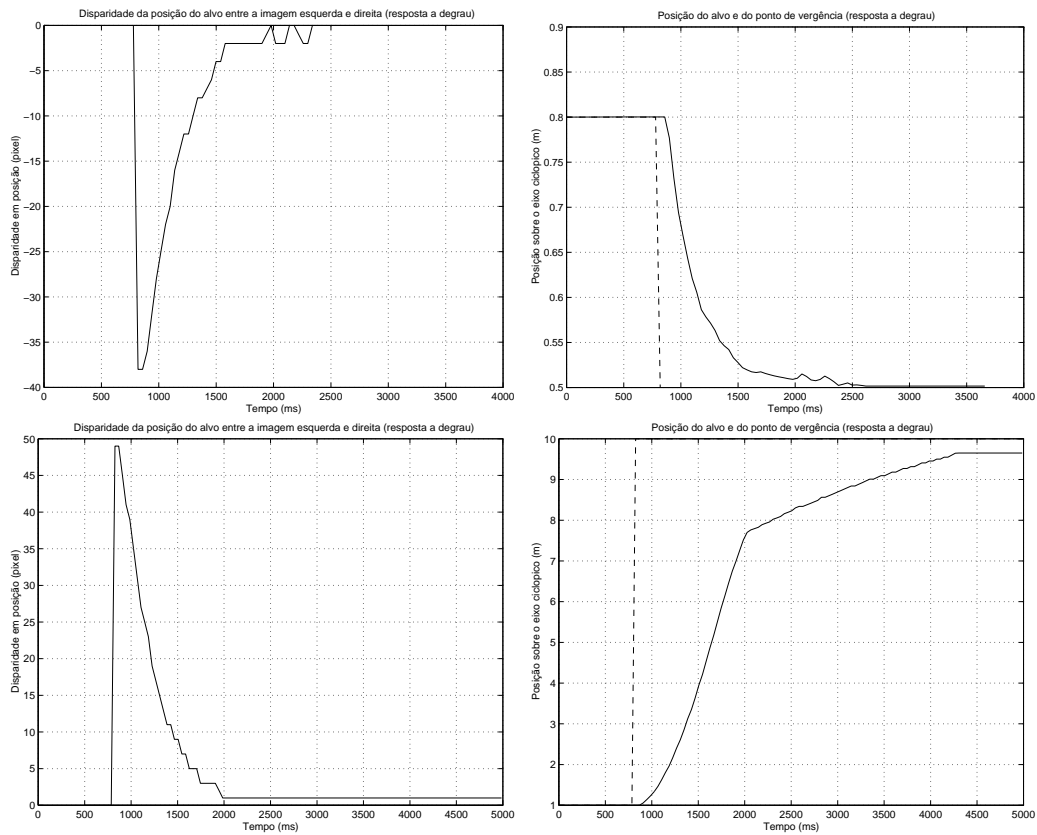


Figura 3.12: Resposta a degrau. Na linha de cima o degrau tem amplitude -0.3 metros, sendo a posiç o de partida 0.8 metros. Na linha de baixo a perturbaç o em degrau tem uma amplitude de 9 metros com posiç o inicial 1 metro. Esquerda: Regulaç o na imagem. Disparidade da posiç o de projec  o do alvo entre a imagem esquerda e direita (linha a cheio [-]). Direita: Resposta servo-mec nica. Posiç o angular do alvo  $\psi$  (linha tracejada [- -]) e do motor  $\beta$  (linha a cheio [-])

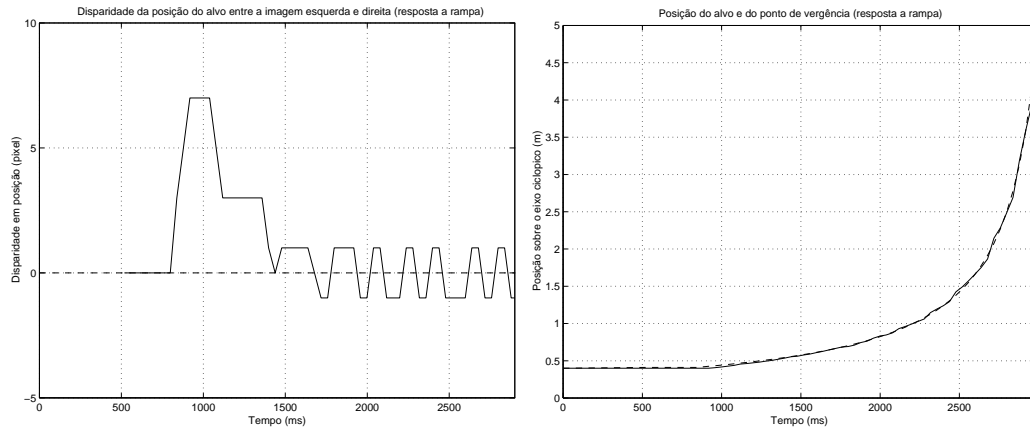


Figura 3.13: Resposta a rampa. Esquerda: Regulação na imagem. Disparidade da posição de projecção do alvo entre a imagem esquerda e direita (linha a cheio [-]). Direita: Resposta servo-mecânica. Posição angular do alvo  $\psi$  (linha tracejada [- -]) e do motor  $\beta$  (linha a cheio [-])

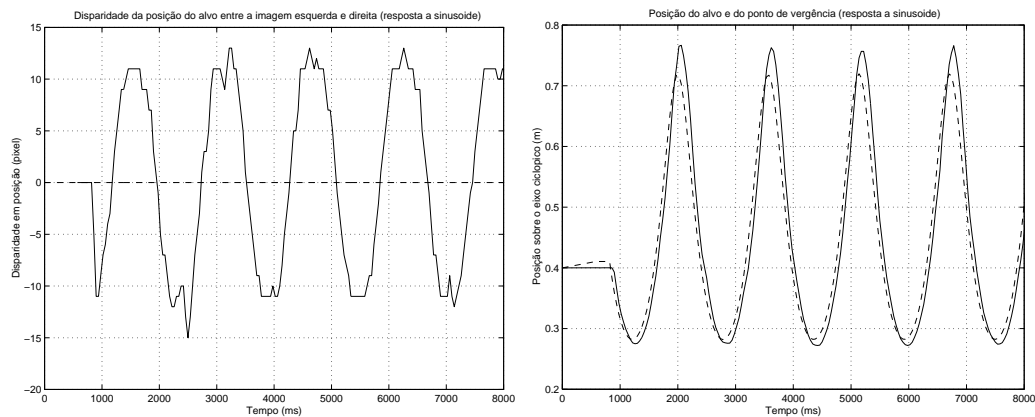


Figura 3.14: Resposta a senoide. Esquerda: Regulação na imagem. Disparidade da posição de projecção do alvo entre a imagem esquerda e direita (linha a cheio [-]). Direita: Resposta servo-mecânica. Posição angular do alvo  $\psi$  (linha tracejada [- -]) e do motor  $\beta$  (linha a cheio [-])



# Capítulo 4

## Arquitectura do Sistema e Controlo de Fixação

Este capítulo foca questões de arquitectura e controlo num sistema de seguimento utilizando visão. O estudo é feito para o seguimento monocular abordado nos capítulos precedentes. Muitos dos conceitos e ideias aqui apresentadas são generalizáveis para o seguimento em vergência bem como para outros problemas de controlo de movimento usando visão.

Os sistemas guiados por visão são sistemas cujas acções decorrem da informação extraída das imagens. No caso particular do seguimento monocular a câmara está montada sobre uma plataforma com dois graus de liberdade: azimute e elevação. Pretende-se controlar o movimento da plataforma de forma ao alvo ser projectado no centro da imagem. A Fig. 4 exhibe um esboço da arquitectura utilizada para implementar controlo visual de movimento em tempo-real. Podem ser identificados três processos distintos e concorrentes: o processamento visual das imagens, o controlo de baixo-nível do actuador, e o controlo de alto-nível de fixação.

Os capítulos precedentes abordam principalmente o processamento visual e a extracção de informação das imagens. A frequência deste primeiro processo é limitada pela frequência de aquisição de imagem (tipicamente 25Hz). Diferentes optimizações foram discutidas e implementadas de forma aos algoritmos serem rápidos, precisos e robustos. Isto é essencial para a implementação de controlo visual de movimento em tempo-real com alto desempenho. O facto das câmaras estarem montadas numa plataforma activa cria dificuldades adicionais ao processamento visual devido ao movimento auto-induzido nas imagens ("egomotion"). Utilizou-se filtragem de Kalman

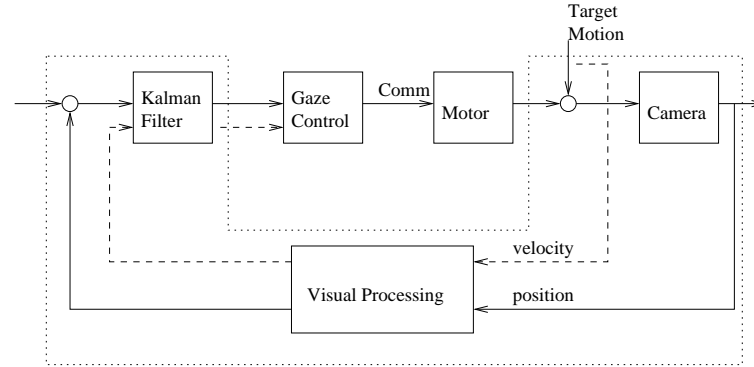


Figura 4.1: Arquitectura do algoritmo de seguimento visual. O controlo é feito utilizando simultaneamente informação de posição e velocidade. A figura mostra os três processos concorrentes: processamento visual (limitado pela linha tracejada), controlo de fixação e controlo do motor.

para estimar os parâmetros 3D do movimento do alvo e limitar o efeito dos erros de medição na imagem de forma a obter um comportamento de seguimento mais robusto e suave.

A função do processo de baixo-nível é controlar os actuadores da plataforma activa. A escolha do tipo de actuadores e as correspondentes estratégias de controlo são discutidas em [45, 21]. O sistema MDOF controla os servomotores através de malhas locais com um controlador PID. Esta malha local de controlo é analisada em detalhe no seguimento do capítulo.

O terceiro processo, o controlo de fixação, tem como objectivo estabelecer a ligação entre a malha de processamento visual e os actuadores do sistema. A entrada do controlador de fixação é a informação extraída das imagens e a saída são os comandos a serem enviados para a malha de controlo dos actuadores mecânicos. Uma das principais funções do controlador de fixação é lidar com o problema do atraso. Atrasos na malha directa e de realimentação afectam substancialmente o desempenho global do sistema de seguimento. Este é um assunto exclusivamente abordado em [19, 20, 46]. A latência introduzida pelo processamento visual e aquisição de imagem é uma das principais razões pela qual os sistemas baseados em visão são difíceis de controlar. Atrasos de comunicação e inércia mecânica dos actuadores também afectam o desempenho do global. O controlador de fixação é desenhado para compensar estes atrasos e aumentar a largura de banda do sistema.

## 4.1 Controlo de Baixo Nível dos Servo-Motores

Muitos sistemas de visão activos têm utilizado motores de passo para gerar o movimento [47]. Estes dispositivos apresentam em geral uma grande facilidade de "interface" e um controlo simples. Estas vantagens terão sido decisivas para a opção feita na construção desses sistemas de visão. No entanto, e comparativamente aos motores DC, os motores de passo têm desempenhos menos brilhantes em termos de aceleração e seguimento suave de velocidade. Estas duas últimas características são requisitos essenciais para um sistema de seguimento em tempo-real de elevado desempenho. A necessidade de elevadas acelerações e bom controlo de velocidade são argumentos fortes para a escolha de motores DC com caixa de engrenagem [1, 45]. A utilização de caixas de engrenagem permite obter momentos de rotação elevados com motores pequenos permitindo grandes acelerações e uma boa atenuação das perturbações em torque. No caso particular da cabeça MDOF o movimento é gerado por motores DC equipados com caixas de engrenagem "harmonic drive" e codificadores ópticos no eixo anterior. As caixas de engrenagem "harmonic drive" minimizam as folgas mecânicas apresentando um elevado factor de rigidez entre os eixos de saída do motor e da caixa de engrenagem. A principal desvantagem é a introdução de atrito estático o que conduz a comportamentos não-lineares.

A posição dos motores pode ser controlada directamente usando realimentação visual. Neste caso a frequência de controlo do motor é igual à frequência de aquisição de imagem (25Hz). Este período de amostragem é demasiado grande para o controlo eficiente dum motor DC. A Fig. 4.2 mostra a resposta em frequência do motor em malha aberta onde se pode observar o atraso de fase. O tempo para o processamento visual na malha de realimentação introduz um atraso de fase adicional muito significativo. Este atraso adicional tende a piorar a estabilidade global do sistema de seguimento. Uma solução para lidar com o problema é diminuir os valores de ganho de forma a manter a estabilidade em malha fechada. Nesta situação o sistema tem sempre desempenhos pobres devido à baixa frequência de amostragem e ao baixo ganho de controlo. Em resumo, o controlo dos actuadores utilizando directamente realimentação visual conduz inevitavelmente a sistemas de baixo desempenho. Isto porque o atraso de fase introduzido pelo processamento na malha de realimentação empurra o sistema para a instabilidade. A única forma de manter a estabilidade é diminuir os valores de ganho na

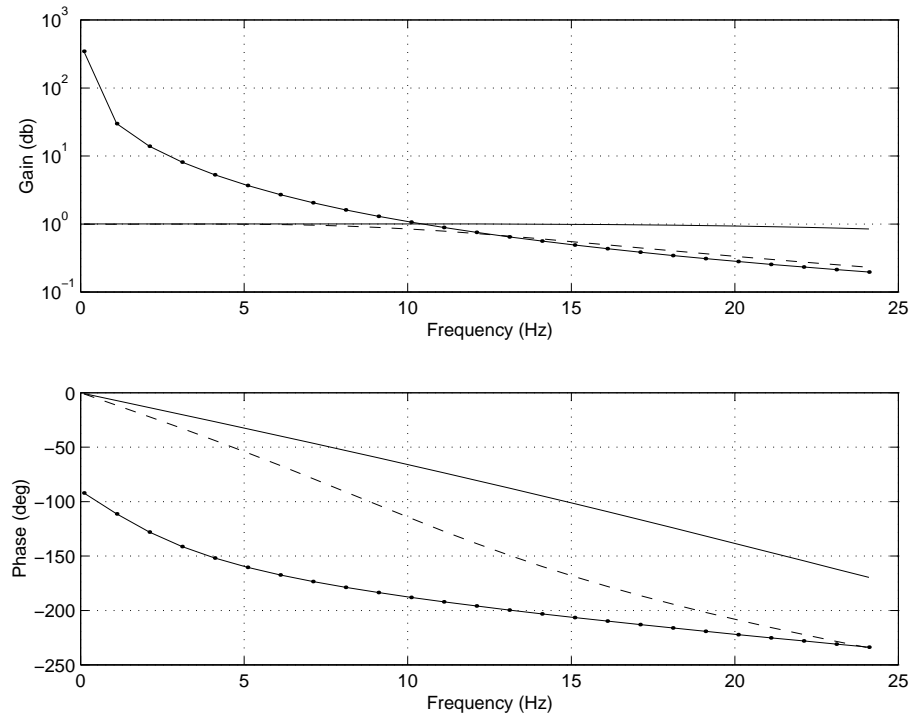


Figura 4.2: Resposta em frequência do motor (ganho em cima e fase em baixo). A linha a cheio com pontos sobrepostos [-.-] mostra a resposta em frequência do motor em malha aberta ("torque mode"). A linha tracejada [- -] refere-se à resposta do motor controlado em malha-fechada com um PID ("position mode"sem o gerador de trajectória). A linha a cheio [-] mostra a resposta em frequência quando o motor é controlado em malha-fechada e o sinal de referência é fornecido pelo gerador de trajectória ("velocity mode" com geração em rampa)

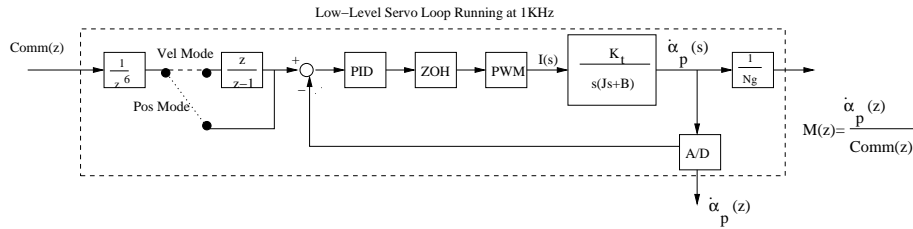


Figura 4.3: Esquema detalhado da malha de baixo-nível de controlo dos servo-motores. A malha tem uma frequência de amostragem de 1KHz. A comunicação é síncrona e tem um atraso de 6 instantes de amostragem (frequência de 166Hz). Os servo-módulos da DCX podem operar em modo de posição e de velocidade. No primeiro o sinal de referência é dado directamente pelos comandos de posição. No segundo existe um gerador de trajectória sendo o sinal de referência a integração do comando de velocidade. A malha de controlo de posição é realimentada pela informação dos codificadores. O sinal digital de actuação é dado por um filtro PID. Um ZOH é utilizado para fazer a conversão D/A estando a sua saída ligada a um amplificador PWM. O último bloco representa a redução feita pela caixa de engrenagem. O factor de rigidez entre o eixo do motor e da caixa é considerado unitário (não há folgas mecânicas).

malha.

Os actuadores do sistema MD0F estão equipados com codificadores ópticos no eixo anterior ("back shaft") que permitem conhecer a posição do motor com uma resolução de  $0.00036^\circ$  (100000 contagens/revolução). A informação de posição pode ser utilizada para implementar uma malha local de controlo do servo-motor com elevada frequência de amostragem. Isto permite melhorar a resposta em frequência do dispositivo mecânico.

O controlo dos graus de liberdade da cabeça MD0F é feito através de placas de controlo comerciais [3]. Cada uma destas placas pode ser equipada com um máximo de seis módulos dedicados que permitem controlar o mesmo número de servo-motores. A Fig. 4.3 esquematiza um desses módulos. Cada módulo dedicado controla o seu eixo implementando uma malha fechada com um filtro digital PID com frequência de amostragem de  $1\text{KHz}$ . A informação de posição do codificador é usada para realimentação. A comunicação entre o PC hospedeiro e cada módulo de servo-controlo é síncrona com uma frequência de 166Hz. Isto significa que o processo utilizador a correr no computador pode enviar comandos e ler a informação dos codificadores

em intervalos de 6ms. A utilização de placas de controlo dedicadas aumenta a modularidade do sistema e liberta tempo de processamento do PC hospedeiro que fica dedicado ao processamento visual e ao controlo de fixação. O uso de placas comerciais de controlo é uma solução com um custo aceitável e arquitecturalmente bem equilibrada para a construção de sistemas de visão activos de tempo-real.

A linha a tracejado da Fig. 4.2 mostra a resposta em frequência da malha fechada de controlo de posição do motor. O filtro PID foi sintonizado por colocação de pólos de forma a se obter uma resposta ligeiramente sub-amortecida. Esta sintonização, embora teoricamente trivial, é dificultada pelas não-linearidades introduzidas pelo atrito estático da caixa de engrenagem. O aumento da frequência de amostragem para  $1KHz$  e a sintonização adequada do controlador PID permite melhorar a resposta em frequência do actuador. Note que a margem de fase é consideravelmente aumentada o que permite aumentar o ganho da malha global de seguimento sem que o atraso de fase introduzido pelo tempo de processamento visual torne o sistema instável. Conclui-se assim que o controlo dos actuadores através de uma malha local com realimentação de posição permite obter um sistema de seguimento com um desempenho superior aquele que se obteria usando realimentação visual directa.

A placa de controlo DCX [3] permite diferentes modos de controlo dos eixos. O parágrafo anterior apresenta o modo de controlo em posição. Neste modo o processo utilizador, a correr no PC hospedeiro, envia uma comando de posição com uma frequência máxima de 166Hz. Este comando é a referência da malha de controlo exibida na Fig. 4.3 que tem uma frequência de amostragem de 1KHz. O tempo de comunicação entre o PC e a placa introduz uma atraso de fase indesejável. A placa DCX permite também o controlo em modo velocidade. Neste caso o processo utilizador envia um comando de velocidade que é integrado por um gerador de trajectórias. A referência da malha de controlo em posição é dada pela saída do gerador de trajectórias. O facto da placa permitir a opção de controlo em posição ou velocidade é uma característica extremamente útil para a operação de sistemas de visão activos. Certos tipos de comportamento, como o movimento em sacada, exigem controlo em posição dos actuadores, enquanto outros, como o seguimento suave, são mais facilmente implementados usando controlo em velocidade.

A linha a cheio da Fig. 4.2 mostra a resposta da malha de de controlo

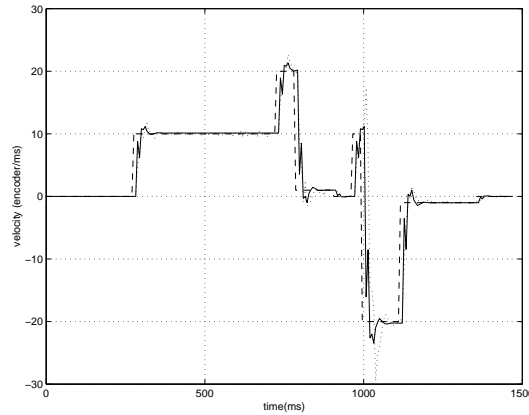


Figura 4.4: Identificação da malha de baixo-nível de controlo do motor. O motor é controlado em velocidade. O sinal de entrada é o comando de velocidade e o sinal de saída é a velocidade medida derivando a informação de posição do codificador em ordem ao tempo. A frequência de amostragem é de 166Hz. A linha tracejada [- -] mostra o sinal de entrada, a linha pontilhada [...] a resposta real do motor e a linha a cheio [-] a resposta prevista pelo modelo obtido

de velocidade do servo-motor. A entrada do sistema  $M(z)$  é o comando de velocidade enviado pelo processo utilizador e a saída é a velocidade medida derivando a informação de posição do codificador. No controlo em modo de posição a referência da malha é actualizada com uma frequência de 166Hz. Em modo de velocidade, embora os comandos sejam igualmente enviados de 6 em 6ms, a referência da malha de controlo de posição é actualizada com uma frequência de 1KHz. Isto explica a melhoria na margem de fase observável na Fig. 4.2. O algoritmo de seguimento da MDOF controla os servo-motores em modo de velocidade.

$$M(z) = \frac{\dot{\alpha}_p(z)}{Comm(z)} = \frac{0.8808z^2 - 0.5927z + 0.5588}{z^4 - 0.3784z^3 - 0.0017z^2 + 0.2162z} \quad (4.1)$$

A função de transferência da malha de controlo do motor é dada por 4.1. A entrada é o comando de velocidade enviado para a placa DCX ( $Comm(z)$ ) e a saída é a velocidade do motor medida através do codificador óptico ( $\dot{\alpha}_p(z)$ ). O modelo foi obtido usando técnicas de identificação lineares baseadas em minimização quadrática. Note que a malha de controlo tem um período de

amostragem de 1ms e que a saída em velocidade está a ser sub-amostrada a uma frequência de 166Hz (frequência de comunicação). Desta feita o modelo obtido não descreve alguns comportamentos de alta-frequência. A Fig. 4.4 compara a resposta do modelo com a resposta real do sistema para uma certa entrada.

## 4.2 Interpolação para Compensação do Atraso Visual

O atraso na captura das imagens é o intervalo de tempo que vai desde o instante em que a imagem é adquirida até ao momento em que ela se encontra disponível para ser processada. O valor estimado para o atraso na aquisição de imagens na cabeça MDOF é de 30ms. O atraso de processamento visual é o tempo que o sistema demora a correr os algoritmos visuais. No caso do seguimento monocular este intervalo de tempo é de 6ms, e no seguimento binocular, em que são processadas duas imagens, o atraso de processamento é de cerca de 11ms. Este capítulo foca o seguimento monocular, mas os conceitos aqui discutidos são generalizáveis ao seguimento binocular.

A frequência de aquisição de imagem é de 25Hz. Em cada 40ms é adquirida uma nova imagem que está pronta para ser processada 30ms depois. O processamento visual demora 6ms a estimar a posição e velocidade do alvo a partir da imagem. Assim, as estimativas dos parâmetros angulares de movimento do objecto à saída do filtro de Kalman (ver Fig. 4) têm um atraso de 36ms. Este atraso, que resulta da soma do tempo de aquisição com o tempo de processamento, designa-se por atraso visual.

Observe a Fig. 4.5. Os gráficos da direita referem-se à resposta a uma perturbação em rampa enquanto os da esquerda mostram a resposta à perturbação não-linear da Fig. 3.7. Assuma que o processamento visual é perfeito e que o erro na estimação dos parâmetros angulares de movimento é nulo. A informação de movimento do alvo é actualizada de 40 em 40ms com um atraso de 36ms. Os dois gráficos de cima comparam o comando ideal com o comando que é realmente enviado para o motor. O tempo de aquisição e processamento visual introduz um atraso de fase na malha de controlo que deteriora o desempenho global do sistema de seguimento. O atraso visual pode ser compensado utilizando interpolação.

A Fig. 4.6 esquematiza detalhadamente o sistema de seguimento monoc-



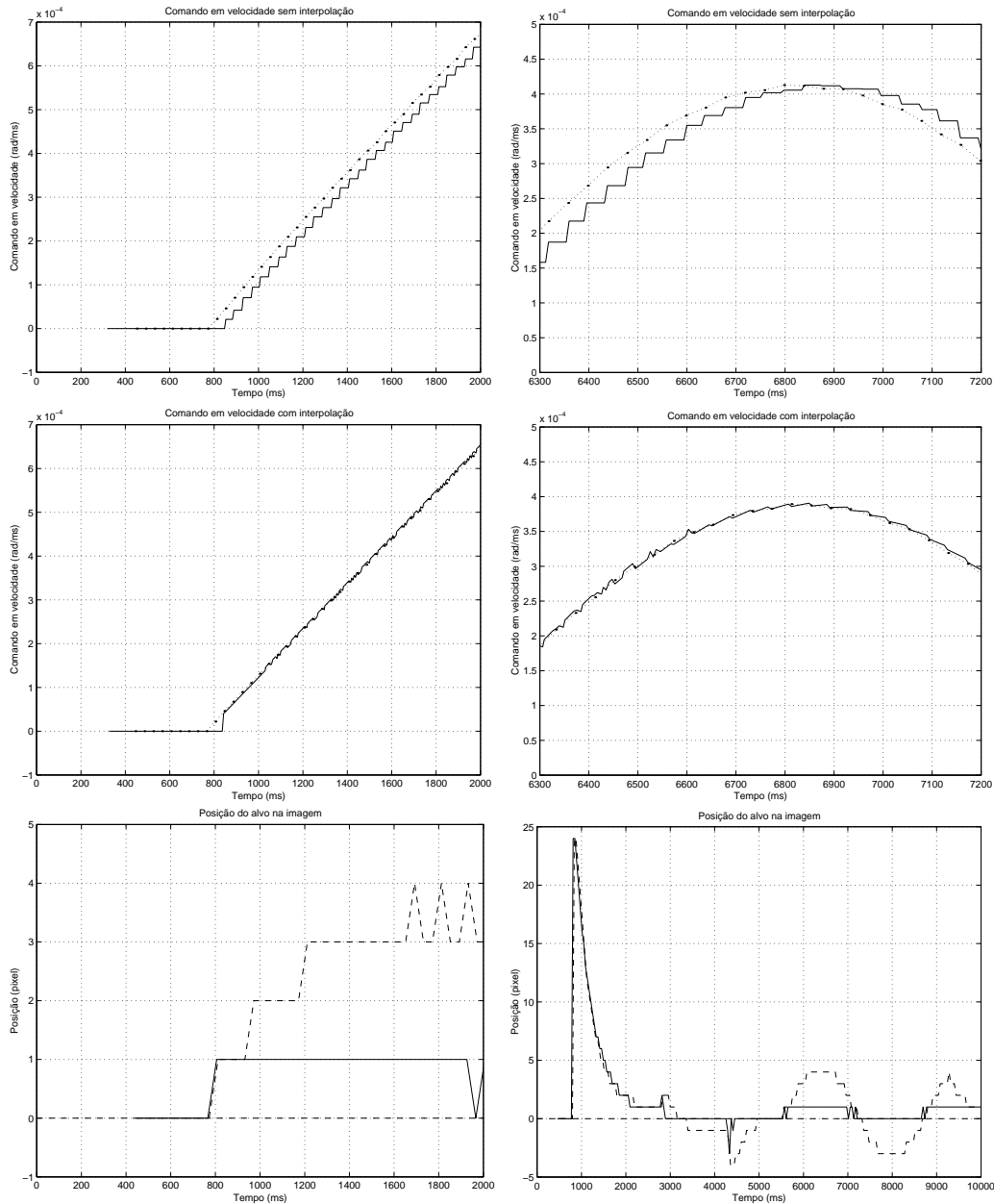


Figura 4.5: Uso de interpolação para compensar o atraso visual. Os gráficos da direita mostram a resposta do sistema a uma perturbação em rampa, os da esquerda a resposta a uma perturbação não-linear (ver Fig. 3.7). Considere que a estimação dos parâmetros de movimento do alvo é perfeita. O atraso na aquisição de imagem e no processamento visual é de 36ms. Em cima: A linha traço-ponto-traço [-.-] mostra o comando em velocidade ideal. A linha a cheio [-] mostra o comando que é enviado para o motor se não houver compensação do atraso visual. No meio: A linha traço-ponto-traço [-.-] mostra o comando em velocidade ideal. A linha a cheio [-] mostra o comando que é enviado para o motor quando se faz compensação do atraso visual usando interpolação. Em baixo: Regulação na imagem. A linha tracejada [- -] mostra a posição do alvo na imagem quando não há compensação do atraso visual. A linha a cheio [-] refere-se à posição do alvo quando se utiliza interpolação. A linha a tracejado horizontal assinala o centro da imagem.

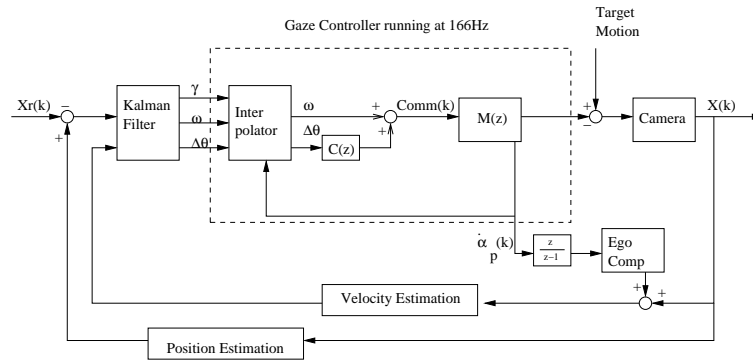


Figura 4.6: Diagrama de blocos do seguimento monocular. A malha de processamento visual envia o vector de estados  $\mathbf{u} = (\Delta\theta, \omega, \gamma)^t$  para o controlador de fixação. O interpolador recebe os parâmetros de movimento angular e faz a compensação do atraso visual.  $C(z)$  é um controlador PD (proporcional-derivativo).  $\text{Comm}(k)$  é o comando enviado para a malha de controlo do motor  $M(z)$

ular. A malha de processamento visual envia para o controlador de fixação estimativas dos parâmetros de movimento angular do alvo  $\mathbf{u} = (\Delta\theta, \omega, \gamma)^t$ . Como a frequência de amostragem é de 25Hz, o vector de estado  $\mathbf{u}$  é actualizado com uma periodicidade de 40ms. O interpolador recebe essa informação com um atraso de 36ms. Os parâmetros actuais de movimento são estimados conhecendo o atraso visual e assumindo que o alvo se movimenta com movimento angular uniformemente acelerado. O modelo de interpolação é similar ao modelo usado no Filtro de Kalman (equação 3.9).

A interpolação pode ser utilizada para compensar quer o atraso visual, quer a baixa frequência de amostragem da malha de processamento de imagem. A frequência de comunicação com a malha de controlo do servo-motor é de 166Hz. O controlador de fixação pode enviar comandos e receber leituras do codificador com uma periodicidade de 6ms. No entanto a informação visual chega com uma frequência de 25Hz. Isto implica que o comando enviado para o servo-motor seja constante durante 40ms. A interpolação é usada para compensar o atraso visual e actualizar o comando de velocidade entre a aquisição de duas imagens. Desta forma tira-se o máximo partido da frequência de comunicação com a malha de servo-controlo.

Os gráficos da segunda linha da Fig. 4.5 comparam o comando ideal com o comando que é enviado para o motor se for utilizada interpolação. A

convergência do comando real para o comando ideal é evidente. Observe o "ripple" no topo da sinusóide devido à variação não-linear dos parâmetros de movimento do alvo que não é descrita pelo modelo de aceleração constante usado na interpolação. A utilização de interpolação melhora claramente a capacidade de regulação do sistema como pode ser observado nos dois últimos gráficos.

O desempenho da interpolação depende de dois factores: a adequação do modelo de movimento escolhido e a qualidade dos resultados da estimação visual. A utilização de um modelo de movimento uniformemente acelerado é aceitável atendendo à frequência de aquisição visual. Embora possa ocorrer pontualmente o efeito de "ripple" observado, trata-se de um fenómeno que dificilmente afecta de forma significativa o desempenho global do sistema. Já a utilização de interpolação numa situação em que os parâmetros de movimento sejam mal estimados pode afectar de forma muito negativa a capacidade de seguimento.

Os dados exibidos na Fig. 4.5 foram obtidos assumindo que a estimação dos parâmetros angulares de movimento do alvo era perfeita. Na prática as medições na imagem são extremamente ruidosas. Quando se discutiu a sintonização do filtro de Kalman concluiu-se que, quando a velocidade angular do alvo variava muito, o processamento visual tinha muitas dificuldades em estimar correctamente o parâmetro de aceleração angular. Se a interpolação partir de um valor errado de aceleração angular é muito natural que, em vez de contribuir para uma maior capacidade de seguimento, acabe por piorar o desempenho global do sistema.

Fizeram-se uma série de testes, envolvendo diferentes tipos de perturbação, com vista a estudar a influência da interpolação no processo de seguimento. Verificou-se que a interpolação pode pontualmente piorar o desempenho do sistema. Isto acontece quando, por alguma razão, a estimativa de movimento proveniente da malha visual está muito desfasada da realidade. A interpolação amplifica o desempenho do processamento visual, tanto no sentido positivo como no negativo. Quando os parâmetros de movimento são correctamente estimados o uso de interpolação melhora o seguimento (como aliás de viu na Fig. 4.5). No caso da estimação ser muito deficiente o interpolador contribui para uma deterioração ainda maior do desempenho.

### 4.3 Controlador de Fixação

A Fig. 4.6 esquematiza toda a malha para controlo do seguimento, que é composta por três processos concorrentes. O processamento visual, com frequência de amostragem de 25Hz, estima o movimento do alvo a partir das imagens. A malha de servo-controlo, com uma frequência de 1KHz, tem por função controlar a velocidade do motor. O controlador de fixação estabelece a ligação entre os dois processos anteriores. A frequência de amostragem do controlador de fixação deve ser igual à frequência de comunicação com a malha de servo-controlo (166Hz). O movimento do alvo entre imagens é estimado por interpolação. O controlador de fixação integra a informação visual com as leituras do codificador óptico actualizando o comando a enviar para o actuador. Desta forma toda a informação disponível é utilizada para otimizar o desempenho global de seguimento.

$$Comm(z) = \omega(z) + C(z).\Delta\theta(z) \quad (4.2)$$

A equação 4.2 calcula o sinal de controlo  $Comm(z)$  enviado para a malha de servo-controlo  $M(z)$ .  $C(z)$  é um controlador PD (proporcional-derivativo) que filtra o erro em posição angular  $\Delta\theta(z)$ . A posição do alvo na imagem permite determinar o erro em posição de seguimento. Note que, para que o erro em regime final seja nulo (sistema de tipo 1), a função de transferência  $C(z)M(z)$  tem que ter um pólo em 1. Fazendo o controlo em modo de velocidade o integrador do gerador de trajectória introduz esse pólo (ver Fig. 4.3). O controlo em modo velocidade, não só melhora a resposta em frequência do motor, como assegura um erro em regime final nulo. Por sua vez, a medição da velocidade na imagem, após compensação do movimento auto-induzido, permite estimar a velocidade angular do alvo. A integração da velocidade do objecto  $\omega(z)$  no comando do motor melhora a resposta transitória do sistema de seguimento. Isto corresponde aquilo que é referido na literatura por "velocity feedforward". Note que a velocidade na imagem deve ser medida após compensação do movimento auto-induzido. Caso contrário a parcela de velocidade em  $Comm(z)$  teria o mesmo papel que a componente derivativa de  $C(z)$ .

Na estratégia de controlo apresentada o comando de velocidade enviado para a malha de servo-controlo é obtido somando a estimativa de velocidade com o erro de posição filtrado por  $C(z)$ . Enquanto a componente de posição é fundamental para a rejeição de perturbações em degrau, a componente de ve-

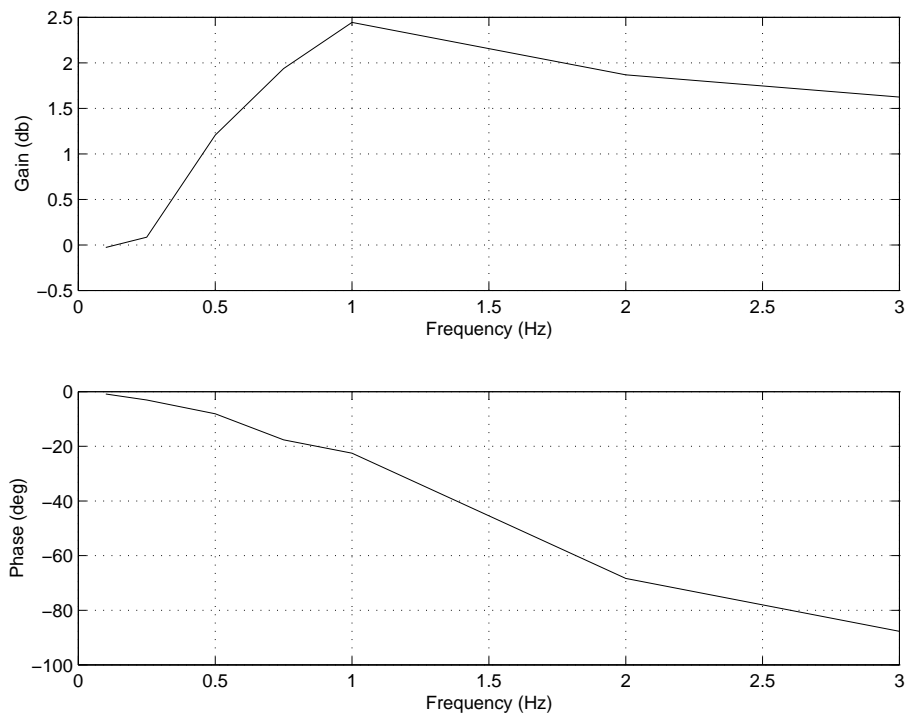


Figura 4.7: Resposta em frequência do algoritmo de seguimento (ganho em cima e fase em baixo).

locidade melhora a resposta transitória do sistema ("velocity feedforward"). O sistema de seguimento esquematizado na Fig. 4.6 é altamente não-linear, principalmente devido ao processamento visual introduzido na malha de realimentação. No entanto, a Fig. 4.7 mostra a resposta em frequência do regulador na compensação de movimento 3D do alvo. Esta resposta em frequência é aproximada e foi obtida perturbando o sistema com trajectórias sinusoidais de diferentes frequências e medindo a amplitude e o atraso de fase das trajectórias descritas pela plataforma. De modo a limitar as não-linearidades introduzidas pelo processamento visual, teve-se o cuidado de dimensionar as trajectórias de perturbação de forma à velocidade induzida na imagem estar dentro dos limites mensuráveis pelo algoritmo multi-resolucional de fluxo ( $\pm 7 \text{pixel/imagem}$  que corresponde a  $\pm 38 \text{deg/s}$ ). A largura de banda do sistema de seguimento é de cerca de 3Hz.

## 4.4 Controlo de Fixação Usando um Controlador DMC

A Fig. 4.3 mostra a resposta em frequência da malha de controlo em velocidade do motor. O atraso de fase observável é devido ao tempo de comunicação e à inércia mecânica do dispositivo (em particular o atrito estático da caixa de engrenagem). A função de transferência da malha de servo-controlo é dada pela equação 4.1. O "deadbeat" é de dois períodos de amostragem o que significa que o  $M(z)$  apresenta uma inércia de reacção de cerca de 12ms. Esta secção discute a utilização de controlo predictivo para lidar com atrasos de comunicação e inércia mecânica dos actuadores.

$$J = \sum_{i=N_1}^{N_2} (y(n+i|n) - w(n+i))^2 + \sum_{j=N_1}^{N_u} \lambda \Delta u(n+j-1)^2 \quad (4.3)$$

Existe uma grande variedade de controladores predictivos [48], mas todos eles têm três elementos em comum: um modelo para predição, uma função objectivo e um processo de minimização para obter a lei de controlo. O modelo de predição é utilizado para estimar a saída do sistema a ser controlado  $y(n+k|n)$  em instantes de tempo futuros. O objectivo é fazer convergir a saída futura do sistema para um sinal de referência  $w(n)$  que traduz o comportamento desejado. Para isso é estabelecida uma função de objectivo  $J$ . A expressão geral de  $J$  é apresentada na equação 4.3.  $N_1$  e  $N_2$  são os limites do

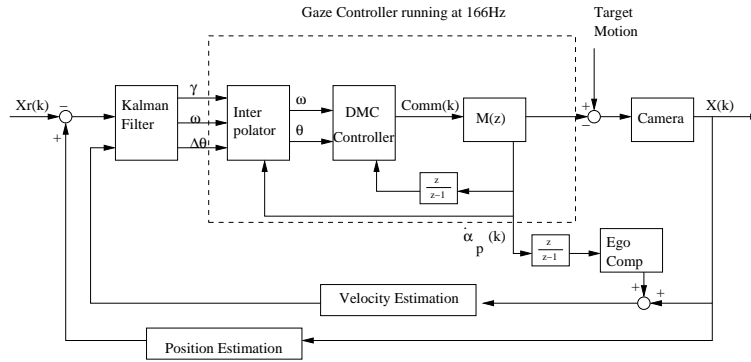


Figura 4.8: Malha de seguimento com o controlo de fixação baseado num controlador DMC.

horizonte de custo,  $N_u$  é o horizonte de controlo,  $u(n)$  é o sinal de controlo,  $\Delta u(n)$  é o incremento de controlo ( $\Delta u(n) = u(n) - u(n-1)$ ) e  $\lambda$  é um factor de penalização que permite controlar a suavidade do controlo. O funcional  $J$  é minimizado de forma a se obter a lei de controlo  $u(n)$ .

#### 4.4.1 O controlador preditivo de Matriz Dinâmica (DMC)

O horizonte de custo é o intervalo de tempo no futuro para o qual é desejável que a saída convirja para a referência  $w(n)$ .  $M(z)$  é o processo a ser controlado e tem um "deadbeat" de 2 períodos de amostragem, o que significa que não é possível forçar a saída no instante de tempo imediatamente a seguir. Assim  $N_1$  deve ser sempre maior ou igual a 2, dado a entrada actual só influenciar a saída a partir desse momento. Considerando que a frequência de amostragem da malha de processamento visual é de 25Hz, o controlador de fixação envia no máximo 7 comandos entre a captura de duas imagens sucessivas. Assim vamos fazer  $N_1 = 2$  e  $N_2 = 8$ . O horizonte de controlo é igual ao horizonte de custo.

Assuma que  $g(n)$  é a resposta a degrau de um processo linear estável sem factor de integração (não tem pólos em  $z = 1$ ). Se  $g(n) = 1$  para  $n > N$  o sistema é completamente descrito pelos primeiros  $N$  instantes de  $g(n)$ . Esta é a ideia em que se baseia um controlador preditivo simples, robusto e intuitivo: o controlador de matriz dinâmica DMC. A Fig. 4.8 esquematiza a malha de seguimento com o controlador DMC.

$$\Delta \mathbf{u} = (\mathbf{G}\mathbf{G}^t + \lambda \mathbf{I})^{-1} \mathbf{G}^t (\mathbf{w} - \mathbf{f}) \quad (4.4)$$

O DMC utiliza os  $N$  primeiros instantes da resposta a degrau de  $M(z)$  para predizer a saída do sistema (neste caso particular  $N = 7$ ). Este controlador assume uma perturbação constante ao longo do horizonte de custo. A perturbação é dada pela diferença entre a saída actual do sistema, medida pelos codificadores, e a saída prevista pelo modelo ( $d(n) = y(n) - y(n|n)$ ). O objectivo do controlo é manter a saída de  $y(n)$  o mais próxima possível do referênciada desejada  $w(n)$  no sentido dos mínimos quadráticos. A acção de controlo para atingir esse objectivo é calculada pela equação 4.4.  $\mathbf{G}$  é designada por matriz dinâmica do sistema,  $\Delta \mathbf{u}$  é o vector de incrementos de controlo e  $\mathbf{w}$  é o vector do sinal de referênciada.  $\mathbf{f}$  é chamado o vector de resposta livre porque representa a parcela da saída que não depende das entradas de controlo futuras.  $\lambda$  é um factor de penalização para o esforço de controlo. Aumentando  $\lambda$  o sistema tende a ficar mais lento e suave. Note que só o primeiro elemento de  $\Delta \mathbf{u}$  é realmente enviado para o motor. O vector de incrementos de controlo é calculado em cada iteração de forma a aumentar a robustez do algoritmo a perturbações no modelo. Para mais detalhes sobre controladores DMC consulte [48].

#### 4.4.2 O sinal de referênciada $w(n)$

O objectivo do controlador preditivo é forçar o motor a ter o mesmo movimento angular do alvo num futuro próximo. Sempre que uma imagem é capturada a malha de processamento visual estima a velocidade e posição angulares do alvo. Considere a frequência de amostragem do controlador de fixação (166Hz).  $P_t(n)$  e  $V_t(n)$  são, respectivamente, a posição e velocidade angular do alvo estimadas pelo interpolador com base na informação visual. O seguimento é perfeito se, no instante de aquisição de imagem seguinte, o sistema compensar o erro em posição e deslocar-se com a velocidade estimada do alvo. Assuma que  $P(n)$  e  $V(n)$  referem-se à posição e velocidade angular do motor. Se o objectivo é chegar à situação de seguimento perfeito em  $M$  instantes de tempo, então tem de se controlar o movimento do motor de forma a que  $P(M) = P_t(M)$  e  $V(M) = V_t(M)$ .

$$\Delta_v = \frac{V_t(M) - V(0)}{M} \quad (4.5)$$



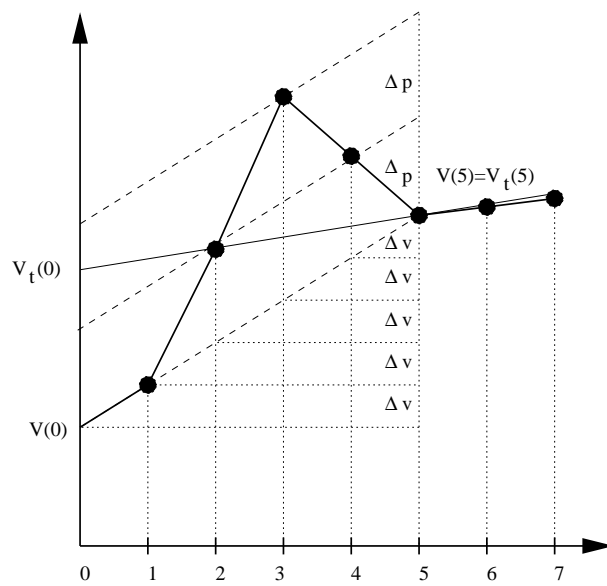


Figura 4.9: O perfil da velocidade de referência  $w(n)$  (linha a cheio grossa [-]). A velocidade do alvo  $V_t(n)$  é assinalada pela linha a cheio fina [-]. Em 5 instantes de tempo a velocidade do motor converge para a velocidade do alvo e o erro em posição é compensado

$$w(i) = V(0) + i.\Delta_v, i = 0, \dots, M \quad (4.6)$$

Assuma que  $P(0)$  e  $V(0)$  são a posição e velocidade actual do motor. Se  $P(0) = P_t(0)$  não existe erro em posição e o controlador só tem de fazer convergir a velocidade do motor para a velocidade prevista do alvo. Nesse caso a velocidade de referência  $w(n)$  é dada pela equação 4.6 (ver Fig. 4.9). Motor acelera ou desacelera de forma linear ( $\Delta_v$  é o incremento) até atingir a velocidade desejada no espaço de tempo previsto.

$$P(n) = P(n-1) + V(n)T \quad (4.7)$$

$$w(n) = \begin{cases} V(0) + i.\Delta_v + (i-1).\Delta_p & \Leftarrow 1 \leq i \leq \frac{M+1}{2} \\ V(0) + i.\Delta_v + (M-i).\Delta_p & \Leftarrow \frac{M+1}{2} \leq i \leq M \end{cases} \quad (4.8)$$

Considere agora o caso geral em que existe também um erro em posição angular. A relação entre a posição e velocidade do motor é dada pela equação 4.7 em que  $T$  é o período de amostragem (6ms). A velocidade de referência  $w(n)$  tem de ser planeada de forma ao motor atingir a velocidade do alvo e recuperar simultaneamente o erro de posição. A Fig. 4.9 mostra o perfil de velocidade  $w(n)$  para atingir esse fim. A ideia é, para além de fazer convergir a velocidade do motor para a velocidade do alvo, obrigar o dispositivo a acelerar/desacelerar e depois desacelerar/acelerar de forma a compensar o erro em posição. A expressão 4.8 calcula a sinal de referência  $w(n)$  assumindo que  $M$  é ímpar e maior que 1.

$$P(M) = P(0) + \sum_{i=1}^M w(i).T \quad (4.9)$$

$$\Delta_p = \frac{P_t(M) - P_t(0) - (M-1).V(0) - \sum_{i=1}^M M i . \Delta_v}{\sum_{i=1}^{\frac{M-3}{2}} 2i + \frac{M-1}{2}} \quad (4.10)$$

O incremento  $\Delta_v$  é dado pela equação 4.5.  $\Delta_p$  é calculado a partir de 4.9. Esta equação determina a posição do motor no instante  $M$  no caso da sua velocidade coincidir com a velocidade de referência. Fazendo  $P(M) = P_t(M)$ , substituindo  $w(i)$  por 4.8 e resolvendo em ordem a  $\Delta_p$ , obtém-se o resultado 4.10. Nesta aplicação em particular considerou-se  $M=5$ .

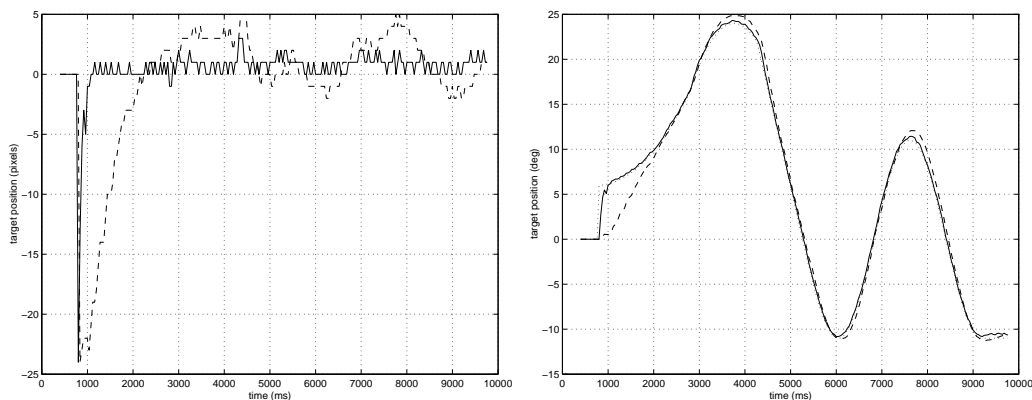


Figura 4.10: Comparação do desempenho da estratégia de controlo inicial com o novo controlador preditivo. Esquerda: Regulação na imagem. Posição de projecção do alvo na imagem usando o controlador DMC (linha a cheio [-]) e o controlador inicial (linha tracejada [- -]). Direita: Resposta servo-mecânica. Posição angular do alvo  $\theta$  (linha pontuada [..]) e do motor  $\alpha_p$  com o controlador preditivo (linha a cheio [-]) e o controlador inicial (linha tracejada [- -])

#### 4.4.3 Desempenho do controlador DMC

A Fig. 4.10 compara o desempenho do controlador DMC com o controlador inicial em compensar a perturbação não linear da Fig. 3.7. Existe uma clara melhoria do desempenho. O erro inicial em posição é imediatamente compensado e o alvo é mantido no centro da imagem. No entanto o controlador DMC é extremamente sensível ao ruído e a perturbações no modelo. Actualmente estamos a trabalhar num controlador preditivo GPC na tentativa de superar esses problemas.



# Capítulo 5

## Conclusões

O ponto de partida deste trabalho foi o algoritmo de seguimento binocular implementado na cabeça robótica MDOF e apresentado em [1, 6, 2, 4, 5, 7]. O primeiro capítulo introduz o sistema de visão activa MDOF bem como o algoritmo de seguimento binocular original que controla o azimute e elevação do pescoço e a vergência simétrica. Em [37] demonstra-se que o seguimento em versão (controlo dos graus de liberdade do pescoço) é similar ao seguimento monocular com a câmara ciclópica imaginária. Assim o seguimento binocular pode ser decomposto em dois sub-problemas: seguimento monocular e seguimento em vergência. Estes dois problemas são o objecto de estudo do presente trabalho.

O seguimento activo de alvos é um problema de controlo de movimento utilizando visão, em que o movimento controlado é o da própria plataforma que suporta as câmaras. A robustez e o desempenho de um sistema deste tipo depende simultaneamente dos algoritmos de processamento visual e de aspectos relacionados com teoria do controlo e arquitectura de sistemas. A dinâmica do sistema não pode ser negligenciada quando se pretendem comportamentos robustos com elevado desempenho. A optimização do sistema deve ser feita recorrendo a modelação e caracterização. Para isso são necessárias ferramentas que permitam avaliar simultaneamente os aspectos visuais e de controlo.

Muitos autores estudam os sistemas de posicionamento controlados por visão como servo-sistemas tradicionais que utilizam um sensor visual na malha de realimentação. A entrada de referência do sistema está relacionada com a posição/velocidade do alvo no espaço 3D, enquanto a saída é dada pela posição/velocidade das juntas da plataforma. Este trabalho defende

que esta perspectiva não é a mais adequada. Primeiro porque o sistema de seguimento pretende manter a posição do alvo estacionária na imagem, e não recuperar os parâmetros do movimento 3D do objecto. O objectivo do seguimento visual está ao nível do plano imagem, fazendo pouco sentido definir as entradas/saídas do sistema no espaço 3D. Segundo porque o movimento no espaço e na imagem relacionam-se por uma transformação projectiva que introduz uma não linearidade na malha de realimentação visual. Assim as ferramentas e os sinais de teste utilizadas na identificação de sistemas lineares não são um estímulo adequado para caracterizar o desempenho do sistema de seguimento se a entrada de referência for definida no espaço 3D .

O seguimento monocular e o seguimento em vergência são modelados como sistemas de regulação cujo objectivo é manter a posição do alvo nas imagens estacionária. Segundo este paradigma a entrada e saída do sistema são, respectivamente, a posição desejada e actual do alvo no plano imagem. O movimento do objecto no espaço 3D é encarado como uma perturbação de dinâmica desconhecida que o sistema tem de compensar. Os sinais de teste para a caracterização e identificação do sistema são definidos no plano imagem. As ferramentas utilizadas na identificação de sistemas lineares, como as entradas em degrau, rampa, parábola e sinusóide, tornam-se estímulos adequados para a avaliação do sistema. Uma perturbação em rampa, por exemplo, corresponde a um movimento do alvo no espaço que induz na imagem uma velocidade constante. O segundo capítulo apresenta técnicas para gerar com rigor os sinais de perturbação do sistema e determina as trajectórias de movimento do alvo que definem sinais em degrau, rampa, parábola e sinusóide na imagem. Este estudo é feito para o seguimento monocular e para o controlo de vergência.

O sistema de seguimento é caracterizado utilizando os sinais de teste deduzidos. Esta avaliação permitiu identificar diversos aspectos que limitam o desempenho global do sistema. O controlo das juntas da plataforma é feito utilizando a informação de posição e velocidade do alvo na imagem após cancelamento do movimento auto-induzido. A informação de posição é essencial para garantir um erro e regime final nulo e a informação de velocidade permite melhorar a resposta transitória do sistema. Constatou-se que a versão original do sistema de seguimento estimava o centro de movimento em vez da posição do alvo. O respectivo algoritmo foi modificado. Verificou-se ainda que o sistema só era capaz de medir velocidades na imagem até  $2 \text{ pixel/imagem}$  e que isso limitava severamente o desempenho

global. Desenvolveu-se um algoritmo de estimação multi-resolução que permitiu aumentar a gama de velocidades mensuráveis para 8 *pixel/imagem*.

A filtragem de Kalman é utilizada para estimar, a partir da informação visual, alguns parâmetros do movimento 3D do alvo necessários ao controlo adequado dos actuadores da plataforma. O presente trabalho projecta dois filtros de Kalman, um para o seguimento monocular e outro para o controlo da vergência.

O filtro projectado para o seguimento monocular assume que o movimento angular do alvo entre duas imagens sucessivas é uniformemente acelerado em torno do centro óptico. A técnica de emulação descrita no capítulo 2 é utilizada para fazer a sintonização do filtro. Uma boa sintonização do filtro passa sempre por estabelecer um compromisso entre suavidade e reacção. O filtro de Kalman tem um efeito suavizador compensando nalguns casos a má prestação do processamento visual. O efeito suavizador é obtido à custa da introdução de uma certa inércia que diminuí a reacção do sistema de seguimento. As melhorias no processamento visual devem ser acompanhadas por uma sintonização menos conservativa do filtro. Desta forma capitaliza-se o acréscimo de capacidade de estimação visual num aumento efectivo de desempenho global do sistema. A estimação da aceleração angular a partir das medições de movimento na imagem é um processo ruidoso. A utilização de um modelo de movimento angular uniformemente acelerado na equação de processo do filtro permite um aumento de desempenho, se a velocidade angular do alvo variar de uma forma aproximadamente linear. Caso contrário a estimação da aceleração é muito deficiente sendo preferível assumir um movimento angular uniforme na equação de processo. Concluiu-se assim que o modelo óptimo a utilizar no filtro depende do tipo de movimento do alvo a seguir.

O filtro de Kalman para o controlo da vergência foi desenhado assumindo que o alvo tem um movimento rectilíneo uniforme ao longo do eixo ciclópico. Utilizaram-se as técnicas de emulação descritas para sintonizar e caracterizar o filtro. Constatou-se que à medida que o alvo se afasta do sistema os valores de ganho do filtro de Kalman tendem a diminuir. Neste caso o movimento projectado no plano imagem diminuí, tornado-se por isso mais difícil de medir. Assim o filtro de Kalman aumenta a sua confiança no modelo de movimento em detrimento da informação visual que lhe chega.

O capítulo 4 foca questões de arquitectura e controlo num sistema de seguimento utilizando visão. Concluiu-se que num sistema deste tipo podem

ser identificados três processos distintos e concorrentes: o processamento visual, o controlo de baixo nível do actuador e o controlo de alto nível de fixação.

Discutiu-se as vantagens de utilizar motores DC com caixa de engrenagem para actuar a plataforma. A necessidade de uma malha de controlo local do servo-motor com elevada frequência de amostragem foi também discutida. O controlo dos actuadores DC foi feito através de placas de controlo comerciais. Estas placas implementam uma malha com um controlador PID digital com frequência 1KHz que utiliza a informação de posição do codificador óptico para realimentação. Esta opção de arquitectura é modular libertando o computador hospedeiro para o processamento visual e controlo de alto nível. Concluiu-se que o controlo em velocidade permitia aumentar a largura de banda do sistema de actuação.

O controlador de fixação faz a ligação entre a malha de processamento visual (25Hz) e a malha de controlo local do actuador (1KHz). A frequência de amostragem deste controlador deve ser igual à frequência de comunicação com a malha servo (166Hz). A aquisição de imagem e o processamento visual introduzem um atraso que deteriora o desempenho global do sistema. Discutiu-se a utilização de interpolação no controlo de fixação de forma a lidar com este problema. Concluiu-se que os resultados desta aproximação dependem muito da qualidade das medições no plano imagem. O controlo de fixação deve utilizar simultaneamente informação de posição e velocidade. A informação em posição permite controlar o erro de seguimento em regime final. O sistema tem de ser de tipo 1 de forma a este erro ser nulo. A velocidade medida na imagem está relacionada com a velocidade angular do alvo. A componente de velocidade no comando do motor permite melhorar a resposta transitória global, correspondendo aquilo que é conhecido na literatura por "velocity feedforward". Uma conclusão importante é que a velocidade na imagem deve ser medida após cancelamento do movimento auto-induzido. Caso contrário a componente de velocidade obtida é equivalente a uma componente derivativa do erro de posição. No fim do capítulo 4 é feito o projecto de um controlador preditivo para o controlo de fixação. Demonstra-se que o controlo preditivo é uma solução para lidar com a inércia mecânica dos actuadores.



# Bibliografia

- [1] Jorge Manuel M. de C. P. Batista. *Sistemas de Visão Activa: Comportamento e Calibração*. PhD thesis, Universidade de Coimbra, Coimbra, Portugal, Outubro 1999.
- [2] J. Batista, J. Dias, H. Araújo, and A. Almeida. The isr multi-degree-of-freedom active vision robot head: Design and calibration. In *in Proc. of M2VIP'95 - Second International Conference on Mechatronics and Machine Vision in Practice*, Hong-Kong, September 1995.
- [3] *DCX-AT100 Modular Multi-Axis Motion Controller, Users Manual, Version 1.1a*. Precision MicroControl, 1993.
- [4] J. Batista, P. Peixoto, and H. Araújo. Real-time visual behaviors with a binocular vision system. In *ICRA97 - IEEE International Conference on Robotics and Automation*, New Mexico, Setembro 1997.
- [5] J. Batista, P. Peixoto, and H. Araújo. Real-time vergence and binocular gaze control. In *IROS97-IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Grenoble, France, September 1997.
- [6] J. Batista, P. Peixoto, and H. Araujo. Real-time active visual surveillance by integrating peripheral motion detection with foveated tracking. In *Proc. of the IEEE Workshop on Visual Surveillance*, pages 18–25, Bombay, India, January 1998.
- [7] P. Peixoto, J. Batista, and H. Araujo. A surveillance system combining peripheral and foveated motion tracking. In *ICPR'98 - 14th International Conference on Pattern Recognition*, pages 18–25, Brisbane, Australia, August 1998.

- 
- [8] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Trans. on Robot. and Automat.*, 8(3):313–326, June 1992.
- [9] P. Allen, A. Timcenko, B. Yoshimi, and P. Michelman. Automated tracking and grasping of a moving object with a robotic hand-eye system. *IEEE Trans. on Robot. and Automat.*, 9(2):152–165, 1993.
- [10] P. I. Corke and M. C. Good. Dynamic effects in visual closed-loop systems. *IEEE Trans. on Robotics and Automation*, 12(5):671–683, October 1996.
- [11] P. I. Corke. *Visual Control of Robots: High-Performance Visual Servoing*. Mechatronics. John Wiley, 1996.
- [12] Richard Hartley. Self-calibration from multiple views with a rotating camera. In *Proc. of European Conference of Computer Vision*, 1992.
- [13] D. W. Murray, P. McLauchlan, I. Reid, and P. Sharkey. Reactions to peripheral image motion using a head/eye platform. In *ICCV93 - IEEE International Conference on Computer Vision*, pages 403–411, 1993.
- [14] F. Du. *The Fundamentals fo an Active Vision System*. PhD thesis, Robotics Research Group, Dept. of Engineering Science, Oxford, U.K., Outubro 1994.
- [15] E. D. Dickmanns, B. Mysliwetz, and T. Christians. An integrated spatio-temporal approach to automatic visual guidance of autonomous vehicles. *IEEE Trans. on Systems, Man and Cybernetics*, 20(6):1273–1284, November/December 1990.
- [16] L. Matthies, T. Kanade, and R. Szeliski. Kalman-filter based algorithms for estimating depth from image sequence. *International Journal of Computer Vision*, 3(3):209–238, 1989.
- [17] J. L. Crowley and H. I. Christensen editors. *Vision as a process*. Springer-Verlag, 1995.
- [18] H. I. Christensen, J. Horstmann, and T. Rasmussen. A control theoretic approach to active vision. *Asian Conference on Computer Vision*, pages 201–210, December 1995.

- [19] Alexandre Bernardino and Jose Santos-Victor. Binocular tracking: Integrating perception and control. *IEEE Trans. on Robotics and Automation*, 15(6), December 1999.
- [20] C. Brown. Gaze controls with interactions and delays. *IEEE Trans. on Systems, Man and Cybern.*, 20(2):518–527, 1990.
- [21] P. M. Sharkey and D. W. Murray. Delay versus performance of visually guided systems. *IEE Proc.-Control Theory Appl.*, 143(5):436–447, September 1996.
- [22] P. M. Sharkey and D. W. Murray. Bandwidth issues for active vision systems. In *in Proc. of M2VIP'95 - Second International Conference on Mechatronics and Machine Vision in Practice*, Hong-Kong, September 1995.
- [23] J. Hill and W. T. Park. Real time control of a robot with a mobile camera. In *in Proc. of 9th ISIR*, pages 233–246, Washington DC, March 1979.
- [24] Y. Aloimonos, I. Weiss, and A. Bandyopadhyay. Active vision. *International Journal of Computer Vision*, 1(4):333–356, January 1988.
- [25] João P. Barreto, P. Peixoto, J. Batista, and H. Araujo. Evaluation of the robustness of visual behaviors through performance characterization. In Markus Vincze and Gregory D. Hager, editors, *Robust Vision for Vision-Based Control of Motion*. IEEE Press, 1999.
- [26] João P. Barreto, P. Peixoto, J. Batista, and H. Araújo. Performance characterization of visual behaviors in an active vision system. In *SIRS98–6th Int. Symposium on Intelligent Robotic Systems*, pages 233–246, Edinburgh, Scotland, July 1998.
- [27] João P. Barreto, P. Peixoto, J. Batista, and H. Araújo. Control performance issues in a binocular active vision system. In *IROS98–IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 233–246, Victoria, Canada, October 1998.
- [28] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994.

- [29] B. K. P. Horn and B. G. Schunk. Determining optical flow. *Artificial Intelligence*, (17):185–203, 1981.
- [30] P. Anandan. *Measuring Visual Motion from Image Sequences*. PhD thesis, University of Massachusetts, Amherst, MA, USA, 1987.
- [31] H. Nagel. On the estimation of optical flow: Relations between different approaches and some new results. *Artificial Intelligence*, (33):299–324, 1987.
- [32] David J. Heeger. Optical flow using spatio temporal filters. *International Journal of Computer Vision*, pages 279–302, 1988.
- [33] Maximo Tistarelli. Multiple constraints to compute optical flow. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(12):1243–1250, December 1996.
- [34] Lucas B. and Kanade T. An iterative image registration technique with an application to stereo vision. In *Proc. DARPA Image Understanding Workshop*, 1981.
- [35] E. Simoncelli, E. Adelson, and D. Heeger. Probability distributions of optical flow. In *CVPR91 - IEEE Conference on Computer Vision and Pattern Recognition*, Maui, Hawaii, June 1991.
- [36] João P. Barreto, J. Batista, and H. Araújo. Solution for visual control of motion: Active tracking applications. In *MED2000 – IEEE Mediterranean Conference on Automatic Control*, pages 233–246, Patras, Greece, July 2000.
- [37] João P. Barreto, P. Peixoto, J. Batista, and H. Araújo. Improving 3d active visual tracking. In Henrik I. Christensen, editor, *Computer Vision Systems, Lectures Notes in Computer Science 1542*. Springer, 1999.
- [38] João P. Barreto, P. Peixoto, J. Batista, and H. Araújo. Mathematical analysis for visual tracking assuming perspective projection. In *SIRS2000–8th Int. Symposium on Intelligent Robotic Systems*, pages 233–246, Reading, England, July 2000.
- [39] J.R. Bergen, P.J. Burton, R. Hingorani, and S. Peleg. Computing two motions from three frames. In *Technical Report, David Sarnoff Research Center*, April 1990.

- [40] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *International Journal of Computer Vision*, (2):283–310, 1989.
- [41] M. Okutomi and T. Kanade. A locally adaptive window for signal matching. *International Journal of Computer Vision*, (7:2):143–162, 1992.
- [42] Stan Birchfield. Derivation of kanade-lucas-tomasi tracking equation. In *Technical Report*, January 1997.
- [43] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the American Society of Mechanical Engineers. Journal of Basic Engineering*, pages 35–45, Mars 1960.
- [44] Yaakov Bar-Shalom and Xiao-Rong Li. *Multitarget-Multisensor Tracking: Principles and Techniques*. Yaakov Bar-Shalom, 1995.
- [45] Athanasios Papoulis. *Probability, Random Variables and Stochastic Processes*. MacGraw-Hill International Editions, 1991.
- [46] P. M. Sharkey, P. F. MaLauchlan D. W. Murray, and J. P. Brooker. Hardware development of the yorick series of active vision systems. *Microprocessors and Microsystems*, (21):363–375, 1998.
- [47] P. Krautgartner and M. Vincze. Performance evaluation of vision-based control tasks. In *IEEE International Conference on Robotics and Automation*, Leuven, Belgium, May 1998.
- [48] K. Pahlavan, T. Uhlin, and J. Eklund. Integrtaring primary ocular processes. *Proc. 2nd European Conf. on Computer Vision*, pages 526–541, 1992.
- [49] E. F. Camacho and C. Bordons. *Model Predictive Control*. Springer-Verlag, 1999.