VALDIR GRASSI JUNIOR

SISTEMA DE VISÃO OMNIDIRECIONAL
APLICADO NO CONTROLE DE ROBÔS MÓVEIS

Dissertação apresentada à Escola Politécnica
da Universidade de São Paulo para obtenção
do Título de Mestre em Engenharia.

SÃO PAULO

2002

VALDIR GRASSI JUNIOR

SISTEMA DE VISÃO OMNIDIRECIONAL
APLICADO NO CONTROLE DE ROBÔS MÓVEIS

Dissertação apresentada à Escola Politécnica
da Universidade de São Paulo para obtenção
do Título de Mestre em Engenharia.

Área de Concentração:
Engenharia Mecânica

Orientador:
Prof. Doutor
Jun Okamoto Junior

SÃO PAULO
2002

FICHA CATALOGRÁFICA

Grassi Jr., Valdir

Sistema de Visão Omnidirecional Aplicado no Controle de Robôs Móveis. São Paulo, 2002.
85p.

Dissertação (Mestrado) – Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos.

1.Visão Omnidirecional 2.Rastreamento Visual 3.Controle Servo Visual I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos II.t.

Aos meus pais e irmãos, pessoas a quem tanto amo e que serão sempre muito especiais para mim

AGRADECIMENTOS

Quero agradecer inicialmente à FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) por tornar possível a realização deste trabalho através do apoio financeiro concedido. Também à RECOPE/MANET/FINEP pela aquisição de alguns equipamentos utilizados neste trabalho.

Ao prof. Jun por ter um dia me convidado para trabalhar com o robô móvel ainda durante a graduação, e por ter me incentivado a continuar trabalhando nesta área, mostrando o quanto pode ser divertido trabalhar com robôs móveis. Sou grato também pela sua amizade e orientação.

Ao prof. Arthur Vieira Porto do Departamento de Engenharia Mecânica da USP de São Carlos por possibilitar a construção do espelho hiperbólico utilizado neste trabalho. Ao Roberto Verzini da Engefilme ind. e com. Ltda. por ter custeado o espelhamento da cúpula de vidro para fabricação do espelho esférico. Ao técnico Adilson por ter usinado as peças de fixação do sistema de visão omnidirecional.

Aos colegas do grupo de pesquisa (GPA) e da sala MC-01 que ajudaram e contribuíram de diversas formas diferentes a este trabalho, inclusive através do companheirismo e do ambiente alegre que proporcionaram ao laboratório.

Aos amigos cristãos que conheci aqui em São Paulo e que me incentivaram durante o mestrado, oraram por mim, e foram uma extensão da minha família aqui nesta cidade. Também aos amigos que encontrei na Poli durante a graduação e que permanecem até hoje formando uma “família monstro” muito divertida e ímpar.

Por fim quero agradecer imensamente aos meus pais, Valdir e Raquel, e aos meus irmãos, Vagner e Viviane, por todo o incentivo e amor. Também agradeço muito a Deus que têm sido minha esperança e suporte.

ERRATA

PÁGINA	LINHA	ONDE SE LÊ	LEIA-SE
Abstract	2 ^a	For this reason, they can be useful to	This type of system is very well suited for tasks such as
Abstract	3 ^a	This type of system does	Such systems do
Abstract	4 ^a	It is required	It requires
Abstract	5 ^a	once	As
Abstract	6 ^a	An effective	One effective
Abstract	7 ^a	with the combined	through the
Abstract	8 ^a	Several shapes	Several different shapes
Abstract	8 ^a	used mounting	used, mounting
Abstract	9 ^a	optics	optical
Abstract	9 ^a	common	commonly
Abstract	11 ^a	system is mounted	system was mounted
Abstract	12 ^a	tracking	tracking in real time
Abstract	14 ^a	with real time data	with data
Abstract	15 ^a	. This visual data	, that
3	2 ^a	constante	constantes
7	5 ^a	livre	livres
7	17 ^a	construir de um sistema	construir um sistema
14	2 ^a	substituir	substituir
25	7 ^a	características,	características
25	15 ^a	do o alvo	do alvo
30	17 ^a	Na formulação proposta	Nesta formulação proposta,
65	12 ^a	Fig.(5.10)	Fig.(5.9)
65	14 ^a	Fig.(5.10)	Fig.(5.9)

RESUMO

Sistemas de visão omnidirecional produzem imagens de 360° do ambiente podendo ser utilizados em navegação, tele-operação e controle servo visual de robôs. Este tipo de sistema dispensa o movimento da câmera para determinada direção de atenção mas requer processamento não convencional da imagem, uma vez que a imagem adquirida se encontra mapeada em coordenadas polares não lineares. Uma maneira efetiva de se obter uma imagem em um sistema omnidirecional é com o uso combinado de lentes e espelhos. Várias formas de espelhos convexos podem ser utilizadas montando-se uma câmera com o seu eixo óptico alinhado com o centro do espelho. Dentre as formas usadas, tem-se os cônicos, parabólicos, hiperbólicos e esféricos. Neste trabalho foi implementado um sistema de visão omnidirecional utilizando um espelho hiperbólico. Este sistema de visão desenvolvido é embarcado em um robô móvel e aplicado em uma tarefa de controle. A tarefa de controle de interesse neste trabalho é a de fazer com que o robô mantenha uma distância constante de um determinado alvo móvel. Esta tarefa é realizada com a realimentação em tempo real de informações visuais do alvo obtidas pelo sistema de visão para controle do robô utilizando uma abordagem de controle servo visual.

ABSTRACT

Omnidirectional vision systems can get images with a 360-degree of field of view. For this reason, they can be useful to robotic navigation, tele-operation and visual servoing. This type of system does not require the movement of the camera to the direction of attention of the robot. On the other hand, it is required a non-conventional image processing once the image captured by this vision system is mapped on a non-linear polar coordinate system. An effective way to obtain an image in an omnidirectional system is with the combined use of lenses and mirrors. Several shapes of convex mirrors can be used mounting the center of the mirror aligned with the camera optics axis. The most common used mirror shapes are conic, parabolic, hyperbolic and spherical. In this work a hyperbolic mirror was used to build an omnidirectional vision system. This system is mounted on a mobile robot and used in a control task. The task of interest here is the tracking of a moving target keeping the distance between the robot and the target constant. This task is accomplished with real time data acquisition from the omnidirectional vision system. This visual data is used as feedback to control the mobile robot in a visual servo approach.

SUMÁRIO

RESUMO

ABSTRACT

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE SÍMBOLOS

1 INTRODUÇÃO.....	1
2 VISÃO OMNIDIRECIONAL.....	4
2.1 Sistemas de visão omnidirecional catadióptricos.....	6
2.2 Espelho hiperbólico para sistema de visão omnidirecional	12
2.3 Retificação de imagem adquirida pelo sistema de visão omnidirecional ...	14
3 RASTREAMENTO VISUAL	25
3.1 Métodos de rastreamento visual.....	26
3.2 Rastreamento visual de uma região de referência baseado no cálculo da soma do quadrado das diferenças (SSD).....	31
3.3 Implementação do método de rastreamento visual para o sistema de visão omnidirecional desenvolvido.....	37
4 SERVO VISUAL	39
4.1 Arquiteturas de sistema de controle servo visual.....	39
4.2 Trabalhos na área de controle servo visual	43

4.3 Controle servo visual baseado em imagem.....	44
5 RESULTADOS EXPERIMENTAIS	54
5.1 Equipamento utilizado	54
5.2 Sistema de visão omnidirecional.....	56
5.3 Rastreamento visual do alvo	69
5.4 Controle servo visual do robô móvel	74
6 CONCLUSÃO.....	78
REFERÊNCIAS BIBLIOGRÁFICAS	81

LISTA DE FIGURAS

Figura 2.1 – Câmera giratória para formação de imagem omnidirecional.....	4
Figura 2.2 – Ganho angular vertical.....	7
Figura 2.3 – Imagem de um piso xadrez obtida com (a) um espelho esférico e com (b) espelho projetado projetado por Hicks e Bajcsy. (Figura extraída de Hicks; Bajcsy, 1999).....	9
Figura 2.4 – Três formas de espelho convexo, duas formas possuem centro único de projeção (Figura extraída de Svoboda et al, 1997).....	10
Figura 2.5 – Esquema representando a geração de imagens perspectivas e panorâmicas através da projeção da imagem adquirida em planos escolhidos.	11
Figura 2.6 – Geometria do espelho hiperbólico e da câmera	13
Figura 2.7 – Esquema representando o processo de retificação de uma imagem	16
Figura 2.8 – Geometria do espelho e da câmera usada para retificação da imagem..	18
Figura 2.9 – Plano de projeção perspectiva.....	21
Figura 2.10 – Relações geometria usadas para geração de imagens perspectivas	22
Figura 3.1 – Graus de liberdade de uma transformação projetiva de um plano: (a) translação ao longo do eixo horizontal da imagem; (b) translação ao longo do eixo vertical da imagem; (c) rotação; (d) fator de escala; (e) shear; (f) relação de aspecto; (g) distorção de projeção ao longo do eixo horizontal da imagem; (h) distorção de projeção ao longo do eixo vertical da imagem. (figura extraída de Van der Zwaan; Santos-Victor, 2001).....	29
Figura 3.2 – Câmera de projeção perspectiva virtual localizada no foco da hipérbole	37
Figura 4.1 – Sistema de controle servo visual <i>dynamic look-and-move</i> baseado em posição (Figura extraída de Hutchinson et al., 1996).....	41

Figura 4.2 – Sistema de controle servo visual <i>dynamic look-and-move</i> baseado em imagem (Figura extraída de Hutchinson et al., 1996)	41
Figura 4.3 – Sistema de controle servo visual baseado em posição com estrutura proposta por Weiss (extraído de Hutchinson et al., 1996)	42
Figura 4.4 – Sistema controle servo visual baseado em imagem com estrutura proposta por Weiss (extraído de Hutchinson et al., 1996)	42
Figura 4.5 – Sistema de coordenadas da câmera e do robô.....	47
Figura 4.6 – Região de referência rastreada na imagem, e a coordenada dos quatro pontos utilizados no controle servo visual	51
Figura 5.1 – Esquema do ambiente experimental utilizado	56
Figura 5.2 – Protótipo do sistema de visão omnidirecional construído.	57
Figura 5.3 – Erro de forma da peça	59
Figura 5.4 – Trajetória da ponta da ferramenta na usinagem do espelho.....	60
Figura 5.5 – Espelhos convexos construídos. (a) Espelho hiperbólico de alumínio com 40mm de diâmetro usinado em torno de ultra-precisão. (b) Espelho esférico de 70mm de diâmetro construído em vidro.	61
Figura 5.6 – Robô móvel com sistema de visão omnidirecional.....	63
Figura 5.7 – Imagens adquiridas pelo sistema de visão omnidirecional. (a) Imagem adquirida com o espelho esférico. (b) Imagem adquirida com o espelho hiperbólico.....	64
Figura 5.8 – Imagem adquirida com o espelho esférico retificada pelo método de transformação direta de coordenadas	64
Figura 5.9 – Imagem adquirida com o espelho hiperbólico retificada pelo método desenvolvido que leva em consideração a equação do espelho hiperbólico e a propriedade de centro único de projeção	64
Figura 5.10 – Imagem adquirida com o espelho hiperbólico retificada pelo método de transformação direta de coordenadas.....	65
Figura 5.11 – Imagens adquiridas pelo sistema de visão omnidirecional de uma folha de papel com linhas horizontais equidistantes colocada ao redor do	

sistema. (a) Imagem adquirida com o espelho esférico. (b) Imagem adquirida com o espelho hiperbólico.....	66
Figura 5.12 – Imagem da folha de papel adquirida com o espelho esférico retificada pelo método de transformação direta de coordenadas.....	67
Figura 5.13 – Imagem da folha de papel adquirida com o espelho hiperbólico retificada pelo método desenvolvido que leva em consideração a equação do espelho hiperbólico e a propriedade de centro único de projeção.....	67
Figura 5.14 – Imagem da folha de papel adquirida com o espelho hiperbólico retificada pelo método de transformação direta de coordenadas.....	67
Figura 5.15 – Imagens perspectivas geradas a partir de imagem omnidirecional.....	68
Figura 5.16 – Resultado do método de rastreamento baseado em SSD aplicado em imagens perspectivas obtidas pela retificação da imagem omnidirecional.....	71
Figura 5.17 – Resultados do método de rastreamento baseado em região de referência	73
Figura 5.18 – Fluxograma do programa implementado para controle do robô.....	74
Figura 5.19. Alvo a ser seguido pelo robô móvel.....	75
Figura 5.20 – Sinal de erro obtida com controle proporcional.....	76
Figura 5.21 – Sinal de controle de velocidade de translação obtida com controle proporcional.....	77

LISTA DE TABELAS

Tabela 5.1 – Resultados numéricos do método de rastreamento baseado em SSD aplicado em imagens perspectivas obtidas pela retificação da imagem omnidirecional para os casos mostrados na Fig. (5.22).....	72
---	----

LISTA DE SÍMBOLOS

φ	Ângulo de elevação
ψ	Ângulo radial, ou de entrada dos raios ópticos dentro da câmera
β	Ganho angular
F	Foco da câmera
F'	Foco do espelho convexo
a	Parâmetro do espelho hiperbólico
b	Parâmetro do espelho hiperbólico
(x,y)	Coordenadas de um ponto na superfície do espelho
h	Distância entre o foco da câmera e o topo do espelho
e	Excentricidade do espelho, ou metade da distância entre o foco da câmera e o foco do espelho hiperbólico
r_{topo}	Raio do topo do espelho
y_{topo}	Coordenada y do topo do espelho para $x = r_{topo}$
ϕ	Ângulo de visão vertical do espelho
f	Distância focal da câmera
t_{pixel}	Tamanho do pixel do CCD da câmera em milímetros

r_{p_topo}	Distância radial dos pixels da borda do espelho observada na imagem adquirida pelo sistema de visão omnidirecional
H_{pn}	Largura da imagem panorâmica
V_{pn}	Altura da imagem panorâmica
(u,v)	Coordenadas de um pixel na imagem adquirida pela câmera
(u_{pn},v_{pn})	Coordenadas de um pixel na imagem panorâmica obtida pela retificação da imagem adquirida pelo sistema de visão omnidirecional
r_p	Distância radial de um pixel na imagem adquirida pelo sistema de visão omnidirecional
$D \cdot V_{pn}$	Altura na dimensão vertical da imagem panorâmica que se estende da linha formada pelo mapeamento dos pixels correspondentes a borda do espelho, $x = r_{topo}$, até a linha formada pelo mapeamento dos pixels correspondentes a $y = 0$
A	Representa a distância entre o foco do espelho e o plano de projeção cilíndrico usado para obter a imagem panorâmica
K	Representa uma constante que relaciona a medida de f em milímetros para um equivalente dessa medida em pixel
(u_p,v_p)	Coordenadas de um pixel na imagem perspectiva obtida pela retificação da imagem adquirida pelo sistema de visão omnidirecional
f_p	Representa a distância, em pixels, do centro único de projeções ao plano de projeção perspectiva usado na geração de imagens perspectivas pela retificação da imagem adquirida pelo sistema de visão omnidirecional
θ	Ângulo azimutal

θ_0	Ângulo azimutal usado para definição de um plano de projeção perspectiva
φ_0	Ângulo de elevação usado para definição de um plano de projeção perspectiva
$\mathbf{d}$	Vetor que representa a translação do alvo na imagem
$\mathbf{A}$	Matriz que representa uma transformação do alvo na imagem (rotação, fator de escala e <i>shear</i>)
$I(\mathbf{x}, t)$	Intensidade de um pixel da região de referência localizado na coordenada $\mathbf{x}$ da imagem usada no rastreamento visual, para o instante de tempo t
$\mathbf{x}$	Posição de um pixel na imagem usada pelo método de rastreamento visual
W	Conjunto das posições dos pixels que representam o alvo na imagem
$O(\mathbf{A}, \mathbf{d})$	Função de erro que representa a soma do quadrado das diferenças usada no método de rastreamento visual de uma região de referência
$\omega(\cdot)$	Função de peso positiva arbitrária usada na equação de erro do método de rastreamento visual
$\mathbf{d}', \mathbf{A}'$	Transformações incrementais da região do alvo
$J(\mathbf{x}, t)$	Intensidade do pixel da região de referência transformada por $\mathbf{A}$ e $\mathbf{d}$
J_u, J_v	Gradientes espaciais de um pixel da região de referência transformada
I_u, I_v	Gradientes espaciais de um pixel da região de referência
α	Rotação do alvo na imagem
s_u, s_v	Fator de escala do alvo na imagem
γ	<i>Shear</i> do alvo na imagem

$h_i(\mathbf{x})$	Valor do resíduo do método de rastreamento visual após cálculo de um conjunto de parâmetros de transformação
res	Resíduo final do método de rastreamento visual, ou seja, parâmetro que corresponde a quão bem os parâmetros de transformação calculados correspondem aos valores reais
$\dot{\mathbf{r}}$	Vetor de velocidades de translação e rotação de um ponto
$\mathbf{T}$	Vetor de velocidades de translação de um ponto
$\boldsymbol{\Omega}$	Vetor de velocidades de rotação de um ponto
T	Componente de translação de um ponto ao longo de um eixo do sistema de coordenadas
w	Componente de rotação de um ponto ao redor de um eixo do sistema de coordenadas
$\mathbf{f}$	Vetor de parâmetros característicos da imagem
$\mathbf{J}_v$	Jacobiano da imagem
${}^c\mathbf{P}$	Vetor de coordenadas de um ponto do alvo em relação ao sistema de coordenadas da câmera
${}^c\boldsymbol{\Omega}_c$	Vetor de velocidades de rotação da câmera expresso no sistema de coordenadas da câmera
${}^c\mathbf{T}_c$	Vetor de velocidades de translação da câmera expresso no sistema de coordenadas da câmera
w_c	Componente do vetor de velocidade de rotação da câmera
T_c	Componente do vetor de velocidade de translação da câmera

$\dot{\mathbf{r}}_c$	Vetor de velocidades de translação e rotação da câmera
$\mathbf{f}_{ref}$	Vetor de referência dos parâmetros característicos da imagem
$\mathbf{K}$	Vetor de ganho proporcional
$\mathbf{E}(\mathbf{f})$	Vetor de erro em função dos parâmetros característicos da imagem
z_d	Distância desejada entre o alvo e a câmera
$\mathbf{U}$	Vetor de controle
T_b	Componente do vetor de velocidade de translação do robô
w_b	Componente do vetor de velocidade de rotação das rodas do robô
$\mathbf{J}_b$	Matriz que relaciona a variável de controle do robô com a velocidade de translação da câmera
θ_b	Ângulo de direção das rodas do robô
θ_c	Ângulo de direção azimutal da câmera
θ_r	Diferença entre o ângulo de direção das rodas e o ângulo de direção azimutal da câmera
z_t, x_t	Coordenadas da distância entre a câmera e a posição desejada para o robô expressas no sistema de coordenadas da câmera
$\mathbf{C}$	Pseudo inversa do Jacobiano da imagem calculado na posição de referência
$\mathbf{N}$	Inversa da matriz $\mathbf{J}_b$

Capítulo 1

INTRODUÇÃO

Robôs móveis autônomos devem possuir a capacidade de realizar tarefas sem a necessidade de intervenção humana. Um robô autônomo também deve ser capaz de perceber o ambiente ao seu redor, e então tomar decisões e agir conforme as mudanças sentidas a fim de cumprir seus objetivos. Realizar tarefas complexas de forma totalmente autônoma pode significar a realização coordenada de diversas tarefas mais simples. Cada uma destas tarefas pode possuir objetivos locais que são atingidos ao se utilizar as informações do meio em que o robô se encontra.

Geralmente deseja-se obter a maior quantidade possível de informação do ambiente no qual o robô está inserido, pois assim melhora-se a base de dados do robô para tomada de decisão e controle autônomo. Uma das diferentes formas de se obter isso é através da fusão de dados de diferentes tipos de sensores. Outra maneira é através da utilização de um único sensor como, por exemplo, um sistema de visão computacional capaz de obter uma imagem com a maior quantidade possível de informação do ambiente. Pode-se então utilizar um sistema de visão computacional projetado para adquirir imagens com campo de visão de 360°. Estes sistemas são chamados de sistemas de visão omnidirecional, e podem ser necessários, por exemplo, quando a tarefa a ser executada pelo robô necessita de informações visuais constantes de objetos localizados em diferentes direções.

Neste trabalho um sistema de visão omnidirecional é implementado e utilizado em uma tarefa de controle de um robô móvel. A tarefa de interesse é a de fazer com que o robô siga um objeto móvel utilizando informações visuais.

Quando deseja-se seguir um objeto ou alvo móvel, é necessário que o alvo a ser seguido esteja continuamente no campo de visão do sistema de visão do robô. Isso pode ser feito utilizando-se um sistema de visão omnidirecional embarcado no robô móvel. No entanto, se um sistema de visão convencional fosse utilizado, a câmera

deveria ser continuamente posicionada na direção do alvo seguido. Basicamente, esse posicionamento de uma câmera convencional poderia ser feito de duas formas. A primeira delas seria posicionar a câmera na direção do alvo através do movimento do robô. Neste caso, a direção da câmera seria alterada quando o corpo do robô girasse de um lado para outro, por exemplo, quando o robô fizesse curvas. Entretanto, esta maneira de posicionar a câmera não é válida para todos os tipos de robôs móveis pois o corpo de alguns robôs móveis, como é o caso do robô utilizado neste trabalho, permanece em uma direção fixa mesmo quando o robô realiza curvas. A segunda forma de se posicionar um sistema de visão convencional na direção do alvo seria utilizar um mecanismo do tipo *pan/tilt*. Neste caso, além do esforço de controle necessário para fazer com que o robô siga o alvo, seria preciso um esforço extra para que o controle da câmera pudesse ser feito. Portanto, ao se utilizar o sistema de visão omnidirecional, evita-se o controle do movimento da câmera para fazer com que o alvo permaneça no campo de visão do robô.

Para a tarefa de controle de interesse neste trabalho, uma vez que o alvo a ser seguido pelo robô esteja continuamente no campo de visão da câmera, é necessário obter características da imagem desse alvo que possam ser utilizadas no controle. Essas características são obtidas através de um processamento computacional da imagem. No entanto, apenas uma parte da imagem possui as informações de interesse, ou seja, a representação do alvo. Se o processamento computacional é feito apenas nesta região de interesse, o tempo computacional gasto é menor do que o tempo que se levaria para processar toda a imagem. Essa redução de tempo computacional é importante em aplicações de controle em tempo real, como é o caso deste trabalho. Uma maneira de se ter a região de interesse continuamente delimitada na imagem é através do uso de métodos de rastreamento visual em tempo real. Esses métodos baseiam-se no fato de que em uma sequência de imagens, de uma cena para outra, a representação do alvo possui uma pequena variação de posição e forma. Sendo assim, para que a posição do alvo na imagem atual seja determinada, é preciso fazer um processamento computacional apenas na região próxima a posição do alvo na imagem anterior. Determinada a região ocupada pelo alvo em cada imagem, pode-se obter em tempo real as informações visuais do alvo relevantes para o controle do robô.

O controle do robô deve ser feito de forma que as características visuais do alvo permaneçam constante na imagem. Dessa forma garante-se que o robô está seguindo o alvo. Assim, continuamente são calculadas as características visuais atuais do alvo e essas são comparadas com os valores de referência. Um valor de erro é calculado da diferença entre os valores das características atuais e de referência. Então o controle do robô é feito para que esse erro seja minimizado. Assim, existe uma realimentação em tempo real de informações visuais em uma malha de controle do robô. Essa abordagem de controle é denominada controle servo visual.

Portanto, de forma geral, a abordagem escolhida para a tarefa de controle de um robô para que ele siga um alvo móvel envolve três etapas descritas ao longo deste trabalho. A primeira delas é desenvolver um sistema de visão omnidirecional para ser embarcado no robô. Então utilizando o sistema de visão omnidirecional desenvolvido, realizar as demais etapas que consistem em fazer um rastreamento visual do alvo, e finalmente controlar o robô móvel com uma técnica de controle servo visual.

A dissertação está organizada da maneira descrita a seguir. No Capítulo 2 é apresentada uma introdução a sistemas de visão omnidirecional com ênfase nos sistemas compostos por espelhos convexos e câmeras. A formulação para projeto do espelho convexo escolhido para ser implementado também é apresentada. No Capítulo 3 são discutidos métodos de rastreamento visual e apresentada a formulação do método utilizado neste trabalho. No Capítulo 4 é apresentada uma introdução a sistemas de controle servo visual. Discute-se as possíveis arquiteturas de sistemas de controle servo visual existentes, e apresenta-se a formulação do método de controle servo visual utilizado. No Capítulo 5 são apresentados e discutidos os resultados obtidos com o sistema de visão, rastreamento visual, e com o controle servo visual implementado. Finalmente, o Capítulo 6 traz uma conclusão do trabalho.

Capítulo 2

VISÃO OMNIDIRECIONAL

Sistemas de visão omnidirecional permitem a aquisição de imagens com campo de visão de 360° . No intuito de desenvolver e aplicar este tipo de sistema, muitos trabalhos têm sido realizados por pesquisadores nas áreas de visão computacional, óptica aplicada e robótica. Yagi (1999) apresenta uma revisão bibliográfica onde são apresentadas aplicações de diversos tipos de sistemas de visão omnidirecional.

Uma maneira de se obter uma imagem com campo de visão de 360° é através da combinação, em uma única imagem, de múltiplas imagens adquiridas por uma câmera que gira com velocidade angular constante ao redor de um eixo vertical (Fig.2.1). Assim, ao término de uma volta, depois de feito um processamento nas diversas imagens capturadas, obtém-se uma imagem omnidirecional. Apesar da imagem omnidirecional obtida possuir uma alta resolução, este tipo de sistema não é adequado para aplicações onde o ambiente é dinâmico pois a imagem omnidirecional completa não é obtida em tempo real.

Figura 2.1 – Câmera giratória para formação de imagem omnidirecional

Seguindo essa mesma idéia de combinar múltiplas imagens em uma única imagem omnidirecional, pode-se construir um sistema omnidirecional composto de múltiplas câmeras apontadas para diferentes direções. Este tipo de sistema pode adquirir em tempo real imagens omnidirecionais de alta resolução. No entanto, a montagem de

tal sistema além de possuir dificuldades na calibração das câmeras, pode resultar em um tamanho não compacto para ser usado em um robô móvel.

Uma outra forma de se obter imagens omnidirecionais em tempo real é utilizar uma lente do tipo “olho de peixe” montada em uma única câmera fixa. Cao et al (1986) usa este tipo de sistema omnidirecional para controlar um robô móvel. Entretanto, este tipo de lente possui um custo alto se comparado com lentes convencionais. Além do mais, as imagens adquiridas possuem uma boa resolução apenas na região central. Nessa região da imagem é onde se encontra representada a imagem do teto, quando a câmera montada no robô móvel é apontada para cima, ou onde se encontra a imagem do próprio robô, quando a câmera é apontada para baixo. Portanto, os objetos do ambiente que são de interesse para a aplicação aparecem representados na periferia da imagem, onde a resolução é baixa.

Um sistema de visão capaz de produzir imagens omnidirecionais em tempo real também pode ser obtido através do uso combinado de espelhos convexos e lentes. A este tipo de sistema de visão dá-se o nome de sistema de visão omnidirecional catadióptrico¹. Neste sistema, uma câmera é montada em frente a um espelho convexo de forma a observar a imagem refletida no espelho. A forma do espelho utilizada pode ser projetada para que se obtenha imagens com características apropriadas para a aplicação que se fará do sistema de visão. O campo de visão vertical obtido com tal sistema pode, dependendo da forma de espelho utilizada, ser maior que o ângulo de visão obtido pelos outros tipos de sistemas omnidirecionais mencionados anteriormente. Também pode-se obter sistemas de visão de tamanho adequado para aplicações em robótica móvel.

No restante deste capítulo apresenta-se na seção 2.1 algumas das possíveis formas de espelho utilizadas em sistemas omnidirecionais catadióptricos bem como suas propriedades. Na seção 2.2 apresenta-se a formulação usada para especificar um espelho hiperbólico. Na seção 2.3 são descritos métodos de transformação da imagem adquirida pelo sistema de visão omnidirecional catadióptrico. Estes métodos

¹ A palavra catadióptrico vem da junção do termo catóptrico, que se refere a sistemas refletores (espelhos), e do termo dióptrico, que se refere a sistemas refratores (lentes)

são chamados de métodos de retificação da imagem, e através deles se obtém imagens panorâmicas ou perspectivas.

2.1 Sistemas de visão omnidirecional catadióptricos

Tipicamente, um sistema de visão omnidirecional catadióptrico consiste em se utilizar um espelho convexo fixo em frente à uma câmera que também permanece fixa. A forma do espelho, juntamente com o modelo de câmera e lente utilizado, confere às imagens obtidas propriedades que podem ser úteis para a aplicação que se fará do sistema omnidirecional.

Dentre as formas de espelho convexo, o espelho esférico, quando usado no sistema de visão omnidirecional, é o que resulta na imagem com maior campo de visão vertical. No entanto, à semelhança do sistema de visão com lente do tipo “olho de peixe”, a imagem obtida possui uma boa resolução na região central e uma baixa resolução na região periférica da imagem. Como exemplo de uma aplicação recente de um sistema de visão omnidirecional com espelho esférico, pode-se citar o trabalho de Gaspar et al (2000). Nesse trabalho é proposto um método de navegação de robôs móveis para ambientes internos baseado em visão.

Yagi et al (1994) utiliza um sistema de visão omnidirecional com espelho cônico na navegação de um robô móvel. Neste sistema de visão, as linhas verticais de objetos do ambiente, tais como portas, mesas e armários, são mapeadas como linhas radiais na imagem capturada. Dessa forma, através da detecção das linhas radiais e da análise da variação destas linhas, é feito um controle do robô para evitar colisão com obstáculos.

Chahl; Srinivasan (1997) desenvolveram uma formulação que descreve uma família de superfícies com uma propriedade específica. Ao longo da superfície do espelho é mantida uma relação constante entre a variação do ângulo de elevação ($\delta\varphi$) e a variação correspondente do ângulo radial ($\delta\psi$). Assim, nesse espelho é mantido um ganho angular vertical, $\beta = \delta\varphi / \delta\psi$, localmente constante. O ângulo de elevação (φ) corresponde ao ângulo de incidência da luz na superfície do espelho, e o ângulo

radial (ψ) correspondente ao ângulo de reflexão da luz para dentro do dispositivo de aquisição da imagem, conforme representado na Fig.(2.2). Considerando a imagem obtida pelo espelho como sendo uma imagem polar, o ângulo de elevação e o de azimute em coordenadas do mundo são mapeados linearmente no raio e ângulo da imagem omnidirecional polar. Isso facilita a obtenção de imagens panorâmicas livre de distorções na vertical. Uma imagem panorâmica pode ser obtida quando uma imagem adquirida por um sistema de visão omnidirecional catadióptrico é transformada de coordenadas polares para cartesianas. Isso é feito quando a coordenada do ângulo azimute na imagem original é mapeada no eixo de coordenada horizontal da imagem panorâmica, e a coordenada radial na imagem original é mapeada no eixo de coordenada vertical da imagem panorâmica.

Figura 2.2 – Ganho angular vertical

Normalmente, para uma imagem panorâmica gerada a partir da imagem adquirida pelo sistema de visão, a densidade de pixel é baixa para a região na vertical da imagem panorâmica que corresponde aos pixels do centro do espelho na imagem, e a densidade é alta para a região na vertical da imagem panorâmica que corresponde a pontos próximos da borda do espelho na imagem. Este problema deve ser levado em conta quando dois espelhos concêntricos são utilizados para construir de um sistema de visão estéreo panorâmico. Neste sistema, duas visões do ambiente são adquiridas

em uma mesma imagem, sendo que a visão do ambiente que se encontra na parte interna da imagem terá uma resolução menor que a outra. Isso dificulta a localização de objetos nas imagens. Tendo esta aplicação estéreo como objetivo, Conroy; Moore (1999) desenvolveram um espelho de forma semelhante ao de Chahl; Srinivasan (1997). No entanto, o perfil é projetado para manter uma relação constante entre o ângulo de elevação (φ) e a densidade de pixels. Dessa forma o espelho desenvolvido possui a característica de obter imagens omnidirecionais com resolução de pixel invariante. Assim, quando duas destas superfícies são projetadas de forma concêntrica para visualização de uma mesma região do ambiente, obtém-se um par estéreo de imagens omnidirecionais com a mesma resolução.

Hicks; Bajcsy (1999) também desenvolveram uma superfície com uma propriedade específica. A superfície desenvolvida possui a propriedade de preservar na imagem a geometria de um plano perpendicular ao eixo de simetria do espelho. Dessa forma, quando o sistema é posicionado na direção vertical, as imagens adquiridas pela câmera preservam a geometria de um plano horizontal. Como exemplo, suponha que a câmera é apontada para cima em direção ao espelho e que no chão exista um desenho como o de um tabuleiro de xadrez. Na imagem adquirida pelo sistema de visão omnidirecional aparecerá o padrão xadrez sem deformações, conforme apresentado na Fig.(2.3b). Todas as retas desenhadas no chão continuam retas na imagem omnidirecional, sem a necessidade de fazer qualquer processamento da imagem adquirida pelo sistema de visão. A imagem obtida com o espelho desenvolvido é uma projeção ortográfica escalonada do plano visualizado. Este tipo de visão pode ser chamado de “*birds eye view*”.

Gaspar et al (2000) obtêm uma imagem semelhante a obtida pelo espelho de Hicks; Bajcsy (1999) ao fazer um processamento computacional de uma imagem adquirida por um sistema de visão que utiliza um espelho esférico. Nesta imagem resultante, as linhas de um corredor aparecem como retas paralelas. Essa informação é usada para guiar o robô móvel de forma que este se mantenha no centro do corredor.

Figura 2.3 – Imagem de um piso xadrez obtida com (a) um espelho esférico e com (b) espelho projetado projetado por Hicks e Bajcsy. (Figura extraída de Hicks; Bajcsy, 1999).

Gächter et al (2001) desenvolveram um perfil de espelho com o objetivo de preservar na imagem adquirida pela câmera a dimensão de um objeto quando este se movimenta em um plano vertical. O espelho foi desenvolvido para uma câmera log-polar denominada SVAVISCA. Essa câmera possui os elementos sensores distribuídos circularmente e seguindo uma função de distribuição logarítmica do centro para a periferia. Através da leitura direta dos pixels do sensor desta câmera, obtém-se uma imagem panorâmica sem a necessidade de processamento da imagem. Considerando um cilindro alinhado com o eixo óptico da câmera, para as imagens obtidas com o espelho desenvolvido por Gächter et al (2001), a relação entre distâncias medidas na direção vertical ao longo da superfície cilíndrica permanece a mesma quando medidas na imagem obtida pela câmera. Segundo Gächter et al (2001), seria vantajoso o uso deste tipo de espelho e câmera em aplicações que necessitem do rastreamento visual de objetos móveis, pois neste caso o tamanho do objeto dependeria apenas da distância do objeto à câmera, e não mais de sua posição vertical.

Algumas formas de espelho associadas a determinados tipos de lente fazem com que o sistema de visão omnidirecional resultante tenha um centro único de projeção (*single effective viewpoint* ou *single projection center*). Para este tipo de sistema de visão omnidirecional, todos os raios se encontram em um único ponto após reflexão

no espelho. Este ponto é chamado de centro de projeção. Baker; Nayar (1998) desenvolveram expressões que representam todo o conjunto de sistemas de visão catadióptricos que possuem a propriedade de centro único de projeção. Basicamente a única maneira factível de obter este tipo de sistema é conciliando um espelho parabólico com uma câmera de projeção ortográfica, ou usando um espelho hiperbólico com uma câmera de projeção perspectiva, conforme mostra a Fig.(2.4).

Figura 2.4 – Três formas de espelho convexo, duas formas possuem centro único de projeção (Figura extraída de Svoboda et al, 1997)

O centro único de projeção torna possível a obtenção de imagens perspectivas livres de distorções a partir da imagem adquirida pelo sistema de visão omnidirecional. Essas imagens perspectivas podem ser obtidas através do mapeamento dos pixels da imagem omnidirecional em um plano perpendicular a um raio que passa pelo centro de projeção do sistema. A imagem obtida é como se tivesse sido adquirida por uma câmera perspectiva com o foco localizado no centro único de projeção do sistema. Além da imagem perspectiva, também pode-se obter imagens panorâmicas livres de

distorções na vertical através do mapeamento dos pixels da imagem em um plano cilíndrico ao redor do sistema de visão, conforme mostrado na Fig.(2.5).

Os trabalhos de Nayar (1997) e Peri; Nayar (1997) descrevem o sistema de visão omnidirecional implementado com uma câmera de projeção ortográfica e um espelho parabólico. Também apresentam um programa de visualização de imagem implementado. Este programa é capaz de simular uma câmera de projeção perspectiva virtual localizada no centro único de projeção. A direção de visão, o campo de visão e a magnificação desta câmera virtual podem ser ajustadas pelo usuário, e a imagem perspectiva mostrada é calculada em tempo real a partir da imagem adquirida pelo sistema de visão omnidirecional implementado.

Svododa et al. (1997) implementa um sistema de visão catadióptrico com centro único de projeção composto por uma câmera de projeção perspectiva e um espelho hiperbólico. Svododa et al. (1997) também mostra as curvas epipolares para este tipo de sistema, possibilitando o uso das imagens omnidirecionais adquiridas por esse sistema para cálculo de distâncias baseado em visão estéreo.

Figura 2.5 – Esquema representando a geração de imagens perspectivas e panorâmicas através da projeção da imagem adquirida em planos escolhidos.

Derrien; Konolige (2000) mostram um método para se obter aproximadamente um centro único de projeção para sistemas omnidirecionais catadióptricos, em particular para espelhos esféricos. Conhecido o percurso dos raios ópticos responsáveis pela formação da imagem no sistema de visão com o espelho esférico, o método consiste em mapear estes raios ópticos para um sistema virtual de imageamento que tenha um centro único de projeção. Nesse mapeamento são preservados os ângulos que os raios ópticos na direção dos objetos do ambiente formam com o eixo óptico do sistema. No entanto, as imagens perspectivas obtidas por esse método possuem uma distorção causada pela aproximação do centro único de projeção.

2.2 Espelho hiperbólico para sistema de visão omnidirecional

Dos tipos de sistemas de visão omnidirecional apresentados escolheu-se implementar o sistema catadióptrico formado por um espelho hiperbólico e uma câmera perspectiva. Esse sistema foi embarcado em um robô móvel e utilizado em uma aplicação de controle servo visual.

O espelho hiperbólico foi escolhido pois, baseado em resultados obtidos por Svoboda et al (1997), as imagens obtidas por esse tipo de espelho possuem uma boa distribuição de resolução. Essa distribuição de resolução é melhor que a obtida com um espelho esférico. Além disso, o sistema de visão omnidirecional formado com o espelho hiperbólico e uma câmera convencional de lente perspectiva possui a propriedade de centro único de projeção. Isso permite a obtenção de imagens panorâmicas ou perspectivas livres de distorções. Essa propriedade será usada na aplicação de rastreamento visual e controle servo visual implementada neste trabalho.

O espelho hiperbólico foi projetado de forma a se obter um sistema de visão omnidirecional compacto o suficiente para ser montado sobre o robô móvel. Para especificação da equação que descreve a superfície do espelho, seguiu-se a metodologia apresentada no trabalho de Svoboda et al (1997).

Conforme apresentado em Svoboda et al (1997), um espelho hiperbólico pode ser definido no sistema de coordenadas centrado no ponto focal F' , como mostrado na Fig.(2.5), pela seguinte equação em função dos parâmetros do espelho a e b :

$$y = \sqrt{a^2 \cdot \left(1 + \frac{x^2}{b^2}\right)} - \sqrt{a^2 + b^2} \quad (2.1)$$

Figura 2.6 – Geometria do espelho hiperbólico e da câmera

Na Figura (2.5), h é a distância entre a borda do espelho e o centro da câmera, $e = \sqrt{a^2 + b^2}$ representa a excentricidade do espelho, r_{topo} é a coordenada x do topo do espelho, e $y_{topo} = h - 2e$ é a coordenada y do topo do espelho. O valor máximo do ângulo de visão ϕ do espelho pode ser dado por:

$$\phi = \frac{\pi}{2} + \arctan\left(\frac{h - 2\sqrt{a^2 + b^2}}{r_{topo}}\right) \quad (2.2)$$

Para se chegar à equação final do espelho é necessário estabelecer uma relação entre os parâmetros a e b . Ao substituir na Eq.(2.1) as coordenadas do topo do espelho (r_{topo}, y_{topo}) , obtém-se a seguinte equação que expressa o parâmetro b em função da relação a/b :

$$b = h \cdot \sqrt{\left(\frac{a^2}{b^2} + 1\right)} - \frac{a}{b} \cdot \sqrt{h^2 + r_{topo}^2} \quad (2.3)$$

Para que o sistema de visão omnidirecional seja compacto, deve-se fixar um valor adequado para r_{topo} e h . Considerando-se que a projeção do espelho na imagem é um círculo e não uma elipse, ou seja, considerando-se que os fatores de escala vertical e horizontal da projeção da imagem na câmera são iguais, a altura h pode ser calculada em função da distância focal da câmera (f), da medida em pixel do raio da borda do espelho (r_{p_topo}), e da medida em milímetros do tamanho de cada pixel do CCD da câmera (t_{pixel}). A seguinte equação expressa h em função destes parâmetros:

$$h = \frac{f \cdot r_{topo}}{t_{pixel} \cdot r_{p_topo}} \quad (2.4)$$

Os parâmetros f , t_{pixel} e r_{p_topo} são parâmetros da lente e câmera utilizada, e da imagem adquirida pelo sistema. Portanto deve ser escolhida a dimensão do espelho (r_{topo}) de tal forma que se obtenha um valor para h desejado. Determinado o valor de h , escolhe-se um valor para a relação a/b e então determina-se o valor de a e b através da Eq.(2.3). Determinados a e b , a equação do espelho e o ângulo máximo de visão podem ser obtidos pela Eq.(2.1) e (2.2) respectivamente. Svoboda et al (1997) apresenta algumas formas de espelho e ângulos de visão obtidos para alguns valores de a/b .

2.3 Retificação de imagem adquirida pelo sistema de visão omnidirecional

Um sistema de visão omnidirecional catadióptrico mapeia o ambiente ao seu redor em coordenadas polares na imagem adquirida pelo sistema. Muitas vezes, conforme a aplicação, se faz necessário transformar a imagem polar adquirida pelo sistema de visão omnidirecional em uma outra imagem através de um re-mapeamento dos

pixels. Esse processo de transformação é denominado retificação da imagem. Através dele pode-se obter uma imagem panorâmica, ou se possível, uma imagem perspectiva. Estas imagens são de fácil observação pelo homem. Além disso, quando é possível obter imagens perspectivas a partir da imagem polar adquirida pelo sistema de visão, algoritmos de processamento de imagem desenvolvidos para imagens perspectiva podem ser utilizados.

A seguir são descritos três métodos de retificação de imagens omnidirecionais implementados. Na seção 2.3.1 é descrito um método para geração de imagens panorâmicas que se baseia na transformação direta de coordenadas polares para cartesianas. Na seção 2.3.2 é descrito outro método para geração de imagens panorâmicas através da transformação de coordenadas polares para cartesianas. No entanto, neste último a equação do espelho hiperbólico e a propriedade de centro único de projeção são consideradas para obtenção de uma imagem livre de distorções verticais. Finalmente, na seção 2.3.3 é descrito um método para geração de imagens perspectivas que utiliza a equação do espelho hiperbólico e sua propriedade de centro único de projeção.

2.3.1 Transformação direta de coordenadas polares para cartesianas

Conforme mencionado anteriormente, uma imagem panorâmica pode ser obtida quando uma imagem adquirida por um sistema de visão omnidirecional catadióptrico é transformada de coordenadas polares para cartesianas. Isso é feito quando a coordenada do ângulo azimuth na imagem original é mapeada no eixo de coordenada horizontal da imagem panorâmica, e a coordenada radial na imagem original é mapeada no eixo de coordenada vertical da imagem panorâmica, conforme ilustrado pela Fig.(2.7).

No desenvolvimento das equações usadas no processo de retificação, a coordenada do centro da imagem adquirida pelo sistema de visão omnidirecional pode ser considerada como sendo a origem (0,0). Na imagem adquirida, a coordenada radial da borda do espelho medida em pixels vale r_{p_topo} . A altura da imagem panorâmica a ser gerada possui valor dado por V_{pn} , e a largura da imagem panorâmica possui valor dado por H_{pn} . Dessa forma, se os pixels ao longo da dimensão radial da imagem

polar adquirida são mapeados linearmente na vertical da imagem panorâmica, as seguintes equações podem ser usadas para fazer a retificação:

$$u = \frac{v_{pn} \cdot r_{p_topo}}{V_{pn}} \cos\left(\frac{2\pi \cdot u_{pn}}{H_{pn}}\right) \quad (2.5)$$

$$v = \frac{v_{pn} \cdot r_{p_topo}}{V_{pn}} \sin\left(\frac{2\pi \cdot u_{pn}}{H_{pn}}\right) \quad (2.6)$$

Figura 2.7 – Esquema representando o processo de retificação de uma imagem

Nas Equações (2.5) e (2.6), u e v são as coordenadas de um pixel na imagem polar capturada pelo sistema de visão omnidirecional que corresponde ao pixel de coordenadas u_{pn} e v_{pn} na imagem panorâmica.

Um determinado pixel na imagem adquirida pelo sistema de visão pode corresponder a vários pixels na imagem panorâmica. Por isso é necessário fazer um mapeamento inverso, ou seja, para cada pixel na imagem panorâmica, determina-se um correspondente na imagem polar adquirida pelo sistema de visão. Isso garante que todos os pixels na imagem panorâmica gerada sejam preenchidos.

Para fazer a retificação em tempo real da imagem, pode-se usar uma *lookup table*. Uma *lookup table* é uma tabela usada para guardar a correspondência entre os pixels da imagem retificada e os pixels da imagem polar adquirida pelo sistema. Uma vez que a relação entre os pixels é calculada pelas Eq.(2.5) e (2.6), essa relação é guardada em uma *lookup table* para que o cálculo não precise ser refeito para cada nova imagem adquirida.

2.3.2 Transformação de coordenadas polares para cartesianas utilizando equação da superfície do espelho hiperbólico e centro único de projeção

Na formação da imagem adquirida pelo sistema de visão omnidirecional muitas vezes ocorre um mapeamento polar não linear do ambiente. Assim, ao longo da direção radial da imagem adquirida ocorrem distorções não lineares do ambiente de acordo com a forma de espelho utilizada. Neste caso, quando usado o método de retificação descrito na seção anterior, as imagens panorâmicas geradas permanecerão distorcidas ao longo da vertical da imagem. Estas distorções podem ser corrigidas se a equação do espelho for usada no processo de retificação da imagem.

O sistema de visão omnidirecional formado com o espelho hiperbólico e uma câmera convencional de lente perspectiva possui a propriedade de centro único de projeção. O centro único de projeção é o ponto onde os raios ópticos refletidos pelo espelho se interceptam. Para efeito prático, é como se o ambiente pudesse ser visto em todas as direções a partir deste ponto. Utilizando essa propriedade juntamente com a equação do espelho pode-se obter uma imagem panorâmica ao mapear os pixels da imagem polar em um plano cilíndrico ao redor do sistema de visão (Fig.2.5). Assim pode-se corrigir as distorções na vertical da imagem panorâmica causada pela forma do espelho.

Figura (2.8) mostra como os pixels ao longo de um raio na imagem polar capturada pelo sistema de visão omnidirecional podem ser mapeados em uma coluna na imagem panorâmica. Neste método de retificação, considera-se a geometria do espelho hiperbólico e a propriedade de centro único de projeção para o mapeamento dos pixels.

$$\tan \varphi_{\max} = \frac{y_{\text{topo}}}{r_{\text{topo}}} = \frac{D \cdot V_{pn}}{A} \Rightarrow A = \frac{D \cdot V_{pn} \cdot r_{\text{topo}}}{y_{\text{topo}}} \quad (2.7)$$

onde r_{topo} , y_{topo} e D são conhecidas, e A representa a distância entre o foco do espelho e o plano de projeção cilíndrico.

$$\tan \varphi = \frac{(D \cdot V_{pn} - v_{pn})}{A} = \frac{y}{x} \Rightarrow y = \frac{x \cdot (D \cdot V_{pn} - v_{pn})}{A} \quad (2.8)$$

$$\tan \psi_{\max} = \frac{r_{p_topo}}{f \cdot K} = \frac{r_{\text{topo}}}{2e + y_{\text{topo}}} \Rightarrow K = \frac{r_{p_topo} \cdot (2e + y_{\text{topo}})}{f \cdot r_{\text{topo}}} \quad (2.9)$$

onde K representa uma constante que relaciona a medida de f em milímetros para um equivalente dessa medida em pixel.

$$\tan \psi = \frac{x}{(2e + y)} = \frac{r_p}{f \cdot K} \quad (2.10)$$

Substituindo a Eq.(2.8) na Eq.(2.10) obtém-se:

$$\frac{A \cdot x}{A \cdot 2e + x \cdot (D \cdot V_{pn} - v_{pn})} = \frac{r_p}{f \cdot K} \Rightarrow r_p = \frac{f \cdot A \cdot x \cdot K}{A \cdot 2e + x \cdot (D \cdot V_{pn} - v_{pn})} \quad (2.11)$$

Substituindo as Eq.(2.7) e (2.9) na Eq.(2.11), obtém-se:

$$r_p = \frac{x \cdot D \cdot V_{pn} \cdot r_{p_topo} \cdot (2e + y_{\text{topo}})}{(D \cdot V_{pn} - v_{pn}) \cdot x \cdot y_{\text{topo}} + D \cdot V_{pn} \cdot 2e \cdot r_{\text{topo}}} \quad (2.12)$$

Substituindo a Eq.(2.7) na Eq.(2.8) obtém-se:

$$y = \frac{x \cdot (D \cdot V_{pn} - v_{pn}) \cdot y_{\text{topo}}}{D \cdot V_{pn} \cdot r_{\text{topo}}} \quad (2.13)$$

A Equação (2.13) representa uma reta que passa pelo foco da hipérbole e também passa por v_{pn} . Para um dado valor de v_{pn} , iguala-se a Eq.(2.13) com a Eq.(2.1) do espelho e determina-se a raiz do polinômio através do método numérico de Newton,

obtendo-se o valor de x onde a reta e o espelho se cruzam. Substituindo o valor de x obtido na Eq.(2.12), chega-se a distância radial r_p de um pixel na imagem polar que equivale a um pixel de coordenada vertical v_{pn} na imagem panorâmica.

Assim como no primeiro método de retificação apresentado, aqui também é feito um mapeamento inverso, ou seja, para cada pixel da imagem final retificada busca-se um correspondente na imagem capturada pelo sistema de visão.

A formulação apresentada até aqui faz o mapeamento inverso apenas de uma coluna da imagem panorâmica para um raio na imagem capturada pelo sistema de visão, conforme mostrado na Fig.(2.7). As Eq.(2.14) e (2.15) estendem o mapeamento para toda a imagem omnidirecional. Nestas equações, u_{pn} representa qual coluna na imagem panorâmica está sendo mapeada; r_p é o valor obtido através da Eq.(2.13) em função de v_{pn} . Os parâmetros u e v representam a coordenada do pixel na imagem capturada pelo sistema de visão que corresponde ao pixel de coordenadas (u_{pn}, v_{pn}) na imagem panorâmica. Nas Eq.(2.14) e (2.15) assume-se que a origem (0,0) da imagem capturada pelo sistema de visão se encontra no centro da imagem.

$$u = r_p \cos\left(\frac{2\pi \cdot u_{pn}}{H_{pn}}\right) \quad (2.14)$$

$$v = r_p \sin\left(\frac{2\pi \cdot u_{pn}}{H_{pn}}\right) \quad (2.15)$$

A retificação da imagem em tempo real utilizando o método descrito pode também ser implementada utilizando-se uma *lookup table*. Assim, o método iterativo de Newton é utilizado apenas uma vez para determinar a correspondência de cada pixel na imagem panorâmica com os pixels da imagem polar adquirida pelo sistema. Então, a correspondência calculada entre os pixels é guardada na *lookup table*.

2.3.3 Retificação da imagem omnidirecional para geração de imagem perspectiva

A propriedade de centro único de projeção permite a obtenção de imagens perspectivas a partir da imagem polar adquirida pelo sistema de visão omnidirecional. No sistema de visão com espelho hiperbólico, o centro único de

projeção localiza-se no foco da hipérbole F' . Então, pode-se definir um plano de projeção perpendicular a qualquer reta que passa pelo foco da hipérbole, e mapear os pixels da imagem polar neste plano (Fig.2.5). As imagens obtidas por esse processo de retificação são como se fossem imagens adquiridas por uma câmera perspectiva virtual com foco localizado no foco da hipérbole.

Um plano de projeção perspectiva pode ser definido por $(f_p, \theta_0, \varphi_0)$, sendo que f_p representa a distância, em pixel, do foco da hipérbole F' ao plano definido, θ_0 representa o ângulo de direção azimutal do plano, e φ_0 representa o ângulo de elevação do plano conforme mostrado na Fig.(2.9). Fazendo uma comparação destes parâmetros com uma câmera virtual perspectiva montada em um mecanismo *pan/tilt*, f_p seria a distância focal da câmera virtual, e θ_0 e φ_0 seriam os ângulos *pan* e *tilt* respectivamente.

Figura 2.9 – Plano de projeção perspectiva

Figura 2.10 – Relações geometria usadas para geração de imagens perspectivas

A coordenada de um pixel no plano de projeção perspectiva (u_p, v_p) pode ser relacionada com a direção (θ, φ) de um raio de luz que chega ao sensor da câmera e forma um pixel na imagem polar. Esta relação pode ser determinada através de relações geométricas, conforme mostrado na Fig.(2.10), e é expressa pelas duas equações apresentadas a seguir:

$$\tan \varphi = \frac{f_p \sin \varphi_0 + v_p \cos \varphi_0}{f_p \cos \varphi_0} \quad (2.16)$$

$$\tan \theta = \frac{(f_p \cos \varphi_0 - v_p \sin \varphi_0) \cdot \sin \theta_0 - u_p \cos \theta_0}{(f_p \cos \varphi_0 - v_p \sin \varphi_0) \cdot \cos \theta_0 + u_p \sin \theta_0} \quad (2.17)$$

Para um determinado raio de luz com ângulo de elevação φ , é possível encontrar o ponto (x, y) na superfície do espelho onde o raio de luz é refletido. Isto é feito resolvendo-se a equação do espelho (Eq.2.1) para $y = x \cdot \tan \varphi$. Determinado o ponto (x, y) na superfície do espelho, o pixel (u, v) na imagem adquirida pela câmera que corresponde ao raio de luz definido pelos ângulos de direção (θ, φ) pode ser encontrado usando as seguintes equações:

$$u = \frac{x \cdot (2e + y_{topo}) \cdot r_{pixel}}{(x \cdot \tan \varphi + 2e)} \cos \theta \quad (2.18)$$

$$v = \frac{x \cdot (2e + y_{topo}) \cdot r_{pixel}}{(x \cdot \tan \varphi + 2e)} \sin \theta \quad (2.19)$$

Portanto, para gerar uma imagem perspectiva a partir da imagem polar adquirida pelo sistema de visão, primeiro define-se a direção e distância do plano de projeção perspectiva $(f_p, \theta_0, \varphi_0)$. O parâmetro f_p define o zoom da imagem perspectiva. Para cada pixel (u_p, v_p) do plano de projeção, determina-se através das Eq.(2.16) e (2.17) a direção (θ, φ) de um raio que passa pelo foco da hipérbole e pelo pixel (u_p, v_p) . A seguir resolve-se a equação do espelho (Eq.2.1) para $y = x \cdot \tan \varphi$. Finalmente, com a coordenada do ponto (x, y) do espelho, determina-se através da Eq.(2.18) e Eq.(2.19) o pixel (u, v) na imagem polar que corresponde ao pixel (u_p, v_p) na imagem perspectiva.

Apesar do ponto (x, y) ser encontrado usando-se um método computacional iterativo, é possível gerar imagens perspectivas em tempo real usando *lookup tables*. Desta forma, é criada uma *lookup table* relacionando todos os pixels (u, v) na imagem polar adquirida pela câmera com os ângulos (θ, φ) correspondentes. Então definido um plano perspectivo $(f_p, \theta_0, \varphi_0)$, cada pixels (u_p, v_p) na imagem

perspectiva gerada pode ser relacionado em tempo real ao seu pixel correspondente (u,v) na imagem adquirida. Isso é feito através da *lookup table* e da Eq.(2.16) e Eq.(2.17).

Capítulo 3

RASTREAMENTO VISUAL

A tarefa de interesse neste trabalho consiste em controlar um robô móvel utilizando informações visuais capturadas pelo sistema de visão omnidirecional embarcado no robô. Este controle têm como objetivo fazer com que o robô siga um alvo móvel. A medida que o alvo móvel ou o robô se movimentam no ambiente, ocorrem alterações na representação do alvo na imagem. Então, uma vez que as alterações na imagem do alvo possam ser medidas, utiliza-se essa informação para controle do robô móvel. Para se medir as alterações na imagem do alvo, pode-se extrair características, do alvo através de um processamento computacional da imagem. Essas características são relações mensuráveis tais como localização de vértices, centróide, deformações de um padrão local de intensidade de pixels, dentre outros.

A extração de características do alvo deve ser feita em tempo real para o controle em malha fechada do robô possa ser realizado. O tempo computacional gasto na extração de características do alvo pode ser reduzido quando, ao invés de toda a imagem, se utiliza apenas uma pequena região onde está localizado o alvo. Isso é válido até mesmo quando a segmentação do o alvo na imagem pode ser obtida diretamente através da aplicação de um limiar ou gradiente.

Como o movimento do alvo é contínuo, em uma sequência de imagens, a variação de posição e forma do alvo é pequena de uma cena para outra. Métodos de rastreamento visual se baseiam neste fato para determinar em tempo real a região que delimita o alvo na imagem. Para isso, estes métodos utilizam a informação de posição e forma do alvo nas imagens anteriores sem que precisem fazer o processamento de toda a imagem. Com a região do alvo delimitada, pode-se obter em tempo real as informações visuais relevantes para que o controle do robô seja feito.

Na seção 3.1 deste capítulo são apresentados diversos métodos de rastreamento visual discutindo a possibilidade de uso destes em sistemas de visão omnidirecional.

Na seção 3.2 apresenta-se a formulação de um método de rastreamento visual de uma região de referência. Este método foi escolhido para determinar as características visuais do alvo que servirão como realimentação sensorial na malha de controle do robô. Na seção 3.3 descreve-se como este método de rastreamento visual escolhido pode ser usado com o sistema de visão omnidirecional desenvolvido neste trabalho.

3.1 Métodos de rastreamento visual

A escolha de um método de rastreamento visual para utilização em controle de robôs móveis depende muito do tipo de tarefa proposta, do ambiente onde esta tarefa será realizada, e do objeto que deverá ser rastreado visualmente. A seguir apresenta-se uma breve descrição de alguns métodos de rastreamento visual.

3.1.1 Rastreamento de aglomerado de pixels

Para algumas aplicações, o alvo pode ser identificado na imagem como sendo um aglomerado de pixels (*blob*) que possuem uma determinada característica. Para estes casos, pode-se rastrear o alvo ao fazer o rastreamento do aglomerado de pixel. Este método baseia-se na segmentação do alvo de acordo com algum critério como, por exemplo, intensidade dos pixels, cor ou movimento.

Uma vez que o alvo é localizado dentro de toda a imagem, delimita-se uma pequena região em torno do alvo, chamada de janela, que será utilizada para determinar a posição do alvo nas imagens posteriores. A cada nova imagem capturada, aplica-se o critério de segmentação apenas na janela da imagem. Então a posição do alvo na janela pode ser determinada pelo centro de massa (centróide) do aglomerado de pixels segmentado. Com base na posição atual do alvo, a localização da janela de interesse é atualizada.

Ao se utilizar o método de rastreamento de aglomerado de pixel da maneira como foi descrito, obtém-se a cada instante de tempo a posição do centróide do alvo. Essa característica do alvo pode ser utilizada para controle servo visual do robô móvel.

Utilizando uma câmera fixa externa ao robô, Santos-Victor (2001) faz o controle servo visual de um robô móvel. Um método de rastreamento de aglomerado de pixel

é utilizado para rastrear a posição do robô na imagem. O aglomerado de pixels é segmentado através da determinação de movimento na imagem, que pode ser calculada através da diferença entre duas imagens consecutivas. Então, a posição do robô na imagem é estimada a partir do cálculo do centro de massa do aglomerado de pixels. Todos os cálculos são feitos apenas em uma janela que se adapta ao redor da posição do robô. Quando o sistema de rastreamento perde a localização do robô na imagem, o tamanho da janela é aumentado até que o robô volte a ser rastreado.

Carvalho et al (2000) também faz uso do método de rastreamento de aglomerado de pixels em uma tarefa de controle servo visual de um robô móvel. Utilizando uma câmera montada no robô móvel, o controle servo visual é feito para posicionar o robô em relação a um alvo fixo. O alvo utilizado é composto de quatro círculos pretos posicionados em um fundo branco como se estivessem nos vértices de um quadrado. O rastreamento visual utiliza a intensidade dos pixels como critério de segmentação.

O rastreamento de um aglomerado de pixels possui certas limitações quanto ao tipo de alvo utilizado na aplicação de controle. Afinal, este método de rastreamento é usado quando o alvo pode ser identificado diretamente após a aplicação de um ou mais métodos de segmentação. No entanto, da mesma maneira que este método é aplicado para imagens obtidas por sistemas de visão convencionais, o método de rastreamento de aglomerado de pixels pode ser aplicado para qualquer tipo de imagem omnidirecional.

3.1.2 Rastreamento de bordas

Quando o alvo possui uma forma geométrica bem definida e conhecida, pode-se fazer o rastreamento deste alvo a partir do rastreamento de suas características geométricas. Essas características geométricas podem ser, por exemplo, retas, intersecção de retas ou cantos. Estas características são identificadas na imagem através da detecção de bordas, ou seja, mudanças bruscas na intensidade dos pixels.

Bordas são detectadas na imagem utilizando-se um cálculo de gradiente de intensidade de pixels. Como o gradiente só depende da variação local de intensidade, esse tipo de método de rastreamento é praticamente insensíveis a variações de iluminação no ambiente. Entretanto, erros de rastreamento podem ocorrer

dependendo do alvo rastreado e do ambiente utilizado. Afinal, bordas detectadas no ambiente podem ser confundidas com o alvo de interesse.

Um exemplo da aplicação deste método de rastreamento com sistema de visão omnidirecional pode ser visto no trabalho de Gaspar et al (2000). Neste trabalho, através de um processamento da imagem obtida por um espelho esférico, obtém-se imagens do tipo “*bird’s eye view*”. Nestas imagens, a geometria do plano horizontal é mantida inalterada, portanto, em um corredor, as retas formadas pelo encontro das paredes com o chão permanecem retas na imagem omnidirecional. As linhas verticais no ambiente também permanecem retas na imagem. Dessa forma, faz-se o rastreamento destas retas para permitir a navegação do robô.

De modo geral, a geometria do alvo quando projetada na imagem omnidirecional pode sofrer deformações. Assim, seguimentos de retas podem ser projetados em curvas na imagem omnidirecional. Portanto deve-se levar isso em consideração ao utilizar um método de rastreamento de bordas aplicado diretamente na imagem omnidirecional adquirida pelo sistema de visão.

3.1.3 Rastreamento de uma região de referência

Em algumas aplicações o alvo a ser seguido é representado por uma região com um padrão de pixels específico, como por exemplo, o rosto de uma pessoa. Neste caso, o rastreamento é feito baseado na correlação da imagem com o padrão de pixels que representa o alvo. Portanto, o alvo é identificado ao longo da sequência de imagens que vão sendo capturadas como sendo a região que melhor corresponde ao padrão de pixels escolhido.

Para que este método possa ser utilizado em tempo real, deve-se limitar a região de busca do alvo. Assim, em um determinado momento durante o rastreamento, considera-se a região de busca do alvo na imagem como sendo as proximidades da última localização conhecida do alvo.

Alguns métodos de rastreamento de uma região de referência consideram que a região que representa o alvo pode sofrer deformações (ou transformações). Estas deformações podem ser modeladas por translação, rotação, fator de escala, etc.

Algumas destas transformações são mostradas na Fig.(3.1). Nestes métodos de rastreamento, o alvo é determinado pela região na imagem que sujeita a um modelo de deformação minimiza uma função de erro em relação ao padrão de referência. A função usada para medir o erro entre as regiões pode ser, por exemplo, a soma do quadrado das diferenças entre as regiões (*SSD: sum-of-squared-difference*). A cada instante de tempo, os valores de retorno deste tipo de método de rastreamento são os valores atribuídos a cada um dos componentes da transformação sofrida pelo alvo, ou seja, os valores de translação, rotação, fator de escala, etc que minimizam a função de erro.

Figura 3.1 – Graus de liberdade de uma transformação projetiva de um plano: (a) translação ao longo do eixo horizontal da imagem; (b) translação ao longo do eixo vertical da imagem; (c) rotação; (d) fator de escala; (e) shear; (f) relação de aspecto; (g) distorção de projeção ao longo do eixo horizontal da imagem; (h) distorção de projeção ao longo do eixo vertical da imagem. (figura extraída de Van der Zwaan; Santos-Victor, 2001).

Hager; Toyama (1998) descrevem um método de rastreamento em tempo real que utiliza a soma do quadrado da diferença (SSD) como forma de medir o erro entre a região de referência e a região na imagem observada. Neste método um modelo de transformação afim é utilizado para representar a deformação sofrida pela região de referência. Hager; Belhumeur (1998) estendem o método para um modelo genérico de transformações incluindo um modelo para iluminação.

Van der Zwaan; Santos-Victor (2001) apresentam outra formulação que utiliza a soma do quadrado da diferença (SSD) como forma de medir o erro entre as regiões. Um modelo de transformação da região de referência que inclui translação, rotação,

fator de escala, *shear*, relação de aspecto, e distorção de projeção perspectiva é utilizado. O cálculo do fluxo óptico também foi incluído no método de rastreamento em uma etapa de predição. Esse método de rastreamento é utilizado no controle servo visual de um robô móvel submarino para realizar a tarefa de permanecer estacionado sobre uma determinada região cuja referência é o solo marinho em águas rasas.

No trabalho de Chang; Hebert (1998) o método de rastreamento de região de referência com cálculo de erro por SSD é utilizado para imagens omnidirecionais obtidas com um espelho parabólico. Duas abordagens para o rastreamento visual são apresentadas. A primeira é fazer o rastreamento visual em imagens perspectivas geradas a partir da imagem omnidirecional adquirida pelo sistema de visão. A segunda abordagem é fazer o rastreamento diretamente na imagem omnidirecional. Neste último caso, a transformação da região do alvo na imagem omnidirecional não pode ser representada como um modelo de transformação afim devido a transformação não-linear causada pela forma do espelho. Dessa maneira, Chang; Hebert (1998) derivam uma formulação que leva em conta essa transformação não-linear. Na formulação proposta o modelo de transformação do alvo é simplificado como sendo composto apenas pela componente de translação do alvo.

3.1.4 Considerações gerais sobre os métodos de rastreamento visual

Velocidade e robustez são fatores importantes em um método de rastreamento visual em tempo real. Velocidade pode ser definida como sendo o tempo ou a frequência com que o método pode ser executado. Robustez pode ser definida como sendo a capacidade de um sistema de rastreamento de rastrear de forma precisa mesmo quando as circunstâncias visuais não são favoráveis ou ideais.

De forma geral, observa-se que existe uma relação de compromisso entre velocidade e robustez, pois quanto menor o tamanho da janela que limita a busca do alvo dentro da imagem menor a robustez do sistema de rastreamento e maior a velocidade de processamento. Desta forma, estes dois fatores devem ser balanceados dependendo da aplicação, chegando-se a um modo de operação satisfatório.

Van der Zwaan; Santos-Victor (2001) mostram resultados analisando a diminuição do erro no rastreamento do alvo com o aumento da dimensão da região de referência, e a conseqüente diminuição na frequência de rastreamento com o aumento da região de referência.

Toyama; Hager (1999) propõe um sistema de rastreamento visual em tempo real com uma arquitetura robusta e adaptativa. Esse sistema utiliza vários métodos de rastreamento visual dispostos em camadas sendo que alguns métodos são mais rápidos e pouco precisos, e outros métodos mais complexos e precisos. O sistema controla a transição entre os diferentes métodos de forma a executar o método mais adequado segundo as condições visuais do ambiente no qual o alvo está inserido.

Para aumentar a velocidade e robustez do rastreamento, alguns trabalhos utilizam modelos dinâmicos conhecidos do movimento do alvo ou da câmera para estimar a posição do alvo na próxima imagem a ser capturada.

Filtros também podem ser usados para melhorar a estimativa de medições feitas a partir do rastreamento visual. No trabalho de Santos-Victor (2001) onde é feito um controle servo visual de um robô móvel usando uma câmera fixa externa ao robô, a variação de posição do alvo calculada pelo método de rastreamento visual é pequena e por isso leva a erros na medida do ângulo de direção do robô. Então um filtro é utilizado para melhorar a estimativa da medida da direção de movimento do alvo.

3.2 Rastreamento visual de uma região de referência baseado no cálculo da soma do quadrado das diferenças (SSD)

Neste trabalho foi escolhido o método de rastreamento visual baseado em região de referência, conforme proposto por Hager; Toyama (1998), para determinar as transformações sofridas pelo alvo na imagem, e assim fornecer parâmetros de entrada para o sistema de controle servo visual. Para este método de rastreamento, considera-se que a imagem na qual o alvo está sendo rastreado é uma imagem de projeção ortográfica ou de projeção perspectiva. O alvo é assumido como sendo uma superfície plana e rígida, e que, no caso da imagem ser uma projeção perspectiva, a projeção do alvo nesta imagem é relativamente pequena se comparada com a

imagem como um todo. Dessa forma os efeitos de perspectiva são considerados como mínimos e as distorções geométricas da região de referência podem ser modeladas por transformações do tipo afim, ou seja, por uma transformação que possui apenas translação, rotação, fator de escala e *shear* como graus de liberdade. Essa transformação afim pode ser representada por um vetor de translação $\mathbf{d} = [du_p, dv_p]^T$ e uma matriz 2 x 2 positiva definida $\mathbf{A}$. O algoritmo de rastreamento determina a transformação afim sofrida pela região de referência minimizando uma função de erro que é dada pela SSD (*sum-of-square-differences*) entre a imagem de referência e a imagem do alvo. A seguir apresenta-se a formulação para este método de rastreamento conforme descrita por Hager; Toyama (1998).

Na notação utilizada, $I(\mathbf{x}, t)$ representa a intensidade do pixel localizado na posição $\mathbf{x} = [u_p, v_p]^T$ da imagem capturada no instante t . Em um instante t_0 , a região de referência, ou seja, o alvo a ser seguido, está representado em uma região da imagem cujo conjunto das posições dos pixels dessa região pode ser expresso por W . Em um instante posterior, $t > t_0$, a região de interesse será uma transformação do tipo afim da região de referência. Se a iluminação permanecer constante, a relação geométrica entre as duas regiões pode ser encontrada ao minimizar a seguinte função:

$$O(\mathbf{A}, \mathbf{d}) = \sum_{\mathbf{x} \in W} (I(\mathbf{Ax} + \mathbf{d}, t) - I(\mathbf{x}, t_0))^2 \omega(\mathbf{x}) \quad (3.1)$$

Na equação acima, $\mathbf{A}$ e $\mathbf{d}$ representam a transformação afim, $\mathbf{x} = [u_p, v_p]^T$, e $\omega(\cdot)$ é uma função de peso positiva arbitrária.

Supondo que a solução no instante de tempo t , $(\mathbf{A}_t, \mathbf{d}_t)$, é conhecida e que o objetivo seja calcular a solução no instante de tempo $t + \tau$ para valores pequenos e positivos de τ . A solução no instante de tempo $t + \tau$ em função de mudanças incrementais, $\mathbf{A}'$ e $\mathbf{d}'$, que ocorrem no intervalo de tempo τ pode ser dada por:

$$\begin{aligned} \mathbf{A}_{t+\tau} &= \mathbf{A}_t (\mathbf{I} + \mathbf{A}') \\ \mathbf{d}_{t+\tau} &= \mathbf{d}_t + \mathbf{A}_t \mathbf{d}' \end{aligned} \quad (3.2)$$

Substituindo este resultado na Eq.(3.1) obtém-se:

$$O'(\mathbf{A}_{t+\tau}, \mathbf{d}_{t+\tau}) = \sum_{\mathbf{x} \in W} (I(\mathbf{A}_t(\mathbf{I} + \mathbf{A}')\mathbf{x} + \mathbf{A}_t\mathbf{d}' + \mathbf{d}_t, t + \tau) - I(\mathbf{x}, t_0))^2 \omega(\mathbf{x}) \quad (3.3)$$

A região de referência submetida a uma transformação afim pode ser expressa por $J(\mathbf{x}, t) = I(\mathbf{A}_t\mathbf{x} + \mathbf{d}_t, t)$. Então, reescrevendo a Eq.(3.3) em função de $J, \mathbf{A}'$ e $\mathbf{d}'$, obtém-se a seguinte função de erro:

$$O'(\mathbf{A}', \mathbf{d}') = \sum_{\mathbf{x} \in W} (J(\mathbf{x} + \mathbf{A}'\mathbf{x} + \mathbf{d}', t + \tau) - I(\mathbf{x}, t_0))^2 \omega(\mathbf{x}) \quad (3.4)$$

Linearizando J em torno de $(\mathbf{A}', \mathbf{d}') = 0$ resulta em

$$O'(\mathbf{A}', \mathbf{d}') = \sum_{\mathbf{x} \in W} (J(\mathbf{x}, t) + \nabla J(\mathbf{x}, t) \cdot (\mathbf{A}'\mathbf{x} + \mathbf{d}') - I(\mathbf{x}, t_0))^2 \omega(\mathbf{x}) \quad (3.5)$$

Na equação acima, $\nabla J = [J_u, J_v]^t$ é o gradiente espacial da região de referência transformada.

Se a solução no instante t é próxima da solução correta, então $J(\mathbf{x}, t) \approx I(\mathbf{x}, t)$, e portanto, $\nabla J \approx \nabla I = [I_u, I_v]^t$, onde I_u e I_v são os gradientes espaciais da região de referência original. Aplicando esta observação na Eq.(3.5), pode-se obter a seguinte equação:

$$O'(\mathbf{A}', \mathbf{d}') = \sum_{\mathbf{x} \in W} (\nabla I(\mathbf{x}, t) \cdot (\mathbf{A}'\mathbf{x} + \mathbf{d}') + (J(\mathbf{x}, t) - I(\mathbf{x}, t_0)))^2 \omega(\mathbf{x}) \quad (3.6)$$

A partir desta equação, o problema pode ser resolvido pela otimização conjunta dos seis valores desconhecidos em $\mathbf{A}'$ e $\mathbf{d}'$. No entanto, algumas regiões podem não possuir textura suficiente para cálculo completo dos seis parâmetros geométricos. Um exemplo, seria se a região de referência fosse definida em um canto com ângulo reto. Uma translação pura do canto poderia ser interpretada por uma translação, fator de escala, ou por uma combinação linear dos dois. Na solução implementada por Hager; Toyama (1998), dividiu-se o cálculo em duas etapas, uma para determinar a

translação e rotação, outra para determinar o fator de escala e *shear*. Segundo eles, em geral, os parâmetros de translação e rotação variam mais rapidamente que os parâmetros de fator de escala e *shear*. Portanto, em situações ambíguas os parâmetros de translação e rotação são preferidos para representar a mudança na imagem.

Para implementar a solução, Hager; Toyama (1998) representou a matriz $\mathbf{A}'$ em uma rotação diferencial e em uma matriz triangular conforme mostrado pela seguinte equação:

$$\mathbf{A}' = \begin{bmatrix} 0 & \alpha \\ -\alpha & 0 \end{bmatrix} + \begin{bmatrix} s_u & \gamma \\ 0 & s_v \end{bmatrix} \quad (3.7)$$

Então resolve-se a equação Eq.(3.6) para dois grupos de parâmetros, $(\mathbf{d}, \alpha)$ e (s_u, s_v, γ) , de forma seqüencial. O parâmetro α representa a rotação incremental, s_u e s_v representam a mudança de fator de escala na horizontal e vertical da imagem, e γ representa o *shear*. Ao se resolver de forma seqüencial, estabelece-se uma preferência para a interpretação das mudanças na imagem. Embora menos exata que a solução simultânea, segundo Hager; Toyama (1998), a pequena quantidade de variação da imagem entre duas imagens consecutivas faz com que a solução através deste método seja suficientemente precisa para a maioria das aplicações.

Para se determinar a translação e rotação, define-se os seguintes parâmetros:

$$\begin{aligned} g_u(\mathbf{x}) &= I_u(\mathbf{x}, t_0) \sqrt{\omega(\mathbf{x})} \\ g_v(\mathbf{x}) &= I_v(\mathbf{x}, t_0) \sqrt{\omega(\mathbf{x})} \\ g_r(\mathbf{x}) &= (v_p I_u(\mathbf{x}, t_0) - u_p I_v(\mathbf{x}, t_0)) \sqrt{\omega(\mathbf{x})} \\ h_0(\mathbf{x}) &= (J(\mathbf{x}, t) - I(\mathbf{x}, t_0)) \sqrt{\omega(\mathbf{x})} \end{aligned} \quad (3.8)$$

Então o sistema linear usado para calculo da translação e rotação é dado por:

$$\sum_{\mathbf{x} \in W} \begin{bmatrix} g_u g_u & g_u g_v & g_u g_r \\ g_v g_u & g_v g_v & g_v g_r \\ g_r g_u & g_r g_v & g_r g_r \end{bmatrix} \begin{bmatrix} \mathbf{d} \\ \alpha \end{bmatrix} = \sum_{\mathbf{x} \in W} \begin{bmatrix} h_0 g_u \\ h_0 g_v \\ h_0 g_r \end{bmatrix} \quad (3.9)$$

Como as derivadas espaciais são calculadas a partir da imagem de referência, g_u , g_v e g_r são constantes ao longo do tempo, então estes valores e a inversa da matriz do lado esquerdo da Eq.(3.9) podem ser calculadas apenas uma vez ao se iniciar o rastreamento em tempo real do alvo.

Uma vez calculado $\mathbf{d} = [du_p, dv_p]^T$ e α , um valor de resíduo pode ser calculado por:

$$h_1(\mathbf{x}) \stackrel{\Delta}{=} h_0(\mathbf{x}) - g_u(\mathbf{x})du_p - g_v(\mathbf{x})dv_p - g_r(\mathbf{x})\alpha \quad (3.10)$$

Este valor de resíduo indica quão bem a região de referência pode ser representada pela transformação através dos parâmetros de translação e rotação. Se este valor for zero, quer dizer que o alvo sofreu apenas translação e rotação. Caso contrário, o valor de resíduo pode ser atribuído ao segundo grupo de parâmetros, ou a erros de linearização ou ruído. Para recuperar os parâmetros de fator de escala e *shear*, pode-se definir os seguintes parâmetros:

$$\begin{aligned} g_{su}(\mathbf{x}) &= u_p g_u(\mathbf{x}) \\ g_{sv}(\mathbf{x}) &= v_p g_v(\mathbf{x}) \\ g_\gamma(\mathbf{x}) &= v_p g_u(\mathbf{x}) \end{aligned} \quad (3.11)$$

Então o sistema linear usado para cálculo do fator de escala e *shear* é dado por:

$$\sum_{\mathbf{x} \in W} \begin{bmatrix} g_{su} g_{su} & g_{su} g_{sv} & g_{su} g_\gamma \\ g_{sv} g_{su} & g_{sv} g_{sv} & g_{sv} g_\gamma \\ g_\gamma g_{su} & g_\gamma g_{sv} & g_\gamma g_\gamma \end{bmatrix} \begin{bmatrix} s_u \\ s_v \\ \gamma \end{bmatrix} = \sum_{\mathbf{x} \in W} \begin{bmatrix} h_1 g_{su} \\ h_1 g_{sv} \\ h_1 g_\gamma \end{bmatrix} \quad (3.12)$$

Os parâmetros g_{su} , g_{sv} e g_γ podem ser calculados ao se iniciar o método de rastreamento, bem como a inversa da matriz do lado esquerdo da equação acima. O novo valor de resíduo pode ser calculado como sendo:

$$h_2(\mathbf{x}) = h_1(\mathbf{x}) - g_{su}(\mathbf{x})s_u - g_{sv}(\mathbf{x})s_v - g_\gamma(\mathbf{x})\gamma \quad (3.13)$$

Então, o valor que representa o quão bem os parâmetros de transformação calculados representam a transformação real sofrida pela região de referência, pode ser dado por:

$$res = \frac{\sqrt{\sum_{\mathbf{x} \in W} h_2(\mathbf{x})^2}}{|W|} \quad (3.14)$$

Uma restrição do método apresentado é o brilho e o contraste do alvo pode não permanecer constante. A solução dada por Hager; Toyama (1998) para esse problema foi a de normalizar as imagens para ter o primeiro momento de inércia zero e o segundo momento igual a um. Extensões do método de rastreamento baseado em SSD para modelos de deformação do alvo e modelos de luminosidade mais complexos são descritos por Hager; Belhumeur (1998).

Uma outra restrição do método apresentado é que os gradientes da imagem são somente válidos localmente. Desta forma, deve-se reduzir a resolução da região utilizada para o cálculo de forma que movimentos maiores que uma fração de pixel possam ser admitidos. Na implementação feita por Hager; Toyama (1998), todos os cálculos exceto o do fator de escala são calculados a uma resolução reduzida, e então os valores são escalados conforme o fator de resolução usado.

O método descrito nesta seção se encontra implementado em uma biblioteca denominada Xvision² (Hager; Toyama, 1998). Além deste método, outros métodos como, por exemplo, rastreamento de bordas e de aglomerados de pixels também se encontram implementados. A biblioteca possui código aberto e está disponível para *download* para uso não comercial.

² Disponível para download através do endereço <http://www.cs.jhu.edu/CIRL/XVision/>

3.3 Implementação do método de rastreamento visual para o sistema de visão omnidirecional desenvolvido

O modelo de transformada afim usado no algoritmo de rastreamento visual descrito na seção anterior se aplica para imagens de projeção ortográfica, ou imagens de projeção perspectiva quando os efeitos de distorção perspectiva podem ser desprezados. Dessa forma, a formulação apresentada não pode ser utilizada diretamente na imagem omnidirecional adquirida pelo sistema de visão desenvolvido neste trabalho. Por esta razão, o algoritmo de rastreamento visual deve ser aplicado em uma imagem perspectiva que é gerada em tempo real a partir da imagem adquirida pelo sistema de visão. Esta imagem perspectiva é definida pelos parâmetros (f_p, θ, φ) , conforme descrito na seção 2.3.3.

Figura 3.2 – Câmera de projeção perspectiva virtual localizada no foco da hipérbole

Conforme o alvo se movimenta, a translação do alvo na imagem calculada pelo algoritmo de rastreamento pode ser usada para mudar a direção do plano de perspectiva. Isto é feito para prevenir que o alvo saia para fora do campo de visão da

imagem perspectiva gerada, o que resultaria em uma perda do alvo pelo algoritmo de rastreamento. A nova direção do plano de projeção perspectiva (θ_n, φ_n) pode ser calculada por $\varphi_n = \varphi + \arctan(dv_p / f_p)$ e $\theta_n = \theta + \arctan(du_p / f_p)$ onde dv_p e du_p são respectivamente a translação vertical e horizontal do centro do alvo na imagem perspectiva. Atualizando-se continuamente a direção do plano de projeção perspectiva faz com que o alvo esteja sempre centrado na imagem. O efeito é o mesmo que se tivéssemos uma câmera perspectiva virtual montada em um mecanismo *pan/tilt* com seu foco sempre localizado no foco da hipérbole, conforme mostrado na Fig.(3.2).

Conforme mencionado anteriormente, um dos métodos de rastreamento visual utilizados no trabalho de Chang; Hebert (1998) é implementado da forma descrita acima. Eles também utilizam o método de rastreamento para controle servo visual de um robô móvel. No entanto a estratégia de controle do robô implementada por eles difere da escolhida neste trabalho, conforme será mostrado no capítulo seguinte. Chang; Hebert (1998) utilizam um filtro de Kalman para estimar o movimento do alvo na imagem e realizam o controle baseado nesta estimativa. A estimativa do movimento do robô também é incorporada no filtro de Kalman.

Capítulo 4

SERVO VISUAL

Tradicionalmente, quando sistemas de visão eram utilizados para controle de robôs, primeiro o ambiente era visto e analisado, e depois o robô se movia para realizar a tarefa. O controle era portanto feito em uma malha aberta, e conseqüentemente, a precisão das operações dependia diretamente da precisão do sistema de visão e da precisão do robô. No entanto, muitas vezes se espera que o robô opere em ambientes dinâmicos onde uma abordagem em malha aberta não é adequada devido a rápidas mudanças no ambiente. Por isso, quando se pretende aumentar a precisão do movimento do robô, ou então reagir a mudanças freqüentes no ambiente de trabalho, se faz necessário um controle em malha fechada com uma realimentação em tempo real das informações visuais. Este tipo de controle pode ser chamado de controle servo visual.

Nos últimos anos têm se visto um aumento na quantidade de trabalhos publicadas na área de controle servo visual. Isto se deve ao fato dos computadores pessoais terem cruzado o limite que permitia a análise de cenas a uma taxa de amostragem suficientemente elevada para que o controle servo visual pudesse ser implementado. Aplicações de controle servo visual abrangem as áreas de manufatura (agarrar objetos em esteiras e montagem de peças), condução de veículos, pouso de veículos aéreos, e outras.

Hutchinson et al (1996) apresentam uma revisão da literatura na área de servo visual, bem como algumas definições, classificações, história, aplicações apresentadas, e uma bibliografia extensa.

4.1 Arquiteturas de sistema de controle servo visual

Uma classificação para sistemas de controle servo visual foi introduzida por Sanderson; Weiss (1980). Essa classificação se baseia em dois critérios: (1) na

estrutura da malha de controle, se hierárquica ou não, e (2) na forma como o sinal de erro é definido, se em posições 3D do espaço ou em termos de características da imagem. Baseados nestes dois critérios, os sistemas de controle servo visual podem ser divididos em quatro categorias representadas pelas Fig.(4.1), (4.2), (4.3) e (4.4).

Segundo o primeiro critério de classificação, um sistema de controle servo visual pode ser classificado como sendo do tipo *dynamic look-and-move* quando a arquitetura de controle é hierárquica, conforme apresentado nas Fig.(4.1) e (4.2). Nesse caso, o sistema de visão é usado para gerar a uma determinada frequência os sinais de referência que servem como entrada para o controlador interno do robô. Este controlador por sua vez opera com uma malha de controle fechada para estabilizar internamente o robô. Em contra partida, um sistema de controle servo visual pode ser classificado como sendo do tipo servo visual direto (*direct visual servo*) quando a malha fechada de controle interno do robô é eliminada, conforme apresentado nas Fig.(4.3) e (4.4). Neste caso, o controlador servo visual calcula diretamente as entradas que devem ser aplicadas nos atuadores do robô, ou seja, a malha de controle visual é usada para estabilizar o robô.

Normalmente, se utiliza a configuração de controle servo visual que possui a malha de controle hierárquica (*dynamic look-and-move*). Isso ocorre primeiramente devido ao fato das taxas de amostragem conseguidas em um sistema de visão não serem muito altas o que faz do controle direto de um manipulador de robô, por exemplo, um problema com uma dinâmica complexa e não linear. O problema se torna mais simples e de possível implementação quando se utiliza uma malha de controle hierárquica. Assim, o controle interno das posições das juntas do robô é feito utilizando-se uma malha fechada com realimentação a uma frequência alta, enquanto que a malha fechada externa fornece realimentação visual a uma frequência mais baixa. Outro fato que contribui para a utilização da configuração hierárquica de controle, é o de que muitos robôs já possuem uma interface para aceitar velocidade cartesiana ou comandos de posição incremental. Isto simplifica a implementação do sistema de controle servo visual com malha hierárquica.

Baseado no segundo critério de classificação, um sistema de controle servo visual pode ser baseado em posição, ou baseado em imagem. No controle baseado em posição, as características são extraídas da imagem e usadas em conjunto com o modelo geométrico do alvo e com o modelo conhecido da câmera para estimar a posição do alvo em relação a câmera. A realimentação é calculada reduzindo-se os erros calculados no espaço de posição do ambiente, conforme mostrado na Fig.(4.1) e (4.3). No controle servo visual baseado em imagem, os valores de controle são calculados diretamente em função das características do alvo na imagem, conforme apresentado na Fig.(4.2) e (4.4). Este método de controle servo visual pode reduzir o atraso computacional, eliminar erros devido à modelagem dos sensores e eliminar erros de calibração da câmera.

Figura 4.1 – Sistema de controle servo visual *dynamic look-and-move* baseado em posição (Figura extraída de Hutchinson et al., 1996)

Figura 4.2 – Sistema de controle servo visual *dynamic look-and-move* baseado em imagem (Figura extraída de Hutchinson et al., 1996)

Figura 4.3 – Sistema de controle servo visual baseado em posição com estrutura proposta por Weiss (extraído de Hutchinson et al., 1996)

Figura 4.4 – Sistema controle servo visual baseado em imagem com estrutura proposta por Weiss (extraído de Hutchinson et al., 1996)

No controle baseado em posição a tarefa de controle é descrita em termos de posicionamento em coordenadas cartesianas. Dessa forma, este tipo de controle é na maioria das vezes muito dependente da calibração. Além disso, o tempo computacional gasto para obtenção da posição no ambiente a partir da imagem pode ser grande em alguns casos. No controle baseado em imagem, como o erro é descrito em função de parâmetros característicos na imagem, a tarefa de controle se torna menos sensível a calibração da câmera se comparado com o controle baseado em posição. Dependendo da aplicação, a calibração da câmera pode até não ser necessária quando é utilizado um controle servo visual baseado em imagem.

O controle baseado em imagem tende a ser mais apropriados para tarefas onde não há um modelo *a priori* da geometria desta tarefa, já os métodos baseados em posição dependem inerentemente deste modelo.

4.2 Trabalhos na área de controle servo visual

Um dos primeiros trabalhos na área de controle servo visual foi realizado por Shirai; Inoue (1973) apud Hutchinson et al (1996). Sua proposta foi usar realimentação visual para corrigir a posição de um robô a fim de aumentar a exatidão da tarefa de se colocar um prisma quadrado em uma caixa. O tempo de realimentação conseguido na época foi cerca de 10s.

A partir do trabalho de Shirai; Inoue (1973) outros trabalhos foram realizados na área de controle servo visual, e a seguir são mencionados alguns destes trabalhos. Uma relação mais extensa de trabalhos nesta área pode ser encontrado em Hutchinson et al (1996), ou em Corke (1997).

Gilbert et al (1980) descreveram como uma câmera pode ser usada para rastrear foguetes fazendo com que o alvo mantenha-se sempre centrado no plano da imagem através do controle de um sistema de *pan/tilt*. Um hardware desenvolvido especialmente para fazer o processamento de imagem foi utilizado.

Zhang et al (1990) apresentam um controlador servo visual de um robô usado para pegar itens de uma esteira que se movem rapidamente (300 mm/s). Neste trabalho, a câmera é colocada no manipulador do robô e a realimentação visual ocorre a cada 140ms.

Allen et al (1991) usam um sistema de visão estéreo para rastrear um alvo que se move a 250 mm/s. Em um trabalho posterior (Allen et al, 1993) estendem isso para agarrar um trem de brinquedo que se move em trajetória circular.

Corke (1997) realiza o controle de posição de um robô Unimate Puma. Neste trabalho, uma abordagem hierárquica do controle servo visual foi usada. Conseguiu-se um período de amostragem de 14 ou 28ms na malha de realimentação do controle interno do robô, e um período de amostragem de 33 ou 40ms para a malha de realimentação visual. Os tempos dependem do sistema de vídeo utilizado, RS 170 ou CCIR respectivamente.

Chaumette et al (1991) descrevem o controle servo visual de um manipulador utilizando uma câmera montada no efetuador do robô. É feito o posicionamento do efetuador em relação ao alvo fixo especificado. Além disso, também são feitos o rastreamento deste alvo quando ele se move, e a estimação da velocidade do alvo. O controle é realizado fazendo-se a realimentação do erro na forma de características do alvo na imagem. Para a aplicação proposta, o alvo é formado por quatro bolas pretas em contraste com o fundo branco. As bolas pretas são posicionadas como se estivessem nos vértices de um quadrado.

Carvalho et al (2000) descreve um trabalho semelhante ao de Chaumette et al (1991), no entanto, aplicado a dinâmica de um robô móvel não-holonômico. A tarefa de controle implementada foi a de mover o robô para uma posição de onde a imagem visualizada do alvo possui as características desejadas.

Chang; Hebert (1998) realizam o controle servo visual de um robô móvel utilizando um sistema de visão omnidirecional. Na aplicação proposta, o usuário define uma região na imagem e então o robô se aproxima desta região. Um filtro de Kalman é utilizado para estimar o movimento do alvo na imagem e realizam o controle baseado nesta estimativa.

Van der Zwaan; Santos-Victor (2001) realizam o controle de um robô móvel submarino para fazer com que ele estacione sobre uma determinada região indicada pelo usuário. Um controle servo visual baseado em imagem é utilizado

4.3 Controle servo visual baseado em imagem

Das arquiteturas de controle servo visual existentes escolheu-se utilizar a de controle servo visual baseado em imagem com malha de controle hierárquica para implementar a tarefa de controle proposta neste trabalho.

A malha de controle hierárquica é utilizada pois o robô móvel já possui um sistema de controle interno para controlar a velocidade de translação do robô e a direção angular das rodas. Dessa forma, o controle servo visual pode operar com uma frequência menor que a do controle interno do robô.

No controle servo visual baseado em imagem, o sinal de erro é definido diretamente em função de parâmetros característicos na imagem, ao invés de ser definido em termos de coordenadas no ambiente. Assim, não é preciso determinar a posição real no espaço dos elementos envolvidos na tarefa de controle.

A tarefa de controle servo visual implementada neste trabalho consiste em fazer com que o robô siga um alvo móvel em tempo real. Para isso utiliza-se um sistema de visão omnidirecional embarcado no robô. O movimento do alvo ou do robô causa uma mudança na imagem do alvo observada pelo sistema de visão. Então, deve-se definir uma função de erro apropriada de forma que quando o robô mantém a distância do alvo, este erro é zero.

Embora a variável de erro seja definida em termos de parâmetros da imagem, as variáveis de controle do robô são definidas no espaço de coordenadas do ambiente. Por essa razão, é necessário relacionar as mudanças nos parâmetros característicos da imagem com as mudanças na posição do robô. Isso é feito através do Jacobiano da imagem.

Seja $\dot{\mathbf{r}} = [T_x, T_y, T_z, w_x, w_y, w_z]^T$ um vetor que representa a velocidade de um ponto no robô móvel, onde as velocidades de translação são representadas por $\mathbf{T} = [T_x, T_y, T_z]^T$, e as velocidades de rotação são representadas por $\mathbf{\Omega} = [w_x, w_y, w_z]^T$. Seja também $\mathbf{f}$ um vetor que representa os parâmetros característicos da imagem, e $\dot{\mathbf{f}}$ a variação destes parâmetros. O Jacobiano da imagem, $\mathbf{J}_v$, faz a relação entre $\dot{\mathbf{r}}$ e $\dot{\mathbf{f}}$:

$$\dot{\mathbf{f}} = \mathbf{J}_v(\mathbf{r}) \cdot \dot{\mathbf{r}} \quad (4.1)$$

Seja m a dimensão do vetor $\dot{\mathbf{r}}$ e k a dimensão do vetor $\dot{\mathbf{f}}$, o Jacobiano da imagem tem dimensão $k \times m$ e pode ser obtido através da seguinte maneira:

$$\mathbf{J}_v(\mathbf{r}) = \left[\frac{\partial}{\partial \mathbf{r}} \right] = \begin{bmatrix} \frac{\partial f_1(\mathbf{r})}{\partial r_1} & \dots & \frac{\partial f_1(\mathbf{r})}{\partial r_m} \\ \vdots & & \vdots \\ \frac{\partial f_k(\mathbf{r})}{\partial r_1} & \dots & \frac{\partial f_k(\mathbf{r})}{\partial r_m} \end{bmatrix} \quad (4.2)$$

O Jacobiano da imagem usado para a aplicação de controle servo visual deste trabalho foi derivado conforme apresentado por Chaumette et al. (1991), Espiau et al (1992), Hutchinson et al (1996) e Carvalho et al. (2000). Primeiramente determina-se o Jacobiano que relaciona a velocidade de uma câmera de projeção perspectiva montada em um robô móvel com as características visuais de um alvo fixo. Posteriormente determina-se uma matriz de transformação que relaciona a velocidade da câmera perspectiva com os parâmetros de controle do robô. A Figura (4.5) mostra uma câmera de projeção perspectiva virtual localizada no centro de projeção único do sistema de visão omnidirecional do robô. As imagens perspectivas obtidas por esta câmera são resultantes do processo de retificação da imagem omnidirecional. A Figura (4.5) também mostra os sistemas de coordenadas desta câmera (O_c), e o sistema de coordenadas de controle do robô (O_b) no qual o eixo z está sempre alinhado com a direção das três rodas síncronas.

A determinação do Jacobiano da imagem para uma câmera de projeção perspectiva é mostrada a seguir. Como parâmetro característico do alvo na imagem utiliza-se a coordenada de um ponto do alvo na imagem.

Seja ${}^c\mathbf{P} = [x, y, z]^T$, a coordenada de um ponto no alvo expressa em relação ao sistema de coordenadas da câmera. Este ponto é projetado na imagem perspectiva nas coordenadas u_p e v_p segundo a seguinte expressão:

$$\begin{bmatrix} u_p \\ v_p \end{bmatrix} = \frac{f_p}{z} \begin{bmatrix} x \\ y \end{bmatrix} \quad (4.3)$$

Na expressão acima f_p é um valor de distância focal da câmera perspectiva dado em pixel, e u_p e v_p são valores dados em pixels.

Figura 4.5 – Sistema de coordenadas da câmera e do robô

Suponha que o alvo esteja parado e que a câmera de projeção perspectiva se mova em relação ao alvo com velocidade angular ${}^c\boldsymbol{\Omega}_c = [w_{cx}, w_{cy}, w_{cz}]^T$ e velocidade de translação ${}^c\mathbf{T}_c = [T_{cx}, T_{cy}, T_{cz}]^T$. A velocidade do ponto ${}^c\mathbf{P}$ em relação ao sistema de coordenadas da câmera será dado por:

$${}^c\dot{\mathbf{P}} = -{}^c\boldsymbol{\Omega}_c \times {}^c\mathbf{P} - {}^c\mathbf{T}_c \quad (4.4)$$

A derivada das coordenadas do ponto ${}^c\mathbf{P}$ também podem ser expressas por:

$$\begin{aligned} \dot{x} &= -zw_{cy} + yw_{cz} - T_{cx} \\ \dot{y} &= -xw_{cz} + zw_{cx} - T_{cy} \\ \dot{z} &= -yw_{cx} + xw_{cy} - T_{cz} \end{aligned} \quad (4.5)$$

Substituindo a Eq.(4.3) nas expressões acima, obtém-se a derivada do ponto ${}^c\mathbf{P}$ em função dos parâmetros característicos na imagem $\mathbf{f} = [u_p, v_p]^T$:

$$\begin{aligned}\dot{x} &= -zw_{cy} + \frac{v_p z}{f_p} w_{cz} - T_{cx} \\ \dot{y} &= -\frac{u_p z}{f_p} w_{cz} + zw_{cx} - T_{cy} \\ \dot{z} &= -\frac{v_p z}{f_p} w_{cx} + \frac{u_p z}{f_p} w_{cy} - T_{cz}\end{aligned}\tag{4.6}$$

A partir da Eq.(4.3), pode-se determinar a derivada dos parâmetros característicos, $\dot{\mathbf{f}} = [\dot{u}_p, \dot{v}_p]^T$, em função da derivada das coordenadas do ponto ${}^c\mathbf{P}$ que por sua vez podem ser expressas pela Eq.(4.6). A derivada dos parâmetros característicos na imagem é obtida por:

$$\begin{aligned}\dot{u}_p &= f_p \frac{z\dot{x} - x\dot{z}}{z^2} \\ &= \frac{f_p}{z^2} \left(z \left(-zw_{cy} + \frac{v_p z}{f_p} w_{cz} - T_{cx} \right) - \frac{u_p z}{f_p} \left(-\frac{v_p z}{f_p} w_{cx} + \frac{u_p z}{f_p} w_{cy} - T_{cz} \right) \right) \\ &= -\frac{f_p}{z} T_{cx} + \frac{u_p}{z} T_{cz} + \frac{u_p v_p}{f_p} w_{cx} - \frac{f_p^2 + u_p^2}{f_p} w_{cy} + v_p w_{cz}\end{aligned}\tag{4.7}$$

De modo similar:

$$\dot{v}_p = -\frac{f_p}{z} T_{cy} + \frac{v_p}{z} T_{cz} + \frac{f_p^2 + u_p^2}{f_p} w_{cx} - \frac{u_p v_p}{f_p} w_{cy} - u_p w_{cz}\tag{4.8}$$

Então reescrevendo as equações anteriores na forma de matriz, obtém-se:

$$\begin{bmatrix} \dot{u}_p \\ \dot{v}_p \end{bmatrix} = \begin{bmatrix} -\frac{f_p}{z} & 0 & \frac{u_p}{z} & \frac{u_p v_p}{f_p} & -\frac{f_p^2 + u_p^2}{f_p} \\ 0 & -\frac{f_p}{z} & \frac{v_p}{z} & \frac{f_p^2 + v_p^2}{f_p} & -\frac{u_p v_p}{f_p} \end{bmatrix} \begin{bmatrix} v_p \\ -u_p \end{bmatrix} \begin{bmatrix} T_{cx} \\ T_{cy} \\ T_{cz} \\ w_{cx} \\ w_{cy} \\ w_{cz} \end{bmatrix} \quad (4.9)$$

Para esta aplicação, utilizou-se um vetor de características do alvo $\mathbf{f}$ composto pelas coordenadas na imagem perspectiva dos quatro cantos que definem a região do alvo sendo rastreada pelo algoritmo de rastreamento visual. Assim a variação dos parâmetros característicos é dada por $\dot{\mathbf{f}} = [\dot{u}_{p1}, \dot{u}_{p2}, \dot{u}_{p3}, \dot{u}_{p4}, \dot{v}_{p1}, \dot{v}_{p2}, \dot{v}_{p3}, \dot{v}_{p4}]^T$, onde (u_{p1}, v_{p1}) é a coordenada em pixels do primeiro ponto e assim por diante. A expressão dada pela Eq.(4.9) que representa $\dot{\mathbf{f}} = \mathbf{J}_v(\mathbf{r}) \cdot \dot{\mathbf{r}}_c$, pode ser reescrita como sendo:

$$\begin{bmatrix} \dot{u}_{p1} \\ \dot{u}_{p2} \\ \dot{u}_{p3} \\ \dot{u}_{p4} \\ \dot{v}_{p1} \\ \dot{v}_{p2} \\ \dot{v}_{p3} \\ \dot{v}_{p4} \end{bmatrix} = \begin{bmatrix} -f_p/z_1 & 0 & u_{p1}/z_1 & (u_{p1}v_{p1})/f_p & -(f_p^2 + u_{p1}^2)/f_p & v_{p1} \\ -f_p/z_2 & 0 & u_{p2}/z_2 & (u_{p2}v_{p2})/f_p & -(f_p^2 + u_{p2}^2)/f_p & v_{p2} \\ -f_p/z_3 & 0 & u_{p3}/z_3 & (u_{p3}v_{p3})/f_p & -(f_p^2 + u_{p3}^2)/f_p & v_{p3} \\ -f_p/z_4 & 0 & u_{p4}/z_4 & (u_{p4}v_{p4})/f_p & -(f_p^2 + u_{p4}^2)/f_p & v_{p4} \\ 0 & -f_p/z_1 & v_{p1}/z_1 & (f_p^2 + v_{p1}^2)/f_p & -(u_{p1}v_{p1})/f_p & -u_{p1} \\ 0 & -f_p/z_2 & v_{p2}/z_2 & (f_p^2 + v_{p2}^2)/f_p & -(u_{p2}v_{p2})/f_p & -u_{p2} \\ 0 & -f_p/z_3 & v_{p3}/z_3 & (f_p^2 + v_{p3}^2)/f_p & -(u_{p3}v_{p3})/f_p & -u_{p3} \\ 0 & -f_p/z_4 & v_{p4}/z_4 & (f_p^2 + v_{p4}^2)/f_p & -(u_{p4}v_{p4})/f_p & -u_{p4} \end{bmatrix} \begin{bmatrix} T_{cx} \\ T_{cy} \\ T_{cz} \\ w_{cx} \\ w_{cy} \\ w_{cz} \end{bmatrix} \quad (4.10)$$

O Jacobiano da imagem conforme a equação anterior é função da distância entre a câmera e os pontos observados (coordenadas z_1, z_2, z_3 , e z_4). Segundo Chaumette et al. (1991), o valor exato desta distância não precisa ser conhecido e pode-se usar valores aproximados. O controle servo visual ainda convergirá mesmo que for utilizado o valor unitário quando a distância é totalmente desconhecida.

Se o vetor de controle fosse a velocidade da câmera, esta velocidade poderia ser determinada pela inversa do Jacobiano da imagem, $\dot{\mathbf{r}}_c = \mathbf{J}_v^{-1}(\mathbf{r}) \cdot \dot{\mathbf{f}}$. Se fosse feito um controle proporcional, o erro poderia ser calculado por $\mathbf{E}(\mathbf{f}) = \mathbf{f}_{ref} - \mathbf{f}$ onde $\mathbf{f}_{ref}$ é o vetor de parâmetros característicos de referência, e a velocidade da câmera seria

controlada pela lei de controle $\dot{\mathbf{r}}_c = \mathbf{K} \cdot \mathbf{J}_v^{-1}(\mathbf{r}) \cdot \mathbf{E}(\mathbf{f})$ onde $\mathbf{K}$ é um vetor de ganho proporcional adequado.

Como a inversa do Jacobiano da imagem é função dos valores medidos dos parâmetros característicos da imagem, ela deveria ser computada a cada iteração do controle. Para evitar este calculo, Chaumette et al. (1991) utiliza um Jacobiano de imagem constante calculado para a posição de referência do alvo. Este Jacobiano da imagem constante deve ser utilizado quando a posição atual do alvo estiver próxima da posição de referência. Como este é o caso desta aplicação, pois o robô deve seguir o alvo mantendo a distância de referência, o Jacobiano da imagem pode ser dado por:

$$\mathbf{J}_v(\mathbf{r}_{ref}) = \begin{bmatrix} -f_p/z_t & 0 & u_{p1r}/z_d & (u_{p1r}v_{p1r})/f_p & -(f_p^2 + u_{p1r}^2)/f_p & v_{p1r} \\ -f_p/z_t & 0 & u_{p2r}/z_d & (u_{p2r}v_{p2r})/f_p & -(f_p^2 + u_{p2r}^2)/f_p & v_{p2r} \\ -f_p/z_t & 0 & u_{p3r}/z_d & (u_{p3r}v_{p3r})/f_p & -(f_p^2 + u_{p3r}^2)/f_p & v_{p3r} \\ -f_p/z_t & 0 & u_{p4r}/z_d & (u_{p4r}v_{p4r})/f_p & -(f_p^2 + u_{p4r}^2)/f_p & v_{p4r} \\ 0 & -f_p/z_d & v_{p1r}/z_d & (f_p^2 + v_{p1r}^2)/f_p & -(u_{p1r}v_{p1r})/f_p & -u_{p1r} \\ 0 & -f_p/z_d & v_{p2r}/z_d & (f_p^2 + v_{p2r}^2)/f_p & -(u_{p2r}v_{p2r})/f_p & -u_{p2r} \\ 0 & -f_p/z_d & v_{p3r}/z_d & (f_p^2 + v_{p3r}^2)/f_p & -(u_{p3r}v_{p3r})/f_p & -u_{p3r} \\ 0 & -f_p/z_d & v_{p4r}/z_d & (f_p^2 + v_{p4r}^2)/f_p & -(u_{p4r}v_{p4r})/f_p & -u_{p4r} \end{bmatrix} \quad (4.11)$$

Na matriz acima, z_d é a distância desejada do alvo à câmera. As características de referência do alvo na imagem, $\mathbf{f}_{ref} = [u_{p1r}, u_{p2r}, u_{p3r}, u_{p4r}, v_{p1r}, v_{p2r}, v_{p3r}, v_{p4r}]^T$, são as coordenadas dos quatro cantos da região de referência usada para iniciar o algoritmo de rastreamento visual, conforme mostrado na Fig.(4.6). E as características visuais atuais do alvo são obtidas como resultado de cada iteração do algoritmo de rastreamento visual, a partir dos parâmetros de deformação da região de referência.

Conforme dito anteriormente, as variáveis de controle do robô são diferentes das variáveis de controle da câmera perspectiva. Como a câmera é montada no robô, é possível relacionar o movimento da câmera, $\dot{\mathbf{r}}_c$, com os parâmetros de controle do

robô usado neste trabalho, $\mathbf{U}_b = [T_{bz}, w_{by}]^T$, onde T_{bz} é a velocidade de translação, e w_{by} é a velocidade de esterçamento das rodas do robô.

Figura 4.6 – Região de referência rastreada na imagem, e a coordenada dos quatro pontos utilizados no controle servo visual

Dos componentes de velocidade da câmera, sabe-se que o componente T_{cy} é zero pois a câmera e o robô se movimentam apenas no plano horizontal. Então se utilizou o método usado por Carvalho et al. (2000) para transformar as variáveis de controle da câmera nas variáveis de controle do robô. Os componentes de rotação da câmera, ${}^c\boldsymbol{\Omega}_c = [w_{cx}, w_{cy}, w_{cz}]^T$, são separados dos componentes de translação. Dessa forma, os componentes de rotação podem ser calculados diretamente pelo Jacobiano da imagem. Então, os componentes de translação T_{cx} e T_{cz} são relacionados aos parâmetros de controle $\mathbf{U}_b = [T_{bz}, w_{by}]^T$ utilizando a seguinte matriz $\mathbf{J}_b$ (Carvalho et al., 2000):

$$\begin{bmatrix} T_{cx} \\ T_{cz} \end{bmatrix} = \mathbf{J}_b \begin{bmatrix} T_{bz} \\ w_{by} \end{bmatrix}, \text{ para } \mathbf{J}_b = \begin{bmatrix} -\sin\theta_r & -z_t \\ -\cos\theta_r & x_t \end{bmatrix} \quad (4.12)$$

Na equação acima, θ_r é a diferença entre a direção de movimento θ_b e a direção azimutal da câmera perspectiva θ_c , assim $\theta_r = \theta_b - \theta_c$. Os valores z_t e x_t são as

coordenadas expressas no sistema de coordenadas da câmera da distância entre a câmera e a posição desejada para o robô. Para evitar reconstrução 3D durante o controle servo visual, a estratégia usada por Carvalho et al. (2000) foi considerar $z_t = z_d$ para todo instante de tempo, e fazer uma aproximação simples a cada instante de tempo do valor de x_t baseado diretamente na informação obtida pelo sensor:

$$\hat{x}_t = \frac{z_d}{4f_p} \sum_{i=1}^4 (u_{pi} - u_{pir}) \quad (4.13)$$

A inversa da matriz $\mathbf{J}_b$ é dada por:

$$\mathbf{J}_b^{-1} = \begin{bmatrix} J_{b1} & J_{b2} \\ J_{b3} & J_{b4} \end{bmatrix} = \frac{1}{x_t \cdot \sin\theta_r - z_t \cdot \cos\theta_r} \begin{bmatrix} x_t & -z_t \\ -\cos\theta_r & -\sin\theta_r \end{bmatrix} \quad (4.14)$$

Finalmente, se o vetor de controle for dado por $\mathbf{U} = [T_{bz}, w_{by}, w_{cx}, w_{cy}, w_{cz}]^T$, uma lei de controle proporcional seria escrita como sendo:

$$\mathbf{U} = \mathbf{K} \cdot \mathbf{N} \cdot \mathbf{C} \cdot \mathbf{E}(\mathbf{f}) \quad (4.15)$$

Na equação anterior, a matriz $\mathbf{C}$ representa a pseudo inversa do Jacobiano da imagem calculado na posição de referência, $\mathbf{J}_v^{-1}(\mathbf{r}_{ref})$, o vetor $\mathbf{E}(\mathbf{f}) = \mathbf{f}_{ref} - \mathbf{f}$ representa o erro medido em função dos parâmetros característicos do alvo na imagem, e a matriz $\mathbf{N}$ é definida como sendo:

$$\mathbf{N} = \begin{bmatrix} J_{b1} & 0 & J_{b2} & 0 & 0 & 0 \\ J_{b3} & 0 & J_{b4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.16)$$

Quando se segue um alvo móvel utilizando uma lei de controle proporcional como a descrita acima, enquanto o alvo se move com velocidade constante, o erro permanece constante até que o alvo pare de se mover. Este erro pode ser eliminado se um

controle proporcional derivativo for usado (Chaumette et al. 1999). Neste caso, a lei de controle seria dada por $U = N \cdot C \cdot (K_1 \cdot E(f) + K_2 \cdot \dot{E}(f))$.

No controle servo visual implementado, as componentes de rotação da câmera, com exceção da componente w_{cz} , são controladas diretamente através do rastreamento visual para que o alvo não saia do campo de visão da imagem perspectiva e para que a posição do centro do alvo na imagem permaneça sempre a mesma. Como a posição do centro do alvo na imagem permanece sempre constante, a distância $\hat{x}_l$ dada pela Eq.(4.13) pode ser considerada como sendo zero. A componente de rotação da câmera w_{cz} que compensa a rotação do alvo na imagem perspectiva, não foi controlada. Este controle não é necessário que seja feito pois não têm nenhum efeito na posição do robô em relação ao alvo. Escolheu-se também implementar o controle da direção das rodas θ_b de modo que o valor de θ_r permanecesse sempre zero, ou seja, fazer com que o sistema de coordenadas da câmera permanecesse sempre alinhado com o sistema de coordenadas do robô. Conseqüentemente, a velocidade de translação é controlada pela informação de escala do alvo na imagem que é fornecida pelo algoritmo de rastreamento visual. Dessa forma, simplificou-se o controle servo visual implementado neste trabalho. No entanto, utilizando a formulação descrita neste trabalho, poderia ser empregada uma estratégia de controle diferente para a rotação da câmera e direção das rodas deste robô, ou ainda uma outra tarefa de controle servo visual poderia ser definida, como por exemplo, uma tarefa de *docking*.

Capítulo 5

RESULTADOS EXPERIMENTAIS

Neste capítulo são apresentados e discutidos os resultados obtidos com o sistema de visão omnidirecional implementado neste trabalho. Primeiramente, na seção 5.1, é descrito o equipamento utilizado para a realização dos experimentos. A seguir, na seção 5.2, é apresentada uma descrição do sistema de visão omnidirecional implementado. Também são apresentadas imagens obtidas com o sistema de visão e imagens resultantes da aplicação dos métodos de retificação implementados. Na seção 5.3 são apresentados alguns resultados da utilização do sistema de visão omnidirecional no rastreamento visual de um alvo móvel. E, por fim, na seção 5.4 são apresentados resultados da aplicação do sistema de visão omnidirecional para controle servo visual de um robô móvel.

5.1 Equipamento utilizado

O robô móvel a ser controlado, e no qual foi embarcado o sistema de visão omnidirecional, foi desenvolvido no Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos da Escola Politécnica da USP. Este robô possui um motor de passo e um motor de corrente contínua responsáveis, respectivamente, pelo esterçamento e translação das três rodas síncronas do robô. O controle do motor de passo é feito em malha aberta, enquanto o controle do motor de corrente contínua é feito em malha fechada com realimentação da velocidade de translação a cada 7,5ms. A velocidade máxima de translação que pode ser atingida pelo robô é de 0,4m/s, e a variação da direção das rodas é realizada a uma velocidade de 0,52 rad/s. Um hardware eletrônico embarcado no robô é responsável pelo controle dos motores. Este hardware possui uma interface RS232c pela qual é possível enviar comandos para estabelecer a velocidade de translação do robô e a variação do ângulo de direção das suas rodas. Informações mais detalhadas sobre as características mecânicas, eletrônicas, e do controle de baixo nível do robô móvel podem ser encontradas

respectivamente nos trabalhos de Izzo; Miyahara; Okamoto Jr (1994), Tachibana; Fugiwara; Okamoto Jr. (1995), e Grassi Jr; Okamoto Jr (1998).

Um computador padrão PC104 com processador Pentium 266Mhz operando com sistema operacional Linux também vai embarcado no robô móvel. Este computador possui uma interface de rede conectada a um rádio de ethernet e um canal serial conectado por RS232c ao hardware eletrônico de controle de baixo nível do robô.

As imagens obtidas pelo sistema de visão embarcado no robô são transmitidas para um computador externo ao robô através de um sistema de transmissão de vídeo sem fio. Foi utilizado o sistema de transmissão chamado *Eagle 1000* do fabricante *Trango Systems* que opera a 2.4GHz e é composto por um rádio transmissor e um rádio receptor. O rádio transmissor vai embarcado no robô e conectado à câmera do sistema de visão. O rádio receptor vai conectado à placa de aquisição de imagens do computador externo.

O computador externo ao robô responsável pela aquisição e processamento da imagem possui um processador Pentium 4 de 1,4 GHz, 256 Mbytes de memória, uma placa de aquisição de imagens *Matrox Meteor*, e opera com sistema operacional Linux. Neste computador são executados os programas de retificação da imagem, rastreamento visual e controle servo visual do robô.

Quando é feito o controle servo visual, as variáveis de controle calculadas a cada 33ms são enviadas do computador externo ao robô através da ethernet. A comunicação por ethernet sem fio é feita entre um *Wireless Access Point* conectado a rede ethernet, e um rádio ethernet (*Ethernet Client*) embarcado no robô e conectado a interface de rede do PC104. Para estabelecer esse link de ethernet sem fio foram utilizados *Wireless Access Point* e *Ethernet Client* do fabricante *Aironet* que operam a 2.4GHz e a 11Mbps, padrão 802.11. Uma vez recebidas as referências de velocidade de translação e direção das rodas do robô, o PC104 envia estes valores para o hardware de controle interno do robô através da interface RS-232c.

A Figura (5.1) mostra o ambiente experimental completo utilizado neste trabalho.

Figura 5.1 – Esquema do ambiente experimental utilizado

5.2 Sistema de visão omnidirecional

O sistema de visão omnidirecional implementado é composto de uma câmera com lente perspectiva, um espelho, e peças para fixação deste espelho e da câmera na posição correta, conforme mostrado na Fig.(5.2).

5.2.1 Câmera

A câmera utilizada no sistema de visão é uma câmera industrial modelo CV-M50 monocromática com CCD de 1/2". De acordo com os dados fornecidos pelo fabricante, sua área de *scanning* é de 6,45mm (H) x 4,84mm (V), e para o padrão de vídeo EIA (30 quadros por segundo), o número efetivo de pixels é de 768 (H) x 494 (V). Apesar deste número efetivo de pixels, trabalhou-se com a placa de aquisição adquirindo imagens de 640 (H) x 480 (V) pixels a uma taxa de 30 quadros por segundo.

Figura 5.2 – Protótipo do sistema de visão omnidirecional construído.

5.2.2 Lente

Uma lente convencional do tipo perspectiva foi montada na câmera. A lente utilizada possui uma distância focal variável de 6 à 12mm com ajuste feito manualmente. No entanto, trabalhou-se com a distância focal fixa em 12mm. No sistema de visão omnidirecional implementado, para que imagens focalizadas pudessem ser obtidas para a distância utilizada entre a câmera e o espelho, colocou-se um anel espaçador de 1.5mm entre a câmera e a lente.

5.2.3 Espelho hiberbólico

Um espelho hiperbólico foi projetado conforme a metodologia apresentada na seção 2.2.1. De acordo com as especificações da câmera e da lente utilizadas no sistema de visão, pode-se considerar os seguintes parâmetros aproximados: $f = 12mm$, e $t_{pixel} = 0,01mm$. A imagem adquirida pela câmera possui tamanho de 640 x 480 pixels. Assim, deseja-se que o raio do espelho na imagem assuma valor de $r_{pixel} = 240$ pixels. Por fim, para tornar o sistema compacto, escolheu-se como parâmetro de projeto um espelho com raio $r_{topo} = 20mm$. Substituindo estes valores na Eq.(2.4), obtém-se a distância entre o foco da câmera e o topo do espelho $h = 100mm$.

Utilizando-se uma análise, apresentada em Svoboda et al (1997), da variação do ângulo de visão e da forma do espelho em função da relação de a/b , optou-se por construir um espelho com a relação $a = 2b$. Portanto, os valores calculados através das Eq.(2.3) e (2.2) foram: $b = 19,646$, $a = 39,292$, e $\alpha = 121,3^\circ$. Substituindo estes valores de a e b na Eq.(2.1), obtém-se a forma final do espelho no sistema de coordenadas centrado no foco do espelho F' , descrita pela seguinte função:

$$y = 39,3 \sqrt{\left(1 + \frac{x^2}{186}\right)} - 43,9 \quad (5.1)$$

Uma vez determinada a equação do espelho, é possível sua fabricação a partir de um cilindro maciço de alumínio. Dessa forma, pode-se torneare o alumínio em CNC e depois polir a superfície da peça até que ela se torne reflexiva como um espelho, ou pode-se ao invés de polir, revestir a superfície com uma camada de níquel ou cromo. Esses métodos de fabricação são mencionados nos trabalhos de Chahl; Sirinivasan (1997), Hicks; Bajcsy (1999), e Ishiguro (1998).

O Departamento de Engenharia Mecânica da USP de São Carlos possui um torno CNC de ultra-precisão. Este torno consegue dar às peças usinadas o mesmo acabamento de uma peça polida, resultando em uma superfície reflexiva como um espelho. Assim, utilizando-se o torno de ultra-precisão, evita-se o polimento manual da superfície da peça, ou o depósito de níquel ou cromo. Normalmente, o polimento manual e o depósito de materiais na superfície da peça imprimem erros na sua forma. Quando as dimensões da peça são reduzidas, esses erros podem chegar a ser significativos.

Embora as dimensões do espelho hiperbólico permitissem que este fosse fabricado utilizando-se qualquer um dos processos de fabricação mencionados, optou-se pela fabricação do espelho utilizando o torno de ultra-precisão do Departamento de Engenharia Mecânica da USP de São Carlos.

Portanto, o espelho hiperbólico utilizado neste projeto foi fabricado em alumínio em duas etapas: (1) torneamento de desbaste em uma máquina convencional a fim de dar

a forma básica ao espelho e (2) torneamento e acabamento da superfície final em um torno CNC de ultra-precisão. Após o acabamento da peça, a superfície do alumínio já se encontra espelhada podendo ser usada no sistema de visão. Para evitar que a superfície da peça risque, pode-se revestir a superfície com uma camada fina de cristal de rocha. Embora esse revestimento não tenha sido feito para o espelho construído, o cilindro de acrílico usado no sistema de visão omnidirecional ajuda a manter a superfície espelhada protegida.

Para que a superfície do espelho pudesse ser usinada no torno de ultra-precisão, foi necessário calcular adequadamente os pontos programados no torno. A superfície usinada é formada pela interpolação linear dos pontos programados. A diferença entre a curva teórica e a superfície usinada é chamada de erro de forma, conforme mostrado na Fig.(5.3). Dependendo da quantidade e posição dos pontos, a superfície pode apresentar um erro de forma que prejudique a formação de uma imagem contínua. O cálculo adequado da quantidade e posição destes pontos garante um bom resultado final da superfície, mantendo o erro de forma dentro de um limite desejado.

Figura 5.3 – Erro de forma da peça

Utilizou-se o método proposto por Palma e Porto (1995) para calcular as coordenadas dos pontos que devem ser programados no torno garantindo um valor de erro de forma desejado. Para o cálculo das coordenadas dos pontos, também é feita a compensação do raio da ponta da ferramenta. Esta compensação é necessária pois, ao

longo da usinagem, o contato da ferramenta com a peça não ocorre sempre no ápice da ponta da ferramenta que é usado como referência para zerar o sistema de coordenadas do torno CNC, conforme mostrado na Fig.(5.4).

Figura 5.4 – Trajetória da ponta da ferramenta na usinagem do espelho

A Figura (5.5.a) mostra o espelho hiperbólico construído em alumínio com diâmetro de 40mm.

5.2.4 Espelho esférico

Além do espelho hiperbólico, um espelho esférico foi construído. Assim é possível fazer uma comparação entre as imagens obtidas com o espelho esférico e as imagens obtidas com o espelho hiperbólico.

Normalmente, quando se usa um sistema de visão omnidirecional catadióptrico em um robô móvel de modo que a câmera fique apontada para cima, o reflexo da câmera e do robô no espelho aparecerão na região central da imagem. Quanto maior for o espelho e quanto mais afastado ele estiver da câmera, menor será a região ocupada pela imagem do robô. No entanto, muitas vezes deseja-se que um sistema de visão omnidirecional para robôs móveis seja compacto quando comparado com as dimensões do robô.

Para especificação do tamanho do espelho esférico, procurou-se levar em consideração o tamanho máximo que o sistema de visão poderia ter quando construído. Dessa forma, escolheu-se um diâmetro de 70mm para o espelho esférico. A distância aproximada entre o foco da câmera e o topo do espelho pode ser calculado pela Eq. (2.4). Utilizando-se os parâmetros $f = 12mm$, $t_{pixel} = 0,01mm$, $r_{pixel} = 240$, e $r_{topo} = 70mm$, determina-se que $h = 145mm$. Quando as dimensões da câmera e das peças de fixação são incluídas neste valor, a altura total do sistema de visão resultante é de 410mm. Para efeito de comparação, o robô utilizado possui uma forma cilíndrica com aproximadamente 310mm de altura, e 300mm de diâmetro.

O espelho esférico foi obtido pela deposição de uma camada de metal na parede interior (parte côncava) de uma cúpula semi esférica de vidro. O espelho esférico construído, devido ao processo de fabricação da cúpula de vidro, apresenta uma irregularidade de forma no ápice da cúpula. Apesar de não se poder garantir uma boa precisão de forma na fabricação da cúpula de vidro, as imagens obtidas com este espelho apresentaram um resultado adequado para os propósitos deste trabalho. Além do mais, a irregularidade na cúpula de vidro fica na região do espelho onde são refletidos o robô e a câmera, portanto esse defeito não provoca prejuízos para aplicações em robótica móvel. Utilizando-se essa mesma cúpula de vidro, um resultado melhor poderia ser obtido se o depósito da camada metálica fosse feito na parede exterior (parte convexa) da cúpula de vidro. Assim, a luz refletiria apenas no espelho ao invés de refletir também no vidro.

Figura 5.5 – Espelhos convexos construídos. (a) Espelho hiperbólico de alumínio com 40mm de diâmetro usinado em torno de ultra-precisão. (b) Espelho esférico de 70mm de diâmetro construído em vidro.

A Figura (5.5.b) mostra o espelho esférico construído em vidro com 70mm de diâmetro.

5.2.5 Peças de fixação

Para sustentar o espelho sobre a câmera na altura determinada, utilizou-se um cilindro transparente de acrílico. A sustentação do espelho foi feita dessa forma para que a peça de sustentação não aparecesse na imagem adquirida pelo sistema de visão. Se não fosse dessa forma, ou seja, se parte da imagem fosse ocupada pelo reflexo da peça de sustentação, haveria uma redução da área útil na imagem. Além disso, neste caso, a peça de sustentação poderia obstruir o alvo em um determinado momento, resultando em uma possível falha em algoritmos de rastreamento visual de alvos móveis.

Observou-se qualitativamente que dentro do laboratório, onde a iluminação é feita apenas por luz artificial, não houve interferência significativa do acrílico na qualidade da imagem adquirida pelo sistema de visão. Ou seja, a reflexão da luz na parede do cilindro de acrílico não causa alteração significativa na imagem.

Também foi construída uma peça para fixar o sistema de visão omnidirecional no robô móvel, peças para fixação da câmera na posição vertical, e uma base para fixação do espelho no cilindro de acrílico. A câmera e o espelho devem ser montados de forma que o eixo óptico da câmera esteja alinhado com o eixo de simetria do espelho.

O tamanho resultante do protótipo do sistema de visão omnidirecional construído que utiliza o espelho hiperbólico foi de 285mm de comprimento. O tamanho resultante do protótipo do sistema de visão que utiliza o espelho esférico foi de 410mm de comprimento. Se uma câmera de dimensões menores fosse usada, o tamanho total de ambos os protótipos poderia ser diminuído.

A Figura (5.6) mostra o robô móvel com o sistema de visão omnidirecional montado.

Figura 5.6 – Robô móvel com sistema de visão omnidirecional

5.2.6 Métodos de retificação de imagem

O protótipo do sistema de visão foi montado no robô móvel, e testes foram feitos utilizando-se os espelhos hiperbólico e esférico. Os métodos de retificação de imagem também foram implementados e testados em imagens adquiridas pelo sistema.

A Figura (5.7.a) mostra uma imagem capturada pelo sistema de visão omnidirecional com o espelho esférico. A Figura (5.8) mostra o resultado obtido através do processo de retificação direta de coordenadas descrito no capítulo 2 e aplicado na imagem mostrada na Fig.(5.7.a).

A Figura (5.7.b) mostra uma imagem capturada pelo sistema de visão omnidirecional com o espelho hiperbólico. As Figuras (5.9) e (5.10) mostram, respectivamente, o resultado obtido através da retificação pelo método que usa a equação do espelho, e o resultado obtido pelo método de retificação direta de coordenadas, ambos descritos no capítulo 2.

As duas imagens da Fig.(5.7) foram adquiridas com os objetos do ambiente na mesma posição. Assim pode-se fazer a comparação entre as duas imagens.

Figura 5.7 – Imagens adquiridas pelo sistema de visão omnidirecional. (a) Imagem adquirida com o espelho esférico. (b) Imagem adquirida com o espelho hiperbólico

Figura 5.8 – Imagem adquirida com o espelho esférico retificada pelo método de transformação direta de coordenadas

Figura 5.9 – Imagem adquirida com o espelho hiperbólico retificada pelo método desenvolvido que leva em consideração a equação do espelho hiperbólico e a propriedade de centro único de projeção

Figura 5.10 – Imagem adquirida com o espelho hiperbólico retificada pelo método de transformação direta de coordenadas

Observando-se a imagem capturada pelo sistema de visão omnidirecional com o espelho esférico, Fig.(5.7.a), verifica-se que os objetos na periferia da imagem são representados por uma quantidade de pixels menor que a usada para representar os mesmos objetos na periferia da imagem capturada com o espelho hiperbólico, Fig.(5.7.b). Isso acontece pois devido à sua forma, o espelho esférico produz imagens com uma baixa resolução na periferia (Derrien; Konolige, 2000). O mesmo resultado também pode ser observado ao comparar-se a Fig.(5.8) e a Fig.(5.10) que correspondem ao resultado da retificação, obtida pelo mesmo método, das imagens capturadas com os dois espelhos.

As Figuras (5.9) e (5.10) foram obtidas pela retificação da imagem adquirida com o espelho hiperbólico através de dois métodos de retificação distintos. Ao comparar estas duas imagens, pode-se perceber que na Fig.(5.10) houve uma maior preservação da proporção na imagem da pessoa no centro, e da cadeira à esquerda. As distorções devido a forma do espelho foram eliminadas na Fig.(5.10), restando apenas as distorções devido ao efeito de perspectiva, e as distorções horizontais devido ao plano cilíndrico de projeção. Assim, comprova-se que, ao se usar a equação da superfície do espelho hiperbólico juntamente com sua propriedade de centro único de projeção, é possível obter-se uma imagem panorâmica livre das distorções de forma do espelho.

O mesmo resultado discutido acima pode ser observado em outro par de imagens capturados com o sistema de visão omnidirecional. A Figura (5.11) mostra imagens obtidas quando uma folha de papel quadriculada é enrolada ao redor do cilindro de

acrílico. As linhas horizontais eqüidistantes da folha de papel são mapeadas como círculos na imagem omnidirecional, e as linhas verticais são mapeadas como segmentos de reta radiais.

Comparando-se as duas imagens da Fig.(5.11), percebe-se um maior número de linhas circulares na imagem obtida com o espelho esférico. Isso se deve ao fato do ângulo de visão do espelho esférico ser maior. No entanto, também se percebe que, na periferia das imagens, o espaçamento entre as linhas é maior na imagem adquirida pelo espelho hiperbólico. Esse resultado vem do fato do espelho esférico possuir uma resolução periférica menor que o espelho hiperbólico. O mesmo pode ser observado ao se comparar as imagens retificadas pelo método de mapeamento direto de coordenadas, Fig.(5.12) e (5.14).

Figura 5.11 – Imagens adquiridas pelo sistema de visão omnidirecional de uma folha de papel com linhas horizontais eqüidistantes colocada ao redor do sistema. (a) Imagem adquirida com o espelho esférico. (b) Imagem adquirida com o espelho hiperbólico

A Figura (5.13) mostra a imagem adquirida pelo espelho hiperbólico retificada pelo método que considera a equação do espelho e o centro de projeção único. Observa-se que as linhas horizontais mantêm um espaçamento quase constante. Isso mostra que esse método de retificação elimina os efeitos de distorção na vertical da imagem panorâmica causados pela forma do espelho. As linhas horizontais na imagem não possuem espaçamento completamente constante pois ainda resta o efeito de

perspectiva. Esse efeito de perspectiva vêm do fato de se usar o centro único de projeção na retificação da imagem.

Figura 5.12 – Imagem da folha de papel adquirida com o espelho esférico retificada pelo método de transformação direta de coordenadas

Figura 5.13 – Imagem da folha de papel adquirida com o espelho hiperbólico retificada pelo método desenvolvido que leva em consideração a equação do espelho hiperbólico e a propriedade de centro único de projeção

Figura 5.14 – Imagem da folha de papel adquirida com o espelho hiperbólico retificada pelo método de transformação direta de coordenadas

Além da retificação para obtenção de uma imagem panorâmica, também foi implementado o método de retificação para obtenção de imagens perspectivas. A

Figura (5.15) mostra uma imagem omnidirecional adquirida pelo sistema de visão e imagens perspectivas obtidas a partir desta imagem adquirida.

As imagens perspectivas foram obtidas definindo-se os ângulos de direção do plano de projeção perspectiva e uma distância focal em pixel que representa a distância do plano de projeção ao foco do espelho. O valor dessa distância focal para as imagens obtidas na Fig.(5.15) foi de 170 pixels. Ao mudar este valor, o efeito obtido é o de zoom da imagem perspectiva.

Figura 5.15 – Imagens perspectivas geradas a partir de imagem omnidirecional

Para implementação dos métodos de retificação que utilizam a equação do espelho hiperbólico, é utilizado um método iterativo de solução de equação para se determinar a correspondência entre os pixels da imagem retificada e os pixels da imagem adquirida pelo sistema de visão. Esse método iterativo consome um tempo computacional que impede o seu uso em tempo real. Apesar disso, implementou-se os métodos de retificação para obter imagens panorâmicas ou perspectivas em tempo real. Para isso foram utilizadas *lookup tables*, ou seja, tabelas que guardam a correspondência entre os pixels depois do cálculo ser realizado. Dessa forma, o

cálculo que utiliza o método iterativo é realizado apenas uma vez quando o programa é iniciado, e então se utiliza a *lookup table* para formação em tempo real da imagem panorâmica ou perspectiva. As frequências de retificação obtidas foram maiores que a frequência de aquisição de imagem. Isso significa que todos os quadros de imagem adquiridos são transformados em um quadro de imagem retificada.

Na execução dos métodos de retificação em tempo real utilizando o computador externo ao robô obteve-se um valor de 46,1 Hz para a frequência média de atualização de imagens panorâmicas de tamanho 800 (H) x 240 (V) pixels. Já a frequência média obtida na atualização de imagens perspectivas de tamanho 128 (H) x 96 (V) foi de 95,8 Hz. O tempo médio obtido para cálculo da *lookup table* de tamanho 800 x 240 utilizada na obtenção de imagens panorâmicas foi de 2,9 segundos. O tempo médio obtido para cálculo da *lookup table* de tamanho 1440 x 484 utilizada na obtenção de imagens perspectivas foi de 12,1 segundos.

A *lookup table* utilizada no método de retificação para obtenção de imagens perspectivas relaciona os ângulos de azimuth e de elevação com os pixels da imagem adquirida (seção 2.3.3). Assim, para uma tabela de 1440 x 484, utiliza-se 1440 pontos para mapear os ângulos de azimuth que variam de 0 à 360°, e utilizam-se 484 pontos para mapear os ângulos de elevação que variam de 31° à -90°. Portanto, a *lookup table* utilizada possui uma resolução de 0,25° para os ângulos de azimuth e de elevação.

Os algoritmos de retificação foram implementados visando avaliar a qualidade dos espelhos, validar a metodologia utilizada para a construção dos espelhos, e verificar a propriedade de centro único de projeção do espelho hiperbólico. O método de retificação para geração de imagens perspectivas é usado na solução adotada para o controle servo visual do robô móvel pois o rastreamento visual do alvo é feito em tempo real nas imagens perspectivas geradas, conforme mostrado na próxima seção.

5.3 Rastreamento visual do alvo

O método de rastreamento visual baseado em região de referência e soma da diferença dos quadrados foi implementado neste trabalho utilizando-se a biblioteca

Xvision (Hager; Toyama 1998) disponível para download e de uso livre para fins acadêmicos.

Aplicou-se o método de rastreamento nas imagens perspectivas geradas em tempo real a partir da imagem adquirida pelo sistema de visão omnidirecional com o espelho hiperbólico. O método de rastreamento foi implementado de forma que à medida que o alvo se movimenta, os ângulos de direção do plano de projeção perspectiva são alterados para que o alvo permaneça sempre no centro da imagem perspectiva, conforme descrito na seção 3.3.

O algoritmo de rastreamento implementado pode rastrear o alvo a uma frequência média de 87,8 Hz. Este resultado foi obtido para uma janela perspectiva de 128 x 96 pixels, $f_d = 170$ pixels, região do alvo de tamanho 40 x 40 pixels, redução de resolução de 2, e considerando apenas translação, rotação e mudança de escala no modelo de transformação do alvo. Ao mudar-se o tamanho da região do alvo para 50 x 50 pixels obteve-se uma frequência de rastreamento de 82,9 Hz, e para uma região de 80 x 80 pixels obteve-se uma frequência de rastreamento de 65,5 Hz.

A Figura (5.16) mostra alguns resultados obtidos com o método de rastreamento visual implementado. A Figura (5.16.a) representa o instante em que o rastreamento foi iniciado, e a Fig.(5.16.b) e Fig.(5.16.c) mostram instantes consecutivos. Na Figura (5.16.a), a imagem maior representa a imagem adquirida pelo sistema de visão, abaixo dela à esquerda se encontra a imagem perspectiva gerada, e à direita se encontra a imagem do alvo sendo rastreado reduzida de um fator de 2. A tabela (5.1) mostra os valores de saída do método de rastreamento visual, ou seja, os valores de translação, rotação, e escala do alvo, além dos valores dos ângulos de direção da imagem perspectiva. Estes valores de saída do rastreamento visual são usados como informação visual para o controle servo visual do robô móvel.

Figura 5.16 – Resultado do método de rastreamento baseado em SSD aplicado em imagens perspectivas obtidas pela retificação da imagem omnidirecional.

Tabela 5.1 – Resultados numéricos do método de rastreamento baseado em SSD aplicado em imagens perspectivas obtidas pela retificação da imagem omnidirecional para os casos mostrados na Fig. (5.16)

	Coordenadas do centro da janela		Rotação (radianos)	Fator de escala	φ (graus)	θ (graus)
	u_p (pixels)	v_p (pixels)				
(a)	64.991	41.9848	0.0297684	1.00343	0.147174	-19.8579
(b)	65.2626	42.0287	0.0628379	0.639316	-37.3551	-18.5041
(c)	66.093	41.7768	0.0477359	1.07126	-51.9687	-18.5483
(d)	65.7152	41.9972	-0.142448	1.09901	-93.3769	-24.525

O modelo de transformação afim, utilizado no método de rastreamento visual para descrever as transformações sofridas pelo alvo, não vale para as imagens omnidirecionais adquiridas diretamente pelo sistema de visão omnidirecional. Isso porque o modelo de transformação afim não representa de forma correta as transformações sofridas pelo alvo na imagem omnidirecional. Por esse motivo o método de rastreamento foi aplicado nas imagens perspectivas, obtidas através de uma retificação local da imagem omnidirecional, conforme foi apresentado nessa seção. Entretanto, para o alvo utilizado, seria possível aplicar o método de rastreamento visual diretamente na imagem omnidirecional se o objetivo fosse apenas limitar uma região de processamento dentro da imagem omnidirecional. Neste caso, os parâmetros de deformação do alvo calculados pelo método de rastreamento não poderiam ser utilizados como entrada para o controle servo visual, já que eles não representam corretamente as deformações que o alvo sofre na imagem. Então, para se fazer o controle servo visual do robô, seria necessário extrair características do alvo através do processamento da região limitada pelo rastreamento visual.

Para ilustrar o funcionamento do rastreamento visual aplicado na imagem omnidirecional, a Fig.(5.17) mostra algumas respostas obtidas. Utilizou-se como parâmetros do rastreamento visual, uma janela de referência de 50x50 pixels e um fator de redução de resolução de 2. A Fig.(5.17.a) mostra a imagem inicial capturada e a Fig.(5.17.e) o alvo dado como referência sujeito a redução de resolução. As Fig.(5.17.b), Fig.(5.17.c) e Fig.(5.17.d) são referentes a instantes posteriores e

representam a imagem obtida pela câmera. Para cada uma destas imagens, as Fig.(5.17.f), Fig.(5.17.g) e Fig.(5.17.h) mostram o alvo rastreado. Observe que a imagem do alvo rastreado tenta manter a mesma orientação e tamanho da janela de referência, verificando o rastreamento quando se tem translação, rotação e mudança de escala do alvo.

Figura 5.17 – Resultados do método de rastreamento baseado em região de referência

5.4 Controle servo visual do robô móvel

Para fazer com que o robô móvel seguisse um alvo móvel definido, implementou-se um controle servo visual baseado em imagem com estrutura hierárquica. Utilizando a configuração de hardware mencionada na seção 5.1 foi possível fazer com que a malha de controle servo visual operasse a uma taxa de 33ms, sem qualquer perda de quadro de imagem. O controlador interno do robô opera a uma taxa de 7.5ms.

Figura 5.18 – Fluxograma do programa implementado para controle do robô

Na solução implementada, para o alvo parado, primeiramente especifica-se a região do alvo em uma imagem perspectiva gerada a partir da retificação local da imagem omnidirecional. Esta região é definida como sendo a região de referência para o rastreamento visual, e os quatro cantos desta região são usados como referência para

o controle servo visual. A partir daí o alvo pode ser movimentado, e a Figura (5.18) mostra um diagrama do funcionamento do loop de controle servo visual implementado. As transformações sofridas pela região do alvo na imagem perspectiva são calculadas pelo método de rastreamento visual. O parâmetro de translação do alvo na imagem calculado pelo rastreamento visual é usado para atualizar o ângulo de direção do plano de projeção perspectiva de forma que o centro do alvo permaneça sempre na mesma posição da imagem perspectiva. A direção das rodas do robô também é alterada para que esteja sempre na direção do alvo. O parâmetro de fator de escala do alvo é usado para calcular o sinal de erro que serve como entrada do controlador servo visual. O controlador servo visual utiliza o sinal de erro para gerar o sinal de controle de velocidade de translação do robô. Os sinais de controle de direção e velocidade das rodas são enviadas para o controlador interno do robô através da ethernet. Estes valores são tidos como referência para o controlador de baixo nível interno ao robô. O parâmetro de rotação do alvo na imagem, que é calculado pelo método de rastreamento visual, não foi usado pois não têm nenhum efeito no controle de posição do robô em relação ao alvo.

Um experimento foi realizado iniciando o controle com o alvo a uma distância de um metro do robô, conforme mostrado na Fig.(5.19). Após iniciado o controle, o alvo foi movido em linha reta com uma velocidade aproximadamente constante e então foi parado. Este experimento foi realizado utilizando o controlador servo visual implementado como controlador proporcional.

Figura 5.19. Alvo a ser seguido pelo robô móvel

O experimento apresentado foi repetido várias vezes e as Fig.(5.20) e (5.21) mostram um dos resultados obtidos. A Figura (5.20) mostra os resultados obtidos para o sinal de erro e a Fig.(5.21) mostra o sinal de controle de velocidade de translação para o experimento realizado com o controle proporcional. Como o único parâmetro calculado pelo rastreamento visual usado para gerar o sinal de erro é o fator de escala do alvo, na Fig.(5.20) pode-se considerar que $e_1 = e_4$, $e_2 = e_3$, $e_5 = e_6$, e $e_7 = e_8$ (vide também a Fig.4.6). Na Fig.(5.21) é apresentado apenas o sinal de controle da velocidade de translação, pois nos experimentos realizados o robô se move em linha reta, portanto a direção das rodas permanece a mesma durante o movimento.

Na Fig.(5.20) o sinal de erro permanece com valores menores que 1 pixel enquanto o robô está parado. O erro aumenta quando o alvo começa a se mover, permanece quase constante e depois diminui quando o alvo pára e o robô se aproxima dele. Quando o alvo pára de se mover, o sinal de erro reduz para aproximadamente ± 1 pixel. Isto é devido a resolução do controle de velocidade da malha interna do robô. Ou seja, o sinal de controle de velocidade gerado a partir deste erro não é suficiente para fazer com que o robô se mova. A Fig.(5.21) mostra o perfil de velocidade do robô. O comportamento deste gráfico segue o sinal de erro e mostra a velocidade máxima de aproximadamente 170 mm/s.

Figura 5.20 – Sinal de erro obtida com controle proporcional

Figura 5.21 – Sinal de controle de velocidade de translação obtida com controle proporcional

Capítulo 6

CONCLUSÃO

Neste trabalho diversos tipos de sistemas omnidirecionais foram estudados e implementou-se um sistema composto por uma câmera fixa disposta em frente a um espelho convexo de perfil hiperbólico. Este espelho foi projetado para ter 40mm de diâmetro resultando em um sistema omnidirecional com um tamanho adequado para o robô móvel utilizado nesta aplicação. Um espelho esférico também foi construído para que pudesse ser feita uma comparação da imagem obtida pelo espelho hiperbólico e pelo esférico. As imagens adquiridas com o uso do espelho hiperbólico mostraram ter uma boa distribuição da resolução se comparadas com as imagens adquiridas com um espelho esférico.

Algoritmos de retificação da imagem omnidirecional foram implementados utilizando-se a propriedade de centro único de projeção do espelho hiperbólico. Os algoritmos implementados permitem a obtenção de imagens panorâmicas e perspectivas em tempo real. A imagem panorâmica é obtida através do mapeamento em um plano cilíndrico ao redor do sistema os pixels da imagem capturada pela câmera. Quando estes pixels da imagem adquirida pelo sistema são mapeados em um plano de projeção perspectiva definido em uma determinada direção e distância do sistema, obtém-se a imagem perspectiva. Assim o efeito obtido com o sistema de visão omnidirecional desenvolvido através do método de retificação para obtenção de imagens perspectivas é o de se ter uma câmera virtual de projeção perspectiva localizada no centro único de projeção. Pode-se então obter com o sistema de visão omnidirecional a mesma funcionalidade de uma câmera perspectiva montada em um sistema de controle *pan/tilt*.

Utilizando o sistema de visão desenvolvido, foi proposta uma solução de controle servo visual para permitir que um robô móvel siga um alvo conhecido dentro de uma sala. O sistema de controle utiliza informações visuais extraídas das imagens capturadas pelo sistema de visão omnidirecional embarcado no robô para realimentar

em tempo real a malha de controle do robô. Para obter as informações visuais do alvo foi escolhido um método de rastreamento visual que utiliza uma região definida como referência e faz o rastreamento baseado na minimização do erro calculado pela soma do quadrado da diferença (SSD) entre a região de referência definida e a imagem atual. O método de rastreamento retorna as transformações sofridas pela região de referência tornando possível determinar se houve rotação, translação ou mudança no fator de escala da região de referência considerada como alvo. Este método de rastreamento considera o alvo como sendo uma superfície plana e rígida, e modela as deformações sofridas pelo alvo como transformações afins. O método foi implementado e aplicado em imagens perspectivas geradas em tempo real a partir da imagem adquirida pelo sistema de visão omnidirecional. A posição na imagem perspectiva dos quatro cantos da região rastreada é tida como referência para o controle servo visual do robô. A direção do plano de projeção perspectiva é constantemente alterada de forma que o centro do alvo permaneça sempre na mesma posição na imagem perspectiva. A biblioteca XVision descrita no trabalho de Hager; Toyama (1998) foi utilizada para implementação do método de rastreamento visual.

O controle servo visual do robô é feito através de um controle servo visual baseado em imagem com malha de controle hierárquica. Este método de controle não requer que a posição real do alvo ou do robô no ambiente seja conhecida e também não requer calibração da câmera. O controle servo visual foi implementado à semelhança do trabalho de Carvalho et al (2000). No entanto, utilizou-se a imagem perspectiva obtida a partir de uma imagem omnidirecional ao invés de um sistema de posicionamento de câmera do tipo *pan/tilt*. Também o método de rastreamento do alvo e a tarefa de controle implementados aqui diferem dos apresentados em Carvalho et al (2000), que utiliza um método de rastreamento de aglomerado de pixels (*blob*) e implementa uma tarefa de *docking* do robô.

O controle servo visual pode ser implementado com um período de amostragem de 33ms. Dessa forma, não ocorre perda de nenhum quadro de imagem capturado pelo sistema de visão. Testes realizados apresentaram os resultados esperados e o robô pode seguir o alvo móvel em tempo real.

Como trabalhos futuros, outros métodos de rastreamento visual ou controle servo visual poderiam ser implementados utilizando o sistema de visão omnidirecional desenvolvido. Uma abordagem alternativa para o rastreamento visual do alvo seria utilizar um método que permitisse a detecção de efeitos de distorção perspectiva na região do alvo. Outra abordagem seria utilizar um método de rastreamento que permitisse o rastreamento do alvo direto na imagem omnidirecional adquirida pelo sistema de visão. Por fim, na implementação do controle servo visual poderia ser utilizado um controlador proporcional derivativo, ou até mesmo um filtro para estimar a velocidade do alvo no ambiente.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALLEN, P. K.; TIMCENKO, A.; YOSHIMI, B.; MICHELMAN, P. Real-time visual servoing. **IEEE Transactions on Robotics and Automation**, v.9, n.2, p152-165, 1993.
- ALLEN, P. K.; YOSHIMI, B.; TIMCENKO, A. Real-time visual servoing. In: IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, Sacramento, California, abril 1991. **Proceedings**. Piscataway, NJ: IEEE Computer Press, 1991. p.851-856.
- BAKER, S.; NAYAR, S.K. A Theory of Catadioptric Image Formation. In: INTERNATIONAL CONFERENCE ON COMPUTER VISION, 6., Bombay, India, 1998. **Proceedings**. 1998. p.35-42.
- CAO, Z. L.; RONING, J. J.; HALL, E. L. Omnidirectional vision navigation integrating beacon recognition with positioning. In: MOBILE ROBOTS CONFERENCE, Cambridge, Massachusetts, U.S.A, outubro 1986. **SPIE Proceedings**, vol. 727. 1986 p.213-220.
- CARVALHO, J. R. H.; RIVES, P.; SANTA-BÁRBARA, A.; BUENO, S. S. Visual servo control of a class of mobile robot. In: IEEE INTERNATIONAL CONFERENCE ON CONTROL APPLICATIONS, Anchorage, Alaska, USA, 2000. **Proceedings**. Piscataway, NJ: IEEE Computer Press, 2000. p.1-6.
- CHAHN, J.S.; SRINIVASAN, M.V. Reflective Surfaces for Panoramic Imaging. **Applied Optics**, v.36, n.31, p.8275-8285, 1997.
- CHANG, P.; HEBERT, M. Omni-directional Visual Servoing for Human-Robot interaction. In: IEEE/RSJ INTERNATIONAL CONFERENCE ON

- INTELLIGENT ROBOTIC SYSTEMS (IROS), Victoria, B.C., Canada, outubro 1998. **Proceedings**. Piscataway, NJ: IEEE Computer Press, 1998. p.1801-1807
- CHAUMETTE, F.; RIVES, P.; ESPIAU, B. Positioning of a Robot to an Object, Tracking it and Estimating its Velocity by Visual Servoing. In: IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, Sacramento, California, 1991. **Proceedings**. Piscataway, NJ: IEEE Computer Press 1991. p.2248-2253.
- CONROY, T.L.; MOORE, J.B. Resolution Invariant Surfaces for Panoramic Vision Systems. In: IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV'99), Kerkyra, Grécia, 1991. **Proceedings**. Piscataway, NJ: IEEE Computer Press, 1991. p.392-397.
- CORKE, P. **Visual Control of Robots**: High-Performance Visual Servoing, Taunton: Research Studies Press Ltda, 1997. 353p. (Robotics and Mechatronics Series, 2)
- DERRIEN, S.; KONOLIGE, K. Approximating a Single Viewpoint in Panoramic Imaging Devices. In: IEEE WORKSHOP ON OMNIDIRECTIONAL VISION (OMNIVIS'00), Hyatt Regency, Hilton Head Island, 2000. **Proceedings of ICRA'00**. Piscataway, NJ: IEEE Computer Press, 2000. p. 83-90.
- ESPIAU, B.; CHAUMETTE, F.; RIVES, P. A New Approach to Visual Servoing in Robotics. **IEEE Transactions on Robotics and Automation**, v.8, n.3, p.313-326, 1992.
- GÄCHTER, S.; PAJDLA, T.; MIČUŠÍK, B. M. Mirror Design for an Omnidirectional Camera with a Space Variant Imager. In: WORKSHOP ON OMNIDIRECTIONAL VISION APPLIED TO ROBOTIC ORIENTATION AND NONDESTRUCTIVE TESTING (NDT), Budapest, Hungary, 2001. **Proceedings**. 2001. Disponível em: <ftp://cmp.felk.cvut.cz/pub/cmp/articles/pajdla/gaechter-OV2001.ps.gz>. Acesso em: 19 nov. 2001.

- GASPAR, J.; WINTERS, N.; SANTOS-VICTOR, J. Vision-based Navigation and Environmental Representations with an Omni-directional Camera. **IEEE Transactions on Robotics and Automation**, v.16, n.6, p.890 -898, 2000.
- GILBERT, A.; GILES, M.; FLACHS, G.; ROGERS, R.; YEE, H. A real-time video tracking system. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v.2, n.1, p.47-56, 1980.
- GRASSI JR, V.; OKAMOTO JR, J. **Desenvolvimento de Linguagem de Programação para Robôs Móveis**. São Paulo: Departamento de Engenharia Mecânica da Escola Politécnica da USP, 1998. (Trabalho de Formatura)
- HAGER, G.; BELHUMEUR, P. Efficient region tracking with parametric models of geometry and illumination. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v.20, n.10, p1025–1039, 1998.
- HAGER, G.; TOYAMA, K. The “XVision” system: A general purpose substrate for real-time vision applications. **Computer Vision and Image Understanding**, v.69, n.1, p.23–27, 1998.
- HICKS, A.; BAJCSY, R. Reflective Surfaces as Computational Sensors. In: WORKSHOP ON PERCEPTION FOR MOBILE AGENTS AT CVPR’99, Fort Collins, Colorado, USA, 1999. **Proceeding of CVPR99**. Piscataway, NJ: IEEE Computer Press, 1999
- HUTCHINSON, S.; HAGER, G.; CORKE, P. A tutorial on visual servoing, **IEEE Transactions on Robotics and Automation**, v.12, n.5, p.651-670, 1996.
- ISHIGURO, H. Development of low-cost and compact omnidirectional vision sensors and their applications. In: INTERNATIONAL CONFERENCE INFORMATION SYSTEMS, ANALYSIS AND SYNTHESIS, 1998. **Proceedings**. 1998. p. 433-439.

- IZZO, A. R.; MIYAHARA, A. Y; OKAMOTO JR., J. **Projeto do Acionamento Mecânico da Base Móvel**. São Paulo: Departamento de Engenharia Mecânica da Escola Politécnica da USP, 1994. (Trabalho de Formatura)
- NAYAR, S.K. Catadioptric Omnidirectional Camera. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), Porto Rico, 1997. **Proceedings**. Piscataway, NJ: IEEE Computer Press, 1997. p.482-488.
- PALMA, J.G.; PORTO, A.J.V. Sistema de Apoio a Programação Comando Numérico para Usinagem de Ultraprecisão. São Carlos, 1995. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo.
- PERI, V.N.; NAYAR, S.K. Generation of Perspective and Panoramic Video from Omnidirectional Video. In: DARPA IMAGE UNDERSTANDING WORKSHOP, 1997. **Proceedings**. 1997. p.243-245.
- SANDERSON, A. C.; WEISS, L. E. Image-based visual servo control using relational graph error signals, IEEE INTERNATIONAL CONFERENCE ON CYBERNETIC AND SOCIETY, 1980. **Proceedings**. 1980. p.1074-1077.
- SANTOS-VICTOR, J. Visual Servoing of Cellular Robots, In: EUROPEAN CONTROL CONFERENCE (ECC2001), Porto, Portugal, setembro 2001. **Proceedings**. 2001. Disponível em: <www.isr.ist.utl.pt/labs/vislab/publications/ps/01-ecc-cellular.pdf>. Acesso em: 19 nov. 2001
- SHIRAI, Y.; INOUE, H. Guiding a robot by visual feedback in assembling tasks, **Pattern Recognition**, v.5, p.99-108, 1973.
- SVOBODA, T.; PAJDLA, T.; HLAVAC, V. **Central Panoramic Cameras: Geometry and Design**. Praga: Czech Technical University, Faculty of Electrical Engineering, Center for Machine Perception, dez. 1997. (Research Report K335/97/147). Disponível em: <<ftp://cmp.felk.cvut.cz/pub/cmp/articles/svoboda/TR-K335-97-147.ps.gz>>. Acesso em: 19 nov. 2001.

TACHIBANA, L. S.; FUJIWARA, R.; OKAMOTO JR., J. **Projeto Base Móvel: 2ª Fase de Desenvolvimento do Hardware Eletrônico**. São Paulo: Departamento de Engenharia Mecânica da Escola Politécnica da USP, 1995. (Trabalho de Formatura)

TOYAMA, K.; HAGER, G. Incremental focus of attention for robust visual tracking. **International Journal of Computer Vision**, v.35, n.1, p.45–53, 1999.

VAN DER ZWAAN, S.; SANTOS-VICTOR, J. Real-time Vision-based Station Keeping for Underwater Robots. In: IEEE OCEANS 2001, Honolulu, novembro 2001. **Proceedings**. 2001. Disponível em: <www.isr.ist.utl.pt/labs/vislabs/publications/ps/01-oceans-sleeping.pdf>. Acesso em: 19 nov. 2001

YAGI, Y. Omnidirectional Sensing and Its Applications. **IEICE Transactions on Information and Systems**, v.E82-D, n.3, p.568-579, 1999.

YAGI, Y.; KAWATO, S.; TSUJI, S. Real-Time Omnidirectional Image Sensor (COPIS) for Vision-Guided Navigation. **IEEE Transactions on Robotics and Automation**, v.10, n.1, p.11-22, 1994.

ZHANG, D. B.; GOOL, L. V.; OOSTERLINCK, A.. Stochastic predictive control of robot tracking systems with dynamic visual feedback. In: IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, Cincinnati, Ohio, maio 1990. **Proceedings**. Piscataway, NJ: IEEE Computer Press, 1990. p.610-615.