

# 1 Introdução

Após o crescimento individual das áreas de robótica e visão computacional, percebe-se hoje em dia uma fusão de ambas. A visão computacional é uma ciência que tem como objetivo extrair informações de imagens capturadas por dispositivos como câmera de vídeo, *scanner*, etc. Por sua vez, a robótica está sendo muito explorada no campo industrial, a fim de garantir uma melhora significativa na eficácia e qualidade do trabalho, além da possibilidade de substituir o trabalho humano em locais perigosos e de difícil acesso.

Sistemas robóticos podem ser entendidos como dispositivos eletromecânicos equipados com sensores e atuadores, controlados através de um sistema computacional. Estes sensores são utilizados para medir diferentes grandezas no ambiente de trabalho, como por exemplo posição, velocidade e força.

Atualmente, a maior parte dos robôs industriais é programada para seguir uma trajetória pré-definida. Isso é suficiente quando o robô trabalha num ambiente fixo, onde os objetos de interesse estão sempre a uma distância pré-definida do robô. Entretanto, se a posição do robô é alterada, todas as trajetórias devem ser reprogramadas para que este ainda seja capaz de realizar as devidas operações (Augustson, 2007).

A visão estereoscópica ou estéreo diz respeito à reconstrução da informação em três dimensões. É a reprodução artificial da visão binocular natural. A visão monocular permite examinar a posição e a direção dos objetos dentro do campo da visão humana em um único plano. Permite reconhecer nos objetos a forma, as cores e as dimensões. A fotografia simples é uma reprodução da visão monocular. Por outro lado, a visão binocular permite a percepção de profundidade, que é dada pela diferença de ângulos com que as imagens são percebidas.

A estereoscopia é utilizada para a obtenção da posição de um ponto no espaço tridimensional. Para isso, no entanto, é necessário determinar projeções de pontos da cena em um par de imagens. Diversos algoritmos para este fim foram

propostos na literatura, tais como o algoritmo proposto por (Tomasi e Kanade, 1991), e (Harris e Stephens, 1988).

Um dos mais recentes algoritmos utilizados na área de visão computacional para este fim é o algoritmo SIFT (*Scale Invariant Feature Transform*), proposto por David Lowe (2004). Este algoritmo é invariante a rotação, a iluminação, a escala e a posição da câmera, e garante um bom desempenho para diferentes tipos de condições de captura das imagens.

Com o objetivo de explorar as vantagens da visão computacional e da robótica, a PETROBRAS patrocina projetos de identificação de pose de manipuladores durante intervenções submarinas. O manipulador TA-40 (Figura 1), acoplado a um ROV (*Remote Operating Vehicle*), tem como objetivo principal operar válvulas em painéis submarinos sob grandes profundidades. Um sistema baseado em visão computacional seria capaz de determinar automaticamente a posição de um objeto alvo. Já o ROV é o veículo responsável pela locomoção do manipulador, capaz de operar em grandes profundidades.



**Figura 1 - Manipulador TA-40**

Inspirado na aplicação acima, o presente trabalho tem como objetivo comparar diferentes técnicas de controle servo-visual através do desenvolvimento de um sistema robótico experimental composto de uma mesa automatizada XYθ e

uma câmera acoplada à sua extremidade. Esta câmera é responsável pela captura de imagens que servirão de base para o sistema de visão computacional determinar e controlar a posição da mesa robótica. O projeto é composto de quatro partes distintas: a primeira referente à mecânica do robô em questão, onde uma mesa XYθ comercial deverá ser automatizada e controlada por um computador. A segunda refere-se à eletrônica utilizada como interface entre o robô e o computador. A terceira parte refere-se ao *software* de visão computacional, responsável por determinar, a partir de imagens obtidas pela câmera, a posição do alvo a ser alcançada pelo robô. Será utilizado o algoritmo SIFT para a determinação dos pontos-chaves nas imagens. A última parte diz respeito às diferentes técnicas de controle a serem aplicadas no posicionamento do robô. Através do controle, espera-se que o robô seja capaz de alcançar diferentes configurações em relação aos objetos alvos. Espera-se que o robô atinja seu alvo utilizando a realimentação de sensores de posição (*encoders*), e através de imagens capturadas ao longo de seu percurso. Smith e Papanikolopoulos (1996) chamam o controle com realimentação exclusiva por sensores de posição de controle “cego”, uma vez que, determinada a posição do alvo, o robô se movimenta sem o auxílio da câmera. Já o controle com realimentação por imagem utiliza vários *frames* da câmera, e por isso é capaz de compensar erros de ruídos e de posicionamento durante o percurso. Espera-se, no entanto, que este apresente um desempenho computacional inferior devido ao tempo de processamento de cada imagem.

Diversos autores já desenvolveram pesquisas nestas áreas. Inoue e Shirai (1971) utilizaram um manipulador robótico de 7 graus de liberdade, com uma câmera na extremidade do último elo (sistema *eye-in-hand*), para encaixar um objeto dentro de um orifício do mesmo formato. Através de programação própria, o controle determina visualmente a posição desejada e a real, e através apenas de imagens o objeto alcança o alvo desejado. Allota e Colombo (1999) implementaram um sistema robótico com câmera também na extremidade. As características visuais foram obtidas através de contornos, e com o desenvolvimento de um controle 2D/3D baseado em imagens. O sistema foi capaz de realizar tarefas de posicionamento e movimentação de objetos. Houshangi (1990) desenvolveu um sistema para capturar objetos em movimento usando uma câmera fixa e pré-calibrada.

A contribuição deste trabalho está na comparação entre as principais arquiteturas em um mesmo sistema experimental, especialmente desenvolvido para este propósito. Além disto, este trabalho apresenta uma formulação eficiente para o controle servo-visual usando SIFT quando os pontos de referência se encontram em um mesmo plano.

Esta dissertação está dividida em seis capítulos, descritos da seguinte forma: o capítulo 2 apresenta a teoria necessária para a compreensão do trabalho, dividida entre os assuntos referentes à área de controle e robótica e os assuntos ligados a área de visão computacional. O capítulo 3 apresenta a integração entre as áreas de visão computacional e controle robótico, descrevendo todo o equacionamento e métodos utilizados. O capítulo 4 descreve o procedimento experimental realizado para validar o projeto e para a extração de resultados. As equações apresentadas no capítulo 3 são adaptadas de forma a torná-las compatíveis com o experimento. O capítulo 5 apresenta os resultados alcançados usando as diferentes técnicas de controle. É feita uma comparação entre as técnicas implementadas, apontando as vantagens e desvantagens de cada uma. O capítulo 6 apresenta as conclusões do trabalho.