



## SISTEMA DE RASTREAMENTO POR VISÃO EM ROBÔS MÓVEIS COM OTIMIZAÇÃO POR PROJETO FATORIAL

JOSÉ MAURÍCIO S. T. MOTTA, SÉRGIO ROBERTO GONSALVES TOURINO

Universidade de Brasília, Faculdade de Tecnologia  
Departamento de Engenharia Mecânica  
Campus Universitário Darcy Ribeiro, 70910-900 – Brasília – DF – Brasil

(Recibido 20 de marzo de 2006, para publicación 15 de enero de 2008)

**Resumo** – Este artigo apresenta um método de otimização baseado em análise por projeto fatorial aplicado a um sistema de rastreamento por visão embarcado em um robô móvel, com o objetivo de melhora da resposta do sistema. Uma revisão da técnica de projeto fatorial e detalhes de implementação para identificação e otimização de parâmetros de sensores são apresentados. O método foi aplicado a um modelo que utiliza cálculo de fluxo ótico por correlação SSD (soma dos quadrados das diferenças). Resultados mostram quais os mais importantes parâmetros e suas influências sobre o tempo de processamento das imagens e o erro nas estimativas de posição, incluindo interações entre os parâmetros. Uma técnica de obtenção de valores ótimos de parâmetros para as imagens é implementada e discutida.

**Palavras Chave** – Navegação robótica, rastreamento por visão, projeto fatorial, otimização de parâmetros.

### 1. INTRODUÇÃO

O rastreamento por visão em robôs móveis é uma abordagem cada vez mais utilizada no controle de navegação visto que emula a visão humana. Um sistema ideal de controle seria obtido através de um modelo tridimensional completo do ambiente, planejamento de caminho e execução dos passos para mover o robô neste trajeto. Entretanto, diversos desafios ainda não foram totalmente resolvidos, como reconhecimento de objetos, desvio de obstáculos e fusão sensorial [1-2].

A navegação robótica baseada em rastreamento visual tem sido alvo de grande esforço em pesquisas nos últimos anos [1-4]. Um sistema robusto de rastreamento pode prover informações sobre o movimento 3D relativo de um objeto com relação ao observador, e pode simplificar a obtenção da forma ou localização de um objeto.

Vários pesquisadores têm publicado abordagens diferentes para o problema de rastreamento em robôs e muitas destas tem sido apresentadas para o rastreamento de um objeto móvel em uma seqüência de imagens, baseados em fluxo ótico, correlação de imagens e *snakes* [5-7].

Plakas e Trucco [8,9] apresentam um modelo de câmera não-calibrada para uso em ambiente submarino. O algoritmo de visão baseia-se na reconstrução projetiva de pontos da imagem através do cálculo da distância relativa entre os pontos do sistema estéreo. O algoritmo usa correlação SSD de Tomasi-Kanade-Shi [10,11] e uma forma de rejeição automática de características espúrias por meio de cálculo de resíduos. O sistema mostra-se robusto em ambientes aquáticos artificiais. Entretanto, distâncias absolutas não podem ser obtidas através das imagens sem se conhecer a posição absoluta de pelo menos cinco pontos característicos.

Espiau [12] apresenta uma investigação que fortalece as conclusões de Plakas e Trucco, mostrando resultados de um sistema controlado por visão com pequena influência dos erros de calibração da câmera. A conclusão principal foi que sistemas de visão em malha fechada podem compensar erros de calibração da câmera, significando que tal etapa nesses sistemas tem menor importância.

Spindler [13] apresenta uma estimação do movimento bidimensional aparente induzido por uma seqüência de imagens de um veículo submarino a fim de compensar este movimento, utilizando um modelo “affine” e fluxo ótico baseado em gradiente.

A pesquisa relacionada com este artigo apresenta um esquema de otimização baseado em estatística experimental, denominada projeto fatorial, para investigar parâmetros em um sistema de rastreamento, neste caso utilizando correlação SSD e fluxo ótico. Testes foram realizados utilizando um simulador validado pela comparação de sua curva de velocidade com a curva real de um robô móvel, Nomad XR4000, com câmera monocular. Um controlador por lógica nebulosa foi utilizado para cancelar o movimento aparente das imagens capturadas. Este artigo apresenta somente o projeto dos principais parâmetros do sistema de rastreamento por visão.

## 2. PROJETO FATORIAL

Projeto fatorial é uma técnica para projetar um experimento ou estimar em quanto mudanças nos parâmetros de entrada afetam o sistema [14, 15]. O guia ISO para a Expressão de Incertezas em Medições [16] menciona a utilidade das técnicas de análise de variância para determinar incertezas em medições.

Antes de conceber o projeto de um experimento e aplicar testes de desempenho é necessário considerar o tempo gasto para realizá-los. No projeto fatorial, o número total de corridas ( $N$ ) é determinado utilizando-se a expressão  $N=L^V$ , onde  $L$  é o número de níveis em cada parâmetro e  $V$  é o número de parâmetros experimentais investigados. Como um exemplo, quando estudando quatro parâmetros com três níveis cada, o total de corridas é  $3^4=81$  [15].

Para gerar um projeto fatorial genérico seleciona-se um número fixo de níveis (ou versões) de cada parâmetro (ou fator) e realizam-se experimentos com todas as combinações. Um nível de um parâmetro refere-se aos valores discretos do domínio daquele parâmetro. Por exemplo, se um experimento na qual as temperaturas estudadas têm os valores de 20°C, 50°C e 100°C então o experimento tem 3 níveis associados à variável temperatura (parâmetro). Se há  $l_1$  níveis para o primeiro parâmetro,  $l_2$  níveis para o segundo, ..., e  $l_n$  níveis para o  $n$ -ésimo parâmetro, o arranjo completo de  $l_1 \times l_2 \times \dots \times l_n$  corridas do experimento é denominado projeto fatorial  $l_1 \times l_2 \times \dots \times l_n$ .

Se apenas dois níveis do parâmetro  $n$  são considerados, então este é referido como um projeto fatorial do tipo  $2^n$ . Esse tipo de projeto fatorial: a) requer um menor número de corridas por fator estudado e, mesmo que não seja possível obter informações completas sobre os efeitos de cada fator, permite que experimentos futuros sejam melhor planejados; b) seu uso combinado com o conceito de projeto fatorial fracional (comentado adiante) permite que menos experimentos sejam realizados mesmo com um grande número de fatores; c) pode ser usado em blocos de experimentos cuja complexidade pode seguir os requisitos estabelecidos durante os experimentos e d) a interpretação dos resultados pode ser realizada ou utilizando o senso comum ou através de cálculos elementares.

Projetos fatoriais possuem uma importante capacidade de apresentar as interações entre parâmetros, mas isto não significa que estas interações são significativas. Os efeitos principais de um parâmetro tendem a ser maiores que os efeitos de interações entre dois parâmetros, e estes maiores que os efeitos entre três, e assim por diante.

O número de corridas ou experimentos realizados em um projeto fatorial  $2^n$  cresce geometricamente à medida que  $n$  cresce. Entretanto, em muitos casos a informação desejada dos experimentos pode ser obtida com apenas uma fração das corridas totais, levando ao conceito de projeto fatorial fracional.

Projetos fatoriais fracionais desprezam pequenos efeitos provenientes de interações de altas ordens para reduzir os experimentos. Neste caso define-se um projeto fatorial fracional com base no expoente original menos o fator de fração. Assim, para um projeto fatorial do tipo  $2^5$  usando 1/4 dos experimentos mostra-se que  $1/4 \cdot 2^5 = 2^3$ .  $2^{-2} = 2^{5-2}$ , referindo-o como um projeto fatorial do tipo  $2^{5-2}$ . Entretanto, quanto maior o fracionamento menor a confiança nos resultados.



Fig. 1. Robô móvel Nomad XR4000.

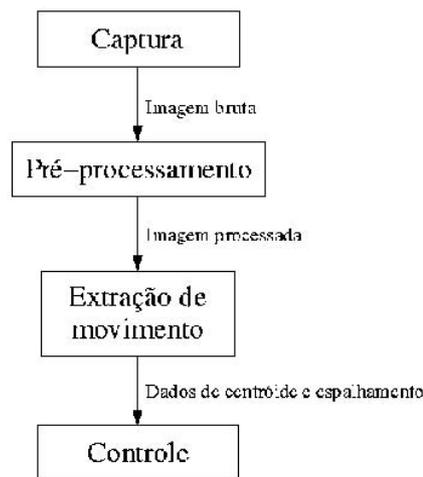


Fig. 2. Módulos do sistema de rastreamento por visão.

### 3. SISTEMA DE RASTREAMENTO BASEADO EM VISÃO

O robô móvel no qual o sistema de rastreamento baseado em visão é montado pode ser visto na Fig. 1, compreendendo sistemas de controle, sensores, comunicação e programação, com o objetivo de aplicação em pesquisa e desenvolvimento em manipulação robótica, visão computacional, navegação por sensores e aprendizado. Possui um subsistema de visão monocular com uma câmera colorida e uma placa de captura com 45 Mbytes/s de transferência, gerando imagens de  $256 \times 256$  píxeis nesta aplicação.

Programas para o subsistema de visão não são disponibilizados pelo fabricante. O sistema de rastreamento desenvolvido neste trabalho compreende os quatro módulos principais apresentados na Fig. 2.

Este artigo apresenta a concepção do segundo e terceiro blocos e o projeto de seus parâmetros. O pré-processamento inclui filtragem, limiarização da imagem e cálculo do gradiente. O passo de cálculo do movimento inclui estimação do fluxo ótico e correlação entre imagens em seqüência. A estimativa da posição é a saída do sistema.

#### 4. CÁLCULO DO FLUXO ÓTICO

O cálculo do fluxo ótico é implementado neste trabalho com base no método da correlação SSD, através da equação abaixo [17]:

$$S(C, d) = \sum_{j=-n}^n \sum_{i=-n}^n W(i, j) [I(C + (i, j), t) - I(C + (i, j), t + 1)]^2 \quad (1)$$

onde  $W(i, j)$  representa uma máscara de pesos,  $I(C, t)$  é a intensidade da imagem em  $C$  no tempo  $t$  e  $d$  é restrito a uma vizinhança com forma quadrada igual a  $(2n + 1)^2$ , centrada em  $C$ .

$S$ : correlação entre imagens no tempo  $t$  e  $t + 1$

$W$ : máscara de pesos

$C$ : posição da imagem  $(x, y)$

$d$ : vizinhança do ponto  $(x, y)$  na imagem

$n$ : semi-largura da vizinhança  $d$

$I$ : valor da intensidade da imagem

$t$ : tempo de captura da imagem prévia

A velocidade desse cálculo é influenciada principalmente por dois fatores: os tamanhos da janela de correlação e da janela de busca.

A janela de busca define a região de busca na imagem dentro da qual a janela de correlação calcula a máxima similaridade entre pontos de duas imagens capturadas dentro de um intervalo de tempo. Esta janela define um conjunto de  $N$  pontos na imagem que serão temporariamente o ponto central  $C(x, y)$  da janela de correlação. A correlação local em  $(x, y)$  é calculada e seu valor armazenado, e uma nova correlação é calculada no próximo ponto dentro da janela de busca. O ponto  $P(x, y, t)$  com a maior correlação é definido como o ponto central da janela de busca,  $B(x, y, t + 1)$ , da próxima imagem.

A janela de correlação é uma matriz quadrada de pontos vizinhos ao ponto central  $C(x, y)$ . Estes pontos são utilizados para o cálculo do valor de correlação entre duas imagens em seqüência.

Uma vez que todos os pontos com maior correlação entre duas imagens são calculados o campo de velocidade de cada posição  $(x, y)$  pode ser estimado ( $v = (v_x, v_y)$ ), produzindo o fluxo ótico. A velocidade é obtida pelo cálculo aproximado da derivada temporal, utilizando diferenças finitas. Os dados de velocidade são usados para antecipar a próxima posição do ponto característico, localizando a nova posição central da janela de busca na próxima imagem.

#### 5. ESTIMAÇÃO DO MOVIMENTO NA IMAGEM

Para calcular o movimento na imagem pontos característicos foram definidos por toda a imagem, utilizando como base uma seleção dos pontos com maiores valores de gradiente da imagem limiarizada. Para uniformizar a distribuição dos pontos característicos na imagem esta foi subdividida em retângulos e o número de pontos característicos em cada um foi limitado. Na seqüência, o excesso de pontos foi eliminado limitando o seu número na imagem. A Fig. 3 mostra uma imagem com diversos pontos característicos indicados por quadrados. O centróide de todos os pontos característicos selecionados na imagem foi calculado, e um valor de espalhamento foi definido como uma dispersão em torno da média, como:

onde:

$(x_c, y_c)$ : centróide dos pontos característicos

$(x_i, y_i)$ : posição do ponto  $i$

$(x_e, y_e)$ : valor de espalhamento dos pontos

$N$ : número total de pontos no conjunto

O movimento da imagem na direção  $x$  foi estimado pela posição do centróide dos pontos característicos e o movimento na direção  $z$  pelo valor de espalhamento indicando que o objeto se aproxima para a câmera quando o mesmo aumenta e se afasta quando este valor diminui.

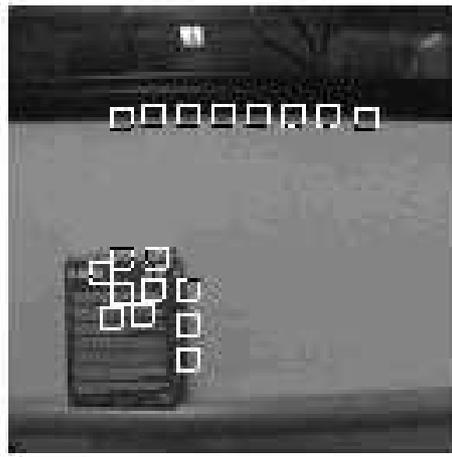
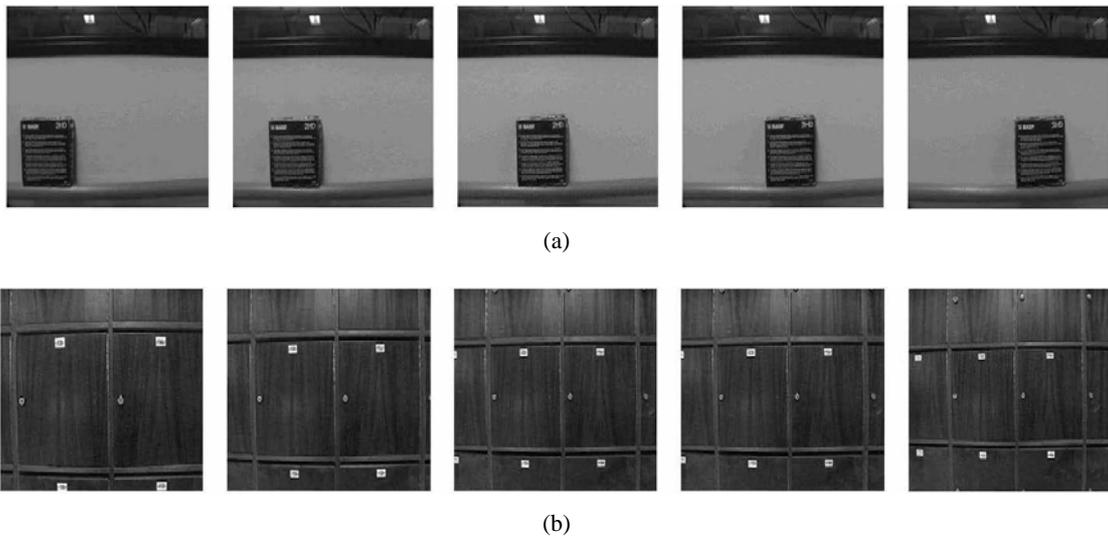


Fig. 3. Pontos característicos em uma imagem.

Fig. 4. Movimentos nas imagens: (a) direção  $x$ ; (b) direção  $z$ .

### 5.1. Análise dos Resultados

A fim de verificar a efetividade do método para calcular o movimento nas imagens testes foram realizados com imagens reais do laboratório, tanto para a direção  $x$  como para a  $z$ , separadamente.

A Fig. 4 apresenta imagens capturadas dos dois tipos de movimento, direções  $x$  e  $z$ , respectivamente. A Tabela 1 mostra os valores de centróide e espalhamento dos pontos característicos para esses movimentos.

Pode ser observado da tabela que durante o movimento o número de pontos característicos (NPI) reduz-se. Isso ocorre porque o algoritmo descarta pontos sem movimento significativo, reduzindo o processamento. A desvantagem dessa estratégia é que os valores de centróide e espalhamento sofrem mudanças abruptas de valores. Apesar dessas variações o sistema de controle não é afetado consideravelmente.

Observa-se pela Tabela 1(a) que o valor de  $y_c$  mantém-se relativamente constante, mostrando que o movimento é predominantemente na direção  $x$ . Verifica-se também que, após a estabilização do valor de

Tabela 1. Valores de centróide e espalhamento dos pontos característicos para movimentos laterais (a) e de afastamento (b).

NPI	$x_c$	$y_c$	$x_e$	$y_e$	NPI	$x_c$	$y_c$	$x_e$	$y_e$
19	48	62	332	504	19	48	62	332	504
19	51	62	267	511	19	46	57	379	568
19	55	63	256	512	8	40	75	40	23
17	56	62	200	444	8	38	75	30	22
17	60	63	218	448	8	40	75	30	21
16	67	64	188	405	8	42	74	29	21
16	71	63	205	398	8	42	76	17	20
16	75	64	218	403	8	43	75	17	24
16	80	64	235	397	8	44	75	15	21
16	84	63	251	395	8	45	74	15	21

(a)

(b)

Tabela 2. Parâmetros utilizados no estudo por projeto fatorial.

Variável (índice)	Uso	(-)	(+)
CORRELATION_SIZE (1)	Largura da janela de busca	10	30
ROWS_CORREL (2)	Largura da janela de correlação	5	15
MAX_FEAT_NUM (3)	Máximo de pontos de interesse	50	200
THRES_PERCENT (4)	Limiar para pontos de interesse	0,1	5,0
UNGROUP_SIZE (5)	Distância para apagar os pontos	5	20

NPI, os valores de  $x_e$  e  $y_e$  sofrem pequena variação, mostrando que a distância objeto-câmera não se altera consideravelmente.

A Tabela 1(b) mostra que o período transiente do valor de NPI é bem menor, garantindo uma maior estabilidade dos valores de centróide e espalhamento. É possível observar que os valores de  $x_c$  e  $y_c$  permanecem relativamente constantes durante a seqüência, indicando que não há movimento na direção  $x$ . Por outro lado os valores de  $x_e$  e  $y_e$  apresentam uma clara redução, mostrando que o objeto afasta-se em relação ao robô.

Pode ser observado desses testes que o algoritmo para estimação do movimento é capaz, na condição de pequenos deslocamentos lineares, de seguir o movimento dos pontos característicos e gerar valores de posição de centróide e de espalhamento como previsto.

## 6. OTIMIZAÇÃO DOS PARÂMETROS DO SISTEMA

Para observar os efeitos dos parâmetros de maior importância na efetividade do sistema de rastreamento, foi realizado um estudo utilizando a metodologia de projeto fatorial. Com isso pode-se também desenvolver uma forma de otimização para acelerar o processamento das imagens. Foram utilizados cinco parâmetros no estudo, mostrados na Tabela 2.

Tabela 3. Resumo das conclusões obtidas com os experimentos.

Variável de Saída	Variáveis que Influenciam
$t_1$	1, 12, 3 e 4
$t_f$	1, 2 e 12
$e_x$	1, 2 e 12
$e_y$	1, 2, 12 e 4

### 6.1. Projeto de Experimentos e Análise dos Resultados

Para simplificar a realização dos experimentos e permitir a repetição de cada variação (obtendo dados estatísticos) um projeto fatorial fracional  $2^{5-1}$  com 3 repetições foi realizado (total de 48 corridas). O parâmetro 5 (UNGROUP\_SIZE) foi fracionado durante os experimentos por ter sido verificado anteriormente uma baixa influência sobre os dados estudados (posteriormente comprovado). A repetição das variações foi necessária, pois o tempo de processamento, uma das variáveis de saída, não se apresentou constante para uma mesma configuração dos parâmetros. A ordem das corridas foi definida aleatoriamente para os experimentos.

A seqüência de experimentos foi feita observando-se os seguintes valores de saída:  $t_1$  (tempo de processamento da 1ª imagem),  $t_f$  (tempo de processamento das imagens seguintes da seqüência),  $e_x$  (erro na direção  $x$ ) e  $e_y$  (erro na direção  $y$ ).

O tempo de processamento das imagens foi medido através de funções internas da linguagem de programação destinadas para esse propósito, como a função *clock()*, que retorna o número de unidades de tempo do processador desde o início do programa. Armazenando-se o instante de tempo inicial basta subtrair do tempo final para obter a duração do processamento das imagens.

Os erros de posição em  $x$  e  $y$  foram calculados utilizando como referência o próprio sistema operando com valores elevados de janela de busca e correlação (respectivamente valores de 50 e 30), de forma que o erro dessa referência seria mínimo em relação ao erro dos testes.

A análise dos resultados obtidos levou às conclusões resumidas na Tabela 3. A convenção adotada indica que 12 representa as interações entre as variáveis 1 e 2. Observa-se que os parâmetros 1 e 2 (largura da janela de correlação e de busca, respectivamente) são os principais fatores que influenciam as saídas  $t_1$ ,  $t_f$ ,  $e_x$  e  $e_y$ .

Para destacar a importância dos parâmetros 1 e 2 foi realizada uma análise de sensibilidade a fim de se obterem valores ótimos para os mesmos, para redução do tempo de processamento sem perda considerável da robustez do sistema, mantendo valores baixos para  $e_x$  e  $e_y$ . Para este propósito, foram realizados testes variando-se a janela de busca entre os valores de 10 e 20 e a janela de correlação entre 5 e 15, valores estes utilizados como limites nas corridas do projeto fatorial realizado anteriormente. Os resultados são apresentados nos gráficos (a), (b) e (c) da Fig. 5. Os testes foram realizados para se obterem os “melhores” valores para os parâmetros 1 e 2, mantendo-se os demais parâmetros com os seguintes valores: (3) 2,0; (4) 100 e (5) 5.

Através de uma análise dos gráficos foram obtidos, por inspeção visual, os valores de 9 para largura da janela de correlação e de 17 para largura da janela de busca. Estes valores apresentam baixo tempo de processamento e baixo erro, simultaneamente, ainda que não ideais isoladamente.

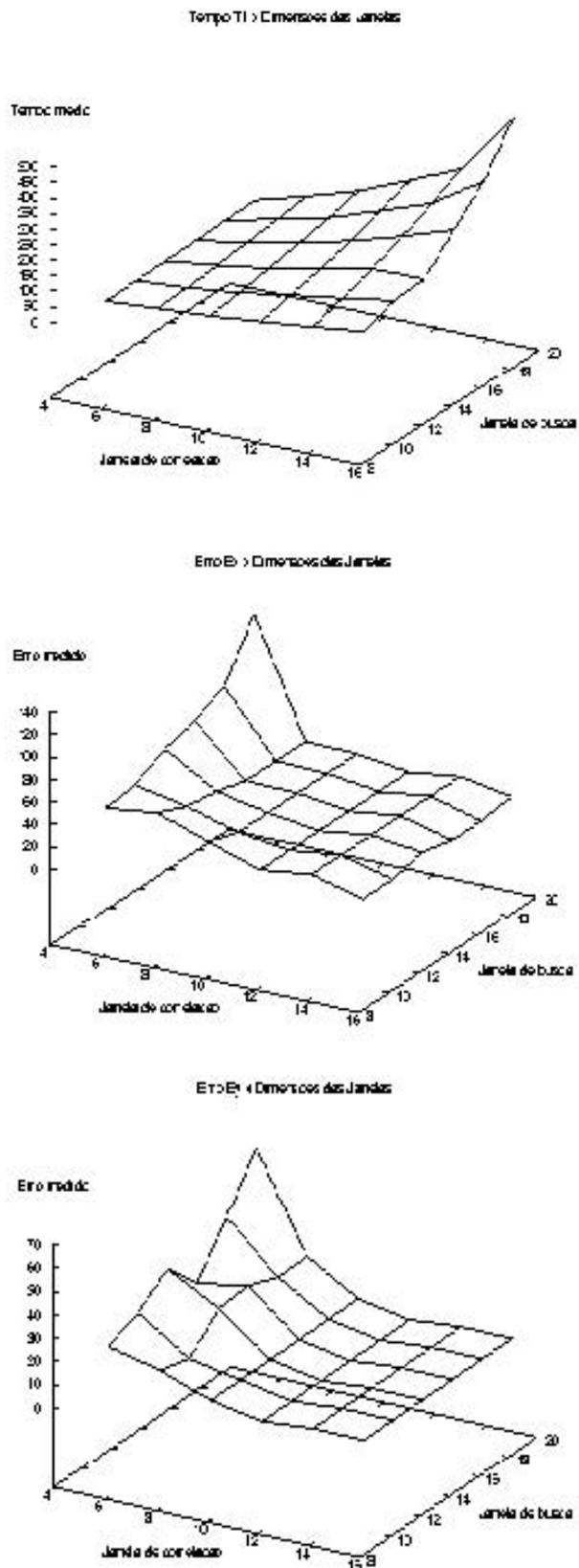


Fig. 5. Resultados da análise de sensibilidade: (a) Tempo x dimensões das janelas; (b) Erro  $e_x$  x dimensões das janelas; (c) Erro  $e_y$  x dimensões das janelas.

## 7. CONCLUSÕES

Os estudos realizados através da metodologia de projeto fatorial foram capazes de agregar um maior conhecimento sobre os principais parâmetros responsáveis pelo desempenho do sistema de rastreamento por imagens: as larguras da janela de busca e de correlação.

Com base nos resultados obtidos através do estudo fatorial foi feita uma análise de sensibilidade nos dois parâmetros, obtendo-se assim valores “ótimos” para as condições de utilização do sistema robótico, abrangendo de forma simultânea um baixo tempo de processamento por imagem (da ordem de 55ms) e baixos valores de erro na localização dos pontos característicos das imagens, aspectos desejáveis para o sistema de controle do robô.

O processo de "escolha" dos valores ótimos através de gráfico pode ser automatizado facilmente através de técnica de análise de variância.

## REFERÊNCIAS

- [1] Trucco, E., Verri, A., *Introductory Techniques for 3-D Computer Vision*, Prentice-Hall, New Jersey, USA (1998)
- [2] Klette, R., Schlüns, K., A. Koschan, A., *Computer Vision: Three-Dimensional Data from Images*, Springer-Verlag, Singapore (1998)
- [3] Corke, P., Hutchinson, S.A., “Real-Time Vision, Tracking and Control”, *Proceedings of 2000 IEEE International Conference on Robotics & Automation*, San Francisco-CLA, USA (2000)
- [4] Kara, R., Wira, P., Kihl, H., “Robot Vision Tracking with with a Hierarchical CMAC Controller”, *Proceedings of Fourth International Conference on Knowledge-Based Intelligent Engineering Systems & Allied Technologies*, Brighton, UK (2000)
- [5] Santos-Victor, J., Sentiero, J., “A 3D Vision System for Underwater Vehicles: Na Extended Kalman-Bucy Filtering Approach”, *Acoustic Signal Processing for Ocean Exploration*, 567-580 (1993)
- [6] Kass, M., Witkins, A., Terzopoulos, D., “SNAKES: Active Contour Models”, *International Journal of Computer Vision*, **1**(4), 321-331 (1988)
- [7] Santos-Victor, J., Sandini, G., “Visual Behaviours for Docking”, *Computer Vision and Image Understanding*, **67**(3), 223-238 (1997)
- [8] Plakas, K., Trucco E., Fusiello, A., “Uncalibrated Vision for 3D Underwater Applications”, *Proceedings of IEEE Oceans '98*, 272-276, Nice, France (1998)
- [9] Plakas, K., Trucco, E., “Developing a real-time, robust, video tracker”, *Proceedings of OCEANS 2000 MTS/IEEE Conference and Exhibitions* **2**, 1345-1352 (2000)
- [10] Tomasi, C., Kanade, T., “Detection and tracking of point feature”, *Technical Report CMU-CS-91-132*, Carnegui-Melon University, Pittsburg, PA (1991)
- [11] Shi, J., Tomasi, C., “Good features to track”, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 593-600 (1994)
- [12] Espiau, B., “Effect of Camera Calibration Errors on Visual Servoing in Robotics”, *Proceedings of 3<sup>rd</sup> International Symposium on Experimental Robotics*, Kyoto, Japan (1993)
- [13] Spindler, F., Bouthemy, P., “Real-time estimation of dominant motion in underwater video images for dynamic positioning”, *Proceedings of IEEE International Conference on Robotics & Automation*, Louvain, Belgium, 1063-1068 (1998)
- [14] Coleman, D.F., Montgomery, D.C., “A systematic approach to planning for a designed industrial experiment”, *Technometrics*, **35**, 1-27 (1993)
- [15] Montgomery, D.C., *Design and analysis of experiments*, Wiley, New York (1991)
- [16] ISO Guide to the Expression of Uncertainty in Measurement (GUM), Geneva: International Organization for Standization (ISSO) (1995)
- [17] Anandan, P., Bergen, J.R., “Hierarchical model-based motion estimation”, *Proceeding of European Conference on Computer Vision*, Santa Margherita, Ligure, Italy (1992)

## **A ROBOT NAVIGATION VISUAL TRACKING SYSTEM OPTIMIZED BY USING EXPERIMENTAL FACTORIAL DESIGN**

**Abstract** - This article reports the development of an optimization scheme based on factorial design analysis applied to a vision tracking system embedded on a mobile robot aiming at speeding up the vision system responses for navigation. A review of factorial design is presented along with implementation details to identify and optimize sensor parameters. The method was applied to a model that uses SSD (Sum of Squared Differences) correlation for feature tracking and a gradient-based optical flow estimation for position calculation. Results show which are the most important parameters and the influence they have on time expense for image processing and position error estimation, including parameter interactions. A technique for finding optimal parameter values for each image is implemented and discussed.

**Keywords** – Robot Navigation, Visual Tracking, Factorial Design, Parameter Optimization.