

CALIBRAÇÃO DE IMAGENS ESTEREOSCÓPICAS DE OBJETOS EM 3 DIMENSÕES PARA GERAÇÃO AUTOMÁTICA DE TRAJETÓRIA DE MANIPULADORES ROBÓTICOS INDUSTRIAIS

MILENA F. PINTO¹, LINDOLPHO O. A. JUNIOR¹, LUIS C. G. LOPES¹,
AURÉLIO G. MELO¹.

1. *Centro Federal de Educação Tecnológica de Minas Gerais*
Rua José Peres 558 - Centro - Leopoldina - MG - Brasil

milena.pinto@member.isa.org, lindolpho@leopoldina.cefetmg.br,
gamboa@leopoldina.cefetmg.br, aureliokta@gmail.com

Abstract— This project aims to perform the calibration of images captured by a system composed of two set of stereoscopic vision system composed by two cameras each one, as part of a project to identify and analyze the trajectory of a standard object in 3D space. Also, future implementation of automatic generation for transmission and dynamics of an industrial robotic manipulator, guided by the trajectory of a standard object, is proposed. Calibration method is proposed and discussed. Results of calibration method are presented.

Keywords— Image Recognition, Calibration, Trajectory Identification, Matlab, OpenCV.

Resumo— Este projeto visa realizar a calibração de imagens capturadas por um sistema composto por dois conjuntos de câmeras estereoscópicas como parte de um projeto de identificação e análise da trajetória de um objeto padrão no espaço 3D. Este trabalho tem como principal objetivo à aplicação futura de geração automática da cadeia cinemática e dinâmica de um manipulador robótico industrial guiado pela trajetória do objeto e identificado pelo sistema que se propõe. É realizada a discussão dos métodos de calibração e apresentação dos resultados obtidos.

Palavras-chave— Reconhecimento de imagens, Calibração, Identificação de Trajetória, Matlab, OpenCV.

1 Introdução

Observando a descrição de Heinen (2002) pode-se constatar que das diversas áreas de pesquisa em robótica autônoma, existem tópicos de fundamental importância, como o planejamento de trajetória e navegação de robôs manipuladores.

Devido ao custo e imprecisão de métodos como identificação por rádio frequência e GPS, é crescente o uso de técnicas de visão computacional para este fim.

O emprego do processamento digital de imagens permite extrair e identificar informações das mesmas, além de melhorar a qualidade visual de certos aspectos estruturais, facilitando assim a percepção humana e a interpretação automática por meio de máquinas (Pedrini, 2008).

De acordo com Batchelor (1997) a utilização de processamento de imagens apresenta uma ampla gama de vantagens em relação a outros métodos de sensoriamento, pois permite uma grande flexibilidade na sua utilização.

A análise quantitativa e a interpretação de imagens representam atualmente um ponto de apoio importante em diversas disciplinas científicas. Tal é o caso, por exemplo, na ciência dos materiais, na biofísica, na medicina, na física da matéria condensada, etc. (Albuquerque, 2000).

Dentre as vantagens do sensoriamento através de imagens, pode-se citar que este método não é destru-

tivo, pode ser utilizado em tempo real, ser instalado em local de difícil acesso e não necessita de contato com o objeto de sensoriamento.

Conforme Filho (1999), a iluminação é uma das principais fontes de influência no processo de captura de imagens, sendo sua não uniformidade o principal foco de interferência. Desta forma, a viabilidade da utilização destas ferramentas pode ser associada à possibilidade do uso de mecanismos que permitam controlar a luminosidade no ambiente de detecção, destacando as regiões de interesse. (Pavim, 2005)

A aquisição das imagens está condicionada ao uso de lentes e sensores, os quais ocasionam deformações na imagem final. A fim de minimizar estas imperfeições, são utilizadas técnicas computacionais de calibração, permitindo então obter uma imagem fidedigna a realidade.

A calibração é necessária ainda ao processo de cálculo da posição de um objeto no espaço, onde se deve conhecer uma série de parâmetros da câmera e da imagem.

O objetivo desse trabalho é a calibração de imagens estereoscópicas no espaço 3D para um objeto padrão, utilizando o conceito de processamento digital de imagem, para que em trabalhos futuros sejam obtidos os vetores de posição instantânea para orientação de trajetórias. Pretende-se que este trabalho possa ser utilizado no cálculo de trajetórias para um manipulador robótico industrial.

2 Metodologia

Para identificação de trajetórias desenvolveu-se inicialmente um algoritmo para o reconhecimento de um objeto no espaço utilizando a biblioteca de visão computacional OpenCV (Intel, 2004).

Devido à dificuldade do reconhecimento de formas geométricas complexas, selecionou-se como alvo de reconhecimento objetos padrão na forma cilíndrica, pois a simplicidade de sua forma contribuirá neste momento, para estabelecer formas de análise futura de elementos complexos de estudo.

Em seguida, transferiu-se o algoritmo de reconhecimento para uma biblioteca do *software matlab®* (Matlab, 2005) a fim de tornar viável o reconhecimento de trajetórias. A biblioteca foi desenvolvida utilizando-se do recurso *mex* do *Matlab®*

A biblioteca *mex* é uma interface externa que permite a comunicação de linguagens de baixo nível, tais como C, C++ e Fortran, com o *Matlab®* (Matlab, 2011).

A opção pela utilização dessa ferramenta deu-se pela capacidade de execução de códigos C existentes como rotinas diretas do *Matlab®*. Realizou-se também, a interface dos aplicativos desenvolvidos a fim de melhorar os resultados obtidos.

O reconhecimento de trajetórias no espaço em 3D não envolve apenas a identificação da forma nas imagens, é preciso que seja feita uma tradução entre a posição virtual da imagem com sua posição no mundo real.

Para realizar essa tradução construiu-se um protótipo com a finalidade de produzir os testes e realizar a validação do modelo.

Este protótipo consiste de um servomecanismo acoplado a um disco com intuito de movimentar o objeto a ser rastreado dentro de uma trajetória conhecida. Desenvolveu-se ainda, uma interface de comunicação com o computador objetivando a tradução dos sinais de controle para o servomecanismo cujo protótipo, juntamente com a placa de captura, pode ser visto na Figura 1.

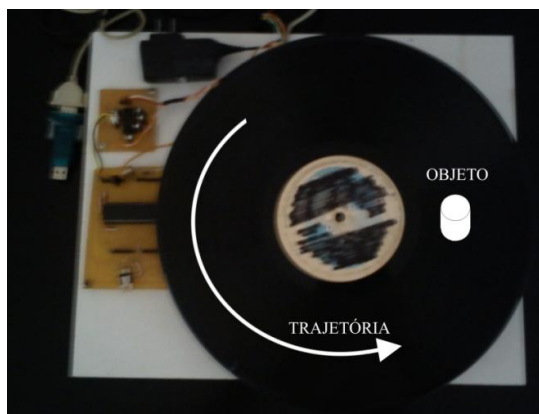


Figura 1. Protótipo de teste e validação do modelo em 2D.

Para reconstrução da localização 3D da imagem, serão utilizadas linhas de projeção obtidas

através de imagens capturadas por dois conjuntos de câmeras, cada um composto por um par de câmeras constituindo um sistema de visão estereoscópico, como pode ser observado na Figura 2.

Uma vez capturada a imagem pelo conjunto de câmeras 1, realiza-se o processo de localização do objeto para o plano α , permitindo determinar as coordenadas dos pontos de interesse (x,y) . Este processo será descrito na seção 4. A partir da imagem obtida pelo segundo conjunto de câmera é possível determinar a posição do mesmo objeto, agora para o plano β , (y,z) , tornando possível o cálculo da sua posição no espaço tridimensional.

Os conjuntos de câmeras 1 e 2 devem ser posicionados de forma que a posição de sua lente, seja totalmente paralela ao plano de captura.

Utilizando-se de marcações de referência é possível definir uma matriz de transformação que permitirá o cálculo das dimensões reais a partir do número de pixels medido, bem como o cálculo das distorções e erros de medição.

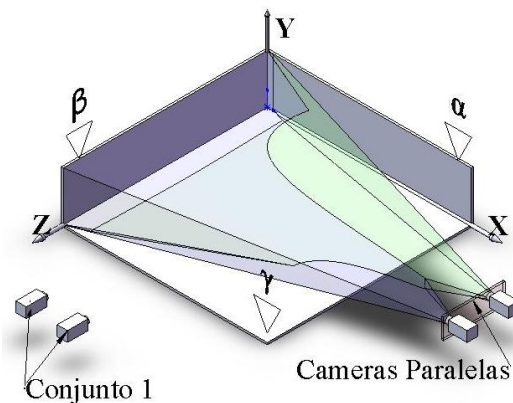


Figura 2. Linhas de Projeção 3D e localização real da imagem.

Um fator de grande influência na qualidade das medições é a resolução da câmera, ou seja, quanto maior a densidade de pixels, maior será a precisão das mesmas. (Panich, 2010)

É preciso considerar ainda que quanto mais distante o objeto encontra-se da objetiva, menor será a quantidade de pixels representando o mesmo, assim, há uma tendência ao acréscimo da taxa de erro com o aumento da distância do alvo. (Kytö et. All., 2011)

As imagens capturadas pelas câmeras possuem distorções devido a imperfeições nas lentes e nos sensores de captura, e assim, para minimizar estes efeitos, é necessário realizar um tratamento nas imagens capturadas. (Bradski, 2008)

Este tipo de processo é utilizado para obter os parâmetros da câmera empregada, como distância focal, além de outros dados necessários ao reconhecimento da posição da imagem no espaço tridimensional.

Uma vez realizado o tratamento, será utilizado o algoritmo de reconhecimento para encontrar o objeto cuja posição é desejada, e a partir desta e dos dados de calibração, calcular a localização real do objeto padrão.

3 Reconhecimento de Imagem

Utilizou-se um algoritmo de reconhecimento de imagem baseado na implementação da transformada de Hough encontrada na biblioteca OpenCV. Esta biblioteca oferece uma série de construções da transformada. Desta forma selecionamos o método que utiliza como entrada o conjunto de bordas da figura.

A transformada de Hough consiste em método que realiza a detecção de formas geométricas em imagens digitais, cujas definições matemáticas são conhecidas, tais como círculos e retas. (Chaves, 2007).

Contudo, para que a Transformada de Hough seja uma ferramenta utilizável, é necessário que exista um processamento inicial, extraindo as bordas dos objetos em questão. A detecção de bordas é uma eficiente ferramenta no processo de análise de imagens digitais. Sendo assim, o algoritmo selecionado para este fim, foi o de detecção de bordas de Canny, tal como sugerido em Duarte (2005).

Para interpretar as imagens de uma forma eficiente no processo de reconhecimento de imagens é preciso que seja realizada inicialmente uma suavização da imagem, além de um aumento do contraste.

Desenvolveu-se o algoritmo de reconhecimento na linguagem computacional C no software Visual Studio (Microsoft, 2005) de forma a permitir a geração de uma biblioteca utilizável no *Matlab*. Assim, o algoritmo retorna ao *Matlab* por um vetor contendo a posição do objeto, cuja trajetória deve ser reconhecida, bem como a cor e uma marca de tempo.

O *software Matlab* é responsável pela realização da captura da imagem a ser processada e enviá-la para o algoritmo de reconhecimento, recebendo os dados de retorno, e também, armazenando o histórico de posições para posterior geração de um vetor de trajetórias.

Na Figura 3 observam-se os dados dos círculos reconhecidos, bem como seus diâmetros, suas posições e cores do ponto central, respectivamente. Na janela inferior, estão apresentados os círculos reconhecidos e os pontos centrais marcados em verde.

```
>> um
circulos=2
0000 120.000x 98.000 26.000 (255,246,000)
0001 123.000x 152.000 26.420 (000,148,222)
d [0] [1] -> 54.083269
```

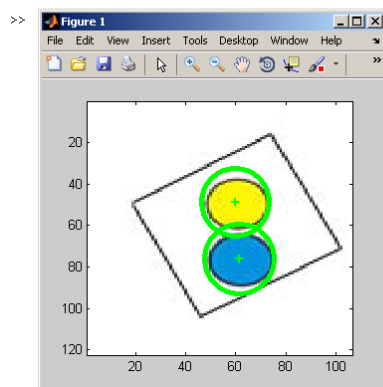


Figura 3. Resultado do Reconhecimento de Imagens

4 Triangulação

Nesta etapa será considerado o procedimento de triangulação descrito por DAVIES (2005).

Uma vez que as imagens já passaram pelo processo de calibração e não possuem maiores distorções, será utilizado o algoritmo de reconhecimento para localizar o mesmo ponto sobre as duas imagens.

A distância do objeto (Z) pode então ser calculada através da diferença entre as coordenadas x (esquerda x^l e direita x^r), da distância entre centro das câmeras T e distância para o plano de captura (f). Podemos observar a disposição dos valores na Figura 4. A coordenada Z pode ser calculada então a partir da Equação 1.

$$Z = \frac{fT}{x^l - x^r} \quad (1)$$

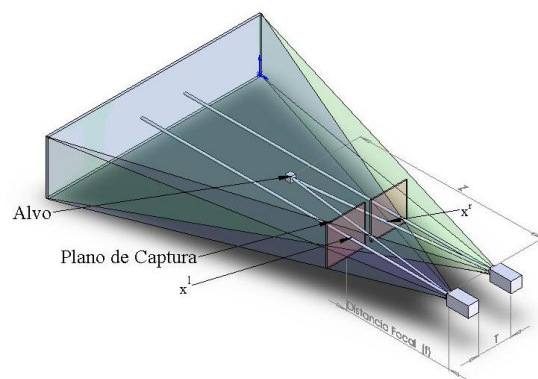


Figura 4. Cálculo da posição do ponto P a partir de duas imagens perfeitamente alinhadas e não distorcidas.

5 Calibração

As câmeras em geral tendem a produzir deformações nas imagens capturadas, sendo as bordas detentoras das distorções mais acentuadas. Estas deformações são causadas por imperfeições nas lentes e nos sensores de captura.

A distância do plano de captura da câmera tem ainda uma influência positiva sobre o nível de distorção da imagem, caso a resolução da câmera seja suficientemente grande para que haja uma boa densidade de pixels na representação do alvo, haverá uma redução do nível de erro em função do aumento da distância do ponto a ser obtido.

Para obter imagens com um mínimo adequado de deformações que permitam a precisão no processo de rastreamento do objeto no espaço 3D, é necessário realizar um processo de calibração da câmera.

Seguindo o processo proposto por Bradski (2008), o primeiro passo do método de calibração é a remoção das distorções radiais e tangenciais existentes nas imagens capturadas.

A distorção radial provoca deformações nas bordas da imagem causando assim, seu arredondamento. Já a distorção tangencial, tende a provocar deformações nos tamanhos reais devido ao fato da

câmera não estar totalmente paralela ao plano de captura.

O segundo passo deste processo é retificar a imagem obtida. Isto envolve o procedimento de ajuste das imagens capturadas pelas duas câmeras, para que então possam se encaixar perfeitamente.

Realizou-se o processo de captura da posição real conforme o fluxograma da Figura 5, e na Figura 6, é feito o detalhamento do processo de calibração.

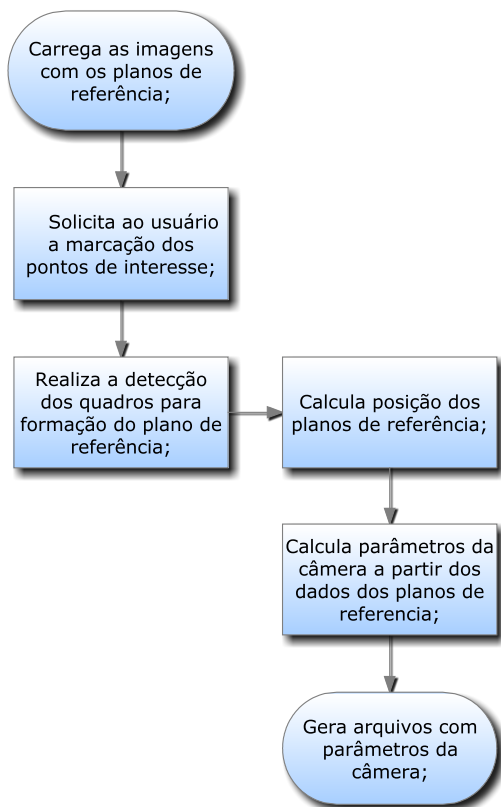


Figura 5. Procedimento de Calibração

Em tese, qualquer imagem com características conhecidas, pode ser utilizada para formação dos planos de referência conforme indicado em Borchart (2010). Devido à facilidade de reconhecimento das formas e dos padrões regulares tabuleiros de xadrez se apresentam como uma boa escolha.

Assim, conforme o fluxograma do processo de calibração apresentado na Figura 5, é realizada a leitura das imagens dos tabuleiros de xadrez e posteriormente solicitado ao usuário que realize a marcação das arestas das casas do tabuleiro.

Posteriormente no processo de cálculo da posição 3D, será necessário obter uma matriz de rotação e translação que permita calcular a posição real dos pontos no plano em que os mesmos foram capturados a partir da imagem obtida. Os coeficientes da matriz foram avaliados a partir da posição de origem dos pontos e das distâncias X e Y entre eles.

Para realização do procedimento de calibração utilizou-se o método descrito por Hartley e Kang (2007). Este processo consiste na obtenção de imagens em diversos planos de referência, e a partir das

mesmas, são extraídas características conhecidas para que sejam calculados os parâmetros das câmeras, e além do mais, destaca-se o fato de que o resultado apresenta boa precisão em relação a outros métodos.

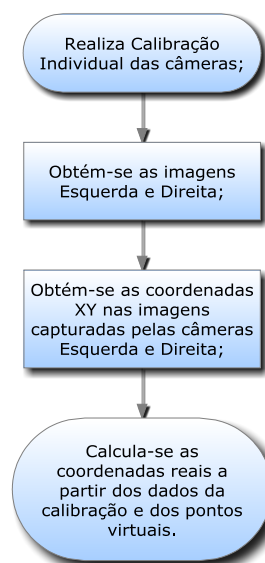


Figura 6. Procedimento de Reconhecimento do Ponto no Espaço

Como desenvolvimentos futuros pretende-se realizar o cálculo da posição real do objeto no espaço, tal como o modelo cinemático inverso do mesmo.

A posição real do objeto pode ser facilmente calculada a partir dos pontos referenciais, e também, da distância focal da imagem, como pode ser visto na Figura 4. A distância será calculada por triangulação, baseada em medidas de distância Euclidiana.

O processo foi representado apenas para o plano α , porém, sua aplicação pode ser estendida para os planos β e γ e assim, é possível determinar o ponto relativo a cada coordenada espacial.

6 Modelo cinemático Inverso

As coordenadas estabelecidas para o objeto padrão serão usadas em um modelo cinemático obtido pela matriz Jacobiana, para geração automática de trajetória de um robô manipulador, conforme Spong, (2005). O método proposto visa obter a variável de espaço \vec{r} , o qual é capaz de representar completamente uma junta robótica. Inicialmente escreve-se a relação vetorial para o espaço cartesiano e para as juntas do manipulador conforme Equação 1.

$$\vec{r} = \vec{F} \cdot (\vec{q}) \quad (1)$$

Considerando a variação ao longo do tempo de uma variável no espaço tem-se o exposto na Equação 2.

$$\vec{r} = [x \ y \ z \ \phi \ \theta \ \varphi]^t \quad (2)$$

Assim, as componentes cinemáticas diferenciais desse vetor \vec{r} podem ser dadas pela Equação 3.

$$\frac{d\vec{r}}{dx} = \left[\frac{dx}{dx} \quad \frac{dy}{dx} \quad \frac{dz}{dx} \quad \frac{d\theta}{dx} \quad \frac{d\varphi}{dx} \right]^t \quad (3)$$

A partir deste ponto, identificam-se as velocidades lineares e angulares. A relação das mesmas com as juntas do manipulador é dada pela diferenciação de funções vetoriais, onde J é o Jacobiano da função vetorial \vec{F} como exposto pela Equação 4.

$$d\vec{r} = J \cdot d\vec{q} \quad (4)$$

Portanto, assumindo n juntas em variáveis cartesianas, pode-se obter analiticamente a relação entre as diferenciais cartesianas e as diferenciais das juntas conforme representado na matriz da Equação 5.

$$J = \begin{bmatrix} \frac{\partial r_1}{\partial q_1} & \frac{\partial r_1}{\partial q_2} & \dots & \frac{\partial r_1}{\partial q_n} \\ \frac{\partial r_2}{\partial q_1} & \frac{\partial r_2}{\partial q_2} & \dots & \frac{\partial r_2}{\partial q_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial r_m}{\partial q_1} & \frac{\partial r_m}{\partial q_2} & \dots & \frac{\partial r_m}{\partial q_n} \end{bmatrix} \quad (5)$$

7 Resultados

No processo de captura estereoscópica utilizou-se duas câmeras idênticas Logitech QuickCam, com resolução de 1024 x 768 pixels, a fim de que, posteriormente, possa ser realizado o processo de cálculo da posição real. Realizou-se o procedimento de forma que as imagens dos planos fossem obtidas de forma praticamente simultânea.

As câmeras foram calibradas individualmente com o objetivo de calcular os parâmetros de cada uma. Na Figura 7 temos o resultado do algoritmo de calibração antes e após o processo de remoção das distorções.

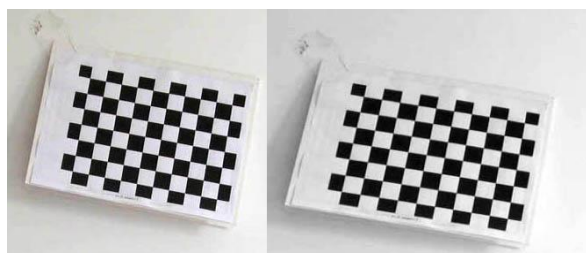


Figura 7. Imagens antes e após o processo de calibração

Na Tabela 1 os resultados do processo de calibração para as câmeras são apresentados, além do erro médio percentual de cada parâmetro.

Durante o processo de calibração percebeu-se que o erro está fortemente ligado à precisão em que são reconhecidas as arestas dos quadrados, bem como ao foco da imagem.

Um alto número de planos capturados pode não fornecer um erro médio baixo, caso o reconhecimento das arestas não seja realizado corretamente.

Tabela 1. Parâmetros da Câmera

Parâmetro	Esquerda	Direita	Erro (%)
Distância Focal X	1380.8	1272.6	5,18
Y	1323.6	1356.7	
Centro Ótico X	614.9	594.9	4,8
Y	318.6	321.2	
Erro de Conversão X	0.14	0.21	4,13
Y	0.19	0.20	

No processo de calibração observou-se ainda que a distorção da imagem é diretamente proporcional a distância de captura, e assim, quanto mais distantes os planos de calibração, maiores serão as taxas de erro.

De forma a mensurar o desempenho do algoritmo foi realizada a medição do tempo de execução. O processo foi realizado com uma média de 0,44 segundos para cada imagem. Isto indica que o algoritmo pode ser bem aplicado em ambientes cujos requisitos de desempenho não sejam extremamente rígidos.

Observou-se, ainda, que o erro de conversão dos pixels apresentou valores baixos, menores que 0,5 pixels em cada direção.

8 Conclusão

O algoritmo de reconhecimento de imagens desenvolvido mostrou-se eficaz para aplicação em sistemas de identificação de objetos em movimento.

Observa-se que a capacidade de reconhecimento está diretamente ligada a calibração das imagens capturadas.

Através dos resultados obtidos é possível afirmar que o processo de calibração é funcional e preciso, e os dados medidos permitirão o cálculo de posição para objetos padrão, em espaços cartesianos 3D.

Agradecimentos

Os autores agradecem ao apoio financeiro do CNPQ a este trabalho, através de Bolsa PIBITI. Também agradecem ao CEFET-MG e ao Grupo de Pesquisa em Controle, Automação e Energia.

Referências Bibliográficas

- Albuquerque, M. P. (2000). Processamento de Imagens: Métodos e Análises. Centro Brasileiro de Pesquisas Físicas – CBPF/MCT.
- Batchelor, B. G., Whelan, P. F. (1997). Intelligent Vision Systems for Industry. 1ª Ed. Springer.
- Borchardt, T. B. (2010) Metodologia Baseada em Warping para Correção de Distorções em Siste-

- mas de Endoscopia. Dissertação de Mestrado. Santa Maria, RS, Brasil, 2010.
- Bradski, G., Kaehler A. (2008) Learning OpenCV. OREILLY.
- Chaves, R. F. (2007). “Uma proposta de um sistema de reconhecimento de íris humana”. FEEC, UNICAMP, Campinas - São Paulo.
- Davies, E. R. (2005). “Machine Vision: Theory, Algorithms e Practicalities.”. Elsevier; Londres.
- Duarte, G. D. (2005) “Uso da Transformada de Hough na Detecção de Círculos em Imagens Digitais”. IFSUL, Pelotas.
- Filho, M. O., Neto, H. V. (1999). “Processamento Digital de Imagens.” BRASPORT, Rio de Janeiro – RJ, Brasil. ISBN 8574520098.
- Hartley R., Kang S. B. (2007). Parameter – Free Radial Distortion Correction with Center Distortion. IEEE.
- Heinen, F. J. (2002). Sistema Híbrido para Rôbos Móveis Autônomos. MCA/Unisinos; São Leopoldo – RS, Brasil.
- Intel. (2004). Open Source Computer Vision Library. In: www.intel.com/research/mrl/research/opencv
- Matlab. (2005). The language of Technical Computing. In: www.mathworks.com/products/matlab
- Matlab. (2011). Mex-files Guide. In: www.mathworks.com/support/tech-notes/1600/1605.html#intro
- Microsoft Visual Studio. (2005). Visual Studio 2005. In: msdn.microsoft.com/pt-br/vstudio/default
- Kytö, M., Nuutinen M., Oittinen P., (2011) Method for measuring stereo camera depth accuracy based on stereoscopic vision. Aalto University School of Science and Technology, Department of Media Technology, Otaniementie, Finland.
- Pavim, X. A. (2005), Contribuições na otimização de um sistema de visão para detecção, medição e classificação automática do desgaste de ferramentas de corte em processos de usinagem. UFSC; Florianópolis – SC, Brasil.
- Pedrini, H., Schwartz, W.R. (2000). Análise de imagens digitais: Princípios, algoritmos e aplicações. 1ª Ed. Thomson.
- Panich, S. (2010). Comparison of Distance Measurement Between Stereo Vision and Ultrasonic Sensor. Department of Mechanical Engineering. Srinakharinwirot University. Bangkok, Thailand.
- Spong, M. W. et. all. (2005). Robot Modeling and Control. Ed. Wiley.