

# Aperfeiçoamento do método de calibração monocular baseada em gabaritos de uma única dimensão

José A. de França<sup>1</sup>, Elaine G. Alves<sup>1\*</sup>, Robinson Hoto<sup>2</sup>

<sup>1</sup>Departamento de Engenharia Elétrica - Universidade Estadual de Londrina  
Caixa Postal 6025 - 86051-990, Londrina, PR, Brasil

<sup>2</sup>Departamento de Matemática - Universidade Estadual de Londrina  
Caixa Postal 6001 - 86051-990, Londrina, PR, Brasil

josealexandre@eeol.org, elainegracindo@gmail.com, hoto@uel.br

**Resumo.** *Recentemente, um novo método de calibração de câmeras foi proposto. Tal método usa um gabarito de uma única dimensão (um bastão com algumas esferas) para determinar os parâmetros intrínsecos da câmera. Apesar de bastante flexível, o método é muito susceptível a ruídos, inviabilizando o seu uso na prática. No presente trabalho, propõe-se uma normalização nos dados de entrada que torna o algoritmo muito mais robusto. Tal normalização não aumenta a complexidade computacional do algoritmo e torna a sua exatidão muito próxima a de outros métodos não-lineares encontrados na literatura. Resultados experimentais em dados sintéticos e imagens reais mostram que as modificações propostas aumentam muito a eficiência do método, tornando-o viável para usos práticos.*

**Palavras-chave:** Calibração de câmeras, Gabarito de calibração, Objeto 1D.

## 1. Introdução

A visão computacional é muito utilizada em aplicações, tais como, automação industrial, robótica e engenharia biomédica. Com auxílio de uma câmera, sem nenhum contato físico, é possível extrair uma variedade de informações de uma cena como, por exemplo, forma, textura, cor, quantidade, dimensão, etc. Quando a informação que se deseja extrair de uma imagem envolve algum tipo de grandeza geométrica (comprimento, ângulos, distâncias, etc), é necessário conhecer exatamente os parâmetros utilizados pela câmera para transformar um ponto no espaço 3D em um ponto no plano da imagem. Tais parâmetros podem ser conhecidos através de um processo chamado de calibração.

O processo de calibração consiste em utilizar a câmera para capturar imagens de um objeto de dimensões e forma bem conhecidos (chamado de gabarito de calibração). Em seguida, utiliza-se a relação entre pontos do gabarito e suas respectivas projeções no plano de imagem para determinar os parâmetros da câmera.

Inicialmente, os algoritmos de calibração utilizavam apenas gabaritos 3D [Tsai 1987]. Geralmente, tais gabaritos são compostos por dois planos ortogonais com um padrão bem conhecido (composto por quadrados ou círculos) em suas faces. Com o passar dos anos, foram propostos novos métodos de calibração que utilizam gabaritos 2D [Agrawal and Davis 2003, Zhang 2000], ou seja, um único plano com um padrão

---

\*A autora agradece a CAPES e a PROPPG-UEL pelo financiamento das pesquisas.

conhecido em uma das faces. Mais recentemente, alguns algoritmos de calibração propostos utilizam objetos de uma única dimensão (pontos em uma reta) para calibrar sistemas de visão que utilizem uma única câmera [Zhang 2004] ou um conjunto binocular [de França and Stemmer 2006]. A principal vantagem desta abordagem é que um objeto 1D pode ser “visto” por mais de uma câmera (ou conjunto binocular) ao mesmo tempo. Isto torna possível, por exemplo, calibrar de uma só vez o sistema de visão de um time de robôs que estejam posicionados em torno do gabarito.

No artigo que deu origem à teoria de calibração com gabaritos de uma dimensão, Zhang (2004) propôs um algoritmo não-linear para resolver o problema de calibração. Tal algoritmo necessita de uma estimativa inicial dos parâmetros da câmera. Esta estimativa pode ser obtida com auxílio de um método linear proposto no mesmo artigo. Infelizmente, esse método linear é bastante susceptível a ruídos normalmente presentes na projeção dos pontos do gabarito. Isto leva a uma estimativa inexata dos parâmetros da câmera, pois, se for mal inicializado, o algoritmo não-linear pode “ficar preso” em um mínimo local. Dessa forma, o método proposto por Zhang (2004) é importante para a área de calibração de câmeras apenas no aspecto teórico.

No presente trabalho, demonstra-se empiricamente que a exatidão do algoritmo de calibração linear proposto por Zhang pode ser bastante melhorada levando-se em conta algumas considerações numéricas, mais especificamente, o condicionamento do conjunto de equações lineares que deve ser resolvido. Uma simples transformação nas projeções dos pontos do gabarito antes da solução deste conjunto de equações aumenta a estabilidade do método significativamente, pois torna o problema melhor condicionado. Além disso, a complexidade computacional inserida no algoritmo pela transformação proposta é desprezível.

## 2. Notação

No decorrer do texto, matrizes e vetores são representados por letras, números ou símbolos em negrito. Constantes são expressas por letras, números ou símbolos em itálico. Além disso, adotou-se a prática notação  $\mathbf{A}^{-T} = (\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$  para toda matriz quadrada e inversível.

Considerando o modelo de câmera *pinhole* [Faugeras and Luong 2001], as coordenadas de um ponto 3D no sistema de coordenadas do ambiente de uma câmera são apresentadas como  $\mathbf{M} = [x, y, z]^T$  e a projeção correspondente no plano de imagem  $I$ , como  $\mathbf{m} = [u, v]^T$ . Além disso, as coordenadas homogêneas de um ponto  $\mathbf{m} = [u, v, \dots]^T$  são representadas por  $\tilde{\mathbf{m}}$ , isto é,  $\tilde{\mathbf{m}} = [u, v, \dots, 1]^T$ . De uma forma mais geral, um ponto qualquer em coordenadas homogêneas é representado por  $\tilde{\mathbf{m}} = [u, v, \dots, t]^T$ . Um índice, se houver, indica a posição do ponto em um conjunto de pontos.

Com a notação adotada, a relação entre um ponto 3D,  $\mathbf{M}$ , e sua projeção,  $\mathbf{m}$ , em uma câmera *pinhole* é dada por

$$\tilde{\mathbf{m}} \simeq \mathbf{A} [\mathbf{I}_3 \mathbf{0}_3] \tilde{\mathbf{M}}, \quad (1)$$

com

$$\mathbf{A} = \begin{bmatrix} \alpha & s_k & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (2)$$

onde “ $\simeq$ ” indica que os dois lados da equação podem diferir por uma constante,  $\alpha$  e  $\beta$  dão a relação entre pixels da imagem e distâncias no ambiente da câmera, respectivamente, nas direções horizontal e vertical. Já o parâmetro  $s_k = f \cot \theta_{sk}$  está relacionado ao ângulo,  $\theta_{sk}$ , que corrige a ortogonalidade da matriz destes sensores. O ponto  $\mathbf{m}_0 = [u_0, v_0]^T$  são as coordenadas do ponto central da câmera.

O estado da arte da fabricação de câmeras permite desenvolver câmeras com  $\theta_{sk}$  muito próximo de  $\pi/2$ . De fato, normalmente,  $\|\pi/2 - \theta_{sk}\|$  é da ordem  $10^{-6}$ . Por isto, neste trabalho, considera-se  $s_k = 0$ . Dessa forma, há apenas quatro parâmetros intrínsecos a serem estimados, ou seja,  $\alpha$ ,  $\beta$ ,  $u_0$  e  $v_0$ .

### 3. Descrição do Método

O gabarito 1D utilizado na calibração deve ser algo como um bastão (ou mesmo um cordão) com diversas esferas (indexadas por  $j$ ) ao longo de sua extensão. A distância entre esferas consecutivas deve ser conhecida com exatidão. Então, no processo de calibração, o gabarito deve deslocar-se no campo visual da câmera. Enquanto isso, uma sequência de imagens (indexada por  $i$ ) deve ser capturada<sup>1</sup>. Zhang (2004) demonstrou que a calibração monocular com este tipo gabarito só é possível se duas condições forem satisfeitas: o gabarito conter três ou mais pontos e um desses pontos,  $\mathbf{M}_1$ , deve ser fixo, como mostra o esboço da figura 1. Dessa forma, a projeção deste ponto,  $\mathbf{m}_1$ , é a mesma em todas as imagens da sequência.

No esquema da figura 1, o comprimento,  $L$ , do gabarito pode ser expresso em função das coordenadas dos pontos extremos  $\mathbf{M}_1$  e  $\mathbf{M}_{2i}$ , ou seja,

$$\|\mathbf{M}_{2i} - \mathbf{M}_1\| = L. \quad (3)$$

Além disso, desde que as posições relativas das esferas são conhecidas, um ponto do gabarito localizado entre  $\mathbf{M}_1$  e  $\mathbf{M}_{2i}$  pode ser expresso por

$$\mathbf{M}_{ji} = \lambda_{1j}\mathbf{M}_1 + \lambda_{2j}\mathbf{M}_{2i}, \quad (4)$$

onde  $\lambda_{1j}$  e  $\lambda_{2j}$  são conhecidos.

Da equação (1), considerando a profundidade (desconhecida) do ponto  $\mathbf{M}_{ji}$  igual a  $z_{ji}$ , tem-se

$$\mathbf{M}_1 = z_1 \mathbf{A}^{-1} \tilde{\mathbf{m}}_1, \quad (5)$$

$$\mathbf{M}_{2i} = z_{2i} \mathbf{A}^{-1} \tilde{\mathbf{m}}_{2i}, \quad (6)$$

$$\mathbf{M}_{ji} = z_{ji} \mathbf{A}^{-1} \tilde{\mathbf{m}}_{ji}. \quad (7)$$

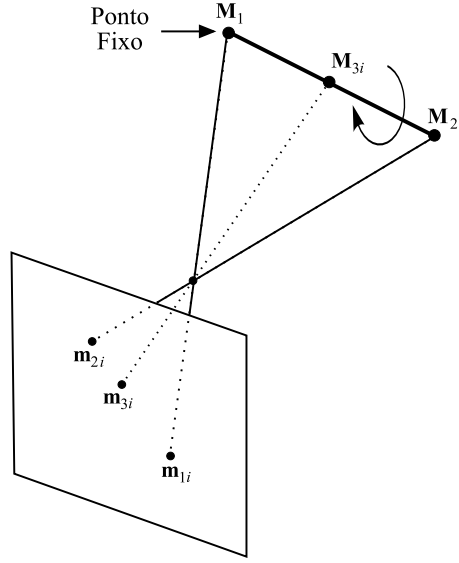
Dessa forma, substituindo as equações anteriores em (4), obtém-se

$$z_{ji} \tilde{\mathbf{m}}_{ji} = z_1 \lambda_{1j} \tilde{\mathbf{m}}_1 + z_{2i} \lambda_{2j} \tilde{\mathbf{m}}_{2i} \quad (8)$$

após eliminar-se  $\mathbf{A}^{-1}$  de ambos os lados.

---

<sup>1</sup>Como explicado ao longo da seção, cada uma das imagens capturadas fornece três equações independentes necessárias a estimação dos parâmetros desconhecidos.



**Figura 1. Esboço do gabarito 1D utilizado na calibração.**

Agora, aplicando-se o produto cruzado em ambos os lados da equação anterior com  $\tilde{\mathbf{m}}_{ji}$ , tem-se

$$z_1 \lambda_{1j} (\tilde{\mathbf{m}}_{1i} \times \tilde{\mathbf{m}}_{ji}) + z_{2i} \lambda_{2j} (\tilde{\mathbf{m}}_{2i} \times \tilde{\mathbf{m}}_{ji}) = \mathbf{0} \quad (9)$$

que pode ser escrita como

$$z_{2i} = -z_1 \frac{\lambda_{1j} (\tilde{\mathbf{m}}_{1i} \times \tilde{\mathbf{m}}_{ji}) \cdot (\tilde{\mathbf{m}}_{2i} \times \tilde{\mathbf{m}}_{ji})}{\lambda_{2j} (\tilde{\mathbf{m}}_{2i} \times \tilde{\mathbf{m}}_{ji}) \cdot (\tilde{\mathbf{m}}_{2i} \times \tilde{\mathbf{m}}_{ji})}. \quad (10)$$

Substituindo (5) e (6) em (3) e considerando  $z_{2i}$  dado por (10), obtém-se

$$z_1 \|\mathbf{A}^{-1} \mathbf{h}_{ji}\| = L, \quad (11)$$

onde

$$\mathbf{h}_{ji} = \tilde{\mathbf{m}}_1 + \frac{\lambda_{1j} (\tilde{\mathbf{m}}_1 \times \tilde{\mathbf{m}}_{ji}) \cdot (\tilde{\mathbf{m}}_{2i} \times \tilde{\mathbf{m}}_{ji})}{\lambda_{2j} (\tilde{\mathbf{m}}_{2i} \times \tilde{\mathbf{m}}_{ji}) \cdot (\tilde{\mathbf{m}}_{2i} \times \tilde{\mathbf{m}}_{ji})} \tilde{\mathbf{m}}_{2i}. \quad (12)$$

A equação (11) é equivalente a

$$\mathbf{h}_{ji}^T \mathbf{B} \mathbf{h}_{ji} = L^2, \quad (13)$$

onde, desde que  $s_k = 0$ ,

$$\mathbf{B} = z_1^2 \mathbf{A}^{-T} \mathbf{A}^{-1} = \begin{bmatrix} B_{11} & 0 & B_{13} \\ 0 & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix}. \quad (14)$$

Levando em consideração que

$$\mathbf{h}_{ji}^T \mathbf{B} \mathbf{h}_{ji} = [a_{ji}^2, b_{ji}^2, 2a_{ji}c_{ji}, 2b_{ji}c_{ji}, c_{ji}^2]^T \mathbf{b} = \mathbf{u}_{ji} \mathbf{b}, \quad (15)$$

com  $\mathbf{h}_{ji} = [a_{ji}, b_{ji}, c_{ji}]^T$  e  $\mathbf{b} = [B_{11}, B_{22}, B_{13}, B_{23}, B_{33}]^T$ , a equação (13) pode ser reescrita como

$$\mathbf{u}_{ji}^T \mathbf{b} = L^2. \quad (16)$$

Com  $n$  imagens, tem-se  $\mathbf{U}_n = [\mathbf{u}_{j1}, \mathbf{u}_{j2}, \dots, \mathbf{u}_{jn}]^T$ . Dessa forma, considerando  $\mathbf{L}^2 = [L^2, \dots, L^2]^T$ , pode-se encontrar  $\mathbf{b}$  resolvendo-se

$$\mathbf{U}_n \mathbf{b} = \mathbf{L}^2, \quad (17)$$

ou seja,

$$\mathbf{b} = (\mathbf{U}_n^T \mathbf{U}_n)^{-1} \mathbf{U}_n^T \mathbf{L}^2. \quad (18)$$

Dado um gabarito com mais de três pontos, tem-se um número maior de equações. Contudo, para cada imagem, apenas uma equação é linearmente independente. Assim, desde que existem 5 incógnitas (os 4 parâmetros de  $\mathbf{A}$  e  $z_1$ ), são necessários, no mínimo, 5 deslocamentos do gabarito para solucionar o problema.

Uma vez que  $\mathbf{B}$  for conhecida, a matriz  $z_1 \mathbf{A}^{-1}$  pode ser obtida de  $\mathbf{B}$  através da decomposição de Cholesky [Golub and Van Loan 1996]. Por sua vez, com  $z_1$  e  $\mathbf{A}$  conhecidos, o ponto  $\mathbf{M}_1$  pode ser obtido a partir de (5) e os pontos  $\mathbf{M}_{2i}$  a partir de (10) e (6). Por último, pode-se usar (4) para obter-se  $\mathbf{M}_{ji}$ .

O método descrito nesta seção não fornece bons resultados na presença de ruído. Assim, Zhang (2004) definiu um critério de otimização que envolve as projeções dos pontos do bastão,  $\mathbf{m}_1$ ,  $\mathbf{m}_{2i}$  e  $\mathbf{m}_{2ji}$ , ou seja,

$$\sum_{i=1}^n \left( \|\mathbf{m}_{1i} - \hat{\mathbf{m}}_1(\mathbf{A}, \mathbf{M}_1)\|^2 + \sum_{j=2}^q \|\mathbf{m}_{ji} - \hat{\mathbf{m}}_{ji}(\mathbf{A}, \mathbf{M}_{ji})\|^2 \right), \quad (19)$$

onde  $n$  é o número de imagens capturadas do gabarito,  $q$  é o número de pontos do mesmo e  $\hat{\mathbf{m}}_{ji}(\mathbf{A}, \mathbf{M}_{ji})$  é a estimativa da projeção de  $\mathbf{M}_{ji}$  de acordo com (1).

Na equação anterior, as incógnitas a serem estimadas são: os 4 parâmetros intrínsecos de  $\mathbf{A}$ , o ponto  $\mathbf{M}_1$  e  $2n$  parâmetros,  $\theta_i$  e  $\phi_i$ , que definem as coordenadas esféricas dos pontos  $\mathbf{M}_{ij}$  de acordo com

$$\mathbf{M}_{2i} = \mathbf{M}_1 + L[\sin \theta_i \cos \phi_i, \sin \theta_i \sin \phi_i, \cos \theta_i]^T, \quad (20)$$

onde  $\theta_i$  e  $\phi_i$  são ângulos que definem a orientação do gabarito.

Como discutido, a minimização da equação (19) envolve a solução de um problema de minimização não-linear. Normalmente, faz-se necessário uma estimativa inicial de todas as incógnitas envolvidas no problema. Evidentemente, para isto, o método linear apresentado pode ser utilizado.

#### 4. Aperfeiçoamento do Método

Na visão computacional, as coordenadas das projeções são dadas em relação ao canto superior-esquerdo da imagem. Aplicando uma transformação projetiva,  $\mathbf{T}_2$ , aos pontos,  $\tilde{\mathbf{m}}$ , da imagem, tem-se  $\tilde{\tilde{\mathbf{m}}} = \mathbf{T}_2 \tilde{\mathbf{m}}$ . Dessa forma, da equação (1), obtém-se  $\tilde{\tilde{\mathbf{m}}} \simeq$

$\mathbf{T}_2 \mathbf{A} [\mathbf{I} \ 0] \widetilde{\mathbf{M}}$ . Esta relação implica que  $\widehat{\mathbf{A}} = \mathbf{T}_2 \mathbf{A}$  representa os parâmetros intrínsecos da câmara que correspondem as projeções  $\widetilde{\mathbf{m}}$ .

Como discutido na seção anterior, para o caso da calibração com um gabarito 1D, tem-se a equação (11) que leva a (17). De forma semelhante, considerando os pontos  $\widetilde{\mathbf{m}}$  e a matriz  $\widehat{\mathbf{A}}$ , obtém-se um sistema de equações da forma  $\widehat{\mathbf{U}}_n \widehat{\mathbf{b}} = \mathbf{L}^2$ . Além disso, é fácil demonstrar que  $\widehat{\mathbf{U}}_n = \mathbf{U}_n \mathbf{S}$ , onde  $\mathbf{S}$  é uma matriz  $4 \times 4$  que depende apenas dos elementos de  $\mathbf{T}_2$ .

Com o resultado anterior, seria de se esperar que o problema de calibração com um gabarito 1D chegasse a exatamente a mesma solução caso fossem realizados os passos a seguir.

1. Transformar os pontos fazendo  $\widetilde{\mathbf{m}}_{ji} = \mathbf{T}_2 \widetilde{\mathbf{m}}_{ji}$ .
2. Encontrar  $\widehat{\mathbf{b}}$ , resolvendo (18) e utilizando os pontos  $\widetilde{\mathbf{m}}_{ji}$ .
3. Fazer  $\mathbf{b} = \mathbf{S} \widehat{\mathbf{b}}$ .

Contudo, isto seria verdadeiro se e somente se  $\widehat{\mathbf{b}} = \mathbf{S}^{-1} \mathbf{b}$  fosse solução do problema

$$\min_{\mathbf{x}} \|(\mathbf{U}_n \mathbf{S}) \mathbf{x} - \mathbf{L}^2\|.$$

No entanto, baseado nos resultados experimentais da próxima seção, isto não é verdadeiro. Dessa forma, a solução do problema de calibração com um gabarito 1D é extremamente dependente da matriz  $\mathbf{T}_2$ . Sendo assim, é necessário encontrar a transformação  $\mathbf{T}_2$  que forneça a melhor estimativa do vetor  $\mathbf{b}$  e, por tanto, da matriz  $\widehat{\mathbf{A}}$ . Diversos autores já propuseram soluções para este problema [Izquierdo and Guerra 2003, Torr 1995]. Em particular, neste trabalho utilizou-se a transformação proposta por Hartley (1997), ou seja,  $\mathbf{T}_2$  deve ser tal que o centróide do conjunto de pontos  $\mathbf{m}_{ji}$  esteja na origem e a distância média de um ponto à origem seja  $\sqrt{2}$ . Esta mudança produz melhora significativa na exatidão da matriz  $\mathbf{A}$ . Além disso, o custo computacional acrescentado ao problema devido a esta transformação é desprezível.

Uma vez que os parâmetros intrínsecos da câmara tenham sido estimados no sistema de coordenadas definido por  $\mathbf{T}_2$ , a matriz  $\mathbf{A}$  (do sistema de coordenadas original) pode ser obtida fazendo

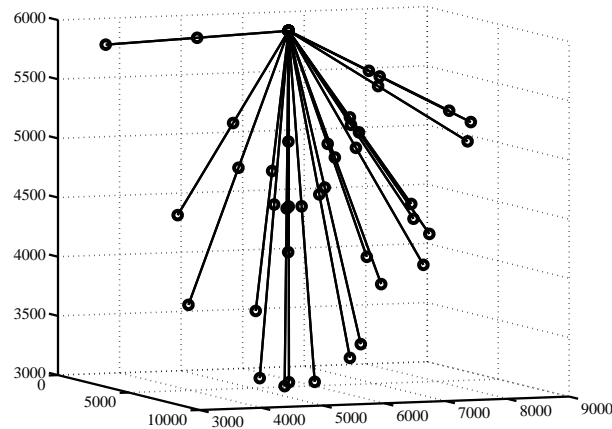
$$\mathbf{A} = \mathbf{T}_2^{-1} \widehat{\mathbf{A}}. \quad (21)$$

## 5. Resultados Experimentais

Na avaliação do método proposto, foram utilizados dados sintéticos e dados adquiridos de imagens reais. Os resultados obtidos com cada tipo de dado são discutidos nas seções a seguir.

### 5.1. Dados Sintéticas

Os dados sintéticos foram obtidos supondo uma câmara com os seguintes parâmetros:  $\alpha = 842$ ;  $\beta = 879$ ;  $s_k = 0$ ;  $u_0 = 358$ ;  $v_0 = 207$ . A resolução da imagem é feita igual a  $640 \times 512$  pixels. Além disso, foi simulado um bastão de 30 cm de comprimento e três pontos co-lineares e equidistantes. Em tal simulação, 20 imagens sintéticas foram geradas. Entre uma imagem e outra, os ângulos da equação (20),  $\theta_i \in [-\pi/2, \pi/2]$  e



**Figura 2. Esboço do movimento realizando pelo gabarito 1D de três pontos durante uma simulação.**

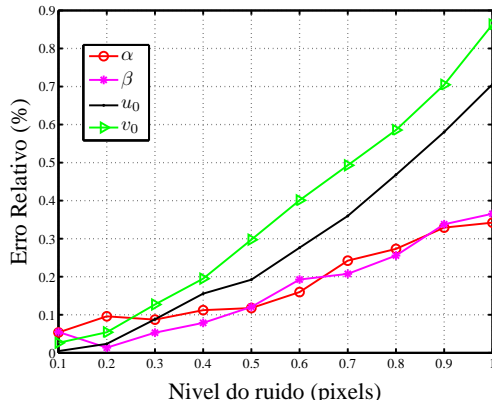
$\phi_i \in [-\pi/2, \pi/2]$ , variaram aleatoriamente mas de acordo com uma distribuição uniforme (veja-se a figura 2 para um exemplo dos dados simulados).

Os dados obtidos na simulação foram utilizados para avaliar o desempenho do método com respeito ao nível de ruído presente nos dados. Para isto, ruído gaussiano de média zero e desvio padrão  $\sigma$  foi acrescentado aos pontos projetados nas imagens sintéticas. Tal ruído variou de 0,1 a 1,0 pixel. Para cada nível de ruído foram realizadas 250 simulações e a mediana de cada parâmetro intrínseco foi retido. Essas medianas foram comparadas com os parâmetros da câmera simulada. Em cada simulação, os parâmetros intrínsecos da câmera foi estimado pelo método de Zhang (2004) e pelo proposto no presente artigo. Os resultados da comparação são apresentados nas figuras 3 e 4. Nestas, os erros são apresentados de forma relativa (em porcentagem).

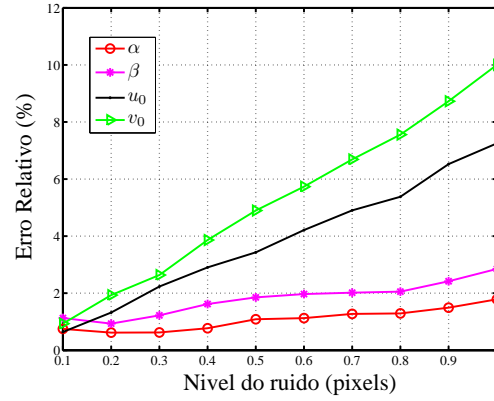
Analisando os resultados observa-se que a exatidão do método que utiliza a normalização proposta degrada-se rapidamente com o aumento do nível de ruído presente nos dados. Por outro lado, a exatidão do método de Zhang varia mais lentamente. Contudo, a exatidão do método proposto é da ordem de 10 vezes maior que a do método de Zhang. Em especial, para  $\sigma = 0,5$  (valor maior do que os encontrados na prática), tem-se um erro para  $\alpha$  e  $\beta$  em torno de 0,1 %. De fato, com a normalização proposta os resultados obtidos são melhores do que, até mesmo, os alcançados com o método não-linear proposto por Zhang e resumido na equação (19). Isto pode ser comprovado observando a figura 5. Esta mostra o resultado da minimização não-linear que utilizou um valor inicial obtido sem auxílio da normalização proposta. Observa-se que a solução não-linear reduz pela metade o erro obtido na solução linear sem normalização (figura 4). Contudo, a solução não-linear apresenta um erro bem maior que o obtido pelo método linear, cujos dados foram normalizados (figura 3). Evidentemente, o algoritmo de minimização não-linear ficou preso em um mínimo local devido a inicialização ruim dos parâmetros a serem estimados.

## 5.2. Dados Reais

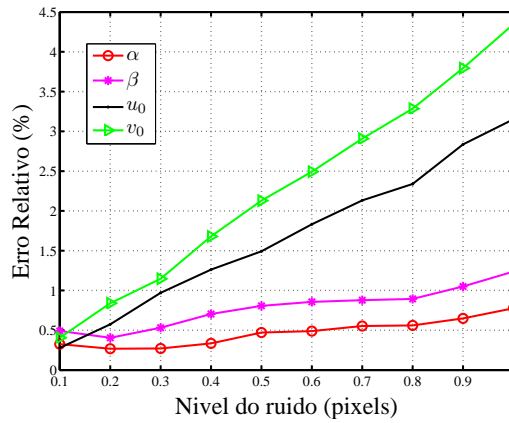
Para adquirir os dados reais, foi utilizado um gabarito com três esferas de plástico branco, com 2 cm de diâmetro cada uma. Tais esferas foram fixadas em um bastão de metal. A distância entre os centróides de esferas vizinhas foi feita, com auxílio de uma



**Figura 3.** Erro relativo obtido na estimação dos parâmetros intrínsecos da câmera pelo método linear com dados normalizados.



**Figura 4.** Erro relativo obtido na estimação dos parâmetros intrínsecos da câmera pelo método linear com dados sem normalização.



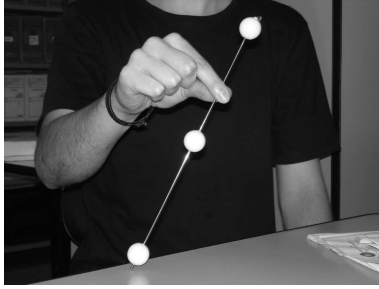
**Figura 5.** Erro relativo obtido na estimação dos parâmetros intrínsecos da câmera pelo método não-linear inicializado com parâmetros obtidos com dados sem normalização.

régua comum, igual 15 cm. Manualmente, o gabarito foi movido em diferentes posições a uma distância de aproximadamente 50 cm da câmera. Durante o movimento, 20 imagens foram adquiridas (uma dessas é apresentada na figura 6). Devido as imagens serem consecutivas, desde que a localização das esferas seja marcada manualmente na primeira imagem e os respectivos centróides conhecidos, é simples identificá-las nas outras imagens com o auxílio de um algoritmo de rastreamento<sup>2</sup> baseado em um algoritmo *blob* gaussiano [Lindeberg 1994, Cap. 7]. A localização do centróide de cada esfera foi utilizada como dado de entrada para o algoritmo proposto. O resultado deste experimento é apresentado tabela 1. O resultado referenciado como **NORMALIZADO** foi obtido com o método linear e a normalização proposta. O resultado **ZHANG** é a solução não-linear do problema de calibração 1D sem normalização.

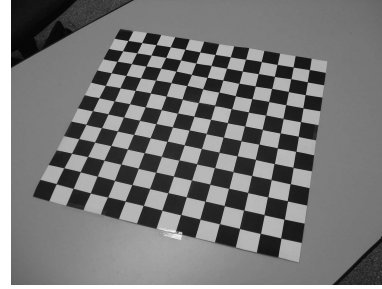
A título de comparação, os parâmetros intrínsecos da câmera utilizada nos testes também foram estimados pelo método proposto por Zhang (2000). O gabarito utilizado

<sup>2</sup>Do inglês: *tracking*.





**Figura 6. Gabarito 1D utilizado na avaliação do método proposto.**



**Figura 7. Gabarito plano utilizado na comparação do método proposto.**

**Tabela 1. Parâmetros intrínsecos do conjunto binocular utilizado para gerar os dados sintéticos.**

Método	Parâmetros			
	$\alpha$	$\beta$	$u_0$	$v_0$
<b>NORMALIZADO</b>	844,9312	876,9088	355,5348	208,4789
<b>ZHANG</b>	852,1759	894,4997	365,9896	203,3809
<b>BOUGUET</b>	842,4040	879,5474	358,0411	207,2355

neste método é um plano com um padrão quadriculado. Neste trabalho, utilizou-se um gabarito planar com 121 pontos (figura 7). Durante a calibração, foram realizados 10 deslocamentos e uma sequência com 10 imagens foram capturadas. Em seguida, a implementação de Bouguet (2004) do algoritmo de Zhang (2000) foi utilizada para estimar os parâmetros da câmera. O resultado desta calibração está apresentado na terceira linha da tabela 1 e é referenciado por **BOUGUET**. Observa-se que o resultado **ZHANG** está bastante diferente do obtido pelo algoritmo Bouguet (2004). Por outro lado, mesmo sendo obtido por um algoritmo linear, o resultado **NORMALIZADO** é muito próximo do **BOUGUET**.

## 6. Conclusões e Trabalhos Futuros

Neste trabalho aperfeiçoou-se a técnica de calibração monocular com um gabarito de uma única dimensão. Tal técnica apenas requer que a câmera observe, a partir de diferentes orientações, um bastão graduado, ou seja, pontos colineares, cujas posições relativas são conhecidas. Durante a captura das imagens a câmera deve estar imóvel, pois uma das extremidades do gabarito deve permanecer em um ponto fixo durante o movimento. No entanto, tal movimento não precisa ser conhecido.

A metodologia de calibração consiste em um algoritmo linear que fornece uma estimativa inicial dos parâmetros intrínsecos da câmera. Para que seja possível uma solução linear, uma substituição de variáveis é necessária. Tal substituição torna o sistema muito susceptível a ruídos. Contudo, demonstrou-se experimentalmente neste trabalho que uma simples normalização dos dados de entrada tornam o problema melhor condicionado. Tal normalização é essencial para que a técnica possa ser utilizada em situações práticas.

Com as modificações sugeridas neste trabalho, baseado nos resultados obtidos tanto em simulações quanto em imagens reais, conclui-se que um bastão com apenas três

pontos colineares constitui um prático gabarito de calibração. Tal gabarito pode, inclusive, estar visível para várias câmeras ao mesmo tempo. Isto permite a calibração de vários conjuntos binoculares em bloco. Além disso, dependendo da exatidão pretendida, a solução obtida de forma linear, mas com a normalização proposta, pode dispensar a necessidade de refinamento por métodos não-lineares bem mais complexos. Isto é possível porque a solução linear proposta já fornece uma exatidão bem menor que 1 %.

O trabalho de pesquisa sobre o tema de calibração com um gabarito 1D ainda continua. Faz-se necessário provar matematicamente o efeito da normalização dos dados no resultado final e a escolha da melhor transformação  $T_2$ . Além disso, deve-se examinar a influência da variação do número de imagens e da quantidade de pontos do gabarito na solução final do problema.

## Referências

- Agrawal, M. and Davis, L. (2003). Camera calibration using spheres: A semidefinite programming approach. In *IX IEEE International Conference on Computer Vision*, pages 782–789.
- Bouguet, J. T. (2004). *Camera Calibration Toolbox for Matlab*. MRL - Intel Corp.
- de França, J. A. and Stemmer, M. R. (2006). Stereo calibrations with a free-moving stick. In *VII Induscon - Conferência Internacional de Aplicações Industriais*, Recife, Brazil.
- Faugeras, O. and Luong, Q.-T. (2001). *The Geometry of Multiple Images: The Laws That Govern the Formation of Multiple Images of a Scene and Some of Their Applications*. MIT Press, Cambridge, MA.
- Golub, G. H. and Van Loan, C. F. (1996). *Matrix Computations*. The Johns University Press.
- Hartley, R. (1997). In defence of the eight point algorithm. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(6):580–593.
- Izquierdo, E. and Guerra, V. (2003). Estimating the essential matrix by efficient linear techniques. *IEEE Trans. on Circuits and Systems for Video Technology*, 13(9):925–934.
- Lindeberg, T. (1994). *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers.
- Torr, P. H. S. (1995). *Outlier Detection and Motion Segmentation*. PhD thesis, University of Oxford, Engineering Dept, Inglaterra.
- Tsai, R. Y. (1987). A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf cameras and lenses. *IEEE Journal Robotics and Automation*, RA-3(4):323–344.
- Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334.
- Zhang, Z. (2004). Camera calibration with one-dimensional objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26:892–899.