

Stereo Calibration with a Free-Moving Stick

José A. de França

Universidade Estadual de Londrina
Departamento de Engenharia Elétrica
Caixa Postal 6025
86051-990, Londrina - PR, Brazil
<http://research.eeol.org/>
Email: josealexandre@eeol.org

Marcelo R. Stemmer

Universidade Federal de Santa Catarina
Departamento de Automação e Sistemas
Caixa Postal 476
88040-970, Florianópolis-SC, Brazil
<http://www.das.ufsc.br/~marcelo/>
Email: marcelo@das.ufsc.br

Abstract— We propose a new method to stereo rig calibration. The method consists of upgrading a 3D projective representation to Euclidean using a free-moving 1D calibration apparatus (points aligned on a line). For higher accuracy, a nonlinear technique based on the maximum likelihood criterion is used to estimate the stereo rig parameters (the intrinsic and extrinsic parameters). Experiments with real images validate the new method and show that the results of the technique are comparable to the performance of the most complex methods.

Index Terms - Camera calibration, calibration taxonomy, calibration apparatus, 1D objects, stereo calibration.

I. INTRODUÇÃO

Deteção e localização de obstáculos estão entre as tarefas mais básicas da robótica móvel. Tais tarefas são essenciais para, por exemplo, proteger o robô de colisões e guia-lo dentro de ambientes desconhecidos ou dinâmicos. Para solucionar tais problemas, algumas abordagens encontradas na literatura necessitam recuperar informações 3D do ambiente do robô. Muitas vezes, tais informações são obtidas com o auxílio de transdutores como, por exemplo, sonares e *laser rangefinders*. Em outros casos, são utilizados sistemas baseados em visão monocular em conjunto com um conhecimento prévio do ambiente ou com uso de marcas artificiais¹. Contudo, a visão computacional pode recuperar as informações 3D completamente, sem fazer suposições sobre o ambiente do robô, apenas com o auxílio de um conjunto de visão binocular. Contudo, geralmente, antes que o sistema de visão binocular possa ser utilizado, este tem que ser calibrado. Tal calibração consiste em estimar os parâmetros intrínsecos e extrínsecos das câmeras que compõem o par estéreo. Normalmente, a calibração é realizada antes da operação normal do sistema. Neste caso, insere-se um objeto 3D (chamado gabarito) de forma e tamanho conhecidos no campo visual da câmera. Em seguida, o algoritmo de calibração retira as informações necessárias e estima os parâmetros da câmera. Apesar de extremamente exata, a calibração com gabarito é pouco flexível, pois exige que a operação do sistema seja interrompida sempre que uma nova calibração for necessária.

Para tornar o processo de calibração mais flexível, Faugeras [1] e Hartley et al. [2] desenvolveram a técnica de cali-

bração projetiva. Esta não necessita de gabarito de calibração. Contudo, o modelo utilizado para a câmera não permite a recuperação de informações importantes da cena como, por exemplo, distâncias, ângulos e paralelismo [3]. Dessa forma, a técnica não pode ser utilizada em todas as aplicações.

Uma alternativa mais geral e que serve a maioria das aplicações é a calibração automática. Este conceito foi inicialmente proposto por Maybank e Faugeras [4] e dispensa o uso de um gabarito de calibração. Ao invés disso, a câmera deve se movimentar com relação a uma cena estática. Dessa forma, a imagem do cone absoluto [5] fornece restrições que possibilitam a recuperação de todos os parâmetros do modelo da câmera. Por outro lado, a teoria de calibração automática, apesar de possuir um custo computacional extremamente elevado, leva-nos a resultados pouco exatos. Estes fatos contribuem para que essa técnica ainda não seja largamente utilizada na prática.

No caso da visão binocular surgem novas restrições que facilitam a tarefa de calibração. Dentre essas, a mais importante é a restrição epipolar [6]. Esta relaciona pontos na imagem capturada por uma das câmeras a pontos na imagem da outra câmera do mesmo par estéreo.

Quando o sistema não está calibrado, a restrição epipolar é completamente caracterizada pela matriz fundamental. Tal matriz concentra uma série de informações sobre o sistema binocular e torna todas as operações realizadas menos complexas. Contudo, antes de ser útil, a matriz fundamental deve ser estimada a partir de um conjunto inicial de pontos. Veja, por exemplo, os trabalhos de [7]–[10].

Como proposto por Hartley et al. [2], uma vez que a matriz fundamental é conhecida, esta permite realizar uma calibração projetiva. Tal calibração pode ser atualizada para uma calibração euclidiana se uma transformação projetiva, que tem características já bem conhecidas [11], for estimada. Neste trabalho, propõe-se um novo método de calibração de um conjunto binocular que estima tal transformação usando, como gabarito, um bastão graduado. Basicamente, o bastão funciona com um gabarito de uma única dimensão. Por utilizar um objeto de calibração de simples confecção e manuseio, o método proposto acrescenta flexibilidade ao problema. Além disso, os resultados obtidos mostram que o método obtém resultados bem mais exatos, em comparação as técnicas automáticas.

¹Do inglês: *landmarks*

O método proposto é uma extensão do recente trabalho de Zhang [12], que propôs uma técnica de calibração de uma única câmera com um gabarito 1D. Por sua vez, tal técnica surgiu da tendência atual, que pode ser comprovada examinando artigos recentes como, por exemplo, os trabalhos de Agrawal e Davis [13], Zhang [14] e de Sturm e Maybank [15], de aumentar a flexibilidade da tarefa de calibração reduzindo as dimensões do gabarito.

II. FUNDAMENTAÇÃO TEÓRICA

O modelo de câmera utilizado neste trabalho é o *pinhole*. Neste, a relação entre um ponto 3D, \mathbf{M} , e a sua projeção \mathbf{m} no plano da imagem é dada por²

$$\tilde{\mathbf{m}} \simeq \tilde{\mathbf{P}}\tilde{\mathbf{M}}, \quad (1)$$

onde \mathbf{P} é a matriz de projeção da câmera e é igual a

$$\mathbf{P} \simeq \mathbf{A}[\mathbf{R} \ \mathbf{t}], \quad (2)$$

na qual

$$\mathbf{A} = \begin{bmatrix} f & f \cot \theta_{sk} & u_0 \\ 0 & f n_{x/y} & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3)$$

\mathbf{R} é uma matriz 3×3 que representa a rotação da câmera em relação ao sistema de coordenadas do ambiente, \mathbf{t} é um vetor que equivale a translação do centro da câmera com relação ao sistema de coordenadas do ambiente.

Na equação (3), f é a distância focal, enquanto que $n_{x/y}$ é inserido para levar em consideração que a matriz de sensores da câmera não é um quadrado perfeito. Já o parâmetro $s_k = f \cot \theta_{sk}$ está relacionado ao ângulo, θ_{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.

Normalmente, referem-se a matriz \mathbf{A} como os “parâmetros intrínsecos da câmera”. Enquanto que o vetor \mathbf{t} e a matriz \mathbf{R} possuem toda informação sobre os “parâmetros extrínsecos da câmera”.

Agora, considerando um conjunto binocular, sem perda de generalidade, assumimos que a câmera da esquerda está na origem do sistema de coordenadas do ambiente. Então, considerando as equações (1) e (2), temos

$$\tilde{\mathbf{m}}_1 \simeq \mathbf{A}[\mathbf{I} \ \mathbf{0}]\tilde{\mathbf{M}} \quad \text{e} \quad \tilde{\mathbf{m}}_2 \simeq \mathbf{A}'[\mathbf{R} \ \mathbf{t}]\tilde{\mathbf{M}}. \quad (4)$$

A equação (4) pode ser re-escrita em função das matrizes de projeção das câmeras, \mathbf{P} e \mathbf{P}' , ou seja,

$$\tilde{\mathbf{m}}_1 \simeq \tilde{\mathbf{P}}\tilde{\mathbf{M}} \quad \text{e} \quad \tilde{\mathbf{m}}_2 \simeq \tilde{\mathbf{P}}'\tilde{\mathbf{M}}, \quad (5)$$

com,

$$\mathbf{P} \simeq \mathbf{A}[\mathbf{I} \ \mathbf{0}] \quad \text{e} \quad \mathbf{P}' \simeq \mathbf{A}'[\mathbf{R} \ \mathbf{t}]. \quad (6)$$

Do trabalho de Hartley et al. [2], sabe-se que, se apenas a matriz fundamental do sistema for conhecida, é possível obtermos matrizes de projeção, \mathbf{P} e \mathbf{P}' , que permitem realizar

²Coordenadas homogêneas de um ponto $\mathbf{M}_\alpha = [x_\alpha, y_\alpha, \dots]^T$ são representadas por $\tilde{\mathbf{M}}_\alpha$, isto é, $\tilde{\mathbf{M}}_\alpha = [x_\alpha, y_\alpha, \dots, 1]^T$.

uma calibração projetiva³, isto é, dado uma correspondência $(\mathbf{m}_1, \mathbf{m}_2)$, o ponto 3D obtido com as matrizes \mathbf{P} e \mathbf{P}' está no espaço \mathcal{P}^3 e não no Euclidiano.

Na calibração projetiva, as matrizes de projeção das câmeras são as seguintes

$$\mathbf{P} \simeq [\mathbf{I} \ \mathbf{0}] \quad \text{e} \quad \mathbf{P}' \simeq [\mathbf{H} \ w\mathbf{e}_2], \quad (7)$$

onde [16]

$$\mathbf{H} = \mathbf{H}_\infty + \mathbf{e}_2\mathbf{w}^T, \quad (8)$$

$\tilde{\mathbf{e}}_2$ é o epipolo da câmera da direita, w é uma constante qualquer não-nula e \mathbf{w} a um vetor 3D qualquer, ou seja, \mathbf{H} tem infinitas soluções. Contudo, uma solução possível pode ser obtida fazendo

$$\mathbf{H} = \left(-\frac{[\tilde{\mathbf{e}}_2]_\times}{\|\tilde{\mathbf{e}}_2\|^2} \right) \mathbf{F}. \quad (9)$$

De Hartley et al. [2] e Faugeras [1] temos que as matrizes \mathbf{P} e \mathbf{P}' estão relacionadas com as matrizes de projeção \mathbf{P} e \mathbf{P}' por uma transformação projetiva desconhecida, \mathbf{T}_{PE} , ou seja,

$$\mathbf{P} = \mathbf{P}\mathbf{T}_{PE} \quad \text{e} \quad \mathbf{P}' = \mathbf{P}'\mathbf{T}_{PE}. \quad (10)$$

Além disso, como demonstrado por [11], \mathbf{T}_{PE} tem nove graus de liberdade e é dada por

$$\mathbf{T}_{PE} = \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{w}^T & w \end{bmatrix}, \quad (11)$$

onde $\mathbf{W}_\infty = [\mathbf{w}^T, w]^T$ [veja as equações (7) e (8)] é um vetor 4×1 e é igual a equação do plano no infinito no sistema de coordenadas definido por \mathbf{P} e \mathbf{P}' .

III. MÉTODO PROPOSTO

Nas seções seguintes, um novo método de calibração de um conjunto binocular é proposto. Este utiliza, como gabarito de calibração, um bastão graduado, ou seja, um objeto de uma única dimensão. Por usar um gabarito de dimensões reduzidas, deve-se verificar se o mesmo possui restrições suficientes para que seja possível estimar-se todos os parâmetros do conjunto binocular.

A. Exequibilidade Matemática

O método de calibração proposto estima \mathbf{T}_{PE} utilizando, como gabarito, um bastão de comprimento conhecido que desloca-se livremente a frente das câmeras. Assim, inicialmente, deve ser verificado se um gabarito na forma de um bastão fornece restrições suficientes para que todos os parâmetros de \mathbf{T}_{PE} sejam conhecidos unicamente.

³Do inglês: *weak calibration*

1) *Solução com Apenas Dois Pontos:* Utilizando, simplesmente, uma haste de tamanho conhecido como gabarito, tem-se apenas dois pontos 3D, M_{1i} e M_{2i} (figura 1(a)), para auxiliar na calibração. Como as projeções desses pontos pertencem, necessariamente, ao mesmo plano, as equações em (5) fornecem apenas três restrições para cada ponto 3D. Assim, se a haste deslocar-se livremente, para n deslocamentos, temos $6n$ equações independentes. Por outro lado, M_{2i} pode ser expresso em função das três coordenadas de M_{1i} e duas incógnitas, θ_i e ϕ_i , ou seja,

$$M_{2i} = M_{1i} + L[\text{sen } \theta_i \cos \phi_i, \text{sen } \theta_i \text{sen } \phi_i, \cos \theta_i]^T, \quad (12)$$

onde L é o comprimento da haste de calibração e θ_i e ϕ_i são ângulos que definem a orientação da haste. Dessa forma, para n deslocamentos livres da haste de calibração e considerando os graus de liberdade de T_{PE} , tem-se um total de $9 + 5n$ incógnitas. Assim, com $n = 9$ deslocamentos, é possível estimar T_{PE} unicamente.

2) *Solução com Três Pontos:* Se um ou mais pontos, M_{ji} , $j \geq 3$, forem acrescentados a haste de calibração (figura 1(b)), considerando que as distâncias $d^2(M_{1i}, M_{ji})$ e $d^2(M_{2i}, M_{ji})$ são conhecidas, cada ponto M_{ji} pode ser escrito em função das coordenadas dos pontos extremos, M_{1i} e M_{2i} , ou seja,

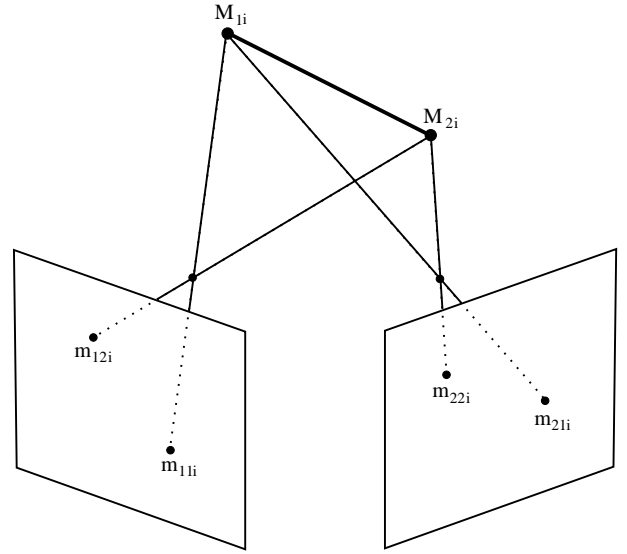
$$M_{ji} = \lambda_{1j}M_{1i} + \lambda_{2j}M_{2i}, \quad (13)$$

onde λ_{1j} e λ_{2j} são constantes conhecidas que dependem de $d^2(M_{1i}, M_{ji})$ e $d^2(M_{2i}, M_{ji})$. Assim, a inserção de um ou mais pontos a haste de calibração não acrescenta nenhuma incógnita ao sistema. Por outro lado, para n deslocamentos do bastão, os pontos M_{1i} e M_{2i} fornecem $6n$ equações independentes. Contudo, desde que o terceiro ponto deve ser colinear a M_{1i} e M_{2i} , o ponto M_{3i} fornece apenas n equações independentes. Dessa forma, tem-se um total de $6n + n = 7n$ equações e, como antes, apenas $9 + 5n$ incógnitas. Assim, a calibração é possível com apenas cinco deslocamentos e cinco pares de imagens estéreo.

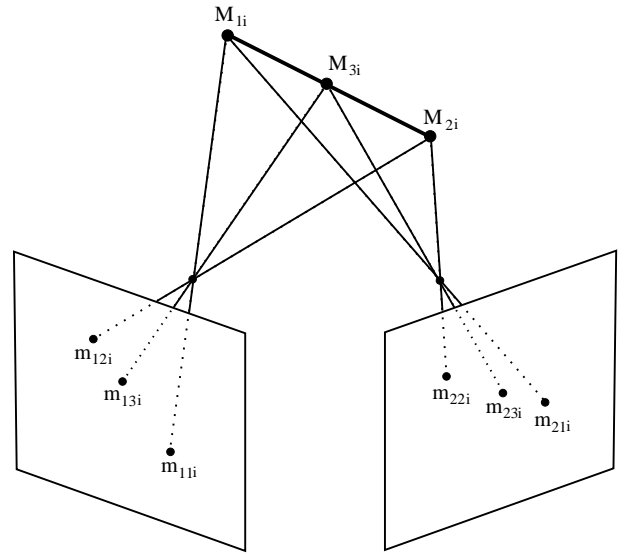
3) *Solução com Quatro ou Mais Pontos:* Considere um bastão com quatro pontos, M_{ji} , $j = 1, 2, 3$ e 4, com posições relativas conhecidas. O ponto M_{4i} , entre os pontos extremos da haste, deve satisfazer a restrição de razão cruzada com os pontos M_{1i} , M_{2i} e M_{3i} . Dessa forma, M_{4i} não fornece equações independentes, pois uma vez que as projeções de M_{1i} , M_{2i} e M_{3i} tenham sido obtidas, a projeção de M_{4i} é completamente determinada. Com isso, uma quantidade de pontos superior a três não acrescenta nenhuma restrição ao problema.

B. Solução linear

Em um deslocamento (indexado por i) de um bastão de calibração com q pontos marcados, temos uma correspondência ($m_{1ji} \leftrightarrow m_{2ji}$) para cada ponto M_{ji} , onde $j = 1, 2 \dots q$, pertencente ao gabarito. Como tais correspondências podem ser obtidas com exatidão, após n deslocamentos, a matriz fundamental do sistema pode ser estimada com excelente precisão usando $q \times n$ correspondências.



(a)



(b)

Fig. 1. Esquemas de calibração estereoscópica com uma haste, utilizando: (a) um bastão de comprimento conhecido e (b) um bastão com três pontos, cujas posições relativas são conhecidas.

Uma vez que a matriz fundamental for conhecida, ela pode ser utilizada para obtermos \mathcal{P} e \mathcal{P}' . Dessa forma, uma reconstrução projetiva, M_{ji} , de cada ponto M_{ji} pode ser obtida. Além disso, da equação (10), temos

$$\widetilde{M}_{ji} \simeq T_{PE} \widetilde{M}_{ji}.$$

A equação anterior pode ser transformada em uma igualdade

se considerarmos que \mathbf{T}_{PE} é dada por (11), ou seja,

$$\mathbf{M}_{ji} = \frac{1}{\widetilde{\mathbf{M}}_{ji}^T} \mathbf{A}^{-1} \mathbf{M}_{ji}. \quad (14)$$

Como o comprimento, L , e a graduação (distância entre um ponto e outro) do bastão são conhecidos, as coordenadas dos pontos \mathbf{M}_{ji} , para $j = 3, 4 \dots q$, podem ser expressas como função dos pontos extremos, ou seja,

$$\mathbf{M}_{ji} = \lambda_{1j} \mathbf{M}_{1i} + \lambda_{2j} \mathbf{M}_{2i}, \quad (15)$$

onde λ_{1j} e λ_{2j} dependem da graduação do bastão, mas são conhecidos.

Substituindo (14) na equação anterior e multiplicando ambos os lados da equação por \mathbf{A} , temos

$$\frac{1}{\widetilde{\mathbf{M}}_{ji}^T} \mathbf{M}_{ji} = \lambda_{1j} \frac{1}{\widetilde{\mathbf{M}}_{1i}^T} \mathbf{M}_{1i} + \lambda_{2j} \frac{1}{\widetilde{\mathbf{M}}_{2i}^T} \mathbf{M}_{2i}.$$

Após alguma manipulação algébrica, a equação anterior torna-se

$$\mathbf{u}_{ji}^T \mathbf{W} = 0, \quad (16)$$

onde

$$\mathbf{u}_{ji}^T = \widetilde{\mathbf{M}}_{1i}^T + \frac{\lambda_{1j} (\mathbf{M}_{1i} \times \mathbf{M}_{ji}) \cdot (\mathbf{M}_{2i} \times \mathbf{M}_{ji})}{\lambda_{2j} (\mathbf{M}_{2i} \times \mathbf{M}_{ji}) \cdot (\mathbf{M}_{2i} \times \mathbf{M}_{ji})} \widetilde{\mathbf{M}}_{2i}^T. \quad (17)$$

Considerando um bastão com $q > 2$ pontos, a equação (16) nos leva a

$$\mathbf{U}_{qi} \mathbf{W} = \mathbf{0}, \quad (18)$$

onde

$$\mathbf{U}_{qi} = [\mathbf{u}_{3i}, \mathbf{u}_{4i} \dots \mathbf{u}_{qi}]^T.$$

Para $q \geq 4$, a equação (18) possui apenas duas equações independentes. Por isso, desde que \mathbf{W} possui três graus de liberdade, no mínimo, $n \geq 2$ deslocamentos são necessários para estimar \mathbf{W} unicamente. Além disso, com mais deslocamentos podemos usar redundância para reduzir a influência do ruído resolvendo

$$\begin{bmatrix} \mathbf{U}_{q1} \\ \mathbf{U}_{q2} \\ \vdots \\ \mathbf{U}_{qn} \end{bmatrix} \mathbf{W} = \mathbf{0}. \quad (19)$$

Uma vez que \mathbf{W} foi estimado, considerando apenas a câmara que está na origem do sistema de coordenadas, temos as seguintes relações

$$\begin{aligned} \mathbf{M}_{1i} &= Z_{1i} \mathbf{A}^{-1} \widetilde{\mathbf{m}}_{11i}, \\ \mathbf{M}_{2i} &= Z_{2i} \mathbf{A}^{-1} \widetilde{\mathbf{m}}_{12i}, \\ \mathbf{M}_{ji} &= Z_{ji} \mathbf{A}^{-1} \widetilde{\mathbf{m}}_{1ji}, \end{aligned} \quad (20)$$

onde Z_{xi} é a profundidade, ou seja, a coordenada Z de \mathbf{M}_{xi} , e $\widetilde{\mathbf{m}}_{1ji}$ é a projeção de \mathbf{M}_{xi} na imagem I_1 .

Desde que \mathbf{W} já é conhecido e considerando (3) e (14), temos

$$\begin{aligned} Z_{xi} &= S_w \frac{{}^P Z_{xi}}{\widetilde{\mathbf{M}}_{xi}^T \mathbf{W}} \\ &= S_w \beta_{xi}, \end{aligned} \quad (21)$$

onde ${}^P Z_{xi}$ é a coordenada Z de $\widetilde{\mathbf{M}}_{xi}$ e S_w é uma constante inserida para levar em consideração que \mathbf{W} é um vetor homogêneo. Observe que S_w depende apenas de \mathbf{W} e não dos pontos \mathbf{M}_{xi} .

Agora, como o comprimento, L , e a graduação do bastão são conhecidos, para determinar \mathbf{A}^{-1} , podemos utilizar as equações a seguir

$$\begin{aligned} \|\mathbf{M}_{2i} - \mathbf{M}_{1i}\| &= L, \\ \alpha_{2j} \|\mathbf{M}_{2i} - \mathbf{M}_{ji}\| &= L, \\ \alpha_{1j} \|\mathbf{M}_{ji} - \mathbf{M}_{1i}\| &= L, \end{aligned}$$

onde $\alpha_{xi} = 1/(1 - \lambda_{xj})$.

Se substituirmos (20) nas equações anteriores, temos

$$\begin{aligned} \|S_w \mathbf{A}^{-1} (\beta_{2i} \widetilde{\mathbf{m}}_{12i} - \beta_{1i} \widetilde{\mathbf{m}}_{11i})\| &= L, \\ \alpha_{2j} \|S_w \mathbf{A}^{-1} (\beta_{2i} \widetilde{\mathbf{m}}_{12i} - \beta_{ji} \widetilde{\mathbf{m}}_{1ji})\| &= L, \\ \alpha_{1j} \|S_w \mathbf{A}^{-1} (\beta_{ji} \widetilde{\mathbf{m}}_{1ji} - \beta_{1i} \widetilde{\mathbf{m}}_{11i})\| &= L. \end{aligned}$$

A equação anterior é equivalente a

$$\begin{aligned} \mathbf{Y}_{21i}^T \mathbf{B} \mathbf{Y}_{21i} &= L^2, \\ (\alpha_{2j})^2 \mathbf{Y}_{2ji}^T \mathbf{B} \mathbf{Y}_{2ji} &= L^2, \\ (\alpha_{1j})^2 \mathbf{Y}_{j1i}^T \mathbf{B} \mathbf{Y}_{j1i} &= L^2, \end{aligned} \quad (22)$$

onde $\mathbf{Y}_{xyi} = (\beta_{xi} \widetilde{\mathbf{m}}_{1xi} - \beta_{yi} \widetilde{\mathbf{m}}_{1yi})$ e $\mathbf{B} = S_w^2 \mathbf{A}^{-T} \mathbf{A}^{-1}$ é a matriz do cone absoluto.

Levando em consideração que

$$\mathbf{a}^T \mathbf{B} \mathbf{a} = [a_1^2, 2a_1 a_2, a_2^2, 2a_1 a_3, 2a_2 a_3, a_3^2]^T \mathbf{b}, \quad (23)$$

onde $\mathbf{a} = [a_1, a_2, a_3]^T$ e $\mathbf{b} = [b_{11}, b_{12}, b_{22}, b_{13}, b_{23}, b_{33}]^T$, com b_{ij} igual ao elemento da i -ésima linha e j -ésima coluna de \mathbf{B} , o sistema de equações (22) pode ser representado por

$$\mathbf{V}_i \mathbf{b} = \mathbf{L}^T \mathbf{L}, \quad (24)$$

onde \mathbf{V}_i é uma matriz que depende dos coeficientes dos vetores \mathbf{Y}_{xyi} da equação (22) como indicado por (23) e $\mathbf{L} = [L, L, \dots, L]^T$.

Independente do número de pontos existentes no bastão de calibração, apenas uma das equações em (22) é linearmente independente. Então, como \mathbf{B} tem seis graus de liberdade (um de S_w e cinco de \mathbf{A}^{-1}) são necessários $n \geq 6$ deslocamentos para estimar \mathbf{B} unicamente, resolvendo a seguinte equação

$$\begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \\ \vdots \\ \mathbf{V}_n \end{bmatrix} \mathbf{b} = \mathbf{L}^T \mathbf{L}. \quad (25)$$

Uma vez que \mathbf{B} for conhecida, a matriz $S_w \mathbf{A}^{-1}$ pode ser obtida de \mathbf{B} através da decomposição de Cholesky. Em seguida, se necessário, os pontos \mathbf{M}_{1i} e \mathbf{M}_{2i} podem ser obtidos de (14) e (12) respectivamente.

As projeções, na imagem I_1 , dos pontos do bastão podem ser normalizadas através da transformada \mathbf{T}_2 . Neste caso, os parâmetros intrínsecos no sistema de coordenadas original podem ser obtidos pela equação a seguir

$$\mathbf{A} = \mathbf{T}_2^{-1} \mathbf{A}_n, \quad (26)$$

onde \mathbf{A}_n é a matriz da câmera 1 obtida com os pontos normalizados.

C. Solução não-linear

O método da seção anterior minimiza um critério sem significado geométrico. Este é um dos motivos que contribuem para que o mesmo não forneça bons resultados na presença de ruído. Contudo, desde que o conceito de distância não é válido no espaço \mathcal{P}^3 , os pontos 3D, \mathbf{M}_{ji} , no espaço projetivo, não podem ser utilizados para definição de um critério de otimização. Assim, tal critério deve envolver as projeções dos pontos do bastão, \mathbf{m}_{1ji} e \mathbf{m}_{2ji} , nas imagens I_1 e I_2 . Uma solução possível, que age nas projeções dos pontos em uma imagem I_x , é

$$\mathcal{J}_{xi}(\mathcal{P}_x) = \sum_{j=1}^q \left\| \tilde{\mathbf{m}}_{xji} - \mathcal{P}_x \mathbf{T}_{PE}^{-1} \tilde{\mathbf{M}}_{ji} \right\|, \quad (27)$$

onde \mathcal{P}_x é uma matriz de projeção canônica [equação (7), pág. 2], $\tilde{\mathbf{M}}_{2i}$ é dado por (12) e $\tilde{\mathbf{M}}_{ji}$, $j \neq 1$ e 2 , é dado por (15).

Assim, considerando um bastão com q pontos marcados e n pares de imagens estereo, \mathbf{T}_{PE} pode ser obtida minimizando-se o critério simétrico a seguir

$$\sum_{i=1}^n [\mathcal{J}_{1i}(\mathcal{P}) + \mathcal{J}_{2i}(\mathcal{P}')] . \quad (28)$$

A solução da equação anterior envolve a solução de um problema de minimização não-linear. Normalmente, faz-se necessário uma estimativa inicial de todas as incógnitas envolvidas no problema, ou seja, da matriz \mathbf{A} , do plano \mathbf{W} , dos n pontos \mathbf{M}_{1i} e dos n ângulos θ_i e ϕ_i . Evidentemente, tal estimativa inicial pode ser obtida com o método da seção anterior.

IV. RESULTADOS EXPERIMENTAIS

O conjunto binocular utilizado nos testes é formado por duas câmeras USB, modelo Aiptek DV3100. Tais câmeras possuem um sensor CCD que permite capturar imagens nítidas, com dimensões de 640×480 pixels. Inicialmente, os parâmetros do par estereo foram conhecidos através de um experimento de extrema exatidão. Isto é importante para avaliar a eficiência do método proposto. Além disso, os resultados obtidos são comparados aos de outro método não-linear, o método de Zhang [14].

Nas seções a seguir, os experimentos são descritos em detalhes e os resultados são apresentados.

A. Parâmetros absolutos

Para conhecer os parâmetros do conjunto binocular utilizado nos testes, um gabarito 3D foi confeccionado e suas dimensões foram medidas com uma incerteza de 0,01 mm. Além disso, as faces do gabarito são de cor branca, sendo que duas delas possuem 100 padrões circulares da cor preta. Do trabalho de Heikkilä [17], o centróide de tais padrões pode ser conhecido com uma exatidão de 0,05 pixels. Uma vez que todos os

dados foram extraídos, a calibração foi efetuada utilizando o algoritmo de Tsai [18].

Devido a exatidão dos dados 3D e 2D utilizados, os parâmetros do conjunto binocular obtidos neste experimento e disponíveis na tabela I são considerados “parâmetros reais ou absolutos”. Por isso, estes foram utilizados para avaliar a precisão dos métodos testados.

B. Calibração com um plano

A exatidão do método de proposto foi comparada a do método de Zhang [14]. O gabarito utilizado no método é um único plano com um padrão quadriculado. Na calibração, tal gabarito deve fazer movimentos genéricos. Então, além dos parâmetros intrínsecos das câmeras, a cada deslocamento, as coordenadas 3D dos pontos do gabarito, a matriz rotação, \mathbf{R}_i , e o vetor translação, \mathbf{t}_i , do plano em relação a câmera também devem ser estimados. Para tal estimativa, dado n deslocamentos de um gabarito com m pontos, Zhang [14] propôs a seguinte função a ser minimizada

$$\sum_{i=1}^n \sum_{j=1}^m \left\| \mathbf{m}_{ji} - \hat{\mathbf{m}}(\mathbf{A}, \mathbf{R}_i, \mathbf{t}_i, \mathbf{M}_{ji}) \right\|^2, \quad (29)$$

onde \mathbf{m}_{ji} é a projeção do ponto \mathbf{M}_{ji} , durante o i -ésimo deslocamento, e $\hat{\mathbf{m}}$ é o ponto estimado.

Para n deslocamentos e um gabarito com m pontos, o número de incógnitas que devem ser estimadas é igual a $5 + 5n + 3mn$. Além disso, a sua complexidade ainda é bem mais elevada que a do algoritmo proposto neste trabalho. Isto é devido ao número maior de incógnitas e, principalmente, ao fato da equação 29 estimar os parâmetros intrínsecos de apenas uma câmera do par estereo, ou seja, são necessárias duas minimizações para determinar todos os parâmetros do par estereo.

Neste trabalho, utilizou-se um gabarito planar com 121 pontos. Durante a calibração, foram realizados 10 deslocamentos e uma seqüência com 10 imagens estereo foram capturadas. Em seguida, a implementação de Bouguet [19] do algoritmo de Zhang [14], disponível livremente na internet, foi utilizada para estimar os parâmetros do conjunto binocular. Os resultados são apresentados na tabela I.

C. Calibração com uma haste

Para os testes com uma haste graduada, foi confeccionado um gabarito composto por um fio de cobre com 6 esferas brancas. Cada esfera tem cerca de 1,5 cm de diâmetro. Após o gabarito ser confeccionado, a distância entre uma esfera e outra foi medida com auxílio de uma régua comum.

Para capturar as imagens utilizadas nos testes, o gabarito foi movido a frente das câmeras e uma seqüência com 60 imagens estereo foi capturada por cada câmera do conjunto binocular. Devido as imagens serem consecutivas, desde que a localização das esferas seja marcada manualmente no primeiro par de imagens e os respectivos centróides conhecidos, é simples identificá-las nas outras imagens com o auxílio de um

TABLE I
 RESULTADOS EXPERIMENTAIS DA CALIBRAÇÃO DO CONJUNTO BINOCULAR (θ_{sk} É DADO EM GRAUS).

| Método | Câmera 1 | | | | | Câmera 2 | | | | |
|---------------|----------|---------|---------------|--------|--------|----------|---------|---------------|--------|--------|
| | f | n_x/y | θ_{sk} | u_0 | v_0 | f | n_x/y | θ_{sk} | u_0 | v_0 |
| REAIS | 205,97 | 1,02 | 90,01 | 330,56 | 270,67 | 201,89 | 1,02 | 90,01 | 336,67 | 262,23 |
| ZHANG | 1,56 | 1,69 | 0,50 | 3,76 | 2,89 | 1,86 | 1,58 | 0,53 | 3,51 | 3,01 |
| FRANÇA | 2,19 | 2,30 | 5,41 | 3,53 | 3,60 | 2,33 | 2,27 | 5,78 | 2,76 | 3,69 |

algoritmo de rastreamento⁴, baseado em um algoritmo *blob* gaussiano [20].

Para fácil visualização, os resultados de todos os métodos avaliados são apresentados na tabela I. Observe que apenas os parâmetros obtidos com o gabarito 3D (referenciado na tabela como **REAIS**) são dados de forma absoluta. Para os outros métodos os valores dados estão em percentagem (referente ao erro relativo aos parâmetros **REAIS**).

Na tabela I, os parâmetros estimados com o método proposto é referenciado como **FRANÇA**. Analisando tais parâmetros, nota-se que o parâmetro θ_{sk} é estimado com um erro elevado (mais de 5 %). Contudo, se θ_{sk} for considerado igual a $\pi/2$, ou seja, este parâmetro não for incluído na estimação, o erro cometido é de apenas 0,01 %.

De uma forma geral, a exatidão obtida pelo método de Zhang [14] (referenciado na tabela por **ZHANG**) foi um pouco melhor que a do método proposto. Contudo, este fato foi devido principalmente à confecção do gabarito de calibração. No método de Zhang [14], o gabarito foi confeccionado com ajuda do computador, logo as dimensões do gabarito são conhecidas com exatidão. No caso do métodos proposto, o gabarito foi confeccionado à mão, com auxílio de uma régua ordinária. Levando este fato em consideração, os resultados da tabela I demonstram que a técnica é bastante promissora.

V. CONCLUSÃO

O método de calibração proposto foi desenvolvido com o objetivo de flexibilizar o processo de calibração. Seu único pré-requisito é que o conjunto binocular “observe” um gabarito de uma única dimensão como, por exemplo, um bastão graduado ou um cordão de esferas, enquanto este realiza um movimento. Tal movimento não precisa ser conhecido, mas não deve ser uma translação pura ou, evidentemente, apenas rotacional.

Quando comparado a técnicas de calibração automática ou mesmo às que utilizam gabaritos mais complexos, o método proposto ganha em eficiência, pois possui custo computacional menor que as técnicas automáticas e é bem mais flexível que os métodos clássicos. Além disso, como mostram os experimentos realizados em imagens reais e sintéticas, a exatidão obtida com a técnica proposta é da mesma ordem de grandeza das obtidas com os métodos tradicionais.

AGRADECIMENTOS

Os autores agradecem aos revisores anônimos pelas contribuições para melhoria do texto e a PROPPG-UEL e ao CNPq pelo financiamento das pesquisas.

⁴Do inglês: *tracking*.

REFERENCES

- [1] O. D. Faugeras, “What can be seen in three dimensions with an uncalibrated stereo rig?” in *II European Conference on Computer Vision*, 1992, pp. 563–578.
- [2] R. Hartley, R. Gupta, and T. Chang, “Stereo from uncalibrated cameras,” in *II European Conference on Computer Vision*, 1992, pp. 579–587.
- [3] O. Faugeras and Q.-T. Luong, *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, 2001.
- [4] S. Maybank and O. Faugeras, “A theory of self-calibration of a moving camera,” *The International Journal of Computer Vision*, vol. 8, no. 2, pp. 123–151, 1992.
- [5] O. Faugeras and S. Maybank, “Motion from point matches: Multiplicity of solutions,” *The International Journal of Computer Vision*, vol. 4, no. 3, pp. 225–246, 1990.
- [6] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [7] Z. Zhang, “Determining the epipolar geometry and its uncertainty: A review,” *The International Journal of Computer Vision*, vol. 27, no. 2, pp. 161–195, August 1998.
- [8] A. Bartoli and P. Sturm, “Nonlinear estimation of the fundamental matrix with minimal parameters,” *IEEE Trans. on pattern analysis and machine intelligence*, vol. 26, no. 3, pp. 426–432, 2004.
- [9] Z. Zhang and C. Loop, “Estimating the fundamental matrix by transforming image points in projective space,” *Computer Vision and Image Understanding*, vol. 82, no. 2, pp. 174–180, 2001.
- [10] A. Bartoli, P. Sturm, and R. Horaud, “Projective structure and motion from two views of a piecewise planar scene,” in *8th International Conference on Computer Vision*, vol. 1, Vancouver, Canada, 2001, pp. 593–598.
- [11] R. Horaud, G. Csurka, and D. Demirdjian, “Stereo calibration from rigid motions,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1446–1452, 2000.
- [12] Z. Zhang, “Camera calibration with one-dimensional objects,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 892–899, 2004.
- [13] M. Agrawal and L. Davis, “Camera calibration using spheres: A semidefinite programming approach,” in *IX IEEE International Conference on Computer Vision*, 2003, pp. 782–789.
- [14] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [15] P. F. Sturm and S. J. Maybank, “On plane-based camera calibration: A general algorithm, singularities, applications,” in *IEEE Conf. Computer Vision and Pattern Recognition*, 1999, pp. 432–437.
- [16] Q.-T. Luong and T. Viéville, “Canonical representations for the geometries of multiple projective views,” *Computer Vision and Image Understanding*, vol. 64, no. 2, pp. 193–229, 1996.
- [17] J. Heikkilä, “Geometric camera calibration using circular control points,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1066–1076, 2000.
- [18] R. Y. Tsai, “A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf cameras and lenses,” *IEEE Journal Robotics and Automation*, vol. RA-3, no. 4, pp. 323–344, 1987.
- [19] J. T. Bouquet, *Camera Calibration Toolbox for Matlab*. MRL - Intel Corp., 2004.
- [20] T. Lindeberg, *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, 1994.