
CALIBRAÇÃO BINOCULAR COM GABARITOS 1D SEM RESTRIÇÃO DE MOVIMENTOS

José Alexandre de França*
franca@la2i.com

Maria Bernadete de M. França*
bernadete@eeol.org

Marcelo Ricardo Stemmer†
marcelo@das.ufsc.br

Rodrigo Henrique C. Palácios‡
rodrigopalacios@utfpr.edu.br

*Universidade Estadual de Londrina, Departamento de Engenharia Elétrica, Caixa Postal 6025, Londrina, PR, 86051-980

†Universidade Federal de Santa Catarina, DAS/CTC/UFSC, Caixa Postal 476, Florianópolis, SC, 88040-900

‡Universidade Tecnológica Federal do Paraná, Av. Alberto Carazzai, 1640, Cornélio Procopio, PR

ABSTRACT

Binocular calibration with 1D pattern without movements restriction

In computer vision, the camera calibration is a process needed when the recovery of some information, such as angles and distances, is desired. The present work deals with the problem of camera calibration using one-dimensional patterns. Nowadays, this problem only has a solution if some restrictions to the pattern's movement are imposed or if some angles of the cameras are known in advance. However, a different approach can be applied if, instead of only one camera, a stereo system is considered. In that case, the calibration is possible with a one-dimensional pattern that executes an unknown and unrestricted movement, even without any previous information concerning the cameras. Such method is based on the estimation of a transform which, after the estimation of the system's fundamental matrix, allows updating a projective calibration into a Euclidean calibration. Experiments using both real and synthetic images validate the new method and demonstrate that its accuracy is comparable to other well known calibration methods in the literature.

KEYWORDS: Camera Calibration, Calibration Pattern, 1D Object, Stereo Calibration.

RESUMO

Na visão computacional, a calibração de câmeras é um processo necessário quando deseja-se recuperar informações como, por exemplo, ângulos e distâncias. O presente trabalho trata do problema de calibração de câmeras com gabaritos de uma única dimensão. Atualmente, tal problema só tem solução se forem impostas restrições ao movimento do gabarito ou se alguns parâmetros das câmeras já sejam previamente conhecidos. Contudo, demonstra-se que uma abordagem diferente pode ser aplicada se, ao invés de uma única câmera, um conjunto binocular for considerado. Nesse caso, a calibração é possível com um gabarito 1D que realiza um deslocamento desconhecido e sem restrições, mesmo sem nenhuma informação prévia a respeito das câmeras. Tal método baseia-se na estimação de uma transformação que, após a estimação da matriz fundamental do sistema, permite atualizar uma calibração projetiva para uma calibração euclidiana. Experimentos em imagens reais e sintéticas validam o novo método e mostram que a sua exatidão é comparável a de outros métodos clássicos de calibração, já bem conhecidos na literatura.

PALAVRAS-CHAVE: Calibração de câmeras, Gabarito de calibração, Objeto 1D, Calibração binocular.

Artigo submetido em 05/05/2010 (Id.: 01144)

Revisado em 13/09/2010, 08/12/2010

Aceito sob recomendação do Editor Associado Prof. Marco Henrique Terra

1 INTRODUÇÃO

Matematicamente, no processo de criação da imagem, a câmera realiza um mapeamento entre um espaço 3D (o ambiente no qual a câmera está inserida) e um plano (o plano da imagem). Durante este processo, algumas informações como, por exemplo, ângulos, distâncias e volume, são perdidas. Caso estas informações sejam necessárias, é preciso estimar os parâmetros intrínsecos e extrínsecos da câmera, ou seja, matrizes com propriedades especiais que representam o mapeamento efetuado pela câmera, através de um procedimento chamado calibração. Normalmente, durante este procedimento, a câmera captura 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.

Os primeiros algoritmos de calibração a tornarem-se largamente utilizados baseiam-se em gabaritos 3D (Tsai, 1987; Lenz and Tsai, 1988). Geralmente, tais gabaritos devem ser compostos por dois ou mais planos ortogonais com um padrão bem conhecido em suas faces. Esses métodos tem a vantagem de realizarem a calibração com poucas imagens e com excelente exatidão. Com o passar dos anos, foram propostos novos métodos de calibração que utilizam gabaritos 2D (Sturm and Maybank, 1999; Zhang, 2000). Neste caso, as principais vantagens são a simplicidade do gabarito de calibração (até mesmo uma folha de papel com um padrão conhecido pode ser utilizado) e a fartura de planos em ambientes feitos pelo homem (que possibilita a utilização de algum padrão pré-existente no ambiente da câmera como gabarito de calibração). De fato, a abundância e a facilidade de detectar-se planos no ambiente motivou a proposição de algoritmos de autocalibração baseados em planos (Triggs, 1998). Na autocalibração, não são utilizados gabaritos. Ao invés disso, a câmera realiza alguns deslocamentos enquanto captura imagens. Em seguida, normalmente, basta que alguns pontos sejam rastreados ao longo destas imagens para que seja possível realizar a calibração. No entanto, apesar da praticidade e da abundância de algoritmos de autocalibração já propostos (Maybank and Faugeras, 1992; Hartley, 1997; Mendonça and Cipolla, 1999; Dornaika and Chung, 2001), a calibração automática ainda é muito pouco utilizada na prática. Isso se deve principalmente ao grande número de variáveis que precisam ser estimadas, o que leva a algoritmos inexatos e de grande complexidade computacional.

Mais recentemente, Zhang (2004) propôs a calibração utilizando gabaritos de uma única dimensão (pontos em uma reta). Nesta, o gabarito 1D deve realizar deslocamentos desconhecidos, enquanto a câmera captura imagens do mesmo. A única restrição é que um dos pontos do gabarito deve per-

manecer fixo durante a captura das imagens. No entanto, o trabalho de Zhang não pode ser aplicado na prática, pois o método é muito susceptível a ruído, o que, normalmente, faz com que o resultado da calibração seja inútil. Por isso, apenas com o trabalho de França et al. (2010) a calibração com gabaritos 1D pôde ser utilizada da prática, com uma exatidão comparável a de outros métodos já tradicionais. Isto porque, neste trabalho, demonstrou-se que a exatidão do método original de Zhang (2004) é melhorada significativamente, simplesmente, analisando-se o equacionamento do problema. Assim, é possível melhorar o condicionamento numérico realizando uma simples normalização nos dados.

No mesmo ano em que Zhang propôs seu algoritmo de calibração, Cao and Foroosh (2004) alteraram o método de forma a ser dispensável o conhecimento da distância entre os pontos do gabarito. Contudo, neste caso, é necessário utilizar-se um modelo muito simplificado para câmera, o qual nem sempre é possível ser utilizado na prática.

Outra pesquisa importante para a área foi desenvolvida por Hammarstedt et al. (2005), que identificaram e equacionaram as singularidades do problema, tornando possível identificá-las e, portanto, eliminá-las. Além disso, no mesmo trabalho, são apresentadas as soluções do problema para diversos modelos simplificados de câmeras. Também, no mesmo ano, Wu et al. (2005) mostraram que um gabarito 1D que gira em torno de um ponto central é geometricamente equivalente a um gabarito de calibração 2D com lados desconhecidos. Assim, foi demonstrado que a calibração também é possível quando o gabarito realiza um deslocamento desconhecido sobre a superfície de um plano. Contudo, na prática, isto não trás mais flexibilidade ao problema de calibração, além da praticidade já obtida com o uso de gabaritos 2D.

Todos os métodos mencionados previamente e que abordam a calibração de uma única câmera presumem movimentos especiais do gabarito. Em especial, a maioria necessita que um ponto do gabarito esteja sempre fixo, como acontece, por exemplo, no trabalho de França et al. (2010). De fato, ainda hoje não existem métodos de calibração monocular com gabaritos 1D que se deslocam livremente. Isto, além de ser uma limitação, ainda é mais uma fonte de erros, pois, na prática, nem sempre é fácil manter fixo um dos pontos do gabarito enquanto este movimentar-se.

A principal vantagem de utilizar gabaritos 1D para efetuar a calibração é a possibilidade de calibrar mais de uma câmera ao mesmo tempo. Isso acontece porque os pontos do gabarito 1D podem ser “capturados”, simultaneamente, por câmeras até mesmo em pontos de vista bem distintos. Contudo, apesar de Zhang (2004) ter destacado as vantagens do uso de gabaritos 1D para calibração de múltiplas câmeras, apenas após o trabalho de Kojima et al. (2005) isto tornou-se

possível de fato. Isto porque, para o caso de um conjunto de câmeras, o algoritmo de Kojima estima também os parâmetros extrínsecos de todas elas. Contudo, infelizmente, é necessário que pelo menos uma das câmeras, chamada de “câmera de referência”, já esteja calibrada.

No presente trabalho, uma abordagem diferente é aplicada a calibração de um conjunto binocular. Um conjunto binocular é composto por duas câmeras com posições relativas fixas, ou seja, o conjunto pode mover-se, mas a posição de uma câmera em relação à outra deve permanecer sempre constante. Tal arranjo é utilizado, por exemplo, quando deseja-se recuperar informações 3D das imagens (Hartley and Zisserman, 2000). Ao contrário do trabalho de Kojima et al. (2005), apesar do gabarito de calibração poder mover-se livremente, o método proposto não necessita que os parâmetros intrínsecos da câmera de referência sejam conhecidos. Também, como em todo algoritmo de calibração usando gabaritos 1D, o método proposto tem uma grande vantagem sobre aqueles baseados em gabaritos planos, ou seja, é possível calibrar simultaneamente duas câmeras mesmo se estas estiverem em pontos de vistas opostos. Situação encontrada na prática, por exemplo, em aplicações nas quais deseja-se construir um modelo 3D de uma cena que seja independente de ponto de vista (Carranza et al., 2003; Fujii and Tanimoto, 2002; Kanade et al., 1997). Para que isto seja possível, inicialmente, na técnica proposta, a matriz fundamental do sistema é estimada e uma calibração projetiva do conjunto binocular é realizada. Em seguida, o plano no infinito e os parâmetros intrínsecos da câmera de referência são estimados. Dessa forma, a calibração pode ser atualizada para uma calibração euclidiana. Resultados experimentais realizados em imagens sintéticas e reais mostram que o método proposto produz bons resultados mesmo na presença de ruído e, dessa forma, constitui um método prático e bastante flexível de calibração de um ou mais conjuntos binoculares.

2 NOTAÇÃO

As coordenadas homogêneas de um ponto $\mathbf{m} = [x, y]^T$ são representadas por $\tilde{\mathbf{m}}$, isto é, $\tilde{\mathbf{m}} = [x, y, 1]^T$. Índices, se houverem, indicam a posição do ponto em um conjunto de pontos. Assim, com a notação adotada, a relação entre um ponto 3D, $\tilde{\mathbf{M}}$, e sua projeção, \mathbf{m} , considerando o modelo de uma câmera *pinhole* (Hartley and Zisserman, 2000), é dada por

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

onde “ \simeq ” representa que os dois lados da equação podem diferir por uma constante (trata-se de uma equação homogênea) e

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

onde α e β são fatores de escala para, respectivamente, os eixos x e y , $\mathbf{m}_0 = [u_0, v_0]^T$ é o ponto central da câmera e s_k está relacionado ao ângulo que corrige a ortogonalidade da matriz de sensores. Contudo, atualmente é possível fabricar câmeras com s_k muito próximo de zero. Por isto, neste trabalho, considera-se $s_k = 0$. Dessa forma, há apenas quatro parâmetros intrínsecos a serem estimados.

As câmeras que pertencem a um conjunto binocular são chamadas de “câmera 1” e “câmera 2”. A câmera 1 é tomada como referência. No caso de matrizes, um apóstrofo (ou a ausência dele) é utilizado para indicar a qual câmera esta se refere. Por exemplo, \mathbf{A} está relacionada à “câmera 1” e \mathbf{A}' é sempre associada à “câmera 2”. Além disso, sem perda de generalidade, frequentemente assume-se que a câmera da esquerda está na origem do sistema de coordenadas do ambiente. Então, considerando a equação (1), tem-se

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

onde \mathbf{R} é uma matriz 3×3 que representa a rotação da câmera 1 em relação a 2, \mathbf{t} é um vetor que equivale a translação do centro da câmera 1 com relação a 2 e, geralmente, \mathbf{R} e \mathbf{t} são referenciados como os parâmetros extrínsecos da câmera.

A equação anterior muitas vezes é expressa em função das matrizes de projeção das câmeras. Neste caso,

$$\mathbf{P} \simeq \mathbf{A} [\mathbf{I} \mathbf{0}] \quad \text{e} \quad \mathbf{P}' \simeq \mathbf{A}' [\mathbf{R} \mathbf{t}] \quad (3)$$

e o processo de calibração do conjunto binocular consiste em estimar as matrizes \mathbf{P} e \mathbf{P}' .

3 CALIBRAÇÃO BINOCULAR COM UM GABARITO 1D

Em um deslocamento (indexado por i) de um bastão de calibração com n pontos marcados, como esboçado na figura 1, tem-se uma correspondência ($\mathbf{m}_{1ji} \leftrightarrow \mathbf{m}_{2ji}$) para cada ponto $\tilde{\mathbf{M}}_{ji}$ pertencente ao gabarito, onde $j = 1, 2, \dots, n$. Além disso, como as coordenadas de tais correspondências podem ser obtidas com precisão, após N deslocamentos, a matriz fundamental do sistema pode ser estimada com excelente exatidão, usando $n \times N$ correspondências de pontos. Para isto, basta usar um dos diversos métodos já bem conhecidos na literatura (Hartley and Zisserman, 2000).

Do trabalho de Hartley et al. (1992) e Faugeras (1992), sabe-se que, uma vez que a matriz fundamental do conjunto binocular tenha sido estimada, ela pode ser utilizada para obter-se \mathbf{P} e \mathbf{P}' , as matrizes de projeção canônicas do sistema. Além disso, tais matrizes estão relacionadas com as matrizes de projeção euclidianas, \mathbf{P} e \mathbf{P}' , por uma transformação projetiva, \mathbf{T}_{PE} , ou seja,

$$\mathbf{P} = \mathbf{P}' \mathbf{T}_{PE} \quad \text{e} \quad \mathbf{P}' = \mathbf{P}' \mathbf{T}_{PE}, \quad (4)$$

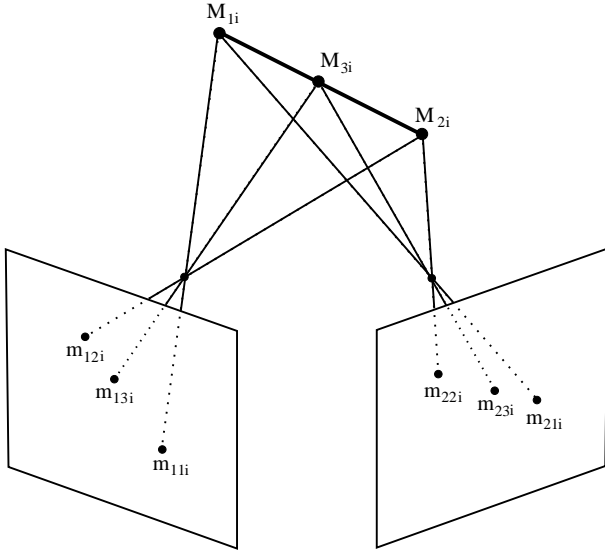


Figura 1: Esquema de calibração binocular utilizando um gabarito com três pontos, cujas posições relativas são conhecidas.

onde

$$\mathbf{T}_{PE} = \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{w}^T & w \end{bmatrix}, \quad (5)$$

com $\mathbf{W} = [\mathbf{w}^T, w]^T$, um vetor 4×1 , representa a equação do plano no infinito no sistema de coordenadas definido por \mathcal{P} e \mathcal{P}' . Dessa forma, para que a calibração do conjunto binocular esteja completa, basta estimar a transformação \mathbf{T}_{PE} . Isto pode ser feito como segue.

Com \mathcal{P} e \mathcal{P}' conhecidas, uma reconstrução projetiva, \mathcal{M}_{ji} , de cada ponto \mathbf{M}_{ji} pode ser obtida, ou seja,

$$\tilde{\mathbf{m}}_{1ji} \simeq \mathcal{P}\tilde{\mathcal{M}}_{ji} \quad \text{e} \quad \tilde{\mathbf{m}}_{2ji} \simeq \mathcal{P}'\tilde{\mathcal{M}}_{ji}.$$

Além disso, da equação (4), sabe-se que \mathcal{M}_{ji} e \mathbf{M}_{ji} estão relacionadas por \mathbf{T}_{PE} , ou seja,

$$\tilde{\mathcal{M}}_{ji} \simeq \mathbf{T}_{PE}\tilde{\mathbf{M}}_{ji}. \quad (6)$$

Assim, considerando que \mathbf{T}_{PE} é dada por (5), a equação (6) pode ser transformada em uma igualdade, ou seja,

$$\mathbf{M}_{ji} = \frac{1}{\tilde{\mathcal{M}}_{ji}^T \mathbf{W}} \mathbf{A}^{-1} \mathcal{M}_{ji}. \quad (7)$$

Além disso, como as distâncias entre os pontos do bastão são conhecidas, as coordenadas dos pontos \mathbf{M}_{ji} podem ser expressas como função dos pontos extremos através da equação

$$\mathbf{M}_{ji} = \lambda_{1j} \mathbf{M}_{1i} + \lambda_{2j} \mathbf{M}_{2i}, \quad (8)$$

onde λ_{1j} e λ_{2j} são constantes conhecidas que dependem das distâncias $\tilde{\mathbf{M}}_{1i} \tilde{\mathcal{M}}_{ji}$ e $\tilde{\mathbf{M}}_{2i} \tilde{\mathcal{M}}_{ji}$. Portanto, substituindo (7)

em (8) e multiplicando ambos os lados da equação por \mathbf{A} , tem-se

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

Agora, após alguma manipulação algébrica, a equação anterior torna-se

$$\mathbf{u}_{ji}^T \mathbf{W} = 0, \quad (9)$$

onde, considerando “ \times ” o produto cruzado e “ \cdot ” o produto escalar entre dois vetores,

$$\mathbf{u}_{ji}^T = \tilde{\mathcal{M}}_{1i}^T + \frac{\lambda_{1j} (\mathcal{M}_{1i} \times \mathcal{M}_{ji}) \cdot (\mathcal{M}_{2i} \times \mathcal{M}_{ji})}{\lambda_{2j} (\mathcal{M}_{2i} \times \mathcal{M}_{ji}) \cdot (\mathcal{M}_{2i} \times \mathcal{M}_{ji})} \tilde{\mathcal{M}}_{2i}^T. \quad (10)$$

Considerando um bastão com $n > 2$ pontos e N deslocamentos do bastão, a equação (9) leva a

$$\begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \\ \vdots \\ \mathbf{U}_N \end{bmatrix} \mathbf{W} = \mathbf{0}, \quad (11)$$

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

A equação anterior permite estimar o vetor \mathbf{W} . Além disso, uma vez que \mathbf{W} seja conhecido, pode-se estimar \mathbf{A} levando em consideração que o comprimento do bastão é conhecido, ou seja,

$$\|\mathbf{M}_{2i} - \mathbf{M}_{1i}\| = L, \quad (12)$$

onde L é o comprimento do gabarito e “ $\|\cdot\|$ ” representa a norma euclidiana de um vetor.

A equação anterior pode ser re-escrita levando em consideração a equação (7), ou seja,

$$\begin{aligned} \mathbf{M}_{2i} - \mathbf{M}_{1i} &= \frac{1}{\tilde{\mathcal{M}}_{2i}^T \mathbf{W}} \mathbf{A}^{-1} \mathcal{M}_{2i} - \frac{1}{\tilde{\mathcal{M}}_{1i}^T \mathbf{W}} \mathbf{A}^{-1} \mathcal{M}_{1i} \\ &= \mathbf{A}^{-1} \mathbf{h}_i, \end{aligned} \quad (13)$$

onde

$$\mathbf{h}_i = \frac{1}{\tilde{\mathcal{M}}_{2i}^T \mathbf{W}} \mathcal{M}_{2i} - \frac{1}{\tilde{\mathcal{M}}_{1i}^T \mathbf{W}} \mathcal{M}_{1i}.$$

Agora, da equação (13), tem-se que equação (12) é equivalente a

$$\mathbf{h}_i^T \mathbf{B} \mathbf{h}_i = L^2, \quad (14)$$

onde $\mathbf{B} \simeq \mathbf{A}^{-T} \mathbf{A}^{-1}$ é a matriz do cone absoluto (Hartley and Zisserman, 2000). Também, levando em consideração que $\mathbf{h}_i^T \mathbf{B} \mathbf{h}_i = [h_{1i}^2, 2h_{1i}h_{2i}, h_{2i}^2, 2h_{1i}h_{3i}, 2h_{2i}h_{3i}, h_{3i}^2]^T \mathbf{b} = \mathbf{V}_i \mathbf{b}$, onde considerou-se $\mathbf{h}_i = [h_{1i}, h_{2i}, h_{3i}]^T$ 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 (14) pode ser representado por

$$\mathbf{V}_i \mathbf{b} = \mathbf{L}^2, \quad (15)$$

onde $\mathbf{L}^2 = [L^2, L^2, \dots, L^2]^T$.

Como \mathbf{B} tem seis graus de liberdade, são necessários $N \geq 6$ deslocamentos para estimar \mathbf{B} unicamente. Assim, com N deslocamentos, pode-se resolver

$$\begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \\ \vdots \\ \mathbf{V}_N \end{bmatrix} \mathbf{b} = \mathbf{L}^2 \quad (16)$$

para que \mathbf{B} seja completamente determinada.

Uma vez que \mathbf{B} for conhecida, a matriz \mathbf{A}^{-1} pode ser obtida de \mathbf{B} através da decomposição de Cholesky (Hartley and Zisserman, 2000). Dessa forma, \mathbf{T}_{PE} está unicamente determinada e o conjunto binocular está completamente calibrado. Em seguida, se necessário, os pontos \mathbf{M}_{1i} e \mathbf{M}_{2i} podem ser obtidos da equação (7).

3.1 Solução não-linear

No método descrito na seção anterior, a matriz \mathbf{T}_{PE} é estimada através da minimização de um critério sem nenhum significado geométrico. Este é um dos motivos que contribuem para que o mesmo não forneça bons resultados na presença de ruído. Contudo, como o conceito de distância não é válido no espaço projetivo, os pontos \mathbf{M}_{ji} não podem ser utilizados para definição de um critério de otimização. Assim, um critério com alguma significância geométrica deve envolver as projeções dos pontos do bastão, \mathbf{m}_{1ji} e \mathbf{m}_{2ji} . Neste trabalho, considerando que as projeções dos pontos do gabarito estão corrompidas de forma independente por um ruído com distribuição gaussiana, propõe-se a solução deste problema através da minimização do critério

$$\mathcal{F}(\mathbf{A}, \mathbf{W}, \mathbf{M}_{ji}) = \sum_{i=1}^N \sum_{j=1}^n \sum_{k=1}^2 \|\mathbf{m}_{kji} - \hat{\mathbf{m}}_{kji}(\mathbf{A}, \mathbf{W}, \mathbf{M}_{ji})\|^2, \quad (17)$$

onde, como já definido, “ $\|\cdot\|$ ” representa a norma euclideana de um vetor,

$$\begin{aligned} \hat{\mathbf{m}}_{1ji}(\mathbf{A}, \mathbf{W}, \mathbf{M}_{ji}) &= \mathcal{P} \mathbf{T}_{PE}^{-1} \widetilde{\mathbf{M}}_{ji}, \\ \hat{\mathbf{m}}_{2ji}(\mathbf{A}, \mathbf{W}, \mathbf{M}_{ji}) &= \mathcal{P}' \mathbf{T}_{PE}^{-1} \widetilde{\mathbf{M}}_{ji} \end{aligned}$$

e \mathbf{M}_{ji} é dado pela equação (8), considerando

$$\mathbf{M}_{2i} = \mathbf{M}_{1i} + L[\sin \theta_i \cos \phi_i, \sin \theta_i \sin \phi_i, \cos \theta_i]^T, \quad (18)$$

onde θ_i e ϕ_i são ângulos que definem a orientação do gabarito.

A minimização da equação (17) envolve a solução de um problema de otimização não-linear. Na visão computacional, o algoritmo de Levenberg-Marquardt (Levenberg, 1944; Marquardt, 1963) é muito utilizado para solucionar tais problemas. Contudo, a complexidade computacional desse algoritmo é η^3 , onde η é número de incógnitas a serem estimadas. A função (17) possui um número excessivo de incógnitas, ou seja, os quatro parâmetros intrínsecos de \mathbf{A} , os quatro elementos de \mathbf{W} , as $3N$ coordenadas dos pontos \mathbf{M}_{1i} e $2N$ ângulos, θ_i e ϕ_i . Assim, o algoritmo Levenberg-Marquardt não é indicado para a calibração binocular com gabaritos 1D. Felizmente, observando as equações (4) e (5), fica claro que uma mudança nos elementos de \mathbf{A} ou \mathbf{W} altera as coordenadas de todas as correspondências de pontos, enquanto que uma alteração em \mathbf{M}_{1i} , θ_i ou ϕ_i reflete-se apenas nos pontos da i -ésima imagem. Assim, o jacobiano da função definida em (17) tem uma estrutura esparsa. Felizmente, para esses casos, existe uma derivação do algoritmo Levenberg-Marquardt, popularizada por Hartley and Zisserman (2000) e referenciada no presente trabalho por Levenberg-Marquardt Dividido, que toma vantagem da estrutura do problema para reduzir a complexidade do algoritmo para apenas η .

A estrutura do algoritmo Levenberg-Marquardt Dividido foi descrita recentemente em (de França et al., 2009). Usando a mesma notação, minimizar (17) é equivalente a resolver o problema

$$\min_{\mathbf{Y}} \|\mathcal{F}(\mathbf{Y}) - \mathbf{X}\|, \quad (19)$$

onde $\mathbf{Y} = [\mathbf{c}^T, \mathbf{d}_1^T, \dots, \mathbf{d}_N^T]^T$ é o vetor de parâmetros a ser estimado e $\mathbf{X} = [\mathbf{X}_1^T, \mathbf{X}_2^T, \dots, \mathbf{X}_n^T]^T$ é um vetor de observações. Mais especificamente, para o caso particular de calibração de um conjunto binocular com gabaritos 1D, $\mathbf{c} = [\alpha, \beta, u_0, v_0, \mathbf{W}^T]^T$, $\mathbf{d}_i = [\mathbf{M}_{1i}^T, \theta_i, \phi_i]^T$ e $\mathbf{X}_i = [\mathbf{a}_1^T, \mathbf{a}_2^T]^T$, onde α, β, u_0 e v_0 são os parâmetros intrínsecos da câmera de referência e $\mathbf{a}_k = [\mathbf{m}_{k1i}^T, \dots, \mathbf{m}_{kni}^T]^T$.

O algoritmo Levenberg-Marquardt Dividido exige uma estimação inicial de todas as incógnitas a serem refinadas. Esta pode ser obtida como segue: a) com os algoritmos lineares dados por (11) e (16), estima-se \mathbf{W} e \mathbf{A} respectivamente; b) os pontos \mathbf{M}_{2i} são conhecidos através de (7), e; c) os ângulos θ_i e ϕ_i são obtidos de (18).

4 RESULTADOS EXPERIMENTAIS

O algoritmo proposto foi implementado no ambiente de programação do MATLAB[®]. Em particular, em todos os experimentos, utilizou-se a implementação do algoritmo Levenberg-Marquardt Dividido apresentada por de França et al. (2009). Também, quando necessário, a matriz funda-

mental do conjunto binocular foi estimada pelo método de Bartoli and Sturm (2004) e a reconstrução projetiva foi realizada através do método ótimo proposto por Hartley and Sturm (1997).

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 a seguir. Além disso, em todos os experimentos, os parâmetros intrínsecos estimados foram comparados aos parâmetros tidos como “reais”. O erro dessa comparação é sempre apresentado de forma relativa, ou seja, as diferenças são apresentadas de forma normalizada em relação ao parâmetro α , como proposto por Triggs (1998).

Os dados sintéticos foram obtidos supondo um conjunto binocular com os parâmetros intrínsecos da tabela 1. A resolução da imagem é feita igual a 512×512 pixels e o gabarito de calibração é suposto igual a 100 cm, com três pontos equidistantes. Além disso, na geração dos dados sintéticos, o ponto $M_{1i} = [X_i, Y_i, Z_i]$, com $\{X_i, Y_i\} \in [-50, 50]$ e $Z_i \in [120, 200]$, e os ângulos da equação (18), $\theta_i \in [\pi/6, 5\pi/6]$ e $\phi_i \in [\pi, 2\pi]$, variaram aleatoriamente, mas de acordo com uma distribuição uniforme. Tais dados foram utilizados para avaliar o desempenho do método com respeito ao nível de ruído presente nos dados. Por isso, nesse experimento, ruído gaussiano de média zero e desvio padrão σ foram acrescentados às projeções nas imagens dos pontos do gabarito, com σ variando de 0,1 a 2,0 pixels. Por sua vez, em cada simulação, as projeções com ruído serviram de dados de entrada para o algoritmo proposto.

Para cada nível de ruído, foram realizadas 200 simulações e a mediana de cada parâmetro intrínseco foi retida. Essas medianas foram comparadas com os parâmetros da tabela 1. Os resultados da comparação são apresentados na figura 2. As figuras 2(a) e 2(b) apresentam, respectivamente, os erros relativos dos parâmetros intrínsecos das câmeras 1 e 2, estimados com o algoritmo linear proposto. Analisando os resultados, observa-se que, para o método linear, o erro aumenta rapidamente com σ . Por exemplo, quando σ triplica de 0,5 para 1,5 pixel, o erro correspondente varia de (aproximadamente) 2,5 para 15 %, ou seja, um fator de seis vezes. Apesar disso, observando as figuras 2(c) e 2(d), que apresentam o refinamento dos resultados das figuras 2(a) e 2(b) obtido pelo método não-linear proposto, nota-se que tal método refina satisfatoriamente os parâmetros de ambas as câmeras. Em especial, para $\sigma = 1,5$ pixel (valor maior do que os encontrados na prática), tem-se um erro da ordem de 0,4 %.

Para adquirir os dados reais, foi utilizado um gabarito com cinco esferas de plástico negro, com 2 cm de diâmetro cada uma. Tais esferas foram fixadas em um bastão de metal. A

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

Câmera 1				Câmera 2			
α_x	α_y	u_0	v_0	α_x	α_y	u_0	v_0
715	712	325	232	700	730	335	222

distância entre os centróides de esferas vizinhas foi medida com auxílio de um paquímetro. Manualmente, o gabarito foi movido em diferentes posições, a uma distância de aproximadamente 100 cm do conjunto binocular. Durante o movimento, 126 pares de imagens foram adquiridos (uma das imagens é apresentada na figura 3). Com auxílio de uma ferramenta de *blob* gaussiano (Lindeberg, 1994, Cap. 7), em cada uma das imagens, as coordenadas dos centróides das esferas negras foram estimadas de forma automática e foram utilizadas como dados de entrada para o algoritmo proposto. O resultado deste experimento é apresentado na tabela 2 e é referenciado como “Proposto”.

Os parâmetros intrínsecos da câmera utilizada nos testes também foram estimados pelo método proposto por Zhang (2000). O gabarito utilizado neste método é um plano com um padrão quadriculado. Neste trabalho, utilizou-se um gabarito com 156 pontos (figura 4). Durante a calibração, foram realizados 22 deslocamentos do conjunto binocular e, com cada câmera, uma seqüência com 22 imagens foi capturada. Em seguida, a implementação de Bouguet (2004) do algoritmo de Zhang (2000) foi utilizada para estimar os parâmetros intrínsecos. Tais parâmetros foram adotados como os parâmetros reais da câmera, são apresentados na tabela 2 e referenciados como “Clássico”.

A última linha da tabela 2 mostra o erro entre o método “Proposto” e o “Clássico” normalizada por α , com proposto por Triggs (1998). Observa-se que tal erro é da ordem de 2% para os parâmetros α e β e da ordem de 4% para as coordenadas do ponto principal. Como os métodos analisados são diferentes, é natural que exista uma diferença entre os dois resultados. Além disso, o método de Zhang (2000), como a maioria dos métodos clássicos, tem dificuldade em estimar os parâmetros u_0 e v_0 (isso explica o erro maior nesses parâmetros). Contudo, a principal razão dessa diferença entre os resultados é a forma rudimentar de construção do gabarito. Isso faz com que os pontos do gabarito não sejam exatamente lineares, invalidando o modelo utilizado. Outra fonte de erros foi a ferramenta utilizada para extração das coordenadas das projeções dos pontos do gabarito, pois, devido a diferenças de iluminação, a ferramenta não conseguiu segmentar as esferas como um círculo perfeito. Dessa forma, os centróides das esferas não foram conhecidos com exatidão. Por isso, considerando todos esses fatores, os resultados obtidos no experimento são satisfatórios.

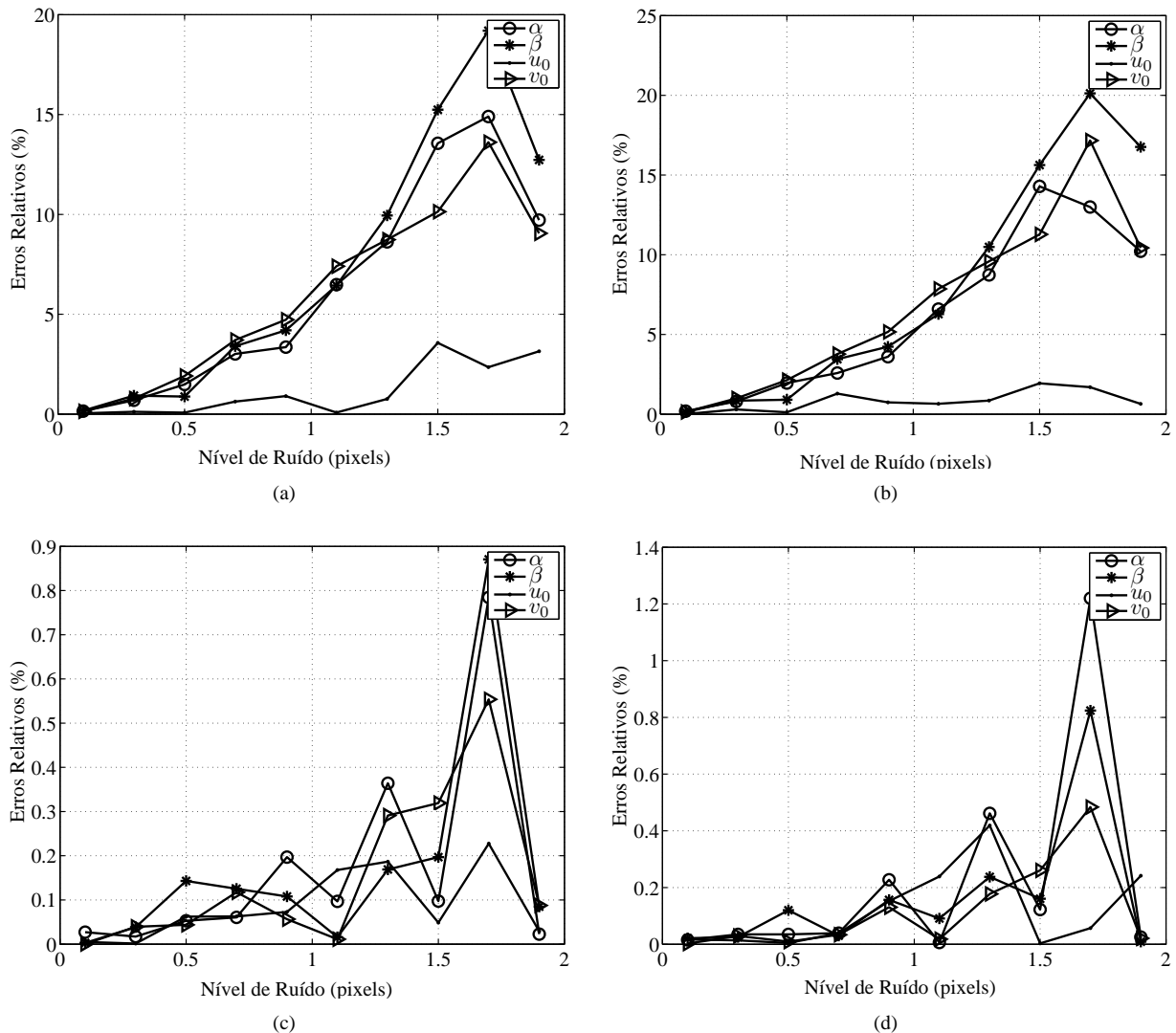


Figura 2: Exatidão com relação ao nível de ruído: a) parâmetros da câmera 1 (método linear); b) parâmetros da câmera 2 (método linear); c) parâmetros da câmera 1 (método não-linear), e; d) parâmetros da câmera 2 (método não-linear).

5 CONCLUSÃO

O presente trabalho demonstra que uma abordagem diferente pode ser aplicada à técnica de calibração com gabaritos 1D se, ao invés de uma única câmera, um conjunto binocular for considerado. Isto porque um conjunto binocular insere restrições extras ao problema de calibração com um gabarito 1D. Tais restrições tornam possível calibrar o conjunto binocular com o gabarito 1D mesmo que este realize um deslocamento desconhecido, sem que nenhum de seus pontos esteja fixo e sem nenhuma informação prévia a respeito das câmeras. Tal método baseia-se na estimação de uma transformação que, uma vez que a matriz fundamental do sistema é conhecida, permite atualizar uma calibração projetiva para uma calibração euclidiana.

Foi proposto um método linear que dá uma estimação inicial dos parâmetros das câmeras e um método não-linear, baseado no critério da máxima verossimilhança, que refina os parâmetros iniciais de forma que estes possam ser utilizados em problemas práticos. Resultados experimentais realizados em imagens sintéticas e reais mostram que o método proposto produz bons resultados mesmo na presença de ruído e, dessa forma, constitui um método prático e bastante flexível de calibrar um ou mais conjuntos binoculares.

No presente trabalho, a distorção radial da câmera foi desprezada. Contudo, esta pode ser facilmente considerada modificando o modelo da câmera. Para isto, a função a ser minimizada pelo algoritmo não-linear deve ser ligeiramente alterada.

Tabela 2: Parâmetros intrínsecos do conjunto binocular estimados através do algoritmo proposto e de um método clássico.

Método	Câmera 1				Câmera 2			
	α	β	u_0	v_0	α	β	u_0	v_0
Proposto	714,90	710,07	480,25	344,44	720,17	713,45	450,50	337,48
Clássico	703,16	696,15	445,54	328,04	710,93	701,32	421,22	364,06
Erro	1,67	1,98	4,94	2,33	1,30	1,71	4,12	-3,74



Figura 3: Gabarito 1D utilizado na avaliação do método proposto.

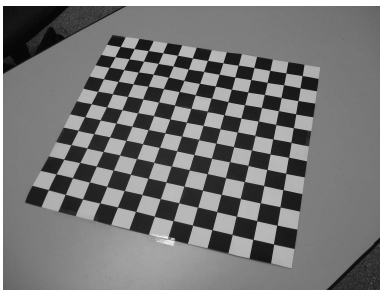


Figura 4: Gabarito plano utilizado na comparação do método proposto.

AGRADECIMENTOS

Os autores agradecem a CAPES e ao FAEPE/UEL pelo financiamento das pesquisas. Também, um agradecimento especial aos comentários anônimos dos revisores que fizeram com que as contribuições do trabalho fossem destacadas mais claramente.

REFERÊNCIAS

Bartoli, A. and Sturm, P. (2004). Nonlinear estimation of the fundamental matrix with minimal parameters, *IEEE Trans. on pattern analysis and machine intelligence* **26**(3): 426–432.

Bouguet, J. T. (2004). *Camera Calibration Toolbox for Matlab*, MRL - Intel Corp. [online] http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.

Cao, X. and Foroosh, H. (2004). Camera calibration without metric information using 1D objects, *International Conference on Image Processing* **2**: 1349–1352.

Carranza, J., Theobalt, C., Magnor, M. A. and Seidel, H.-P. (2003). Free-viewpoint video of human actors, *ACM Transactions on Graphics* **22**(3): 569–577.

de França, J. A., França, M. B. d. M., Koyama, M. H. and da Silva, T. P. (2009). Uma implementação do algoritmo levemberg-marquardt dividido para aplicações em visão computacional, *Semina: Ciências Exatas/Tecnológicas* **30**(1): 51–62.

de França, J. A., Stemmer, M. R., França, M. B. d. M. and Alves, E. G. (2010). Revisiting Zhang's 1D calibration algorithm, *Pattern Recognition* **43**(3): 1180–1187.

Dornaika, F. and Chung, R. (2001). An algebraic approach to camera self-calibration, *Computer Vision and Image Understanding* **83**(3): 195–215.

Faugeras, O. D. (1992). What can be seen in three dimensions with an uncalibrated stereo rig?, *II European Conference on Computer Vision*, pp. 563–578.

Fujii, T. and Tanimoto, M. (2002). Free viewpoint TV system based on ray-space representation, in B. Javidi & F. Okano (ed.), *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 4864 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pp. 175–189.

Hammarstedt, P., Sturm, P. and Heyden, A. (2005). Closed-form solutions and degenerate cases for camera calibration with one-dimensional objects, *Proceedings of X IEEE International Conference on Computer Vision*.

Hartley, R. (1997). Self-calibration of stationary cameras, *International Journal of Computer Vision* **22**(1): 5–23.

Hartley, R., Gupta, R. and Chang, T. (1992). Stereo from uncalibrated cameras, *II European Conference on Computer Vision*, pp. 579–587.

Hartley, R. and Sturm, P. (1997). Triangulation, *Computer Vision and Image Understanding* **68**(2): 146–157.

- Hartley, R. and Zisserman, A. (2000). *Multiple View Geometry in Computer Vision*, Cambridge University Press.
- Kanade, T., Rander, P. and Narayanan, P. J. (1997). Virtualized reality: Constructing virtual worlds from real scenes, *IEEE Multimedia, Immersive Telepresence* **4**(1): 34–47.
- Kojima, Y., Fujii, T. and Tanimoto, M. (2005). New multiple camera calibration method for a large number of cameras, *Proceedings of VIII Videometrics Conference*.
- Lenz, R. and Tsai, R. (1988). Techniques for calibration of the scale factor and image center for high accuracy 3-d machine vision metrology, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **10**(5): 713–720.
- Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares, *Quarterly of Applied Math.* p. 164–168.
- Lindeberg, T. (1994). *Scale-Space Theory in Computer Vision*, 1a. edição edn, Kluwer Academic Publishers.
- Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters, *J. Soc. Indust. Appl. Math.* **11**(2): 431–441.
- Maybank, S. and Faugeras, O. (1992). A theory of self-calibration of a moving camera, *The International Journal of Computer Vision* **8**(2): 123–151.
- Mendonça, P. R. S. and Cipolla, R. (1999). A simple technique for self-calibration, *IEEE Conference on Computer Vision and Pattern Recognition*, p. 1500–1505.
- Sturm, P. F. and Maybank, S. J. (1999). On plane-based camera calibration: A general algorithm, singularities, applications, *IEEE Conference on Computer Vision and Pattern Recognition*, p. 432–437.
- Triggs, B. (1998). Autocalibration from planar scenes, *European Conference on Computer Vision*, pp. 89–105.
- 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.
- Wu, F., Hu, Z. and Zhu, H. (2005). Camera calibration with moving one-dimensional objects, *Pattern Recognition* **40**(1): 755–765.
- 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**(7): 892–899.