

# SERVOVISÃO DIRETA DESACOPLADA

GERALDO SILVEIRA\*, LUIZ G. B. MIRISOLA†, PASCAL MORIN‡

\*CTI Renato Archer, DRVC, Campinas/SP, Brasil

†Universidade Federal do ABC, CMCC, Santo Andre/SP, Brasil

‡Université Pierre et Marie Curie, ISIR, Paris, França

Emails: [Geraldo.Silveira@cti.gov.br](mailto:Geraldo.Silveira@cti.gov.br), [luiz.mirisola@ufabc.edu.br](mailto:luiz.mirisola@ufabc.edu.br), [morin@isir.upmc.fr](mailto:morin@isir.upmc.fr)

**Abstract**— This article addresses the problem of direct vision-based control where the equilibrium is defined via a reference image. Direct methods refer to intensity-based nonmetric techniques to perform that stabilization. Intensity-based strategies provide for higher accuracy, whereas not requiring metric information improves their versatility. However, existing direct techniques either have a coupled error dynamics, or are designed for planar objects only. This paper proposes a new direct technique that decouples the translational motion from the rotational one for the general case of planar and nonplanar targets under general translations and rotations. Furthermore, for the important case of a fronto-parallel planar object, the proposed technique leads to a fully diagonal interaction matrix. The equilibrium is made locally exponentially stable for all those cases. These improvements are theoretically proven and experimentally demonstrated using a 6-DoF robotic arm.

**Keywords**— Visual servoing, intensity-based methods, nonmetric methods.

**Resumo**— Este artigo aborda o problema de controle de robôs baseado em métodos diretos de visão computacional, onde o estado de equilíbrio é definido através de uma imagem de referência. Métodos diretos se referem às estratégias não métricas e baseadas na intensidade dos pixels. Essas técnicas aumentam a precisão e são mais versáteis por não dependerem de qualquer informação métrica do objeto. Todavia, as técnicas diretas existentes possuem uma dinâmica do erro de controle acoplada, ou são restritas a objetos planares. Este artigo propõe uma nova técnica direta que desacopla o movimento translacional do rotacional para o caso geral de objetos planares e não planares sob movimentos translacionais e rotacionais. Além disso, para o caso importante de objetos fronto-paralelos, a técnica proposta conduz a uma matriz de interação completamente diagonal. O estado de equilíbrio é localmente exponencialmente estável para todos estes casos. Todas essas melhorias são teoricamente provadas e experimentalmente demonstradas usando um braço robótico de 6 graus de liberdade.

**Palavras-chave**— Controle baseado em visão, métodos baseados em intensidade, métodos não métricos.

## 1 Introdução

Servovisão se refere ao controle de um robô em relação à cena através da realimentação de imagens. Uma aplicação típica consiste em estabilizar o robô em uma pose definida via uma imagem de referência. Embora existam diversas soluções para este problema (Chaumette e Hutchinson, 2006), a sua maioria: 1) é baseada em extração de primitivas geométricas, como pontos e retas. Logo, elas dependem da existência delas, da sua correta associação, e do ajuste dos parâmetros destes algoritmos; e 2) requer informação métrica (pelo menos aproximada) do objeto para que a lei de controle seja estabilizante. Isto é válido até para os métodos de servovisão baseado em imagem, onde a profundidade intervém na matriz de interação. Estes dois tópicos são discutidos a seguir.

Técnicas de estimação por visão computacional são geralmente classificadas como baseadas em características visuais ou em intensidade. Mesmo com as dificuldades supracitadas, a maioria dos esquemas de servovisão são de fato baseados naquelas primitivas geométricas. Isto se deve provavelmente ao seu relativamente grande domínio de convergência. Já os métodos baseados em intensidade não incluem etapas de extração e associação de características. Elas utilizam diretamente o valor dos pixels para efetuar a estimação dos parâmetros desejados. Portanto, elas utilizam a imagem de forma densa, o que permite grande versatilidade e precisão, além de robustez a mudanças de iluminação arbitrárias (Silveira e Malis, 2010).

Em relação às técnicas de servovisão não métricas, apesar de serem mais versáteis e robustas (Thaler e Goodale, 2010), existem poucos trabalhos nesta classe. Uma razão possível é a dificuldade em definir um erro de controle localmente difeomórfico à pose da câmera. Um exemplo antigo dessas técnicas foi a navegação realizada por um robô terrestre em (Beardsley et al., 1995). Recentemente, foi desenvolvida uma técnica baseada em intensidade, não métrica e geral chamada de Servovisão Direta (SD) (Silveira e Malis, 2012). Ela é geral no sentido em que todos os seis graus de liberdade (g.d.l.) de um robô são estabilizados em relação a objetos planares ou não, sob movimentos rotacionais e translacionais. Não obstante, a dinâmica do erro de controle na SD é acoplada.

Este artigo propõe uma nova técnica de SD desacoplada. A idéia de desacoplamento brevemente apresentada em (de Plinval et al., 2011) para objetos planares é melhorada e estendida para quaisquer superfícies e movimentos, como em (Silveira e Malis, 2012). O movimento translacional é assim desacoplado do rotacional para o caso geral. Além disso, para o importante caso de um objeto planar e fronto-paralelo, é mostrado que a matriz de interação é completamente diagonal. O novo erro de controle é provado localmente difeomórfico à pose da câmera, cujo equilíbrio é mostrado ser localmente exponencialmente estável. Estes avanços são também confirmados experimentalmente com dados sintéticos e reais, para objetos planares e não planares, simulando ou utilizando uma câmera montada em um robô holonômico de 6 g.d.l.

## 2 Fundamentação Teórica

Esta seção define a notação e alguns modelos essenciais. Permita que a norma Euclidiana, uma estimativa, e uma versão transformada da variável  $\mathbf{v}$  seja escrita como  $\|\mathbf{v}\|$ ,  $\hat{\mathbf{v}}$  e  $\mathbf{v}'$ , respectivamente. Um asterisco, e.g.,  $\mathbf{v}^*$ , indica que a variável  $\mathbf{v}$  é definida em relação ao sistema de coordenadas de referência  $\mathcal{F}^*$ . As notações  $[\mathbf{w}]_{\times}$  e  $\text{vex}([\mathbf{w}]_{\times})$  representam, respectivamente, a matriz antisimétrica associada ao vetor  $\mathbf{w} = [w_1, w_2, w_3]^{\top}$  e o seu mapeamento inverso:

$$[\mathbf{w}]_{\times} = \begin{bmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{bmatrix}, \text{vex}([\mathbf{w}]_{\times}) = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}.$$

### 2.1 Geometria Projetiva

A relação geral entre pontos correspondentes  $\mathbf{p} \leftrightarrow \mathbf{p}^*$  em duas imagens é descrita em geometria projetiva por (Faucher et al., 2001):

$$\mathbf{p} \propto \mathbf{G} \mathbf{p}^* + \rho^* \mathbf{e} \in \mathbb{P}^2, \quad (1)$$

onde o símbolo “ $\propto$ ” denota proporcionalidade,  $\mathbf{G} \in \text{SL}(3)$  é uma homografia projetiva induzida por um plano dominante,  $\mathbf{e} \in \mathbb{P}^2$  denota o epipolo, e  $\rho^* \in \mathbb{R}$  é a paralaxe projectiva do ponto 3D cuja projeção na imagem de referência  $\mathcal{I}^*$  é  $\mathbf{p}^*$ , relativamente àquele plano (vide Fig. 1). Esta paralaxe é proporcional à distância deste ponto 3D ao plano e é inversamente proporcional à sua profundidade. O epipolo é proporcional à translação entre o sistema de coordenadas corrente e o de referência; e a homografia  $\mathbf{G}$  pode ser caracterizada como

$$\mathbf{G} \propto \mathbf{G}_{\infty} + \mathbf{e} \mathbf{q}^{*\top}, \quad (2)$$

onde  $\mathbf{q}^* \in \mathbb{R}^3$  é uma representação da linha no infinito do plano na imagem de referência, e  $\mathbf{G}_{\infty} \in \text{SL}(3)$  é a homografia no infinito. Esta homografia é proporcional a uma matriz que é similar a uma matriz de rotação. Um método para estimar todos os parâmetros geométricos  $\{\mathbf{G}, \mathbf{e}, \rho^*\}$  em (1) é apresentado na sequência.

### 2.2 Servovisão Direta: Aspectos da Estimação

Em Servovisão Direta (SD) (Silveira e Malis, 2012), as intensidades dos pixels são diretamente utilizadas na estimação dos parâmetros necessários, sem etapas de extração ou associação de características visuais. O arcabouço para esta estimação é o registro direto de imagens, que consiste em obter os parâmetros que melhor transformam a imagem corrente de forma que a intensidade de cada pixel  $\mathcal{I}(\mathbf{p})$  na imagem transformada esteja o mais próximo possível da intensidade do pixel correspondente na imagem de referência  $\mathcal{I}^*(\mathbf{p}^*)$ .

Portanto, primeiro deve-se definir um modelo adequado de transformações fotogeométricas. Para isso escolhe-se um modelo apropriado de mudanças de iluminação, e.g., (Silveira e Malis, 2010),

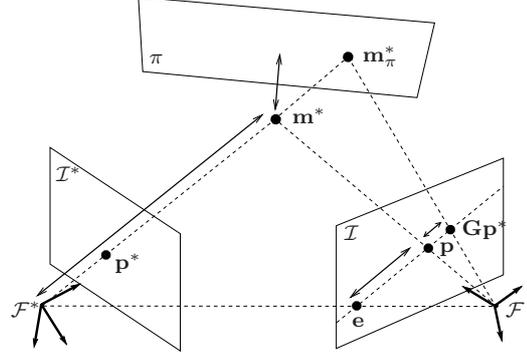


Figura 1: Geometria entre duas imagens.

e um modelo de distorção a partir de (1). Mais formalmente, a ação deste modelo de transformação nos pixels é definida por:

$$\mathcal{I}'_{gh}(\mathbf{g}, \mathbf{h}, \mathbf{p}^*) = \mathcal{S}(\mathbf{p}^*) \cdot \mathcal{I}(\mathbf{w}(\mathbf{g}, \mathbf{p}^*)) + \beta \geq 0, \quad (3)$$

onde o operador “ $\cdot$ ” representa a multiplicação elemento-a-elemento de matrizes, e os parâmetros geométricos e fotométricos estão respectivamente agrupados em  $\mathbf{g} = \{\mathbf{G}, \mathbf{e}, \rho^*\}$  e  $\mathbf{h} = \{\mathcal{S}, \beta\}$ , onde  $\mathcal{S}$  é visto como uma superfície que compensa variações locais e globais de iluminação, e  $\beta \in \mathbb{R}$ .

Um sistema fotogeométrico de registro direto de imagens pode então ser definido como o seguinte problema de otimização não linear:

$$\min_{\substack{\mathbf{g}=\{\mathbf{G}, \mathbf{e}, \rho^*\} \\ \mathbf{h}=\{\mathcal{S}, \beta\}}} \frac{1}{2} \sum_{i=1}^n \underbrace{[\mathcal{I}'_{gh}(\mathbf{g}, \mathbf{h}, \mathbf{p}_i^*) - \mathcal{I}^*(\mathbf{p}_i^*)]^2}_{d_i(\mathbf{g}, \mathbf{h})}, \quad (4)$$

que busca minimizar a norma do vetor das diferenças das imagens  $\mathbf{d}(\mathbf{g}, \mathbf{h}) = \{d_i(\mathbf{g}, \mathbf{h})\}_{i=1}^n$  onde  $n$  é o número de pixels. O problema de otimização não linear em (4) pode ser resolvido por métodos iterativos clássicos como Gauss–Newton. Para uma solução com melhores propriedades de convergência, o leitor pode consultar (Silveira e Malis, 2010).

### 2.3 Servovisão Direta: Aspectos de Controle

Os parâmetros geométricos  $\mathbf{g} = \{\mathbf{G}, \mathbf{e}, \rho^*\}$  estimados conforme descrito na Seção 2.2 podem ser usados em controle baseado em visão. Os erros de controle não métricos translacionais e rotacionais propostos na SD (Silveira e Malis, 2012), i.e.,  $\boldsymbol{\varepsilon}_v \in \mathbb{R}^3$  e  $\boldsymbol{\varepsilon}_\omega \in \mathbb{R}^3$  respectivamente, são:

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{\varepsilon}_v \\ \boldsymbol{\varepsilon}_\omega \end{bmatrix} = \begin{bmatrix} (\mathbf{H} - \mathbf{I}) \mathbf{m}^{*'} + \rho^* \mathbf{e}' \\ \vartheta \boldsymbol{\mu} \end{bmatrix}, \quad (5)$$

onde

$$\mathbf{H} = \mathbf{K}^{-1} \mathbf{G} \mathbf{K}, \quad \mathbf{e}' = \mathbf{K}^{-1} \mathbf{e}, \quad \mathbf{m}^{*'} = \mathbf{K}^{-1} \mathbf{p}^*, \quad (6)$$

e  $\rho^* \in \mathbb{R}$  é a paralaxe projectiva do ponto de controle escolhido  $\mathbf{p}^* \in \mathbb{P}^2$ . A matriz  $\mathbf{K} \in \mathbb{R}^{3 \times 3}$  contém os parâmetros intrínsecos da câmera. O erro rotacional  $\boldsymbol{\varepsilon}_\omega$  em (5) é computado a partir da ho-

mografia  $\mathbf{H} \in \mathbb{R}^{3 \times 3}$  via

$$\mathbf{r} = \frac{1}{2} \text{vex}(\mathbf{H} - \mathbf{H}^\top), \quad (7)$$

$$\vartheta = \begin{cases} \text{real}(\text{asin}(\|\mathbf{r}\|)), & \text{if } \text{tr}(\mathbf{H}) \geq 1, \\ \pi - \text{real}(\text{asin}(\|\mathbf{r}\|)), & \text{caso contrário,} \end{cases} \quad (8)$$

$$\boldsymbol{\mu} = \frac{\mathbf{r}}{\|\mathbf{r}\|}, \quad (9)$$

onde a função  $\text{tr}(\cdot)$  denota o traço de uma matriz. Se  $\|\mathbf{r}\| = 0$ , então  $\boldsymbol{\mu}$  é indeterminado e pode ser escolhido arbitrariamente (e.g.,  $\boldsymbol{\mu} = [0, 0, 1]^\top$ ).

Sejam  $\mathbf{v} = [\mathbf{v}^\top, \boldsymbol{\omega}^\top]^\top \in \mathbb{R}^6$ , respectivamente, as velocidades translacionais e rotacionais da câmera. Foi provado em (Silveira e Malis, 2012) que a lei de controle não métrica

$$\mathbf{v} = \lambda \boldsymbol{\varepsilon}, \quad (10)$$

com  $\lambda > 0$ , estabiliza localmente o equilíbrio  $\boldsymbol{\varepsilon} = \mathbf{0}$  se o ponto de controle (6) é escolhido de forma que a sua paralaxe seja suficientemente pequena.

### 3 Técnica Desacoplada Proposta

Esta seção apresenta a técnica proposta de SD desacoplada. Esta nova técnica é baseada em intensidade, é não métrica, e os movimentos translacionais são desacoplados dos rotacionais. Isto é demonstrado para o caso geral de objetos planares e não planares, sob movimentos quaisquer de translação e rotação. Na sequência, mostramos que a técnica conduz a um sistema completamente desacoplado para o importante caso de um objeto planar fronto-paralelo.

#### 3.1 Erro de Controle e Algumas Propriedades

Como na SD original, o novo erro de controle é construído a partir dos parâmetros estimados  $\mathbf{g} = \{\mathbf{G}, \mathbf{e}, \rho^*\}$  (vide Seção 2.2). Por simplicidade, considere uma versão reduzida daquele erro (5)

$$\boldsymbol{\varepsilon}' = \begin{bmatrix} \boldsymbol{\varepsilon}'_v \\ \boldsymbol{\varepsilon}'_\omega \end{bmatrix} = \begin{bmatrix} (\mathbf{H} - \mathbf{I})\mathbf{m}^{*'} + \rho^* \mathbf{e}' \\ \text{vex}(\mathbf{H} - \mathbf{H}^\top) \end{bmatrix}, \quad (11)$$

cujos erros de controle rotacional é equivalente ao original via  $\boldsymbol{\varepsilon}'_\omega = 2\vartheta^{-1}\|\mathbf{r}\|\boldsymbol{\varepsilon}_\omega$ . Mais ainda, em torno do equilíbrio pode ser mostrado que  $\boldsymbol{\varepsilon}'_\omega \approx 2\boldsymbol{\varepsilon}_\omega$  pois  $\vartheta^{-1}\|\mathbf{r}\| \approx 1$ . Este erro de controle não métrico é geral no sentido que lida com objetos planares ou não, e sob movimentos translacionais e rotacionais entre os sistemas de coordenadas inicial e de referência. Como na relação geral (1), não se assume que  $\rho^* \mathbf{e}' = \mathbf{0}$ .

O novo erro de controle é definido como

$$\bar{\boldsymbol{\varepsilon}} = \mathbf{M} \boldsymbol{\varepsilon}', \quad (12)$$

onde

$$\mathbf{M} = \begin{bmatrix} 2\mathbf{I} & [\mathbf{m}^{*'}]_\times \\ -[\mathbf{c}^{*'}]_\times & \mathbf{I} \end{bmatrix} \quad (13)$$

é uma matriz  $(6 \times 6)$  constante, e  $\mathbf{c}^{*'}$  é um vetor. A escolha deste vetor será detalhada mais adiante. É importante notar que (12) não requer nenhuma informação métrica do objeto, independente da sua forma ou do movimento da câmera.

**Theorem 3.1** *O erro de controle  $\bar{\boldsymbol{\varepsilon}}$  definido em (12) é localmente difeomórfico à pose da câmera em torno da imagem de referência desde que*

$$\mathbf{m}^{*\top} \mathbf{c}^{*' \neq 2, \quad (14)$$

$$\mathbf{m}^{*\top} \mathbf{q}^{*' \neq 2(z^*)^{-1}, \quad (15)$$

onde  $z^* > 0$  é a profundidade do ponto de controle  $\mathbf{p}^*$ , e  $\mathbf{q}^{*' = \mathbf{K}^\top \mathbf{q}^*$  define o plano dominante do objeto. Ademais, a linearização da matriz de interação  $\bar{\mathbf{L}}$  nesta configuração é dada por

$$\bar{\mathbf{L}}|_{\bar{\boldsymbol{\varepsilon}}=\mathbf{0}} = - \begin{bmatrix} \frac{2}{z^*} \mathbf{I} + [\mathbf{m}^{*'}]_\times [\mathbf{q}^{*'}]_\times & \mathbf{0} \\ -\frac{1}{z^*} [\mathbf{c}^{*'}]_\times + [\mathbf{q}^{*'}]_\times & 2\mathbf{I} + [\mathbf{c}^{*'}]_\times [\mathbf{m}^{*'}]_\times \end{bmatrix} \quad (16)$$

**Prova:** Dado que  $\bar{\boldsymbol{\varepsilon}} = \mathbf{M} \boldsymbol{\varepsilon}'$  e  $\mathbf{M}$  é uma matriz constante,  $\bar{\mathbf{L}} = \mathbf{M} \mathbf{L}'$  com  $\bar{\mathbf{L}}$  (resp.  $\mathbf{L}'$ ) a matriz de interação de  $\bar{\boldsymbol{\varepsilon}}$  (resp.  $\boldsymbol{\varepsilon}'$ ). Mostra-se em (Silveira e Malis, 2012) que a linearização de  $\mathbf{L}'$  em  $\boldsymbol{\varepsilon}' = \mathbf{0}$  é

$$\mathbf{L}'|_{\boldsymbol{\varepsilon}'=\mathbf{0}} = - \begin{bmatrix} \frac{1}{z^*} \mathbf{I} & -[\mathbf{m}^{*'}]_\times \\ [\mathbf{q}^{*'}]_\times & 2\mathbf{I} \end{bmatrix}, \quad (17)$$

e (16) segue de (13) e (17). Para concluir a prova, falta mostrar que a matriz acima é inversível quando ambas condições (14) e (15) são satisfeitas. Dado que  $\bar{\mathbf{L}}|_{\bar{\boldsymbol{\varepsilon}}=\mathbf{0}} = \mathbf{M} \mathbf{L}'|_{\boldsymbol{\varepsilon}'=\mathbf{0}}$ ,  $\bar{\mathbf{L}}|_{\bar{\boldsymbol{\varepsilon}}=\mathbf{0}}$  é inversível desde que  $\mathbf{M}$  e  $\mathbf{L}'|_{\boldsymbol{\varepsilon}'=\mathbf{0}}$  sejam ambas invertíveis. Consideremos  $\mathbf{M}$ . Pela fórmula de Schur, e (13),  $\mathbf{M}$  é inversível desde que a matriz

$$\mathbf{M}_0 = 2\mathbf{I} + [\mathbf{c}^{*'}]_\times [\mathbf{m}^{*'}]_\times \quad (18)$$

seja inversível. Considere uma matriz  $\mathbf{Q}$  tal que  $\mathbf{Q} \mathbf{m}^{*' = \|\mathbf{m}^{*'}\| \mathbf{b}_3$  com  $\mathbf{b}_3 = [0, 0, 1]^\top$  o terceiro vetor canônico. Usando o fato de que  $\forall \mathbf{x} \in \mathbb{R}^3$ ,  $\mathbf{Q}[\mathbf{x}]_\times \mathbf{Q}^\top = [\mathbf{Q}\mathbf{x}]_\times$ , verifica-se que, multiplicando (18) pela esquerda por  $\mathbf{Q}$  e pela direita por  $\mathbf{Q}^\top$ ,

$$\mathbf{Q} \mathbf{M}_0 \mathbf{Q}^\top = 2\mathbf{I} + [\mathbf{Q}\mathbf{c}^{*'}]_\times [\|\mathbf{m}^{*'}\| \mathbf{b}_3]_\times. \quad (19)$$

Dado que  $\det(\mathbf{Q} \mathbf{M}_0 \mathbf{Q}^\top) = \det(\mathbf{M}_0)$ , tem-se que

$$\det(\mathbf{M}_0) = 2(2 - \langle \mathbf{Q}\mathbf{c}^{*'}, \|\mathbf{m}^{*'}\| \mathbf{b}_3 \rangle)^2 \quad (20)$$

$$= 2(2 - \langle \mathbf{c}^{*'}, \mathbf{m}^{*' \rangle}^2. \quad (21)$$

Isto mostra que  $\mathbf{M}_0$  (e consequentemente  $\mathbf{M}$ ) é inversível se a condição (14) for satisfeita. Através de um raciocínio similar, mostra-se que  $\bar{\mathbf{L}}|_{\bar{\boldsymbol{\varepsilon}}=\mathbf{0}}$  é inversível se a condição (15) for satisfeita.  $\square$

Uma propriedade interessante da matriz de interação (16) é a sua estrutura bloco-triangular. Esta propriedade é usada a seguir para propor uma lei de controle estabilizadora simples e com propriedades de desacoplamento.

#### 3.2 Lei de Controle e Análise de Estabilidade

Considere uma câmera montada em um robô holonômico de 6 g.d.l., observando um objeto estático, rígido e de forma desconhecida. Permita que as entradas de controle sejam as velocidades translacionais e rotacionais da câmera, agrupadas em  $\mathbf{v} \in \mathbb{R}^6$ . A lei de controle não métrica proposta é

$$\mathbf{v} = \boldsymbol{\Lambda} \bar{\boldsymbol{\varepsilon}}, \quad (22)$$

com uma matriz de ganho diagonal  $\mathbf{A} = \text{diag}(\mathbf{A}_v, \mathbf{A}_w)$ , e  $\mathbf{A}_v = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$ ,  $\mathbf{A}_w = \text{diag}(\lambda_4, \lambda_5, \lambda_6)$ .

**Theorem 3.2** *A lei de controle não métrica (22) garante a estabilidade exponencial local do equilíbrio  $\bar{\mathbf{e}} = \mathbf{0}$  se seguintes condições são satisfeitas:*

$$\text{sign}(\lambda_1) = \text{sign}(\lambda_2) = \text{sign}\left(\frac{2}{z^*} - \mathbf{m}^{*\top} \mathbf{q}^{*'}\right) \quad (23)$$

$$\text{sign}(\lambda_4) = \text{sign}(\lambda_5) = \text{sign}(2 - \mathbf{m}^{*\top} \mathbf{c}^{*'}) \quad (24)$$

$$\lambda_3, \lambda_6 > 0. \quad (25)$$

**Prova:** A partir da matriz de interação (16) e aplicando a lei de controle (22), a linearização do sistema fechado em  $\bar{\mathbf{e}} = \mathbf{0}$  é dada por

$$\dot{\hat{\mathbf{e}}} = \mathbf{A} \bar{\mathbf{e}}, \quad (26)$$

com

$$\mathbf{A} = \bar{\mathbf{L}}|_{\bar{\mathbf{e}}=\mathbf{0}} \mathbf{A} = \begin{bmatrix} \mathbf{A}_{vv} & \mathbf{0} \\ \mathbf{A}_{vw} & \mathbf{A}_{ww} \end{bmatrix} \quad (27)$$

e portanto

$$\mathbf{A}_{vv} = - \left[ \frac{2}{z^*} \mathbf{I} + [\mathbf{m}^{*'}]_{\times} [\mathbf{q}^{*'}]_{\times} \right] \mathbf{A}_v \quad (28)$$

$$\mathbf{A}_{ww} = - [2\mathbf{I} + [\mathbf{c}^{*'}]_{\times} [\mathbf{m}^{*'}]_{\times}] \mathbf{A}_w. \quad (29)$$

Primeiro, determinaremos a expressão analítica dos autovalores de  $\mathbf{A}$ . Como  $\mathbf{A}$  é bloco-triangular, seus autovalores consistem da união dos autovalores de  $\mathbf{A}_{vv}$  e  $\mathbf{A}_{ww}$ :

$$\sigma(\mathbf{A}) = \sigma(\mathbf{A}_{vv}) \cup \sigma(\mathbf{A}_{ww}), \quad (30)$$

com  $\sigma(\cdot)$  denotando o espectro de uma matriz. Considere primeiro  $\mathbf{A}_{vv}$ . Se  $\mathbf{x}$  denota um autovetor de  $\mathbf{A}_{vv}$  associado com o autovalor  $\mu$ , então para qualquer matriz de rotação  $\mathbf{Q}$ ,  $\mathbf{Q}\mathbf{x}$  é um autovetor de  $\mathbf{Q}\mathbf{A}_{vv}\mathbf{Q}^{\top}$  associado ao mesmo autovalor  $\mu$ . Logo,  $\sigma(\mathbf{A}_{vv}) = \sigma(\mathbf{Q}\mathbf{A}_{vv}\mathbf{Q}^{\top})$ . Como na prova do Teorema 3.1, defina  $\mathbf{Q}$  tal que  $\mathbf{Q}\mathbf{m}^{*'} = \|\mathbf{m}^{*'}\| \mathbf{b}_3$  com  $\mathbf{b}_3 = [0, 0, 1]^{\top}$ . Então,  $\mathbf{Q}\mathbf{A}_{vv}\mathbf{Q}^{\top}$  é triangular superior e os termos de sua diagonal correspondem ao seu espectro:

$$\begin{aligned} \sigma(\mathbf{Q}\mathbf{A}_{vv}\mathbf{Q}^{\top}) &= \sigma(\mathbf{A}_{vv}) = \\ & \left\{ -\lambda_1 \left( \frac{2}{z^*} - \mathbf{m}^{*\top} \mathbf{q}^{*'} \right), -\lambda_2 \left( \frac{2}{z^*} - \mathbf{m}^{*\top} \mathbf{q}^{*'} \right), -\lambda_3 \frac{2}{z^*} \right\}. \end{aligned}$$

Aplicando o mesmo procedimento a  $\mathbf{A}_{ww}$  temos

$$\begin{aligned} \sigma(\mathbf{Q}\mathbf{A}_{ww}\mathbf{Q}^{\top}) &= \sigma(\mathbf{A}_{ww}) = \\ & \left\{ -\lambda_4 (2 - \mathbf{m}^{*\top} \mathbf{c}^{*'}), -\lambda_5 (2 - \mathbf{m}^{*\top} \mathbf{c}^{*'}), -2\lambda_6 \right\}. \end{aligned}$$

A prova segue diretamente das expressões de  $\sigma(\mathbf{A}_{vv})$  e  $\sigma(\mathbf{A}_{ww})$ .  $\square$

O Teorema 3.2 provê condições explícitas sobre os ganhos de controle de forma a garantir a estabilidade exponencial local do equilíbrio. Como ambos  $\mathbf{m}^{*'}$  e  $\mathbf{c}^{*'}$  são definidos pelo usuário, a única restrição importante é a escolha de  $\lambda_1$  e  $\lambda_2$ .

Porém, como  $\mathbf{m}^* = z^* \mathbf{m}^{*'}$ , uma consequência direta de (23)–(25) é que a estabilidade é garantida  $\forall \lambda_i > 0$ ,  $i = 1, 2, \dots, 6$ , se as duas condições seguintes forem satisfeitas:

$$\mathbf{m}^{*\top} \mathbf{q}^{*'} < 2, \quad (31)$$

$$\mathbf{m}^{*\top} \mathbf{c}^{*'} < 2. \quad (32)$$

Note que estas condições podem sempre ser satisfeitas. A condição (31), que também existe na SD original, expressa a distância perpendicular entre o ponto de controle escolhido e o plano de referência. Dado que o plano de referência corresponde ao plano dominante do objeto, esta condição pode ser satisfeita se o ponto de controle for escolhido tal que a sua paralaxe  $\rho^*$  for suficientemente pequena. Pode-se simplesmente escolher um ponto com  $\rho^* = 0$ , pois o plano dominante cruza o objeto. Já a segunda condição (32) representa o comprimento da projeção de  $\mathbf{m}^{*'}$  em  $\mathbf{c}^{*'}$  (e vice-versa). Ela pode ser facilmente satisfeita definindo-se, e.g.,  $\mathbf{c}^{*' = \beta \mathbf{m}^{*'}/\|\mathbf{m}^{*'}\|^2, \forall \beta < 2$ . O equilíbrio do sistema (26) pode sempre ser projetado localmente exponencialmente estável  $\forall \lambda_i > 0$ . Ademais, o erro de controle translacional é desacoplado do rotacional, como mostrado em (27).

### 3.3 Uma Matriz Completamente Desacoplada

A matriz de interação obtida (16) é triangular inferior no caso geral. Esta seção apresenta um caso de interesse especial. De fato, ela ocorre com frequência na prática e esta matriz é puramente diagonal. Considere um objeto planar tal que o seu vetor normal escalonado é  $\mathbf{n}^* = [0, 0, 1/d^*]^{\top}$ , i.e., um plano fronto-paralelo ao sistema de referência cuja distância perpendicular é  $d^* > 0$ . Defina o ponto de controle tal que  $\mathbf{m}^{*' = [0, 0, 1]^{\top}$  e use  $\mathbf{c}^{*' = \mathbf{m}^{*'}$ . Como consequências,

$$(z^*)^{-1} = \mathbf{n}^{*\top} \mathbf{m}^{*'} = (d^*)^{-1}, \quad \mathbf{q}^{*'} = \mathbf{n}^*, \quad (33)$$

e as condições de estabilidade (31) e (32) são satisfeitas. Neste caso, a matriz de interação (16) é diagonal e dada por

$$\bar{\mathbf{L}}|_{\bar{\mathbf{e}}=\mathbf{0}} = -\text{diag}((z^*)^{-1}, (z^*)^{-1}, 2(z^*)^{-1}, 1, 1, 2).$$

Portanto, um sistema linear completamente desacoplado (26) com elementos estritamente negativos  $\forall \lambda_i > 0$ . Observe que tal matriz é usada apenas para propósitos de análise.

## 4 Resultados Experimentais

Esta seção apresenta resultados experimentais com dados sintéticos e reais, em simulação e usando uma câmera montada em um robô holonômico com seis g.d.l. Em todo caso, o objetivo de controle consiste em estabilizar o robô de forma que a imagem corrente do objeto coincida com a sua imagem capturada na pose de referência. O erro e a lei de controle são ambos calculados sem extrair características visuais e sem usar informação métrica do objeto. As intensidades dos pixels são diretamente exploradas para estimar todos os parâmetros projetivos necessários. O registro direto de imagens descrito em (Silveira e Malis, 2010) é usado para isso. Os resultados comparam a técnica de SD com a técnica desacoplada proposta, ambas com ganho de controle  $\lambda = 0, 4$  e condição de parada quando a norma dos erros de controle for menor que  $10^{-3}$ .

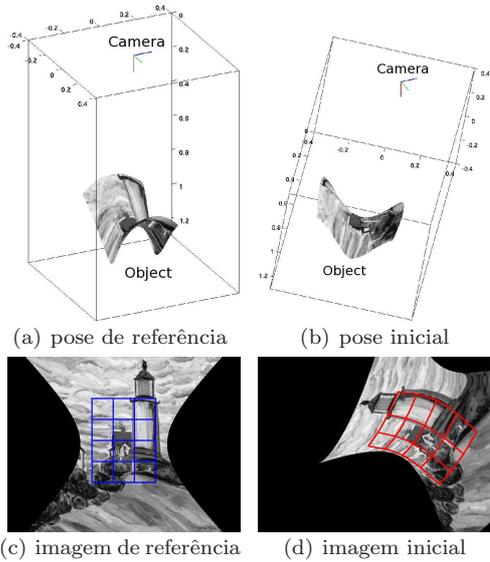


Figura 2: Experimento com dados sintéticos. (Acima) Configurações do sistema de coordenadas da câmera em relação ao alvo, sob diferentes pontos de vista. Apenas a parte explorada do objeto é mostrada. (Abaixo) Imagens vistas pela câmera nestas poses relativas.

#### 4.1 Dados Sintéticos

O primeiro conjunto de experimentos usa um objeto não planar. Este alvo é um parabolóide hiperbólico (i.e., a sela de cavalo), cujo centro é disposto a 1m da pose de referência. O deslocamento da pose inicial da câmera relativa à de referência é de  $[0.17, -0.11, 0.01]^T$  m em translação (norma de 0.2m, i.e., 20% da profundidade do centro), e  $[-0.1, -0.31, 1.04]^T$  radianos em rotação (norma de  $62.4^\circ$ ). Estas informações não estão disponíveis para o cálculo da lei de controle. As distâncias focais são definidas como 500 pixels, fator de obliquidade zero, e o ponto principal é o centro da imagem, que possui  $550 \times 418$  pixels. O período de amostragem é de 30 ms. O ponto de controle é escolhido tal que  $\mathbf{m}^{*f} = [0, 0, 1]^T$ . Por simplicidade, todas as intensidades dos pixels em uma região de interesse de  $200 \times 150$  pixels (i.e., dentro da grade) são exploradas, e não foi imposta nenhuma variação de iluminação. Diversos outros experimentos de estimação podem ser encontrados em (Silveira e Malis, 2010), incluindo diferentes objetos, câmeras coloridas, e variações de iluminação. A configuração desta simulação é mostrada na Fig. 2.

**SD:** Esta seção reporta os resultados obtidos com o erro de controle reduzido (11). Note que estas condições são desfavoráveis para a maioria das técnicas de servovisão 2D, devido à rotação inicial relativamente grande em torno de  $\bar{z}$ . No entanto, a SD completa a tarefa de forma precisa, e após 905 imagens, praticamente sem retroceder a câmera. Os resultados são mostrados na Fig. 3. No entanto, observa-se neste experimento um acoplamento das velocidades translacionais com as rotacionais, em particular em  $v_x$ .

**SD desacoplada:** Esta seção apresenta os resultados obtidos com o erro de controle desacoplado (12) com  $\mathbf{c}^{*f} = \mathbf{m}^{*f}$ . Assim como na SD, a tarefa é completada praticamente sem retroceder a

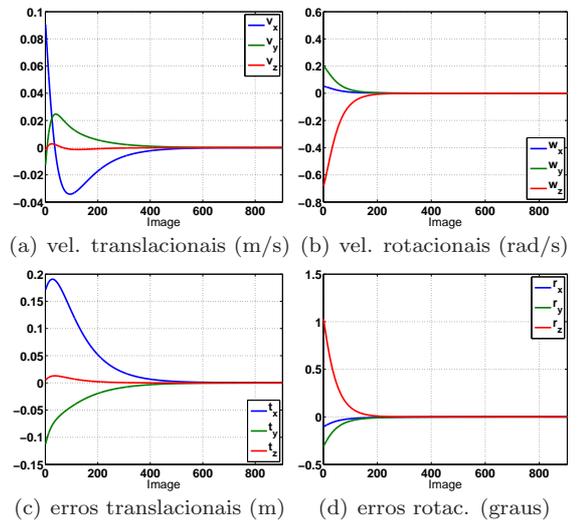


Figura 3: SD em relação a um objeto não planar. (Acima) Entradas de controle. As velocidades translacionais estão acopladas com as rotacionais, em particular  $v_x$ . (Abaixo) Movimento da câmera no espaço 3D até a convergência.

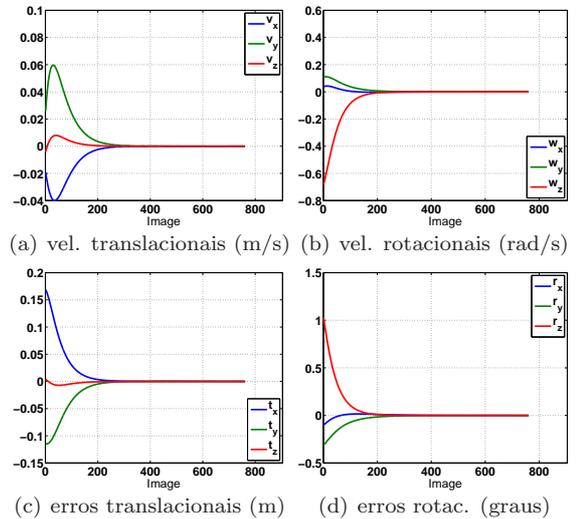


Figura 4: SD desacoplada com um objeto não planar (Acima) Entradas de controle. As velocidades translacionais se comportam de forma desacoplada, comparadas com as da Fig. 3. (Abaixo) Movimento da câmera no espaço.

câmera, mas agora converge após apenas 762 imagens. Portanto, a taxa de convergência é aumentada. O acoplamento das velocidades observado na SD é muito menos severo com esta técnica. Mais ainda, ele desaparece completamente próximo ao equilíbrio. A Fig. 4 mostra esses resultados.

#### 4.2 Dados Reais

Este conjunto de experimentos utiliza um objeto planar, colocado a  $\approx 0.7$  m da pose de referência. O deslocamento da pose inicial do robô relativo à essa pose é de  $[0.13, -0.23, -0.08]^T$  m (norma de 0.27m, ou seja, cerca de 38% da profundidade) em translação, e de  $[-20.3, 2.17, 14.59]^T$  graus (norma de  $25^\circ$ ) em rotação. Estes valores são obviamente desconhecidos pelos algoritmos. Para demonstrar a robustez das técnicas foi utilizada uma webcam descalibrada, de  $320 \times 240$  pixels. Esta câmera é montada no efetuador de um braço robótico de 6

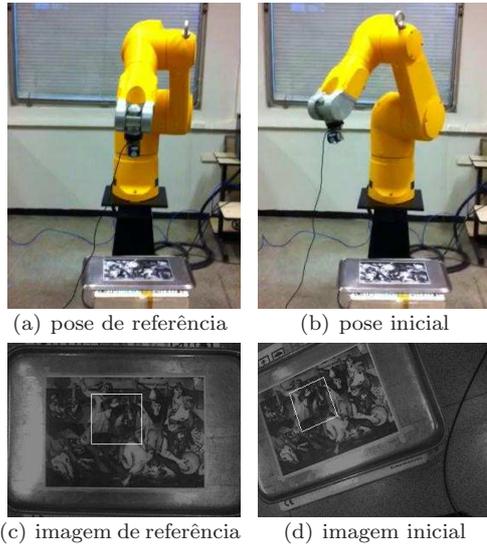


Figura 5: Experimento com dados reais, utilizando uma câmera montada em braço robótico, ambos apenas grosseiramente calibrados. (Acima) Configurações do robô em relação ao objeto. (Abaixo) Imagens respectivas.

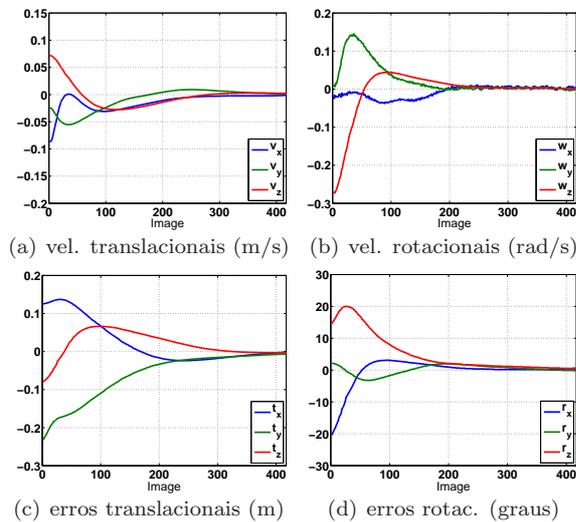


Figura 6: SD com um objeto planar, usando câmera e robô apenas grosseiramente calibrados. (Acima) Velocidades da câmera. Há acoplamento entre elas, em particular  $v_x$  e  $w_y$ . (Abaixo) Movimento no espaço Cartesiano.

g.d.l., e a calibração mão-olho também é apenas aproximada. O framerate é de 30Hz, o limite da webcam. O template de referência tem  $70 \times 70$  pixels de forma a satisfazer restrições de tempo real. O ponto de controle é escolhido como o centro desse template. A Fig. 5 mostra essa configuração.

**SD:** Como nos experimentos anteriores, a SD utiliza o erro de controle reduzido (11). Os resultados deste experimento são mostrados na Fig. 6. Novamente, a tarefa é realizada com precisão, convergindo após 417 imagens. No entanto, há acoplamento entre as velocidades.

**SD desacoplada:** Esses resultados utilizam o erro de controle desacoplado (12) com  $\mathbf{c}^{*l} = \mathbf{m}^{*l}$ . Uma vez mais, a tarefa é efetuada com precisão, convergindo após apenas 309 imagens. A taxa de convergência é portanto aumentada. O acoplamento previamente observado é reduzido, especialmente próximo ao equilíbrio. Vide Fig. 7.

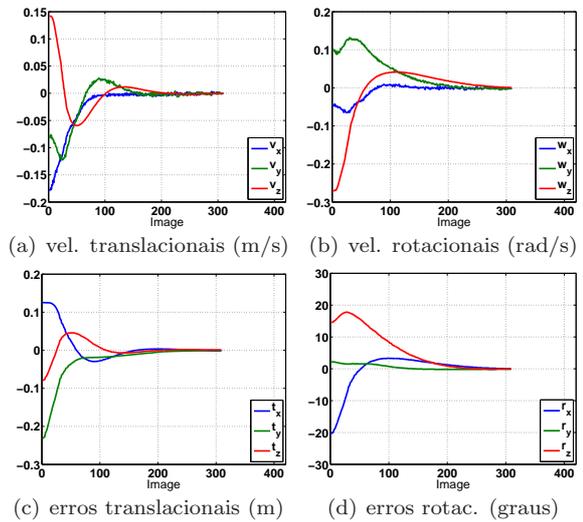


Figura 7: SD desacoplada com um alvo planar usando câmera e robô descalibrados. (Acima) Velocidades da câmera. Um desacoplamento é observado cf. Fig. 6. (Abaixo) Movimento da ponta do braço no espaço Cartesiano.

## 5 Conclusões

Este artigo propõe uma nova técnica de servovição geral, baseada em intensidade, não métrica, e que desacopla a dinâmica do erro de controle. Esta técnica desacoplada é geral no sentido que lida com objetos planares ou não, sob deslocamentos translacionais e rotacionais entre os sistemas de coordenadas de referência e inicial. A técnica proposta explora diretamente as intensidades dos pixels sem extrair ou associar primitivas geométricas, não requer qualquer informação métrica do objeto, e estabiliza exponencialmente localmente o ponto de equilíbrio. Resultados comparativos com uma técnica direta do estado-da-arte usando um braço robótico de 6 g.d.l. confirmam os avanços. Em trabalhos futuros planejamos estender o arcabouço proposto para a estabilização de outros sistemas mecânicos, como robôs não holonômicos.

## Referências

- Beardsley, P. A., Reid, I. D., Zisserman, A. e Murray, D. W. (1995). Active visual navigation using non-metric structure, *IEEE ICCV*, pp. 58–64.
- Chaumette, F. e Hutchinson, S. (2006). Visual servo control part I: Basic approaches, *IEEE Robotics & Automation Magazine* pp. 82–90.
- de Plinval, H., Morin, P., Mouyon, P. e Hamel, T. (2011). Visual servoing for underactuated VTOL UAVs: a linear, homography-based approach, *Proc. IEEE Int. Conf. on Rob. and Automation*, pp. 3004–3010.
- Faugeras, O., Luong, Q.-T. e Papadopoulos, T. (2001). *The geometry of multiple images*, The MIT Press.
- Silveira, G. e Malis, E. (2010). Unified direct visual tracking of rigid and deformable surfaces under generic illumination changes in grayscale and color images, *Int. Journal of Computer Vision* **89**(1): 84–105.
- Silveira, G. e Malis, E. (2012). Direct Visual Servoing: Vision-based estimation and control using only non-metric information, *IEEE Transactions on Robotics* **28**(4): 974–980.
- Silveira, G., Malis, E. e Rives, P. (2006). Visual servoing over unknown, unstructured, large-scale scenes, *Proc. IEEE ICRA, USA*, pp. 4142–4147.
- Thaler, L. e Goodale, M. A. (2010). Beyond distance and direction: The brain represents target locations non-metrically, *Journal of Vision* **10**(3): 1–27.