

# Um observador linear de homografia e aplicações à servovisão rápida baseada em intensidade\*

Geraldo Silveira\* Pascal Morin\*\*

\* Núcleo de Robótica e Visão Computacional, CTI, Campinas/SP, Brasil  
(e-mail: Geraldo.Silveira@cti.gov.br)

\*\* Inst. des Syst. Intell. et de Robot., Sorbonne Université, Paris, França  
(e-mail: pascal.morin@sorbonne-universite.fr)

**Abstract:** This article addresses the problems of fast intensity-based homography estimation and visual servo control. Classical initializations of their real-time respective nonlinear optimizations rely on image correlations or assume small interframe displacements. A new homography-based linear observer is proposed and applied in this work to adequately initialize that optimization without any of those. As for control, another new linear observer is devised from the former and applied to diagonally decouple the control error dynamics. A diagonal error dynamics can drastically reduce time responses and improve system performance. Both observers are not based on any linearization, and are analyzed using Lyapunov theory. Experiments using a camera-mounted 6-DoF robotic arm confirm the effectiveness of the approach to high-speed intensity-based homography estimation and visual servoing.

**Resumo:** Este artigo aborda os problemas da estimação da homografia e da servovisão rápidas baseadas em intensidade dos pixels. Inicializações clássicas das respectivas otimizações não lineares de tempo real dependem de correlações de imagens ou assumem pequenos deslocamentos entre elas. Um novo observador linear da homografia é proposto e aplicado neste trabalho para inicializar aquela otimização sem nenhuma dessas dependências. Para servovisão, outro novo observador linear é então derivado do primeiro e aplicado para desacoplar diagonalmente a dinâmica do erro de controle. Uma dinâmica diagonal pode reduzir drasticamente os tempos de resposta e melhorar o desempenho do sistema. Ambos observadores propostos não são baseados em linearizações e são analisados usando a teoria de Lyapunov. Resultados com um braço robótico de 6 GdL confirmam experimentalmente a eficácia da abordagem de estimação da homografia e servovisão de alta velocidade baseadas em intensidade.

*Keywords:* Robotics; robot vision; homography estimation; vision-based robot control.

*Palavras-chave:* Robótica; visão robótica; estimação da homografia; controle servo visual.

## 1. INTRODUÇÃO

Controle de robôs baseado em visão, também chamado servovisão, refere-se à estabilização de robôs em malha fechada a partir da realimentação de imagens. Uma tarefa típica consiste em estabilizá-lo em uma pose de referência especificada via uma imagem de referência. Um método clássico para executar essa tarefa é definir o erro de controle entre a imagem corrente e a de referência em termos Cartesianos e então aplicar estratégias de controle bem estabelecidas (Wilson et al., 1996). Outro método consiste em definir e regular um erro de controle diffeomórfico à pose da câmera em torno da imagem de referência. A grande maioria dos métodos baseados em diffeomorfismo exige informações métricas (e.g., profundidade do objeto) para fornecer uma lei de controle comprovadamente estabilizante (Chaumette e Hutchinson, 2006). Recentemente, tem surgindo um interesse crescente em estratégias baseadas em *diffeomorfismo não métrico* (Silveira et al., 2020) devido à sua maior versatilidade (nenhuma informação

métrica é necessária) e robustez a erros de parâmetros da câmera (Thaler e Goodale, 2010; Zeller e Faugeras, 1994).

Em todo caso, sistemas de servovisão exploram imagens usando técnicas baseadas em primitivas geométricas ou em intensidade (Szeliski, 2007). A primeira classe executa etapas intermediárias de extração e associação das primitivas (e.g., pontos e retas), e remoção das medidas aberrantes, entre imagens. Assim, ela depende de informações particulares, de uma associação propensa a erros, e de ajustes especiais de parâmetros. Diferentemente, técnicas *baseadas em intensidade* exploram diretamente as intensidades dos pixels sem etapas intermediárias. Portanto, elas não tem essas desvantagens e fazem uso de toda informação visual possível. Essas características as permitem atingir altos níveis de versatilidade, precisão (Irani e Anandan, 1999) e robustez a variações arbitrarias de iluminação (Silveira e Malis, 2010). Por outro lado, seus algoritmos de tempo real possuem um domínio de convergência relativamente menor porque executam otimizações não lineares locais. Algoritmos globais são computacionalmente custosos. Dessa forma, eles necessitam de uma inicialização suficientemente próxima da solução verdadeira.

\* Este trabalho foi em parte financiado pelo Projeto InSAC (FAPESP-2014/50851-0 e CNPq-465755/2014-3).

Existem várias estratégias para inicializar algoritmos de otimização baseada em intensidade de tempo real. As inicializações clássicas assumem pequenos deslocamentos entre imagens ou exploram alguma medida de correlação entre elas. No primeiro caso, simplesmente utiliza-se a última estimativa (ou o elemento identidade, se não disponível) como inicialização, enquanto no segundo caso geralmente implementa-se um preditor de janela deslizante. Nesse caso de preditores, várias soluções candidatas são geradas e avaliadas de acordo com uma dada medida, e a melhor solução é então escolhida. Assim, além de serem computacionalmente custosos, eles restringem apenas alguns Graus de Liberdade (GdL). Por exemplo, se a clássica medida de correlação cruzada de média zero (ZNCC) for utilizada então somente a posição da janela será inicializada, ou seja, apenas 2 GdL serão restritos. Recentemente, vários *observadores baseados em homografia* foram propostos e podem fornecer essa inicialização (Mahony et al., 2012; Eudes et al., 2013; Eudes e Morin, 2014). A matriz de homografia relaciona geometricamente duas vistas de um objeto planar (Faugeras, 1993), e é fundamental em várias aplicações de visão robótica, e.g., rastreamento visual e servovisão (Benhimane e Malis, 2006; Goncalves et al., 2010; Lopez-Nicolas et al., 2010). No entanto, a maioria dos observadores baseados em homografia foram propostos para fusão visual-inercial (Mahony et al., 2012; Eudes et al., 2013; Eudes e Morin, 2014) e, portanto, requerem girômetros e/ou acelerômetros. Uma notável exceção é o estimador proposto por Malis et al. (2009), porém ele presuppõe uma cinemática invariante de velocidade constante no grupo de Lie, o que raramente é verificada na prática.

Este trabalho propõe um novo observador linear da homografia e o aplica para inicializar um método de otimização baseada em intensidade de tempo real sem nenhuma das dependências supracitadas (e.g., necessidade de pequenos deslocamentos e sensores extras), e para auxiliar no desacoplamento diagonal da dinâmica do erro de controle não métrico parcialmente desacoplada em Silveira et al. (2020). Adicionalmente, uma variante do novo observador é também proposta e aplicada para completar a diagonalização e possibilitar uma servovisão diagonal. Sabe-se que uma dinâmica de erro de controle diagonal pode reduzir drasticamente tempos de resposta. Portanto, a combinação dos observadores propostos melhora as propriedades de convergência da estimação e o desempenho da servovisão. Ambos observadores não são baseados em linearizações, utiliza apropriadamente os sinais de controle (i.e., as velocidades da câmera) e são analisados usando a teoria de Lyapunov. Experimentos utilizando um braço robótico de 6 GdL demonstram que a aplicação desses observadores é *crucial* para a estimação da homografia e à servovisão rápidas baseadas em intensidade no contexto considerado.

## 2. PRELIMINARES

Sejam  $S(a)$  e  $\text{vex}(S(a))$  a matriz antissimétrica associada ao vetor  $a \in \mathbb{R}^3$  e seu mapeamento inverso, respectivamente. Assim, o produto vetorial corresponde ao produto  $S(a)b = a \times b$ ,  $\forall a, b \in \mathbb{R}^3$ , e  $\text{vex}(S(a)) = a$ . O traço de uma matriz  $M$  é denotado  $\text{tr}(M)$ ,  $\det(M)$  é seu determinante, e  $\|M\| = \sqrt{\text{tr}(M^T M)}$  sua norma. A matriz identidade e a matriz composta de zeros, ambas de dimensões apropriadas, são representadas por  $I$  e  $0$ , respectivamente. Por

fim, considere a rotação  $R \in \mathbb{SO}(3)$  e a translação  $p \in \mathbb{R}^3$  do sistema de coordenadas de referência  $\mathcal{F}^*$  expresso no sistema corrente  $\mathcal{F}$ . Suas derivadas temporais se escrevem

$$\begin{cases} \dot{R} = -S(w)R, \\ \dot{p} = -S(w)p - v, \end{cases} \quad (1)$$

onde  $v, w \in \mathbb{R}^3$  são as velocidades translationais e rotacionais também expressas em  $\mathcal{F}$ . Se uma dada variável for expressa em  $\mathcal{F}^*$ , um asterisco é empregado, e.g.,  $a^*$ .

### 2.1 Geometria de Duas Vistas

A relação entre pixels correspondentes em duas imagens perspectivas  $\mathbf{p}_i \leftrightarrow \mathbf{p}_i^* \in \mathbb{P}^2$ ,  $i = 1, 2, \dots, \ell$ , de um objeto planar é completamente definida por uma homografia projetiva  $G \in \mathbb{SL}(3)$ . O grupo linear especial  $\mathbb{SL}(3)$  corresponde ao grupo das matrizes  $(3 \times 3)$  com determinante 1. Dada tal homografia e a matriz de calibração de câmera  $K \in \mathbb{R}^{3 \times 3}$ , a homografia Euclidiana  $H$  pode ser obtida via

$$H = K^{-1}GK, \quad (2)$$

a qual pode ser escrita em termos dos elementos de transformação como

$$H = R + pn_s^T, \quad (3)$$

onde  $n_s = n^*/d^* \in \mathbb{R}^3$  é o vetor normal do objeto planar  $n^*$ :  $\|n^*\| = 1$  escalado por sua distância perpendicular  $d^* > 0$ .

### 2.2 Otimização Baseada em Intensidade

O arcabouço baseado em intensidade para a estimação da homografia é tipicamente formulado como o seguinte problema de otimização não linear:

$$\min_{G \in \mathbb{SL}(3), h \in \mathbb{R}^\ell} \frac{1}{2} \sum_{i=1}^{\ell} [\mathcal{I}'(G, h, \mathbf{p}_i^*) - \mathcal{I}^*(\mathbf{p}_i^*)]^2, \quad (4)$$

onde  $\mathcal{I}^*$  denota a imagem de referência e  $\mathcal{I}'$  representa a imagem corrente transformada usando as variáveis de otimização. Dentre elas, se necessário, os parâmetros fotométricos  $h$  podem ser estimados a fim de compensar variações de iluminação. No caso de servovisão de alta velocidade, somente métodos locais de otimização podem ser aplicados para resolver (4), pois os globais são computacionalmente custosos. Portanto, uma inicialização  $G_0 \in \mathbb{SL}(3)$  suficientemente próxima da solução verdadeira é necessária.

A próxima seção propõe um observador linear para fornecer a estimativa  $\hat{H} \in \mathbb{R}^{3 \times 3}$  e, por conseguinte, obter tal inicialização via

$$G_0 = \eta K \hat{H} K^{-1}, \quad (5)$$

onde o fator de escala  $\eta = 1/\det(K \hat{H} K^{-1})^{1/3}$  é aplicado para que  $G_0$  pertença ao grupo  $\mathbb{SL}(3)$ .

## 3. UM NOVO OBSERVADOR LINEAR BASEADO EM HOMOGRAFIA

Considere uma câmera em movimento livre, observando um objeto planar estacionário. Usando (1), a derivada temporal de (3) se escreve

$$\dot{H} = \dot{R} + \dot{p}n_s^T \quad (6)$$

$$= -S(w)R - S(w)pn_s^T - vn_s^T \quad (7)$$

$$= -S(w)H - vn_s^T. \quad (8)$$

O modelo dinâmico é então dado por

$$\begin{cases} \dot{H} = -S(w)H - vn_s^T \\ \dot{\hat{n}}_s = 0 \end{cases} \quad (9)$$

com seu estado composto de  $H$  e  $n_s$ .

O observador proposto é definido como:

$$\begin{cases} \dot{\hat{H}} = -S(w)\hat{H} - v\hat{n}_s^T + \alpha_H \\ \dot{\hat{n}}_s = \alpha_{n_s} \end{cases} \quad (10)$$

com observações  $H$ ,  $v$ , e  $w$ , termos de inovação  $\alpha_H$  e  $\alpha_{n_s}$  definidos abaixo, e condições iniciais  $\hat{H}(0) = H_0$  e  $\hat{n}_s(0) = \hat{n}_{s0}$ . A partir de (4) e (2), os erros do estimador podem ser calculados como se segue:

$$\tilde{H} = \hat{H} - H, \quad (11)$$

$$\tilde{n}_s = \hat{n}_s - n_s. \quad (12)$$

Assim, a dinâmica dos erros de estimação se escreve

$$\begin{cases} \dot{\tilde{H}} = -S(w)\tilde{H} - v\tilde{n}_s^T + \alpha_H, \\ \dot{\tilde{n}}_s = \alpha_{n_s}. \end{cases} \quad (13)$$

Algumas propriedades de estabilidade e convergência do observador proposto (10) são enunciadas abaixo.

**Teorema 1.** Considere o sistema dos erros de estimação (13). Assuma que  $v, \dot{v}$  e  $w, \dot{w}$  são limitadas. Sejam

$$\begin{cases} \alpha_H = -k_1\tilde{H}, & k_1 > 0, \\ \alpha_{n_s} = k_2\tilde{H}^T v, & k_2 > 0. \end{cases} \quad (14)$$

Então, a dinâmica dos erros de estimação satisfazem as seguintes propriedades:

- (1)  $(\tilde{H}, \tilde{n}_s) = 0$  é um ponto de equilíbrio estável;
- (2) Para qualquer condição inicial,

$$\begin{aligned} \lim_{t \rightarrow +\infty} \tilde{H}(t) &= 0, \\ \lim_{t \rightarrow +\infty} v(t)\tilde{n}_s^T(t) &= 0. \end{aligned} \quad (15)$$

**Prova.** Considere a função candidata de Lyapunov

$$V = \frac{1}{2}\|\tilde{H}\|^2 + \frac{1}{2k_2}\|\tilde{n}_s\|^2. \quad (16)$$

Usando o fato que  $\text{tr}(M) = \text{tr}(M^T)$  e  $\text{tr}(ab^T) = a^T b$  para qualquer matriz quadrada  $M$  e quaisquer vetores reais  $a, b$ , a derivada temporal de  $V$  ao longo das soluções do sistema (13) é dada por

$$\dot{V} = \text{tr}(\tilde{H}^T \dot{\tilde{H}}) + \frac{1}{k_2}\tilde{n}_s^T \dot{\tilde{n}}_s \quad (17)$$

$$= \text{tr}(-\tilde{H}^T v\tilde{n}_s^T - k_1\tilde{H}^T \tilde{H}) + \tilde{n}_s^T \tilde{H}^T v \quad (18)$$

$$= -v^T \tilde{H} \tilde{n}_s - k_1 \text{tr}(\tilde{H}^T \tilde{H}) + \tilde{n}_s^T \tilde{H}^T v \quad (19)$$

$$= -k_1 \|\tilde{H}\|^2 \leq 0. \quad (20)$$

Isso implica que a origem do sistema (13) é estável e prova a Propriedade 1.

Iremos agora considerar a Propriedade 2. Dado que  $V$  é decrescente ao longo das soluções de (13),  $\tilde{H}$  e  $\tilde{n}_s$  são limitados. Já que  $v$  e  $w$  são assumidas como limitadas,  $\dot{\tilde{H}}$  e  $\dot{\tilde{n}}_s$  são também limitadas. Assim,  $\ddot{V} = -2k_1 \text{tr}(\tilde{H}^T \dot{\tilde{H}})$  é limitada. Dado que  $V$  tende a um valor constante ao longo das soluções do sistema (i.e., ela é uma função decrescente positiva) e  $\dot{V}$  é uniformemente contínua (i.e., sua derivada temporal é limitada), podemos deduzir por aplicação do Lema de Barbalat que  $\lim_{t \rightarrow +\infty} \dot{V}(t) = 0$  e

portanto  $\lim_{t \rightarrow +\infty} \tilde{H}(t) = 0$ . Isso prova a primeira relação em (15). Vamos agora considerar sua segunda relação.

Já estabelecemos a finitude de  $\tilde{H}, \dot{\tilde{H}}, \tilde{n}_s$  e  $\dot{\tilde{n}}_s$ . Usando a presunção que  $v, \dot{v}$ , e  $w, \dot{w}$  são todas limitadas, podemos deduzir que  $\ddot{\tilde{H}}$  é limitada. Assim,  $\dot{\tilde{H}}$  é uniformemente contínua e dado que  $\tilde{H}$  tende a zero, podemos deduzir por aplicação do Lema de Barbalat que  $\lim_{t \rightarrow +\infty} \dot{\tilde{H}}(t) = 0$ . Isso implica que  $\lim_{t \rightarrow +\infty} v(t)\tilde{n}_s^T(t) = 0$  e prova a segunda relação em (15).  $\square$

Note que a convergência de  $\tilde{n}_s$  para zero não pode ser garantida para qualquer velocidade  $v$ . Não é difícil obter as condições suficientes sobre  $v$  que asseguram tal propriedade de convergência. Entretanto, esse ponto não será endereçado neste artigo, pois a convergência de  $\tilde{n}_s$  para zero não será necessária na sequência.

Finalmente, a estimativa  $\hat{H} \in \mathbb{R}^{3 \times 3}$  pode ser aplicada em (5) para inicializar adequadamente (mesmo sob rápido movimento de câmera) o arcabouço de estimação (4). O valor estimado ótimo  $G \in \mathbb{SL}(3)$  é então aplicado em (2) para obter  $H$ , a ser usada como observação em (11) e requerida pela técnica de servovisão proposta abaixo.

#### 4. UMA NOVA TÉCNICA DE SERVOVISÃO DIAGONAL BASEADA EM HOMOGRAFIA

Sejam os sinais de controle  $u = (v, w) \in \mathbb{R}^6$  as velocidades translacionais e rotacionais da câmera. A técnica de servovisão não métrica parcialmente desacoplada (Silveira et al., 2020) estabiliza exponencialmente localmente todos os 6 GdL de um robô holonômico com uma câmera montada em seu efetuador, independentemente das características do objeto observado (e.g., forma e tamanho), deslocamentos da câmera e suas poses relativas. No caso de um objeto planar, seu erro de controle não métrico é reescrito como

$$\varepsilon = \begin{pmatrix} 2I & S(m^*) \\ -S(c^*) & I \end{pmatrix} \begin{pmatrix} (H - I)m^* \\ \text{vex}(H - H^T) \end{pmatrix} \quad (21)$$

com os escolhidos ponto de controle  $m^* = K^{-1}p^*$  e vetor de controle  $c^* \in \mathbb{R}^3$ . Com o propósito exclusivo de análise, a matriz de interação de (21) linearizada em torno do equilíbrio  $\varepsilon = 0$  é dada por

$$L = - \begin{pmatrix} \frac{2}{z^*}I + S(m^*)S(n_s) & 0 \\ -\frac{1}{z^*}S(c^*) + S(n_s) & 2I + S(c^*)S(m^*) \end{pmatrix}, \quad (22)$$

onde  $z^* > 0$  é a profundidade de  $m^*$ . Essa matriz revela que a dinâmica do erro de controle translacional é desacoplada da rotacional, mas não vice-versa dado que não é diagonal.

A ideia da servovisão não métrica diagonal (Silveira e Mirisola, 2019) para desacoplar diagonalmente a dinâmica do erro consiste em estimar  $\hat{z}^*$  e  $\hat{n}_s$  (para o presente caso de objeto planar), e então definir  $c^*$  tal que se obtenha

$$\left( -\frac{1}{z^*}S(c^*) + S(n_s) \right) v \rightarrow 0. \quad (23)$$

Seu estimador é baseado em uma dinâmica do erro de controle rotacional linearizada. Diferentemente, esta seção propõe outro estimador a partir de (10) que não é baseado em nenhuma linearização, conforme mostrado a seguir.

#### 4.1 Um Outro Observador

A estimativa  $\widehat{z}^*$  necessária para uma servovisão diagonal poderia já ser obtida usando o Observador (10) sem qualquer linearização, i.e., usando  $\widehat{z}^* = 1/(\widehat{n}_s^T m^*)$ . Entretanto, é difícil garantir que  $\widehat{n}_s^T(t)m^* \neq 0, \forall t$ . Esta seção propõe um outro observador linear que endereça esse ponto.

Seja

$$\mu^* = n_s^T m^*. \quad (24)$$

Multiplicando ambos lados de (9) pela direita por  $m^*$ , obtem-se o seguinte modelo dinâmico:

$$\begin{cases} \frac{dHm^*}{dt} = -S(w)Hm^* - v\mu^* \\ \dot{\mu}^* = 0 \end{cases} \quad (25)$$

cujos estado é composto do vetor  $Hm^*$  e o escalar  $\mu^*$ .

Um outro observador pode então ser derivado de (10):

$$\begin{cases} \dot{\widehat{Hm}^*} = -S(w)\widehat{Hm}^* - v\widehat{\mu}^* + \alpha_{Hm^*} \\ \dot{\widehat{\mu}^*} = \alpha_{\mu^*} \end{cases} \quad (26)$$

com observações  $Hm^*, v$ , e  $w$ , termos de inovação  $\alpha_{Hm^*}$  e  $\alpha_{\mu^*}$  definidos abaixo, e condições iniciais  $\widehat{Hm}^*(0) = H_0m^*$  e  $\widehat{\mu}^*(0) = \widehat{\mu}_0^*$ . Defina os erros de estimação como se segue:

$$\widetilde{Hm}^* = \widehat{Hm}^* - Hm^*, \quad (27)$$

$$\widetilde{\mu}^* = \widehat{\mu}^* - \mu^*. \quad (28)$$

Portanto, a dinâmica dos erros de estimação se escreve

$$\begin{cases} \dot{\widetilde{Hm}^*} = -S(w)\widetilde{Hm}^* - v\widetilde{\mu}^* + \alpha_{Hm^*}, \\ \dot{\widetilde{\mu}^*} = \alpha_{\mu^*}. \end{cases} \quad (29)$$

Algumas propriedades desse novo estimador (26) são enunciados na sequência.

**Teorema 2.** Considere o sistema dos erros de estimação (29). Assuma que  $v, \dot{v}$  e  $w, \dot{w}$  são limitadas. Sejam

$$\begin{cases} \alpha_{Hm^*} = -k_1\widetilde{Hm}^*, & k_1 > 0, \\ \alpha_{\mu^*} = k_2(\widetilde{Hm}^*)^T v, & k_2 > 0. \end{cases} \quad (30)$$

Então, a dinâmica dos erros de estimação satisfazem as seguintes propriedades:

- (1)  $(\widetilde{Hm}^*, \widetilde{\mu}^*) = 0$  é um ponto de equilíbrio estável;
- (2) Para qualquer condição inicial,

$$\begin{aligned} \lim_{t \rightarrow +\infty} \widetilde{Hm}^*(t) &= 0, \\ \lim_{t \rightarrow +\infty} v(t)\widetilde{\mu}^*(t) &= 0. \end{aligned} \quad (31)$$

**Prova.** Considere a função candidata de Lyapunov

$$V = \frac{1}{2}\|\widetilde{Hm}^*\|^2 + \frac{1}{2k_2}(\widetilde{\mu}^*)^2. \quad (32)$$

Sua derivada temporal ao longo das soluções do Sistema (29) se escreve

$$\dot{V} = -k_1\|\widetilde{Hm}^*\|^2 \leq 0, \quad (33)$$

o que implica a Propriedade 1.

A prova da Propriedade 2, a qual se baseia na aplicação do Lema de Barbalat, é similar à prova da Propriedade 2 do Teorema 1.  $\square$

A adequação desse novo observador linear é formalmente enunciada com a seguinte propriedade.

**Corolário 3.** Considere o Estimador (26). Assuma que

$$0 < \widehat{\mu}_0^* < \mu^*. \quad (34)$$

Então,  $\widehat{\mu}^*(t) \geq \widehat{\mu}_0^* > 0, \forall t$ .

**Prova.** Dado que  $\widehat{Hm}^*(0) = H_0m^*$ , onde  $H_0$  é medido, então  $V(0) = \frac{1}{2k_2}\widehat{\mu}^*(0)^2$  a partir do Teorema 2. Ademais, como  $\dot{V} \leq 0$  (33), então  $V(t) \leq \frac{1}{2k_2}\widehat{\mu}^*(0)^2, \forall t$ , e portanto,  $\forall t$ ,

$$|\widetilde{\mu}^*(t)| \leq |\widetilde{\mu}^*(0)| \quad (35)$$

$$|\widehat{\mu}^*(t) - \mu^*(t)| \leq |\widehat{\mu}^*(0) - \mu^*(0)| \quad (36)$$

$$|\widehat{\mu}^*(t) - \mu^*| \leq |\widehat{\mu}_0^* - \mu^*|. \quad (37)$$

Deduz-se assim a partir de (37) que,  $\forall t$ ,

$$\widehat{\mu}^*(t) = \widehat{\mu}^*(t) - \mu^* + \mu^* \quad (38)$$

$$\geq \mu^* - |\widehat{\mu}^*(t) - \mu^*| \quad (39)$$

$$\geq \mu^* - |\widehat{\mu}_0^* - \mu^*| \geq \widehat{\mu}_0^*, \quad (40)$$

o que conclui a prova.  $\square$

O ponto importante no corolário acima é a possibilidade de assegurar que a estimativa  $\widehat{\mu}^*$  nunca se torna zero. Essa propriedade depende da Condição (34). O escalar  $\mu^*$  é desconhecido mas, na prática, geralmente não é difícil termos um limite inferior para essa quantidade, a partir da qual  $\widehat{\mu}_0^*$  pode ser escolhido.

#### 4.2 Lei de Controle

Considere um robô holonômico de 6 GdL com uma câmera montada em seu efetuador, observando um objeto planar estacionário de características desconhecidas (e.g., tamanho, textura, situação com relação a ambos  $\mathcal{F}$  e  $\mathcal{F}^*$ ).

**Lema 4.** Considere os observadores baseados em homografia (10) e (26) com (34). Seja

$$c^* = \frac{\widehat{n}_s}{\widehat{\mu}^*}. \quad (41)$$

Então, a dinâmica do erro de controle (21) é localmente ulteriormente diagonal em torno do equilíbrio  $\varepsilon = 0$ :

$$\dot{\varepsilon} = L' u \quad (42)$$

com

$$L' = - \begin{pmatrix} \frac{2}{z^*}I + S(m^*)S(n_s) & 0 \\ 0 & 2I + \widehat{z}^*S(\widehat{n}_s)S(m^*) \end{pmatrix}. \quad (43)$$

Ademais, isto é obtido independentemente das velocidades de controle  $u = (v, w) \in \mathbb{R}^6$ .

**Prova.** A prova consiste em demonstrar que o vetor de controle (41) fornece (23), i.e.,  $L'_{uv}v \rightarrow 0$  com

$$L'_{uv}v = S\left(-\mu^*\frac{\widehat{n}_s}{\widehat{\mu}^*} + n_s\right)v \quad (44)$$

$$= \frac{1}{\widehat{\mu}^*}S(-\mu^*\widehat{n}_s + \widehat{\mu}^*n_s)v \quad (45)$$

$$= -\frac{\mu^*}{\widehat{\mu}^*}S(\widetilde{n}_s)v + \frac{\widetilde{\mu}^*}{\widehat{\mu}^*}S(n_s)v. \quad (46)$$

A partir dos Teoremas 1 e 2, respectivamente, tem-se que

$$\lim_{t \rightarrow +\infty} \widetilde{n}_{s,j}(t)v_i(t) = 0, \quad \forall i, j \in \{1, 2, 3\}, \quad (47)$$

$$\lim_{t \rightarrow +\infty} \widetilde{\mu}^*(t)v(t) = 0. \quad (48)$$

Ademais, a partir do Corolário 3,  $\widehat{\mu}^*(t) \geq \widehat{\mu}_0^* > 0$  de modo que  $\frac{1}{\widehat{\mu}^*}$  é limitada. Portanto, (23) segue de (46).  $\square$



Quanto à convergência não garantida de  $\hat{n}_s$  para zero na Seção 3, também não se pode garantir convergência de  $\hat{\mu}^*$  para zero para qualquer velocidade  $v$ . No entanto, essas convergências não são necessárias para nossos propósitos de servovisão. Dada a convergência desejada (23), a matriz de interação obtida (43) possui uma estrutura diagonal. Note que essa matriz de interação é usada apenas para fins de análise, i.e., não será utilizada para o cálculo do sinal de controle, conforme mostrado abaixo.

**Teorema 5.** Considere a dinâmica do erro de controle diagonal (42). A lei de controle

$$u = \lambda \varepsilon, \quad \lambda > 0, \quad (49)$$

com o vetor de controle (41) preserva a propriedade diagonal e assegura estabilidade assintótica local de  $\varepsilon = 0$ .

**Prova.** Substituindo a lei de controle (49) com o vetor de controle (41) em (42), obtém-se a seguinte dinâmica do erro de controle diagonalmente desacoplada em malha fechada em torno de  $\varepsilon = 0$ :

$$\dot{\varepsilon} = \lambda L' \varepsilon = A \varepsilon. \quad (50)$$

Os autovalores de  $A \in \mathbb{R}^{6 \times 6}$  são  $\lambda(n_s^T m^* - 2/z^*)$  (duplo),  $-2\lambda/z^*$ ,  $\lambda(\hat{z}^* \hat{n}_s^T m^* - 2)$  (duplo), e  $-2\lambda$ . Dado que  $z^*, \lambda > 0$  e também usando o fato que  $z^* n_s^T m^* = 1 = \hat{z}^* \hat{n}_s^T m^*$ , a prova é direta  $\forall m^*$  e  $\forall n_s$ .  $\square$

## 5. RESULTADOS EXPERIMENTAIS

Esta seção apresenta alguns resultados obtidos usando os observadores propostos (10) e (26). Enquanto que o primeiro é aplicado tanto para inicializar a otimização (5) quanto para calcular o vetor de controle (41), o segundo é usado apenas em (41) para diagonalizar a dinâmica da servovisão. O objetivo da combinação desses observadores é, portanto, melhorar as propriedades de convergência do procedimento de estimação e do sistema de controle.

Em relação ao método de estimação, nenhuma primitiva geométrica é extraída ou associada nas imagens, e nem medidas de correlação são aplicadas entre elas. As intensidades dos pixels são diretamente exploradas na otimização não linear (4), a qual é eficientemente resolvida usando os Casos 3.1.1 e 3.2.1 em Silveira e Malis (2010) com um número máximo de 20 iterações. Por sua vez, o objetivo de controle consiste em estabilizar rapidamente a câmera de modo que a imagem corrente do objeto coincida com a sua imagem capturada na pose de referência, ou seja, com a imagem de referência. Embora tenhamos efetuado com sucesso servovisão diagonal baseada em intensidade a velocidades muito mais altas do que as aqui apresentadas, devido a limitações de espaço mostraremos apenas aquelas para as quais a aplicação dos estimadores propostos já é crucial para o êxito das estabilizações.

Em todos os experimentos, as imagens são criadas (se em simulação) ou capturadas a 25Hz, apenas a região de interesse da imagem é explorada e o ponto de controle  $m^*$  é escolhido como o centro dessa região. Os ganhos dos estimadores escolhidos são  $k_1 = k_2 = 20$  para uma rápida estabilização. Por fim, salienta-se que os sinais de controle computados são usados diretamente como observações para os estimadores, ou seja, nenhum outro sensor é usado além de um única câmera convencional.

### 5.1 Dados Sintéticos

Este experimento modela um braço robótico clássico com uma câmera perspectiva montada em seu efetuador, observando um objeto planar estacionário de características desconhecidas. As distâncias focais da câmera modelada são definidas em 500 pixels, sem fator de obliquidade e ponto principal no centro da imagem, que possui  $800 \times 640$  pixels com uma região de interesse de  $300 \times 300$  pixels. O deslocamento inicial entre as poses de referência e a inicial (expressas na primeira) é de  $(0, 16, -0, 04, -0, 21)$  metros e de  $(9, -30, -15)$  graus em translação e rotação, resp. As imagens de referência e a inicial criadas sinteticamente são mostradas na Fig. 1. Obviamente, essas informações métricas não são usadas nem nos procedimentos de estimação nem nos de controle, cujas condições iniciais são fixadas em  $\hat{n}_{s_0} = (0, 0, 1)$ ,  $\hat{\mu}_0^* = \hat{n}_{s_0}^T m^* = 1$  e  $\lambda = 0.6$ .

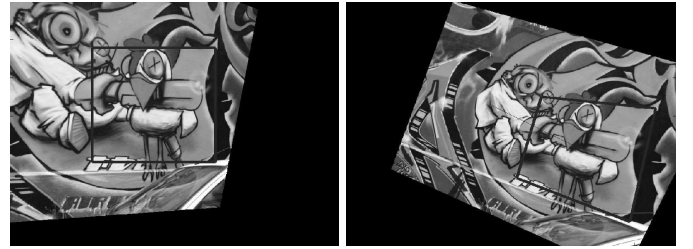


Figura 1. Imagens de referência e inicial, resp., para o experimento de servovisão usando dados sintéticos.

**Nota 1.** Embora a servovisão ainda seja bem sucedida para um ganho de controle muito mais alto (mesmo para  $\lambda = 10$ ), o ganho considerado  $\lambda = 0.6$  já é suficiente para demonstrar a importância dos estimadores propostos. De fato, o método de otimização *falha*, e portanto toda a servovisão também falha, se eles não são aqui aplicados.

A Fig. 2 mostra os resultados da servovisão baseada em intensidade usando os observadores propostos. Note o comportamento desacoplado, especialmente próximo a convergência. A evolução dos erros de estimação medidos é dada na Fig. 3, mostrando rápidas estabilizações.

### 5.2 Dados Reais

Este experimento utiliza um braço robótico industrial de 6 GdL, observando um objeto planar estacionário desconhecido. Uma *webcam* grosseiramente calibrada é montada em seu efetuador e captura imagens de  $320 \times 240$  pixels a 25 Hz. Nesse caso, a região de imagem explorada é de  $150 \times 150$  pixels. As distâncias focais são de cerca de 400 pixels, sem fator de obliquidade e ponto principal como o centro da imagem. A calibração mão-olho é simplesmente considerada como uma inclinação de 20 graus e com um deslocamento de  $-5$  cm e  $5$  cm nas direções  $\vec{y}$  e  $\vec{z}$ , respectivamente. O deslocamento inicial entre as poses de referência e a inicial (expresso na primeira) é de  $(0, 27, 0, 36, -0, 16)$  metros e de  $(5, -40, -30)$  graus em translação e rotação, respectivamente. Esse cenário experimental é mostrada na Fig. 4. Uma vez mais, essas informações métricas não são usadas, obviamente, nem

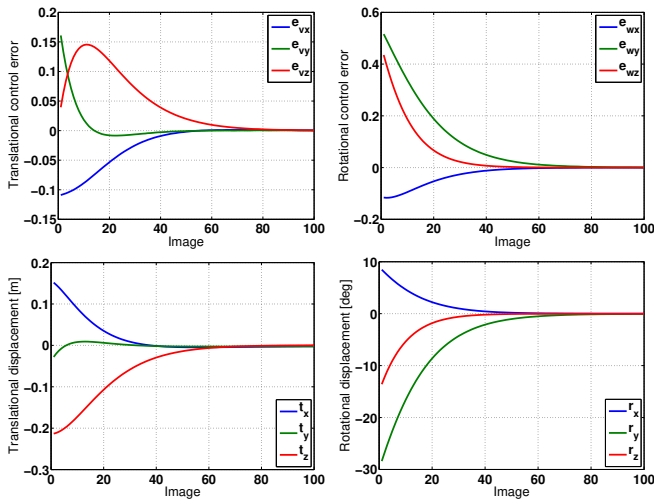


Figura 2. Resultados da simulação usando a servovisão diagonal baseada em intensidade com os observadores propostos. (Acima) Erros de controle. (Abaixo) Movimento da câmera até convergência.

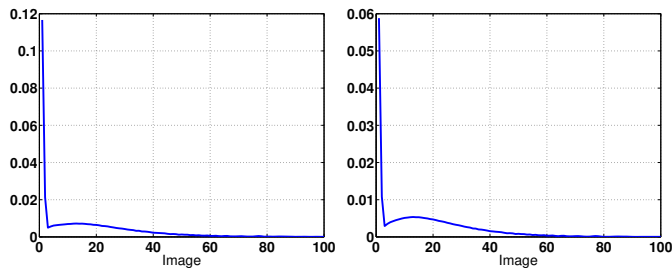


Figura 3. Evolução de  $\|\tilde{H}\|$  e  $\|\tilde{H}m^*\|$ , respectivamente, durante o experimento com dados sintéticos.

nos procedimentos de estimação nem nos de controle, cujas condições iniciais são fixadas em  $\hat{n}_{s_0} = (0, 0, 2)$ ,  $\hat{\mu}_0^* = \hat{n}_{s_0}^T m^* = 2$  e  $\lambda = 0.4$ .

**Nota 2.** Embora a servovisão ainda seja executada com sucesso para um ganho de controle mais alto (i.e., em velocidades mais altas), tal ganho já é suficiente para demonstrar a importância dos estimadores propostos. De fato, o método de otimização *falha*, e portanto toda a servovisão, se eles não são aqui novamente aplicados.

Outrossim, é importante salientar que o contexto é desafiador: frequência de amostragem de apenas 25Hz, deslocamento inicial da câmera relativamente grande (normas de 0,48m e de  $50, 2^\circ$ ), câmera grosseiramente calibrada, e objeto de muitas características desconhecidas. Isso impõe sérias dificuldades para métodos locais de otimização não linear (se não bem inicializados), especialmente para altos ganhos de controle pois induzem grandes deslocamentos entre imagens. A Fig. 5 mostra os resultados da servovisão diagonal baseada em intensidade, usando os observadores propostos. Note também o comportamento desacoplado dos sinais. A evolução dos erros de estimação medidos é dada na Fig. 6, mostrando rápidas estabilizações.

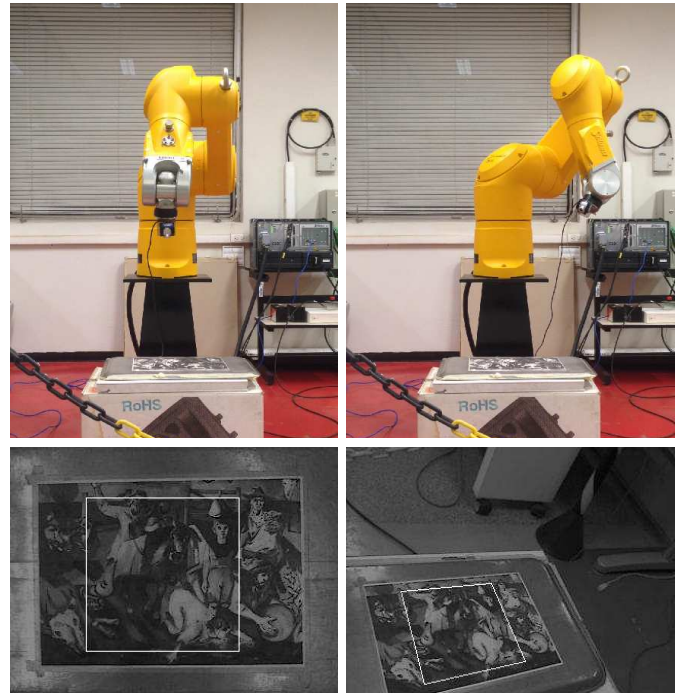


Figura 4. Cenário do experimento usando dados reais: um braço robótico de 6 GdL com uma câmera montada em seu efetuador, observando um objeto planar desconhecido. (Acima) Poses de referência e inicial do robô, resp. (Abaixo) Imagens capturadas respectivas.

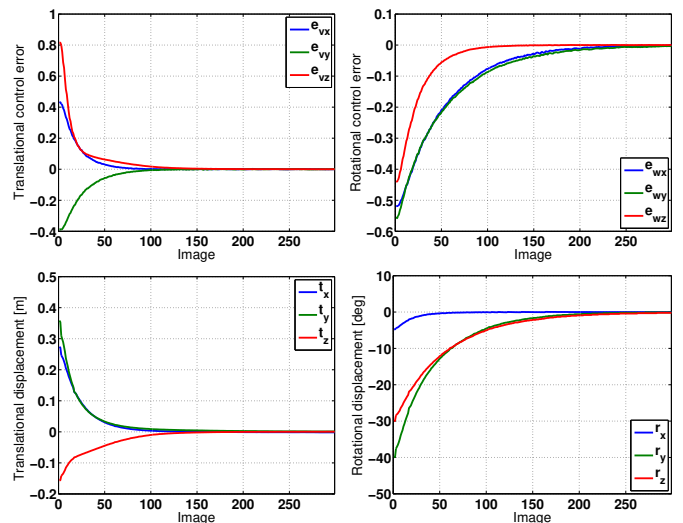


Figura 5. Resultados do experimento real usando a servovisão diagonal baseada em intensidade com os observadores propostos. (Acima) Erros de controle. (Abaixo) Movimento do efetuador até convergência.

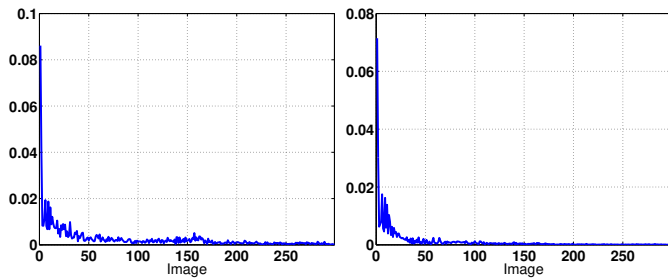


Figura 6. Evolução de  $\|\tilde{H}\|$  e  $\|\tilde{H}m^*\|$ , respectivamente, durante o experimento com dados reais.

## 6. CONCLUSÃO

Este artigo propõe e analisa um novo observador linear da homografia que pode ser usado para filtragem e predição. Esta última aplicação foi implementada para inicializar um procedimento de otimização não linear de tempo real baseada em intensidade, e para auxiliar na diagonalização da dinâmica de servovisão. Uma variante desse observador também foi proposta para completar adequadamente essa diagonalização. Ambos observadores não são baseados em qualquer linearização e faz uso apropriado das velocidades da câmera (no caso, os sinais de controle). A combinação dos observadores propostos permite melhorar as propriedades de convergência tanto do procedimento de estimação quanto do sistema de controle. De fato, os experimentos demonstram que esses estimadores são cruciais para convergência da servovisão no contexto desafiador de deslocamentos iniciais relativamente grandes, de câmera grosseiramente calibrada, de baixas taxas de amostragem, da otimização local baseada em intensidade, e dos altos ganhos de controle. Trabalhos futuros incluem sua extensão para lidar com objetos não planares e o desenvolvimento de novas estratégias de controle não métricas para levar em consideração a dinâmica do robô.

## REFERÊNCIAS

- Benhimane, S. e Malis, E. (2006). Homography-based 2D visual servoing. In *Proc. IEEE ICRA*.
- Chaumette, F. e Hutchinson, S. (2006). Visual servo control part I: Basic approaches. *IEEE Robot. & Autom. Magazine*, 82–90.
- Eudes, A., Morin, P., Mahony, R., e Hamel, T. (2013). Visuo-inertial fusion for homography-based filtering and estimation. In *Proc. IEEE/RSJ IROS*.
- Eudes, A. e Morin, P. (2014). A linear approach to visuo-inertial fusion for homography-based filtering and estimation. In *Proc. IEEE/RSJ IROS*.
- Faugeras, O. (1993). *Three-Dimensional Computer Vision – A geometric viewpoint*. The MIT Press, Cambridge.
- Goncalves, T., Azinheira, J., e Rives, P. (2010). Homography-based visual servoing of an aircraft for automatic approach and landing. In *Proc. IEEE ICRA*.
- Irani, M. e Anandan, P. (1999). All about direct methods. In *Proc. of the Workshop on Vision Algorithms: Theory and practice*.
- Lopez-Nicolas, G., Gans, N.R., Bhattacharya, S., Sagues, C., Guerrero, J.J., e Hutchinson, S. (2010). Homography-based control scheme for mobile robots with nonholonomic and field-of-view constraints. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 40(4), 1115–1127.
- Mahony, R., Hamel, T., Morin, P., e Malis, E. (2012). Nonlinear complementary filters on the special linear group. *International Journal of Control*, 85(10), 1557–1573.
- Malis, E., Hamel, T., R., M., e Morin, P. (2009). Dynamic estimation of homography transformations on the special linear group for visual servo control. In *Proc. IEEE ICRA*, 1498–1503.
- Silveira, G. e Mirisola, L. (2019). Diagonally-decoupled direct visual servoing. In *Proc. IEEE ICRA*, 6611–6616.
- Silveira, G., Mirisola, L., e Morin, P. (2020). Decoupled intensity-based nonmetric visual servo control. *IEEE Transactions on Control Systems Technology*, 28(2), 566–573.
- Silveira, G. e Malis, E. (2010). Unified direct visual tracking of rigid and deformable surfaces under generic illumination changes in grayscale and color images. *International Journal of Computer Vision*, 89(1), 84–105.
- Szeliski, R. (2007). Image alignment and stitching: A tutorial. *Foundations and Trends in Computer Graphics and Vision*, 2(1).
- 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.
- Wilson, W.J., Hulls, C.C.W., e Bell, G.S. (1996). Relative end-effector control using Cartesian position based visual servoing. *IEEE Trans. on Robotics and Automation*, 12(5), 684–696.
- Zeller, C. e Faugeras, O. (1994). Applications of non-metric vision to some visual guided tasks. In *Proc. IEEE Intelligent Vehicles Symposium*, 351–356.