

# Utilização de correspondências entre histogramas normalizados de imagens para a estimação automática de algumas informações de vôo

Felipe Leonardo Lôbo Medeiros

Instituto de Estudos Avançados, Trevo Coronel Aviador José Alberto Albano do Amarante, nº 1, Putim, São José dos Campos, São Paulo

**Resumo** — Neste trabalho foi proposto um algoritmo para a estimação automática da posição, da distância percorrida e do rumo de uma aeronave através do reconhecimento e rastreamento de objetos georreferenciados em imagens obtidas por visada nadir. O algoritmo utiliza correspondência entre histogramas normalizados para o reconhecimento dos objetos, e utiliza triangulação para a estimação da posição da aeronave. As únicas informações necessárias para a estimação da posição são: a altura de navegação da aeronave; e as posições geográficas e azimute dos objetos usados como referência. Resultados são apresentados e analisados.

**Palavras-Chave** — inteligência artificial, visão computacional, estimação de informações de vôo.

## I. INTRODUÇÃO

Devido ao problema de precisão e das falhas do sinal do *Global Positioning System* (GPS) causados, por exemplo, pela vulnerabilidade a interferências eletromagnéticas, técnicas de visão computacional vêm sendo estudadas como alternativas para a estimação automática de posição em sistemas de navegação de aeronaves [1][2][3][4][5].

Neste trabalho, foi proposto um algoritmo para a estimação automática da posição geográfica, da distância percorrida e do rumo de aeronaves através do reconhecimento e rastreamento de objetos georreferenciados em imagens obtidas durante a navegação.

A Seção 2 deste trabalho apresenta o algoritmo proposto. Resultados da aplicação de uma implementação computacional do algoritmo são descritos na Seção 3. Na Seção 4 são apresentadas observações finais referentes aos resultados obtidos.

## II. ALGORITMO PARA A ESTIMAÇÃO AUTOMÁTICA DE POSIÇÃO

O algoritmo é composto por quatro principais processos, sendo que o primeiro processo consiste na aquisição de uma imagem  $I$  através de visada nadir. O algoritmo é descrito na Fig. 1. Nesta figura, a numeração no canto superior à esquerda dos blocos indica o processo ao qual pertence cada bloco. O processo de reconhecimento e de rastreamento de objetos de referência em imagens baseia-se no método de correspondência entre histogramas normalizados de objetos georreferenciados  $R_1$  e  $R_2$  e histogramas normalizados de parcelas (janelas de busca) de imagens obtidas durante a navegação [6][7].

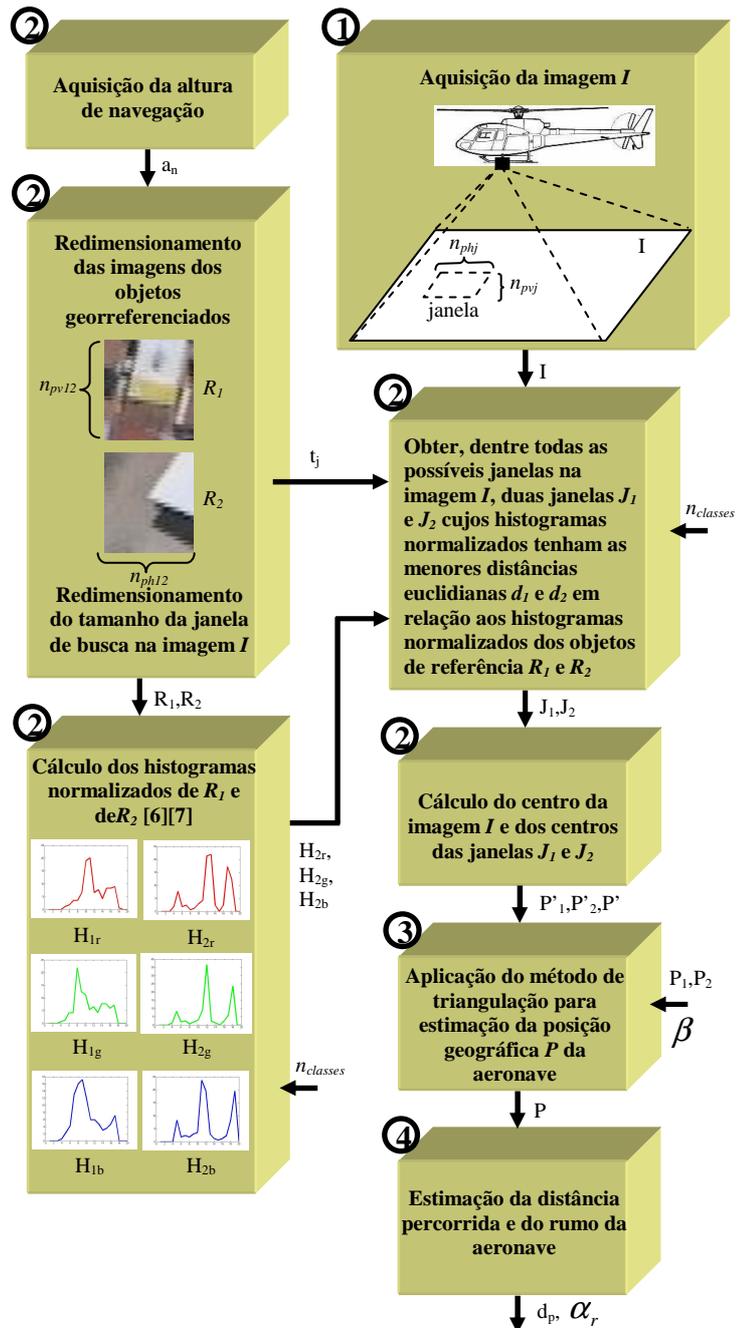


Fig. 1. Algoritmo proposto.

Considerando uma imagem  $M$  do tipo RGB, isto é, uma imagem composta por três matrizes *Red* ( $M_r$ ), *Green* ( $M_g$ ) e *Blue* ( $M_b$ ), o histograma normalizado de cada componente desta imagem é uma distribuição normalizada dos valores dos  $n_l n_c$  píxeis em  $n_{classes}$  classes. O número  $n_{classes}$  é um parâmetro de entrada do algoritmo. Estes histogramas são definidos por (1), (2) e (3).

$$H(M_r) = H_r = \left\{ h_1 = \sum_{l=1}^{n_l} \sum_{c=1}^{n_c} k = \begin{cases} \frac{1}{n_l n_c}, se 0 \leq m_{rlc} < q \\ 0, se m_{rlc} \geq q \end{cases} \right. \\ h_2 = \sum_{l=1}^{n_l} \sum_{c=1}^{n_c} k = \begin{cases} \frac{1}{n_l n_c}, se q \leq m_{rlc} < 2q \\ 0, se m_{rlc} < q e m_{rlc} \geq 2q \end{cases} \\ \dots$$

$$h_{n_{classes}-1} = \sum_{l=1}^{n_l} \sum_{c=1}^{n_c} k = \begin{cases} \frac{1}{n_l n_c}, se (n_{classes} - 2)q \leq m_{rlc} < (n_{classes} - 1)q \\ 0, se m_{rlc} < (n_{classes} - 2)q e m_{rlc} \geq (n_{classes} - 1)q \end{cases} \\ h_{n_{classes}} = \sum_{l=1}^{n_l} \sum_{c=1}^{n_c} k = \begin{cases} \frac{1}{n_l n_c}, se (n_{classes} - 1)q \leq m_{rlc} < n_{classes} q \\ 0, se m_{rlc} < (n_{classes} - 1)q e m_{rlc} \geq n_{classes} q \end{cases}$$

$$H(M_g) = H_g = \left\{ h_1 = \sum_{l=1}^{n_l} \sum_{c=1}^{n_c} k = \begin{cases} \frac{1}{n_l n_c}, se 0 \leq m_{glc} < q \\ 0, se m_{glc} \geq q \end{cases} \right. \\ h_2 = \sum_{l=1}^{n_l} \sum_{c=1}^{n_c} k = \begin{cases} \frac{1}{n_l n_c}, se q \leq m_{glc} < 2q \\ 0, se m_{glc} < q e m_{glc} \geq 2q \end{cases} \\ \dots$$

$$h_{n_{classes}-1} = \sum_{l=1}^{n_l} \sum_{c=1}^{n_c} k = \begin{cases} \frac{1}{n_l n_c}, se (n_{classes} - 2)q \leq m_{glc} < (n_{classes} - 1)q \\ 0, se m_{glc} < (n_{classes} - 2)q e m_{glc} \geq (n_{classes} - 1)q \end{cases} \\ h_{n_{classes}} = \sum_{l=1}^{n_l} \sum_{c=1}^{n_c} k = \begin{cases} \frac{1}{n_l n_c}, se (n_{classes} - 1)q \leq m_{glc} < n_{classes} q \\ 0, se m_{glc} < (n_{classes} - 1)q e m_{glc} \geq n_{classes} q \end{cases}$$

$$H(M_b) = H_b = \left\{ h_1 = \sum_{l=1}^{n_l} \sum_{c=1}^{n_c} k = \begin{cases} \frac{1}{n_l n_c}, se 0 \leq m_{blc} < q \\ 0, se m_{blc} \geq q \end{cases} \right. \\ h_2 = \sum_{l=1}^{n_l} \sum_{c=1}^{n_c} k = \begin{cases} \frac{1}{n_l n_c}, se q \leq m_{blc} < 2q \\ 0, se m_{blc} < q e m_{blc} \geq 2q \end{cases} \\ \dots$$

$$h_{n_{classes}-1} = \sum_{l=1}^{n_l} \sum_{c=1}^{n_c} k = \begin{cases} \frac{1}{n_l n_c}, se (n_{classes} - 2)q \leq m_{blc} < (n_{classes} - 1)q \\ 0, se m_{blc} < (n_{classes} - 2)q e m_{blc} \geq (n_{classes} - 1)q \end{cases} \\ h_{n_{classes}} = \sum_{l=1}^{n_l} \sum_{c=1}^{n_c} k = \begin{cases} \frac{1}{n_l n_c}, se (n_{classes} - 1)q \leq m_{blc} < n_{classes} q \\ 0, se m_{blc} < (n_{classes} - 1)q e m_{blc} \geq n_{classes} q \end{cases}$$

Em que:  $n_l$  e  $n_c$  são os números de linhas e de colunas de uma componente da matriz  $M$ ;  $m_{rlc}$  é uma célula de  $M_r$ ;  $m_{glc}$  é uma célula de  $M_g$ ;  $m_{blc}$  é uma célula de  $M_b$ ;  $q$  é o tamanho do intervalo de valores de cada classe, sendo definido por

$256/n_{classes}$ ; e  $h_{i_{classe}}$  é um componente do histograma, cujo valor é uma normalização da quantidade de células de um componente de  $M$ , cujos valores pertencem a uma classe  $i_{classe}$  do histograma.

A Fig. 2 apresenta um exemplo de histogramas normalizados dos componentes de uma imagem, considerando  $n_{classes} = 256$ . É importante ser observado que o valor de um píxel é um número inteiro pertencente ao intervalo  $[0, \dots, 255]$ . Nos histogramas apresentados neste trabalho, o eixo horizontal representa a distribuição das classes e o eixo vertical representa a normalização do número de píxeis pertencentes a cada classe.

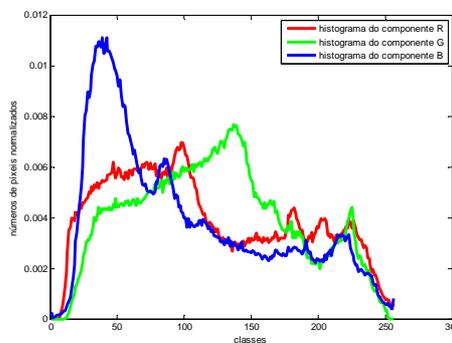


Fig. 2. Histogramas de uma imagem RGB.

O segundo processo estima as posições  $P'_1$  e  $P'_2$  dos objetos  $R_1$  e  $R_2$  e a posição central  $P'$  da aeronave na imagem.

O método de correspondência entre histogramas normalizados foi adotado por permitir que o reconhecimento de objetos seja invariante à rotação da imagem [6][7]. Tal método de correspondência consiste em obter, dentre todas as possíveis janelas de uma imagem  $I$ , duas janelas  $J_1$  e  $J_2$  cujos histogramas normalizados tenham as menores distâncias euclidianas  $d_1$  e  $d_2$  em relação aos histogramas normalizados dos objetos de referência  $R_1$  e  $R_2$ . As distâncias  $d_1$  e  $d_2$  são calculadas por

$$d_1 = \min\{d(H_{1r}, H_{i_{lc,r}}) + d(H_{1g}, H_{i_{lc,g}}) + d(H_{1b}, H_{i_{lc,b}})\} \forall i_{lc} \quad (4)$$

$$d_2 = \min\{d(H_{2r}, H_{i_{lc,r}}) + d(H_{2g}, H_{i_{lc,g}}) + d(H_{2b}, H_{i_{lc,b}})\} \forall i_{lc} \quad (5)$$

Em que:  $i_{lc}$  é a célula central de uma janela da imagem  $I$ ;  $H_{i_{lc,r}}$ ,  $H_{i_{lc,g}}$  e  $H_{i_{lc,b}}$  são os histogramas normalizados dos componentes de uma janela cujo centro é a célula  $i_{lc}$ ; e  $d$  é a distância entre dois histogramas.

A variação de escala causa o problema de diferença de dimensões entre a imagem utilizada para extrair as referências  $R_1$  e  $R_2$ , e a imagem  $I$  usada no processo de

reconhecimento destas referências. Visando à redução do efeito causado pela variação de escala, neste trabalho, a altura de navegação  $a_n$  é utilizada para ajustar as dimensões de  $R_1$ , de  $R_2$  e da janela de busca em uma imagem  $I$ . O número de píxeis na horizontal ( $n_{phj}$ ) e o número de píxeis na vertical ( $n_{pvj}$ ) de uma janela de busca são definidos por

$$n_{phj} = n_{pvj} = \frac{t_j}{t_p} \quad (5)$$

Em que:  $d_j$  é a medida real de um dos lados que delimitam a região representada por uma janela; e  $t_p$  é a medida, em metros, de um dos lados que delimitam a região representada por um píxel.

Os objetos  $R_1$  e  $R_2$  representam regiões quadradas de mesma dimensão da superfície sobrevoada. Cada objeto  $R_1$  ou  $R_2$  está associado com a posição geográfica  $P_1$  ou  $P_2$  do centro da região representada pelo mesmo. O tamanho  $t_p$  é calculado através da resolução do sensor utilizado na captação das imagens. De forma similar, o ajuste das dimensões de  $R_1$  e de  $R_2$  é definido por

$$n_{ph12} = n_{pv12} = \frac{t_{12}}{t_p} \quad (6)$$

Em que:  $n_{ph12}$  é o número de píxeis na horizontal da imagem do objeto  $R_1$  ou do objeto  $R_2$ , como pode ser verificado na Fig. 1;  $n_{pv1}$  é o número de píxeis na vertical da imagem do objeto  $R_1$  ou do objeto  $R_2$ ; e  $t_{12}$  é a medida real de um dos lados que delimitam a região representada por  $R_1$  ou da região representada por  $R_2$ .

No terceiro processo, a posição geográfica  $P$  da aeronave é estimada através de: informações provenientes de um método de triangulação entre as posições  $P'_1$ ,  $P'_2$  e  $P'$  na imagem  $I$ ; posições geográficas  $P_1$  e  $P_2$  das referências; e azimute  $\beta$  do segmento de reta  $\overline{P_1P_2}$ .

O último processo estima a distância percorrida ( $d_p$ ) e o rumo da aeronave ( $\alpha_r$ ) entre as penúltimas e as últimas posições estimadas através do terceiro processo. O rumo de uma aeronave é o ângulo formado pela projeção ortogonal ao solo do movimento da aeronave e um meridiano. O rumo pode se diferenciar da proa da aeronave ( $\alpha_p$ ) devido ao efeito do vento ( $v_v$ ). A proa é o ângulo formado pelo eixo longitudinal da aeronave e um meridiano. A Fig. 3 apresenta um esquema com algumas informações de vôo, incluindo a velocidade em relação ao solo ( $v_s$ ).

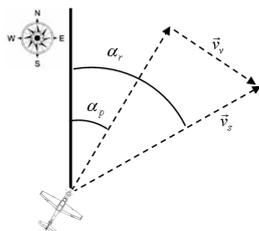


Fig. 3. Esquema com algumas informações de vôo de uma aeronave.

A distância percorrida  $d_p$  corresponde ao somatório das distâncias entre as posições estimadas consecutivamente.

### III. RESULTADOS

Dois trechos de uma videografia obtida por um helicóptero [8] foram utilizados para testar e validar uma implementação computacional do algoritmo proposto. Os testes foram realizados em um computador com um processador de 1.73 GHz e 1 GB de memória RAM.

A Fig. 4 apresenta as imagens dos objetos georreferenciados e seus respectivos histogramas normalizados utilizados como referência para a estimação de posição no primeiro trecho de vídeo. Analisando a Fig. 4, pode-se observar que foram utilizados histogramas com 20 classes para os experimentos com o primeiro trecho de vídeo.

As imagens dos objetos foram obtidas da própria videografia. Porém, é possível a utilização de outras imagens obtidas ou não por diferentes sensores, desde que seja feito um ajuste dos histogramas destas imagens baseado em histogramas de imagens obtidas com o mesmo sensor usado na videografia.

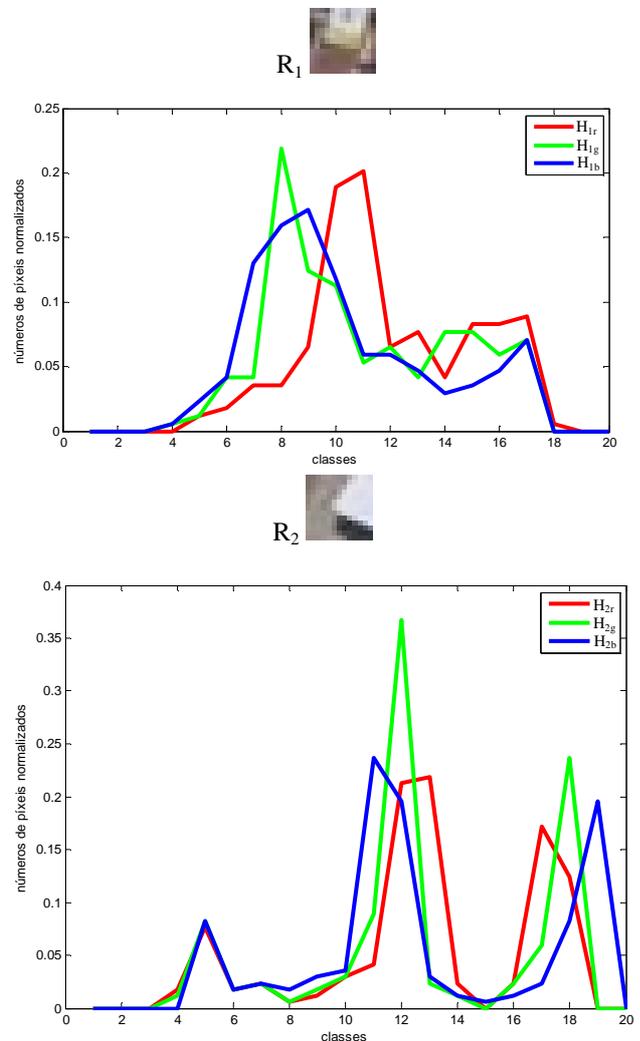


Fig. 4. Imagens dos objetos georreferenciados  $R_1$  e  $R_2$  e seus respectivos histogramas normalizados utilizados como referências para o primeiro trecho de vídeo.

Considerando as posições geográficas  $P_1$  e  $P_2$  de dois objetos de referência, a Fig. 5 apresenta a seqüência de quatro posições geográficas  $P$  da aeronave, estimadas pela implementação do algoritmo no primeiro trecho de vídeo.

Nesta implementação, a taxa de estimação de posição da aeronave foi de 1.8 posição por segundo, considerando uma janela de busca representada por uma matriz quadrada de ordem 20. É importante ser mencionado que há três janelas de busca, uma janela para cada componente da imagem RGB.



Fig. 5. Exemplos de posições estimadas pelo algoritmo no primeiro trecho de vídeo.

Na Tab. 1 são descritas as posições geográficas estimadas da aeronave e correspondentes àquelas apresentadas na Fig. 5. Outra informação apresentada nesta figura é o rumo entre duas posições estimadas. A distância real percorrida pelo helicóptero no primeiro trecho de vídeo é de 161.68 m, e a implementação do algoritmo estimou uma distância percorrida de 156.68 m.

TABELA I ALGUMAS INFORMAÇÕES DE VÔO ESTIMADAS PELO ALGORITMO PARA O PRIMEIRO TRECHO DE VÍDEO

Posição Geográfica da Aeronave (P)	Latitude	Longitude	Rumo
1	-23.19307005197	-45.87270033666	-
2	-23.19278933338	-45.87262594488	14.85°
3	-23.19250463296	-45.87251393711	21.48°
4	-23.19232167174	-45.87235787489	40.47°

Com a finalidade de analisar a precisão dos resultados obtidos, as informações estimadas e apresentadas na Tab. 1 foram comparadas com as respectivas informações reais do veículo. Estas informações reais são apresentadas na Tab. 2.

TABELA II ALGUMAS INFORMAÇÕES DE VÔO REAIS PARA O PRIMEIRO TRECHO DE VÍDEO

Posição Geográfica da Aeronave (P)	Latitude	Longitude	Rumo
1	-23.193086	-45.872814	-
2	-23.192794	-45.872730	14.04°
3	-23.192466	-45.872561	24.43°
4	-23.192270	-45.872380	45.72°

Analisando as duas tabelas, pôde-se calcular o erro médio entre os rumos estimados e os rumos reais, que é igual a 3°, e o erro médio entre as posições geográficas estimadas e as posições reais, que é igual a 9.35 m.

A Fig. 6 apresenta as imagens dos objetos georreferenciados e seus respectivos histogramas normalizados utilizados como referência para a estimação de posição no segundo trecho de vídeo. Analisando a Fig. 4, pode-se observar que foram utilizados histogramas com 40 classes.

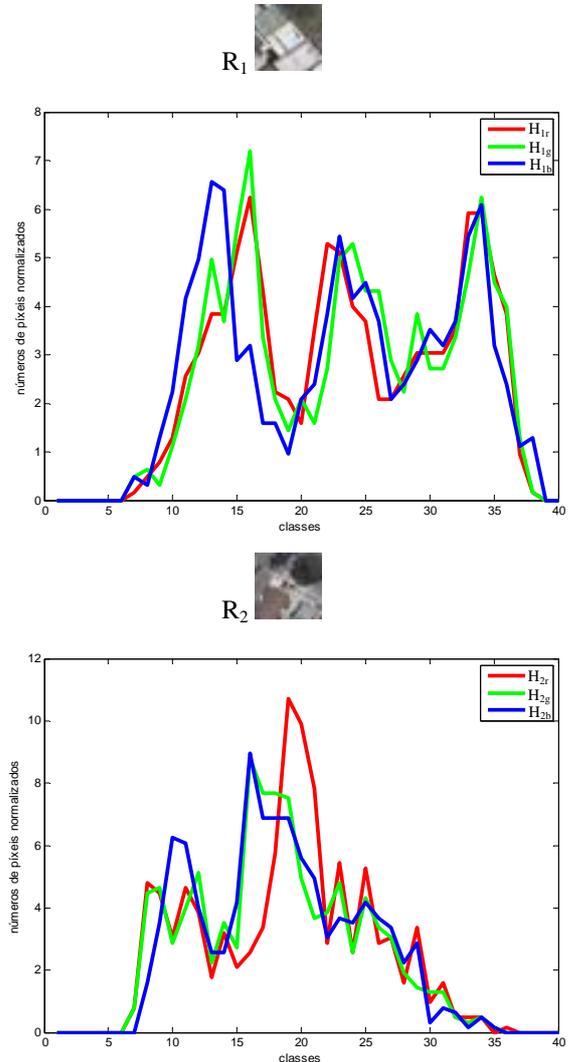


Fig. 6. Objetos georreferenciados  $R_1$  e  $R_2$  e seus respectivos histogramas normalizados utilizados como referências para o segundo trecho de vídeo.

Novamente, considerando as posições geográficas  $P_1$  e  $P_2$  dos dois objetos de referência, a Fig. 7 apresenta a seqüência de quatro posições geográficas  $P$  da aeronave, estimadas pela implementação do algoritmo no segundo trecho de vídeo.

A taxa de estimação de posição da aeronave foi de 0.32 posição por segundo, também considerando uma janela de busca representada por uma matriz quadrada de ordem 40.



Fig. 7. Exemplos de posições estimadas pelo algoritmo no segundo trecho de vídeo.

As posições geográficas estimadas da aeronave e equivalentes àquelas apresentadas na Fig. 7 são descritas na Tab. 3, juntamente com os respectivos rumos entre duas posições consecutivas. A distância real percorrida pelo helicóptero no segundo trecho de vídeo é de 152.95 m, e a implementação do algoritmo estimou uma distância percorrida de 144.1 m.

TABELA III ALGUMAS INFORMAÇÕES DE VÔO ESTIMADAS PELO ALGORITMO PARA O SEGUNDO TRECHO DE VÍDEO

Posição Geográfica da Aeronave (P)	Latitude	Longitude	Rumo
1	-23.18054932171	-45.87085406636	-
2	-23.18003568589	-45.87108808319	335.51°
3	-23.17977285358	-45.87119328118	336.40°
4	-23.17958121165	-45.87133010457	333.82°

Com a finalidade de analisar a precisão dos resultados obtidos, as informações estimadas e apresentadas na Tab. 3 foram comparadas com as respectivas informações reais do veículo. Estas informações reais são apresentadas na Tab. 4.

TABELA IV ALGUMAS INFORMAÇÕES DE VÔO REAIS PARA O SEGUNDO TRECHO DE VÍDEO

Posição Geográfica da Aeronave (P)	Latitude	Longitude	Rumo
1	-23.180508	-45.870917	-
2	-23.179941	-45.871118	340.48°
3	-23.179761	-45.871232	327.66°
4	-23.179525	-45.871383	327.39°

Analisando as duas tabelas, pôde-se calcular o erro médio entre os rumos estimados e os rumos reais, que é igual a 7.21°, e o erro médio entre as posições geográficas estimadas e as posições reais, que é igual a 8.13 m.

#### IV. OBSERVAÇÕES FINAIS

A abordagem proposta neste trabalho mostrou-se promissora no contexto de estimação de localização de uma aeronave sem apoio de ferramentas de localização como, por exemplo, o GPS. Através dos resultados obtidos pela implementação do algoritmo, foi possível constatar uma precisão em torno de 9 metros para a estimação de posições geográficas. Esta precisão corresponde ao erro médio das posições estimadas, o qual foi obtido quando a implementação do algoritmo foi aplicada ao primeiro trecho

de vídeo. Esta precisão de 9 metros pode ser considerada aceitável sabendo-se que a precisão do GPS em sistemas de navegação pode alcançar 15 metros. Também foi possível verificar que as taxas de estimação de posição da aeronave são satisfatórias para aplicações em tempo real. Como trabalho futuro, é sugerida a adaptação do algoritmo para imagens obtidas por visada oblíqua.

#### REFERÊNCIAS

- [1] F. L. L. Medeiros, E. H. Shiguemori, “Estimação automática da posição de uma aeronave através de reconhecimento e rastreamento de objetos em imagens”, XII Workshop Anual de Pesquisa e Desenvolvimento do Instituto de Estudos Avançados (IEAv) (WAI), 2012, p. 95.
- [2] A. E. R. Shabayek, C. Démonceaux, O. Morel, D. Fofi, “Vision Based UAV Attitude Estimation: Progress and Insights” In Journal of Intelligent and Robotic Systems, vol. 65, 2012, p. 295–308.
- [3] S. Ahrens, D. Levine, G. Andrews, J. P. How, “Vision-Based Guidance and Control of a Hovering Vehicle in Unknown, GPS-denied Environments”, IEEE International Conference on Robotics and Automation, 2009, p. 2643–2648.
- [4] E. H. Shiguemori, M. V. T. Monteiro, M. P. Martins, “Landmarks recognition for autonomous aerial navigation by neural networks and Gabor transform”, IS&T/SPIE 19th Annual Symposium Electronic Imaging Science and Technology, 2007, p. 64970-64970.
- [5] L. Merino, J. Wiklund, F. Caballero, A. Moe, J. R. Martínez-de Dios, P.-E. Forssén, K. Nordverg, A. Ollero, “Vision-Based Multi-UAV Position Estimation” In IEEE Robotics & Automation Magazine, vol. 13, 2006, p. 53–62.
- [6] O. Linde, T. Lindeberg, “Composed complex-cue histograms: an investigation of the information content in receptive field based image descriptors for object recognition” In Computer Vision and Image Understanding, vol. 116, 2012, p. 538-560.
- [7] Y. Shan, H. S. Sawhney, B. Matei, R. Kumar, “Shapeme histogram projection and matching for partial object recognition” In IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, 2006, p. 568-577.
- [8] M. P. Martins, F. L. L. Medeiros, M. V. Monteiro, E. H. Shiguemori, M. A. Pizani, L. C. A. Ferreira, “Navegação aérea autônoma por imagens”, VIII Simpósio de Guerra Eletrônica (SIGE), 2006.