

Uso do Fluxo Óptico para detecção de objetos móveis em imagens aéreas

Rodrigo Augusto Rebouças, Mateus Habermann e Elcio Hideiti Shiguemori
Instituto de Estudos Avançados – Divisão de Geointeligência, São José dos Campos - SP

Resumo — O presente artigo aborda a comparação entre duas técnicas de processamento de imagens que buscam localizar objetos móveis em uma sequência de imagens aéreas para aplicação de veículos aéreos não tripulados. Uma das técnicas empregada faz uso do *Speeded Up Robust Features* (SURF) e a outra faz uso da análise do Fluxo Óptico, ambas visam a compensar o deslocamento da câmera. Para detecção de objetos em movimento calcula-se o Fluxo Óptico. Esta análise visa a técnica menos custosa computacionalmente para aplicações que requerem processamento em tempo-real e embarcado.

Palavras-Chave — Processamento de imagens, fluxo óptico, navegação autônoma, VANT.

I. INTRODUÇÃO

O uso de imagens digitais tem crescido nos últimos anos devido, principalmente, ao avanço tecnológico e computacional, como o aumento da capacidade de armazenamento, processamento de dados e o crescente número de pesquisas nas áreas de Visão Computacional e Processamento de Imagens [1] [2]. Seu emprego possui aplicações em diferentes áreas como: sensoriamento remoto [3], segurança [4], navegação autônoma aérea [5], medicina [6] e esportes [7]. Dentre as aplicações do uso de imagens digitais destacam-se as que utilizam imagens aéreas por possibilitarem o monitoramento de uma grande área geográfica. Neste estudo aborda-se o problema de reconhecimento de objetos móveis [8]-[13], onde diferentes trabalhos tratam esse problema, como a que faz uso do algoritmo SIFT [9] [14] para compensar o deslocamento da câmera e operações morfológicas para identificação de veículos em movimento, em outra abordagem calcula-se o fluxo óptico com uso de Unidade de Processamento Gráfico (GPU) [10], em [11] faz-se uso do fluxo óptico, onde são segmentados diversos objetos móveis ao mesmo tempo e é construído o mapa do ambiente, em [12] faz-se uso do fluxo óptico adaptado para objetos semitransparentes, neste caso, a suposição de intensidade constante não é mais válida e em [13] faz-se uso do fluxo óptico para detecção de objetos móveis considerados rápidos.

Especificamente, neste trabalho aborda-se o problema de identificação de objetos móveis em imagens aéreas, tendo como objetivo o desenvolvimento de um método que possibilite a identificação de veículos em movimento com uso de técnicas de reconhecimento de padrões [15], que pode ser empregado como base para a implementação de ferramentas de extração de informação e monitoramento de grandes áreas. São consideradas imagens de uma videografia realizada em uma área contendo veículos em movimento, na cidade de São José dos Campos - SP. Duas técnicas de processamento de imagens são empregadas e comparadas: o *Speed Up Robust Features* (SURF) [16] para detecção de pontos característicos e o Fluxo Óptico com o método de Lucas-Kanade [17] para formação de vetores de deslocamento. Em ambas as técnicas visa-se compensar a diferença das posições de captura do primeiro e segundo *frame*. Em seguida, calcula-se o fluxo óptico para a detecção dos veículos. Este projeto é desenvolvido para aplicação no projeto PITER (Processamento de Imagens em Tempo-Real), que tem como objetivo realizar pesquisas que envolvam o processamento de imagens em tempo real obtidas por um VANT (Veículo Aéreo Não Tripulado) para aplicação na navegação aérea autônoma.

II. METODOLOGIA

O objetivo do estudo é a identificação de objetos móveis em imagens aéreas. São utilizadas imagens de uma videografia realizada em São José dos Campos - SP. As principais técnicas empregadas neste estudo fazem uso do SURF [16] e do Fluxo Óptico para compensar a movimentação da câmera. Em seguida, para análise do deslocamento dos veículos, são usadas técnicas baseadas no Fluxo Óptico. Estas duas abordagens são ilustradas na Fig. 1.

Na Fig. 1 (a), é feita a captura entre dois *frames* consecutivos nos tempos t e $t + \Delta t$. Em seguida é aplicada a técnica SURF [16] para detecção de pontos característicos. Com esta informação é feito o corte em ambas as imagens para se compensar o deslocamento da câmera do VANT. Em seguida, são estimados vetores do Fluxo Óptico com o método de Lucas-Kanade [17] para identificação dos objetos móveis entre os tempos t e $t + \Delta t$. Na Fig. 1 (b) substitui-se o SURF e faz-se uso do Fluxo Óptico.

R. A. Rebouças, rodrigo.pressa@gmail.com, M. Habermann, habermann@ieav.cta.br, E. H. Shiguemori, elcio@ieav.cta.br

Este trabalho foi parcialmente financiado pelo CNPq, através da bolsa de Iniciação Científica PIBIC-IEAV e ao COMAER, através do Projeto PITER, No. V01050303.

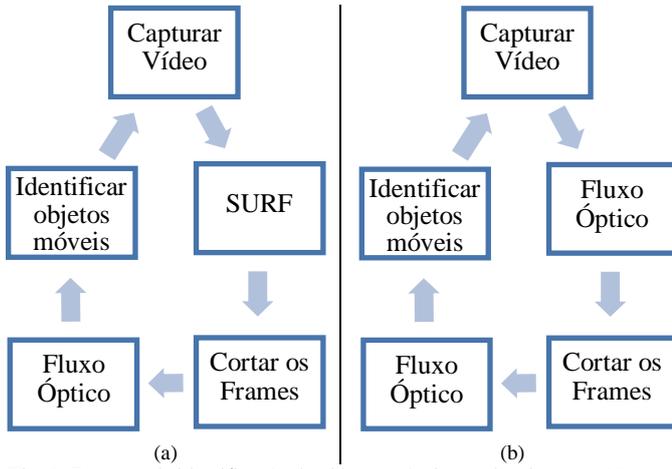


Fig. 1. Processo de identificação de objetos móveis, sendo a imagem (a) com uso do SURF e a (b) com uso do Fluxo Óptico.

III. SURF

Por serem utilizadas imagens aéreas obtidas em diferentes instantes de tempo, faz-se necessário compensar o movimento da câmera. São utilizadas diferentes técnicas para o registro de imagens [18], uma delas com uso do descritor SURF [16] bastante utilizada na literatura por apresentar desempenho computacionalmente superior, quando comparada com outros descritores [19].

O SURF [16] é uma técnica para extrair, descrever e casar pontos-chaves ou pontos de interesse de imagens invariantes a sua iluminação, rotação e escala. Basicamente, ele é dividido em 4 partes: criação da imagem integral, extração dos pontos de interesse, descrição dos pontos de interesse e casamento de padrões (*Template Matching*) desses pontos [20].

A extração dos pontos de interesse é feita com uso do determinante da matriz Hessiana $H(x, \sigma)$, *Fast-Hessian*, como é mostrado em (1).

$$H(p, \sigma) = \begin{bmatrix} L_{xx}(p, \sigma) & L_{xy}(p, \sigma) \\ L_{xy}(p, \sigma) & L_{yy}(p, \sigma) \end{bmatrix} \quad (1)$$

onde $L_{xx}(p, \sigma)$ é a convolução da derivada da Gaussiana de segunda ordem com a imagem no ponto p , com as coordenadas (x, y) e σ é a escala da imagem.

E o determinante $\det(H)$, por custo computacional, é descrito em (2) [16].

$$\det(H) = D_{xx}D_{yy} - (0,9D_{xy})^2 \quad (2)$$

onde D_{xx} é a segunda derivada em relação à x .

Para se reduzir o custo computacional, o processo é calculado em imagens integrais da própria imagem. Em cada um desses pontos é construído um vetor de características com uso de sua vizinhança, formando os descritores. Na Fig. 2 são apresentados os pontos de interesse com o raio de acordo com a sua vizinhança [16].

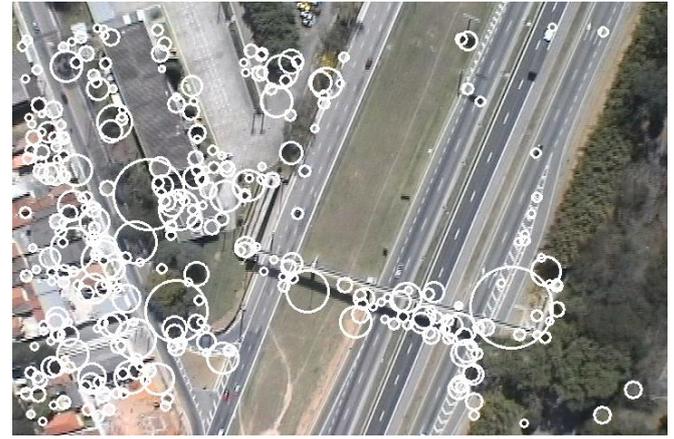


Fig. 2. Pontos de interesse.

Em seguida, é feita a correspondência dos descritores dos pontos de interesse P_1 e P_2 com as coordenadas (x_1, y_1) e (x_2, y_2) . Um ponto de interesse no tempo t é comparado com cada ponto de interesse no tempo $t + \Delta t$ por meio da distância euclidiana $De(P_1, P_2)$ entre os seus vetores de descrição, como é apresentado em (3) [16].

$$De(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (3)$$

Se a distância for menor que a distância do vizinho mais próximo, então atribui-se como casamento, senão segue para o ponto seguinte. Isso é feito com todos os pontos no tempo t para o tempo $t + \Delta t$ [16].

Na Fig. 3 é apresentado um resultado do casamento, porém, para melhor visualização, é apresentado apenas 11% desses casamentos. As linhas que conectam as duas imagens indicam os pontos correspondentes.



Fig. 3. Correspondência entre pontos de interesse.

III. FLUXO ÓPTICO

A segunda técnica utilizada neste trabalho para compensar o movimento da câmera faz uso do Fluxo Óptico com o método de Lucas-Kanade [17], que tem sido utilizada para detecção de movimento [10]-[13].

O Fluxo Óptico é obtido pelo movimento aparente do padrão de brilho na imagem quando a informação varia de acordo com seus pontos de contraste [21], onde calcula-se o deslocamento de objetos de acordo com o movimento relativo entre a câmera e a cena. Tal cálculo é realizado em sequências de imagens, no sentido de que a velocidade instantânea pode ser calculada entre dois *frames* consecutivos.

Assim, esses dois *frames* são analisados nos tempos t e $t + \Delta t$. Como a luminosidade na cena se mantém constante

no decorrer do tempo, tem-se para um objeto em movimento [21]:

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t) \quad (4)$$

Onde $I(x, y, t)$ é a intensidade da imagem no píxel na posição (x, y) e no tempo t , podendo ser expandida a (4) pela série de Taylor, como apresentada em (5) [21].

$$\begin{aligned} I(x + \Delta x, y + \Delta y, t + \Delta t) \\ = I(x, y, t) + \frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y \\ + \frac{\partial I}{\partial t} \Delta t + O^2 \end{aligned} \quad (5)$$

Em que O^2 representa os termos de ordem superior. Então resulta que [21]:

$$\frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t = 0 \quad (6)$$

Ou

$$\frac{\partial I}{\partial x} \frac{\Delta x}{\Delta t} + \frac{\partial I}{\partial y} \frac{\Delta y}{\Delta t} + \frac{\partial I}{\partial t} \frac{\Delta t}{\Delta t} = \frac{\partial I}{\partial x} V_x + \frac{\partial I}{\partial y} V_y + \frac{\partial I}{\partial t} = 0 \quad (7)$$

Em que V_x e V_y são as componentes de velocidade nos eixos x e y , respectivamente [21].

Assim, tem-se que:

$$I_x V_x + I_y V_y = -I_t \quad (8)$$

Ou ainda:

$$\nabla I^T \vec{V} = -I_t \quad (9)$$

Então é empregado o método de Lucas-Kanade [17], no qual píxeis da vizinhança q_1, q_2, \dots, q_n são considerados, como foi mostrado em [13]. Dessa forma, tem-se:

$$\begin{aligned} I_{x(q_1)} V_x + I_{y(q_1)} V_y &= -I_t(q_1) \\ I_{x(q_2)} V_x + I_{y(q_2)} V_y &= -I_t(q_2) \\ &\vdots \\ I_{x(q_n)} V_x + I_{y(q_n)} V_y &= -I_t(q_n) \end{aligned} \quad (10)$$

E pode-se resolver assim:

$$A \cdot v = -b \quad (11)$$

$$\text{Em que } A = \begin{bmatrix} I_{x(q_1)} & I_{y(q_1)} \\ \vdots & \vdots \\ I_{x(q_n)} & I_{y(q_n)} \end{bmatrix}, v = \begin{bmatrix} V_x \\ V_y \end{bmatrix} \text{ e } b = \begin{bmatrix} -I_t(q_1) \\ \vdots \\ -I_t(q_n) \end{bmatrix}$$

E a solução, então, fica:

$$v = (A^T A)^{-1} A^T (-b) \quad (12)$$

O vetor do fluxo óptico é calculado em cada ponto com uso da diferença entre suas coordenadas nas imagens nos tempos t e $t + \Delta t$ e dividido pelo tempo de captura entre elas.

Como a imagem inteira se movimenta, então são construídos vetores de deslocamento na cena inteira, como é apresentada na Fig. 4.

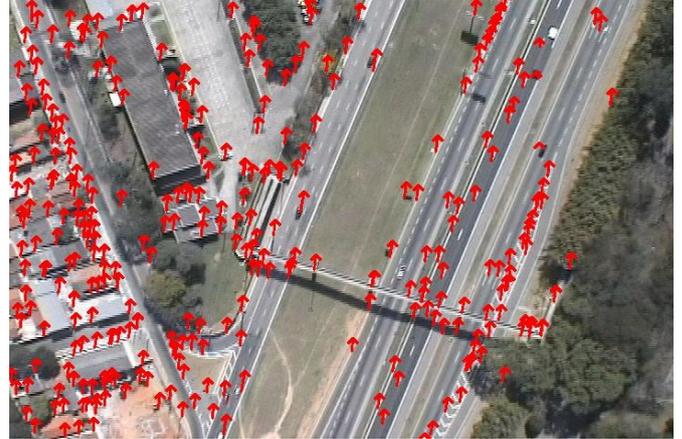


Fig. 4. Vetores de deslocamento entre as imagens nos tempos t e $t + \Delta t$.

IV. IDENTIFICAÇÃO DE DESLOCAMENTO

Ambas as técnicas podem ser usadas para estimar o deslocamento da câmera, como foi apresentado em [8]-[9]. Ou com o SURF [16], onde é feita a diferença das posições de todos os pontos de interesses, ou com uso do Fluxo Óptico com o método de Lucas-Kanade [17], com o cálculo da mediana também para obter a informação de deslocamento no tempo t , para o tempo $t + \Delta t$.

Em seguida é feito o corte em ambas as imagens, como é apresentada na Fig. 5, a fim de identificar apenas os objetos móveis desejados, neste caso, os veículos.

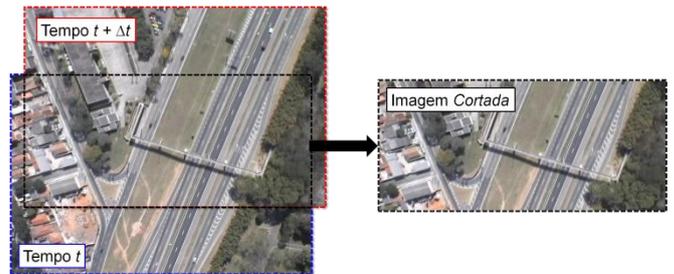


Fig. 5. Corte entre as imagens nos tempos t e $t + \Delta t$.

Para extração de informações desses veículos é calculado outro Fluxo Óptico com o método de Lucas-Kanade [17], como é apresentada na Fig. 6, para extração dos vetores de deslocamento dos veículos Fig. 6 (c) no tempo t Fig. 6 (a), para o tempo $t + \Delta t$ Fig. 6 (b).

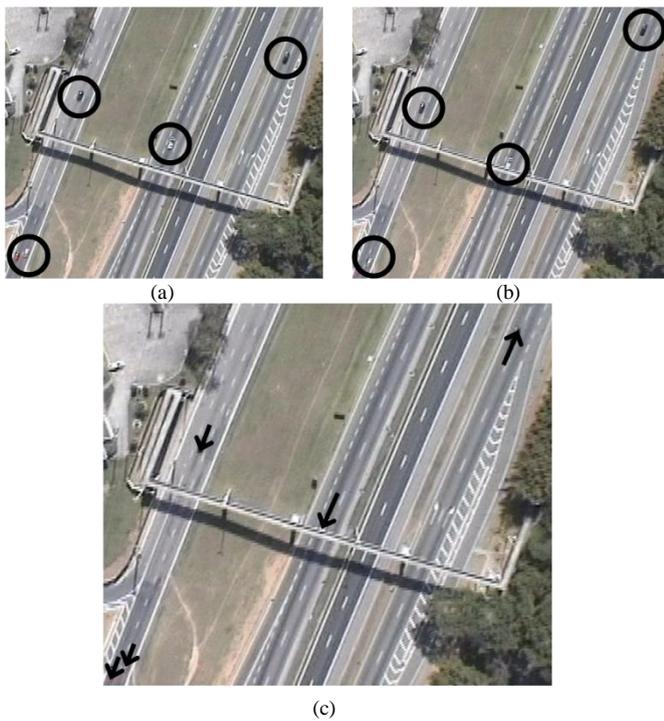


Fig. 6. Tempo t (a) Tempo $t + \Delta t$ (b) Vetor de deslocamento entre eles (c).

V. RESULTADOS

As imagens obtidas dos vídeos podem sofrer degradações, devido à movimentação da câmera, pois são detectados borrões e distorções. Esses ruídos podem ser, erroneamente, considerados em movimento juntamente com os veículos. Na Fig. 7 (a) é apresentada uma imagem capturada do vídeo com os vetores de deslocamento já extraídos, porém como é mostrada na Fig. 7 (b) o recorte da Fig. 7 (a), podem ser visualizados vários vetores de deslocamento que não são veículos. Para compensar a movimentação da câmera, apenas foi considerada como deslocamento a diferença dos vetores de fluxo óptico com pontos com maior deslocamento, como é apresentada na Fig. 7 (c), com o objetivo de identificar apenas os objetos móveis. O quadrado apresentado na Fig 7 (b) e (c) destaca o veículo em movimento.

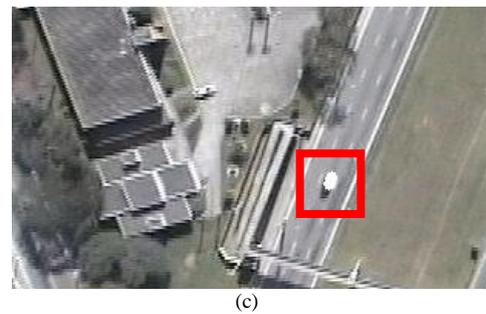
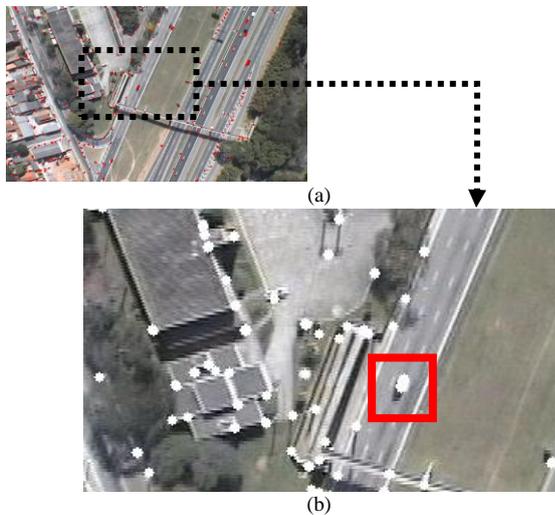


Fig. 7. Cálculo do fluxo óptico entre duas imagens (a), sendo a imagem (b) todos os vetores e a (c) apenas os vetores com maior deslocamento.

Com os resultados obtidos no estudo, avaliou-se o desempenho das técnicas em relação à Precisão (P) e Sensibilidade (S) [22], como é mostrada em (13) e (14).

$$P = \frac{VP}{VP + FP} \times 100 \quad (13)$$

$$S = \frac{VP}{VP + FN} \times 100 \quad (14)$$

São considerados os Veículos em Movimento (VM), os Verdadeiros Positivos (VP), os veículos identificados corretamente, Falsos Positivos (FP), objetos identificados que não são veículos e Falsos Negativos (FN), os veículos que não são identificados. Os frames foram extraídos da videografia realizada em São José dos Campos - SP em intervalos de 0,066 segundos.

Ambas as técnicas apresentam os mesmos resultados, ou seja, obtiveram-se as mesmas coordenadas para o corte dos frames, então sua sensibilidade e precisão são uma só, como mostrado na Tabela I, considerando os 10 primeiros frames.

TABELA I. RESULTADOS DO ESTUDO.

Frames	VM	VP	FP	FN	S (%)	P (%)
1	12	11	2	1	92	85
2	12	10	1	2	83	91
3	12	8	1	4	67	89
4	12	10	0	2	83	100
5	12	10	3	2	83	77
6	13	11	1	2	85	92
7	13	11	1	2	85	92
8	12	9	2	3	75	82
9	12	9	3	3	75	75
10	13	10	1	3	77	91
Média	12	10	2	2	80	87

Para uma melhor análise dos resultados, na Tabela II são apresentados os melhores e piores casos da aplicação em relação à Sensibilidade e Precisão [22], considerando as mesmas variáveis e intervalos da Tabela I.

TABELA II. RESULTADOS DO ESTUDO.

	VM	VP	FP	FN	S (%)	P (%)
Melhores Casos	6	6	0	0	100	100
	13	12	0	1	92	100
	8	7	0	1	88	100
Piores Casos	5	2	1	3	40	67
	14	11	5	3	79	69
	12	9	3	3	75	75

Os resultados são ilustrados nas Fig. 9 à Fig. 14 da seguinte maneira: Os círculos e elipses representam os *VP*; os triângulos representam os *FP*; e os quadrados e retângulos representam os *FN*.

Na Fig. 9, 10 e 11 são mostrados os melhores casos em relação à taxa de sensibilidade e precisão mostrada na Tabela II.



Fig. 9. Vetores de deslocamento entre as imagens nos tempos t e $t + \Delta t$



Fig. 10. Vetores de deslocamento entre as imagens nos tempos t e $t + \Delta t$.

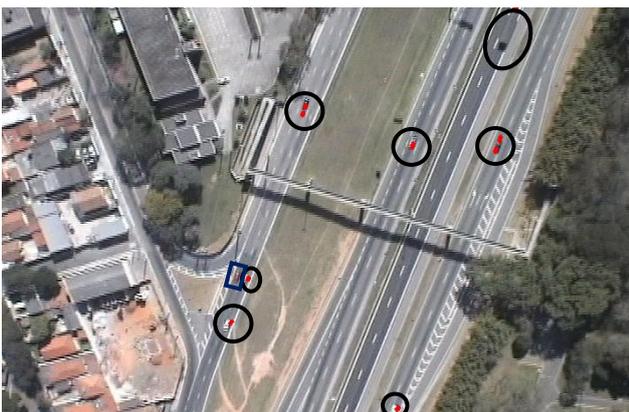


Fig. 11. Vetores de deslocamento entre as imagens nos tempos t e $t + \Delta t$.

Na Fig. 12, 13 e 14 são apresentados os piores casos em relação à taxa de sensibilidade e precisão mostrada na Tabela II.



Fig. 12. Vetores de deslocamento entre as imagens nos tempos t e $t + \Delta t$.



Fig. 13. Vetores de deslocamento entre as imagens nos tempos t e $t + \Delta t$.



Fig. 14. Vetores de deslocamento entre as imagens nos tempos t e $t + \Delta t$.

Para comparação das duas abordagens os métodos foram implementados em um computador com o processador Pentium Dual E2180 2GHz, 2GB de memória RAM no Windows XP SP3. Com os resultados apresentados nas Tabelas I e II, as técnicas apresentam os mesmos resultados, porém os tempos de processamento são diferentes.

Na Tabela III são apresentados os tempos de processamento das técnicas SURF e do Fluxo Óptico nos 10 primeiros *frames* em intervalos de 0,066 segundos.

TABELA III. COMPARAÇÃO DAS TÉCNICAS.

Frames	Tempo de processamento (s)	
	SURF	Fluxo Óptico
1	10,167	0,591
2	10,489	0,565
3	11,055	0,539
4	11,612	0,596
5	11,482	0,604
6	11,258	0,535
7	12,032	0,549
8	12,361	0,567
9	12,777	0,543
10	13,555	0,533
Média	11,679	0,562
Desvio-padrão	1,034	0,027

Com a análise dos resultados, pode-se observar que alguns veículos não foram localizados. Suspeita-se que essas falhas se devem às influências das variações na velocidade, na cor ou tamanho dos veículos. Um outro problema encontrado foi que alguns veículos são detectados por mais de um vetor de deslocamento, principalmente, veículos de grande porte ou que se movimentam em alta velocidade.

VI. CONCLUSÃO

Os resultados obtidos até o momento mostram que as abordagens com uso da técnica SURF [16] para casar pontos característicos e do Fluxo Óptico com o método de Lucas-Kanade [17] para construção dos vetores de movimento são adequadas para compensar o deslocamento da câmera. Os tempos de processamento para obtenção dos resultados com uso do Fluxo Óptico mostraram-se aproximadamente 20 vezes mais rápidos do que com uso do SURF [16]. Assim recomenda-se o método de Lucas-Kanade [17] é mais indicado que com o uso do SURF para aplicações em que se requerem o processamento em tempo-real.

Em seguida é feito outro cálculo do Fluxo Óptico para identificação de movimentos dos objetos móveis utilizados neste estudo mostram-se promissoras.

As técnicas e métodos utilizados para o processamento de imagens digitais apresentam resultados que indicam que estas abordagens podem ser empregadas, porém ressalta-se que este estudo ainda está em desenvolvimento e outras abordagens devem ser estudadas a fim de melhorar a sensibilidade, precisão e velocidade do sistema e obter mais informações sobre os objetos móveis.

REFERÊNCIAS

- [1] G. Bradski, A. Kaehler, "Learning OpenCV: Computer Vision with the OpenCV Library", 1ed., O'Reilly Media, Inc., 2008.
- [2] R. C. Gonzales, R. E. Woods, "Digital Image Processing", 3ed., Person Education, Inc., 2008.
- [3] R. L. Paes, F. Nunziata, E. H. Shiguemori, M. Migliaccio, A. Montuori, "Aplicação de Imagens SAR do Satélite Cosmo-Skymed como Suporte às Operações de Busca da Aeronave da Air France Acidentada no Oceano Atlântico", XI Workshop Anual de Pesquisas e Desenvolvimento do IEAv, p. 62, 2011.
- [4] A. Kandhalu, A. Rowe, R. Rajkumar, Chingchun Huang, Chao-Chun Yeh, "Real-Time Video Surveillance over IEEE 802.11 Mesh Networks", 15th IEEE Real-Time and Embedded Technology and Applications Symposium, p. 205-214, 2009.
- [5] A. Canhoto, E. H. Shiguemori, M. A. P. Domiciano, "Image sequence processing applied to autonomous aerial navigation", Signal and Image Processing Applications (ICSIPA), IEEE International Conference on, p. 496-499, 2009.
- [6] P. Suapang, K. Dejhan, S. Yimmun, "Medical image processing and analysis for nuclear medicine diagnosis", Control Automation and Systems (ICCAS), International Conference on, p. 2448-2451, 2010.
- [7] P. Kannan, S. Deepa, R. Ramakrishnan, "Contrast enhancement of sports images using modified sigmoid mapping function", Communication Control and Computing Technologies (ICCCCT), IEEE International Conference on, p. 651-656, 2010.
- [8] R. A. Rebouças, E. H. Shiguemori, "Acompanhamento de objetos móveis em imagens aéreas", I Simpósio de Ciência e Tecnologia do IEAv, 2012.
- [9] A. M. Rodrigues Neto, E. H. Shiguemori, A. P. A. Castro, "Identificação automática de movimentos a partir de imagen aéreas", V Seminário de Iniciação e Pós-Graduação do IEAv, p. 58-63, 2011.
- [10] S. Ozkan, "Speed-up optical flow calculation using graphical processing units", 20th Signal Processing and Communications Applications Conference (SIU), p. 1-4, Abril 2012.
- [11] Rahul Kumar Namdev, Abhijit Kundu, K. Madhava Krishna, C. V. Jawahar, "Motion segmentation of multiple objects from a freely moving monocular camera", Robotics and Automation (ICRA), IEEE International Conference on, p. 4092-4099, Maio 2012.
- [12] H. Sakaino, "A Semitransparency-Based Optical-Flow Method With a Point Trajectory Model for Particle-Like Video", Image Processing, IEEE Transactions on, vol. 21, p. 441-450, Fevereiro 2012.
- [13] Lei Chen, Hua Yang, T. Takaki, I. Ishii, "Real-time frame-straddling-based optical flow detection", Robotics and Biomimetics (ROBIO), IEEE International Conference on, p. 2447-2452, Dezembro 2011.
- [14] D. G. Lowe, "Distinctive image features from scale-invariant keypoints", International Journal of Computer Vision, vol. 60, p. 91-110, 2004.
- [15] T. E. S. Mello, R. L. Paes, "Simulação de Ambientes Virtuais para Testes de Algoritmos de Visão Computacional", XI Workshop Anual de Pesquisas e Desenvolvimento do IEAv, p. 157, 2011.
- [16] H. Bay, T. Tuytelaars, L. V. Gool, "Surf: Speeded up robust features", Proceedings of the Ninth European Conference on Computer Vision, p. 404-417, 2006.
- [17] B. D. Lucas, T. Kanade, "An Iterative Image Registration Technique with Applications in Stereo Vision", Proceedings of the DARPA Image Understanding Workshop, p. 121-130, 1981.
- [18] H. Pedrini, W. R. Schwartz, "Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações", 1ed., Thomson Learning Edições Ltda, 2007.
- [19] C. L. Rachid, A. A. S. Pereira, "Algoritmos de Busca SIFT e SURF no Uso de Dispositivos Móveis", II Encontro Multidisciplinar da Unipac Ubá, outubro 2009.
- [20] C. Evans, "Notes on the OpenSURF Library", Technical Report CSTR-09-001, University of Bristol, 2009. Disponível em: <<http://www.cs.bris.ac.uk/Publications/Papers/2000970.pdf>>. Acesso em: 20/01/2012.
- [21] A. W. C. Faria, "Fluxo Óptico". UFMG – Universidade Federal de Minas Gerais, ICEx-DCC- Visão Computacional, 2007. Disponível em: <http://www.verlab.dcc.ufmg.br/_media/cursos/visao/2007-1/alunos/alexandrewagner/optical_flow_article.pdf>. Acesso em: 16/02/2012.
- [22] T. Fawcett, "An introduction to ROC analysis", Patter Recognition Letters, vol. 27, n.8, p. 861-874, Junho 2006.