
Relatório Parcial do Projeto 02

Realidade aumentada: desenvolvimento e validação do conceito
na manutenção de aeronaves



Interfaces Avançadas
para Realidade Aumentada



PERÍODO 01/05/2013 – 31/07/2013

Unidade Elaboradora:	IARA – SOFTEX RECIFE	Data: 31/07/2013
Projeto:	Realidade aumentada: desenvolvimento e validação do conceito na manutenção de aeronaves	
Elaborador(es):	Márcio Augusto Silva Bueno	
Outras informações do(s) elaborador(es)		
e-mail:	masb@marciobueno.com	Telefone: (81) 3224-4251

Responsável: Márcio Augusto Silva Bueno
Assinatura:
Aprovado pela EMBRAER por:
Assinatura:

Índice

.....	1
Índice	2
1 Introdução	3
2 Taxonomia	3
3 MAR Baseada em Modelos	4
3.1 Baseada em Arestas	5
3.1.1 Descritores de Arestas.....	7
3.2 Baseada em Fluxo Óptico	8
3.3 Baseada em Textura	9
4 Técnicas Híbridas	11
5 Considerações Finais.....	12
6 Referências.....	13

1 Introdução

Sistemas de Realidade Aumentada sem Marcadores (MAR – *Markerless Augmented Reality*) tem como objetivo sobrepor, em tempo real, objetos 3D virtuais em ambientes reais, visando um aperfeiçoamento da percepção e da interação do usuário com o mundo. A diferença básica entre Realidade Aumentada (RA) baseada em marcadores e sem marcadores é o método utilizado para aumentar (neste caso, a percepção do usuário) o ambiente. A Figura 1 ilustra esta diferença: quando são colocados marcadores fiduciais (vide Figura 1 esquerda) para que os mesmos sejam rastreados - realidade aumentada com marcadores; enquanto que quando os próprios objetos que compõem a cena são rastreados (vide Figura 1 direita) tem-se realidade aumentada sem marcadores (MAR).



Figura 1. Aplicações de Realidade Aumentada: COM Marcadores (esquerda) e SEM Marcadores (direita)

Em MAR, qualquer parte do ambiente real pode ser utilizada como um marcador, sendo possível rastreá-la com a finalidade de aumentar o mundo real. Desta forma, o marcador não será adicionado de forma intrusiva ao ambiente. Além disso, no contexto de MAR são utilizados algoritmos de rastreamento robustos e especializados que extraem características do ambiente, possibilitando o uso destas informações na localização 3D de objetos do mundo real. Entretanto, os sistemas desenvolvidos no contexto de MAR usam técnicas mais complexas e necessitam de um maior poder computacional, para executarem em tempo real quando comparados com as técnicas baseadas em marcadores, ou seja, a complexidade e o desempenho são alguns desafios a serem enfrentados neste contexto.

2 Taxonomia

Existem diversas técnicas e classificações adotadas nesta área. Neste relatório iremos adotar a classificação proposta por [1]: baseadas em modelo e baseadas em reconstrução 3D (SfM: *Struct from Motion*), como pode ser vista na Figura 2.



Figura 2. Taxonomia de MAR monocular em tempo real [1].

No primeiro nível de classificação conforme a Figura 2, têm-se técnicas baseadas em modelos e SfM. Nas técnicas baseadas em modelos existe um conhecimento prévio sobre a cena, que são os modelos 3D utilizados para estimar a pose da câmera (posicionamento e orientação). Nas técnicas baseadas em SfM, as poses da câmera ao longo dos quadros são estimadas sem conhecimento prévio da cena, onde o conhecimento será construído através do rastreamento do ambiente, seguido de uma reconstrução 3D. Métodos baseados em modelos são menos complexos que os baseados em SfM, porém o rastreamento depende da visibilidade do objeto previamente modelado.

As técnicas baseadas em modelos podem ser subdividida em: baseadas em aresta, fluxo óptico e textura (vide Figura 2), e serão detalhadas na próxima seção. Já as técnicas baseada em SfM podem ser classificadas em: SfM em tempo real e Mono SLAM, que são técnicas que reconhecem o ambiente desconhecido através da descoberta das poses de câmeras e da estrutura 3D do ambiente real. Um maior detalhamento destas técnicas de SfM está fora do escopo deste relatório.

3 MAR Baseada em Modelos

A abordagem baseada em modelos usa os modelos CAD (*Computer Aided Design*) 3D relativos aos objetos presentes na cena real [2], para encontrar a pose da câmera e a partir desta poder posicionar corretamente os objetos virtuais na cena. Os modelos 3D são comparados com as imagens das câmeras e geram os seguintes resultados: quais os objetos que estão presentes na cena com suas respectivas posições e orientações. Se o modelo 3D do objeto a ser rastreado não existe, ele pode ser obtido através de um método de digitalização 3D, por exemplo, reconstrução 3D óptica. Este processo na maioria das vezes é realizado *off-line* e consiste em uma preparação para a fase de rastreamento onde os modelos serão correlacionados com as imagens.

A vantagem da utilização de abordagens baseadas em modelos é a possibilidade de interação entre o mundo real e o virtual, seja através de colisão ou oclusão [3], como ilustrados na Figura 3. Estes tipos de interação são implementados na aplicação explorando o conhecimento prévio sobre o posicionamento, orientação e estrutura dos objetos rastreados. Desta forma, o modelo 3D pode ser utilizado em uma simulação física, por exemplo, onde o carro virtual colide com a maquete do castelo (Figura 3 esquerda), ou em

conjunto com um algoritmo de visibilidade onde somente as partes remanescentes do carro virtual são sobrepostas à imagem (Figura 3 direita).



Figura 3. Um carro virtual colidindo com um castelo real (esquerda) e sendo ocluído por ele (direita).

Abordagens baseadas em modelos requerem um conhecimento prévio da cena real, uma vez que são necessários os modelos 3D dos objetos a serem rastreados. Devido ao caráter *off-line* do processo de aquisição dos modelos 3D, estas técnicas em geral não funcionam em tempo real e não podem ser utilizadas em ambientes totalmente desconhecidos. Após a fase de aquisição de modelos, a etapa de inicialização do rastreador é realizada, podendo ser manual ou automática, considerando a existência de um período prévio de treinamento. A inicialização é um aspecto crítico para aplicações de RA, de forma que qualquer falha no rastreamento irá forçar o usuário a ter que inicializar o sistema, muitas vezes manualmente. Outro aspecto inerente às aplicações de RA é que os objetos a serem rastreados devem estar sempre presentes na cena, para que possam ser recuperadas as poses da câmera e consequentemente aumentar o mundo real (através do correto posicionamento de objetos virtuais).

Técnicas de RA baseadas em modelos podem ser classificadas em três categorias (vide Figura 2): a primeira consiste em técnicas que utilizam somente informações a respeito das arestas dos objetos durante o rastreamento [3][4]; a segunda se baseia no fluxo óptico da sequência de imagens [5], enquanto a terceira utiliza informação sobre a textura do objeto para realizar o rastreamento [6][7]. Os aspectos relacionados a cada uma destas categorias serão descritos a seguir.

3.1 Baseada em Arestas

Nesta categoria a pose da câmera é estimada através de um casamento do modelo 3D aramado com as informações de arestas extraídas do mundo real. Este casamento é feito através da projeção do modelo na imagem, minimizando o deslocamento entre o modelo projetado e a imagem do objeto.

Para realizar a tarefa é necessária uma estimativa inicial da pose do objeto. Nos métodos baseados em arestas esta inicialização é feita manualmente. Uma vez que a primeira pose foi obtida, ela será utilizada para projetar o modelo no próximo quadro do vídeo. Assumindo que o deslocamento da câmera entre quadros consecutivos é relativamente pequeno, a estimativa da pose anterior serve como inicialização para prever a próxima pose.

Técnicas baseadas em arestas foram as primeiras abordagens a implementar rastreamento de objetos 3D em tempo real [2]. Devido à sua baixa complexidade, estas

técnicas são fáceis de serem implementadas e têm um bom desempenho. Por causa da utilização apenas de informação sobre arestas, estas técnicas são capazes de rastrear objetos especulares afetados pela iluminação do ambiente. Contudo, métodos baseados em arestas usualmente não suportam movimentos rápidos de câmera, tendo em vista que o modelo projetado estaria inicialmente muito longe da sua posição correta e o algoritmo de minimização não seria suficiente para estimar a pose correta, ocasionando uma falha no rastreamento. Outro problema desta técnica é decorrente de erros de casamento que podem ser causados por elementos do *background* ou sombras na imagem.

Conforme a taxonomia de MAR da Figura 2, métodos baseados em arestas podem ser divididos em duas subcategorias: a primeira compreendendo as técnicas que amostram alguns pontos de controle ao longo das arestas do modelo 3D aramado e comparam as projeções dos pontos com os gradientes presentes na imagem [3], e a segunda abrangendo as técnicas que detectam arestas na imagem e as associam com a projeção do modelo 3D [4].

A primeira tarefa realizada pelas técnicas que utilizam amostragem de pontos é a escolha dos pontos de controle que serão considerados na fase de casamento com os gradientes. Tomando como estimativa a pose anterior, as arestas visíveis do modelo são determinadas e somente os pontos de controle que fazem parte destas arestas serão utilizados. Após esta seleção, os pontos são projetados no plano da imagem utilizando a estimativa anterior da pose da câmera. As arestas das imagens são detectadas através do cálculo dos gradientes nas direções x e y da imagem. Após esta etapa é realizada uma busca em torno da vizinhança dos pontos projetados visando encontrar os pontos correspondentes nas arestas da imagem [8] (vide Figura 4). Por fim, o movimento relativo da câmera é estimado baseado na variação das posições dos pontos de controle.

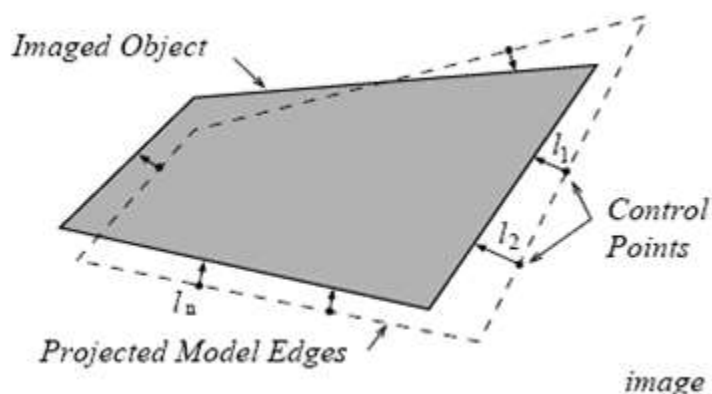


Figura 4. Casamento entre os pontos de controle projetados e a imagem [8].

Técnicas de amostragem de pontos são bastante eficientes, pois o processamento envolvido é simples. As técnicas também são genéricas com relação à forma dos objetos, podendo compreender também objetos curvos. A desvantagem de abordagens baseadas em amostragem de pontos é a falta de robustez no casamento dos pontos, que pode levar a uma estimativa de pose incorreta e à instabilidade no rastreamento. Este problema pode ser solucionado utilizando estimadores robustos que não levam em consideração o ruído na entrada (oclusões e *backgrounds* muito texturizados). Um exemplo de aplicação de MAR no qual um brinquedo (objeto virtual) é colocado sobre uma cadeira pode ser visto na Figura 5. Os pontos de controle e o sistema de coordenadas da cadeira podem ser facilmente distinguidos na imagem. Esta aplicação está descrita em [3].



Figura 5. Aplicação de MAR com amostragem de pontos [3].

Na técnica de detecção explícita de arestas, estas são extraídas da imagem através de um operador de detecção de linhas, como a transformada de Hough. A Figura 6 ilustra o processo de comparação das arestas da imagem com a projeção do modelo 3D aramado, com a finalidade de calcular a pose da câmera atual.

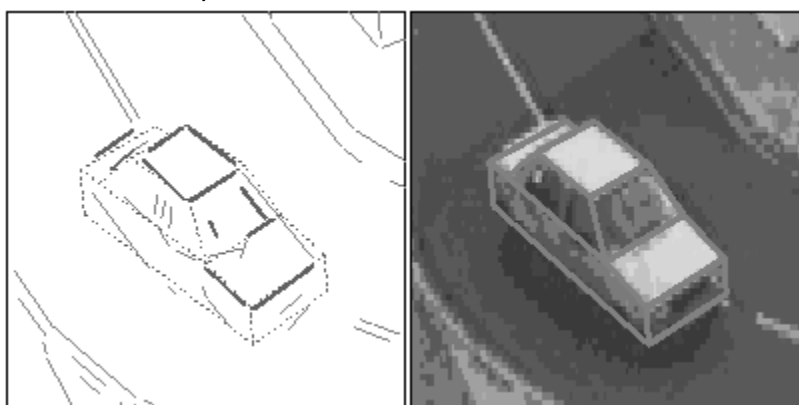


Figura 6. Linhas detectadas na imagem (esquerda) e pose estimada pelo casamento com o modelo 3D projetado (direita) [4].

Métodos de detecção explícita de arestas são mais robustos que os métodos de amostragem de pontos, porém não são genéricos com relação à forma dos modelos, uma vez que suporta apenas a utilização de objetos formados por linhas retas. Pode-se dizer também que as técnicas baseadas em amostragem são mais eficientes do ponto de vista computacional que as técnicas de detecção explícita de arestas.

3.1.1 Descritores de Arestas

Uma subárea de MAR é a baseada em arestas- trata de descritores de arestas. Os descritores são estruturas de dados que contêm informações suficientes para realizar um casamento por comparação direta. Desta forma, dois descritores iguais pertenceriam a uma mesma aresta.

O conceito de descritores foi inicialmente aplicado às características pontuais e utilizado em detecções baseadas em textura. Exemplos de descritores de características pontuais são o SIFT [18], FAST [19] e GLOH [20].

Existem outras abordagens para a detecção e extração de descritores de arestas. Uma delas [21] usa estatística através do conceito de florestas randomizadas, nesta abordagem um treinamento para a classificação das arestas baseada nos seus descritores é

realizado. A classificação é parte do casamento das arestas da imagem com as do modelo e é utilizado diretamente para o rastreamento 3D.

Outro tipo de descritor [25] de arestas é composto de um histograma de orientações, entre outras informações, e seu casamento é realizado através de programação dinâmica, com o algoritmo “*Dynamic Time Warping*” (DTW). Uma especialização do algoritmo DTW, o casamento de Smith-Waterman, utilizado com frequência no alinhamento de sequência de proteínas, apresentou resultados mais eficientes no casamento deste tipo de descritor de arestas. Este casamento pode ser ilustrado através da Figura 7, onde cada cor representa o *matching* entre arestas nas duas imagens.

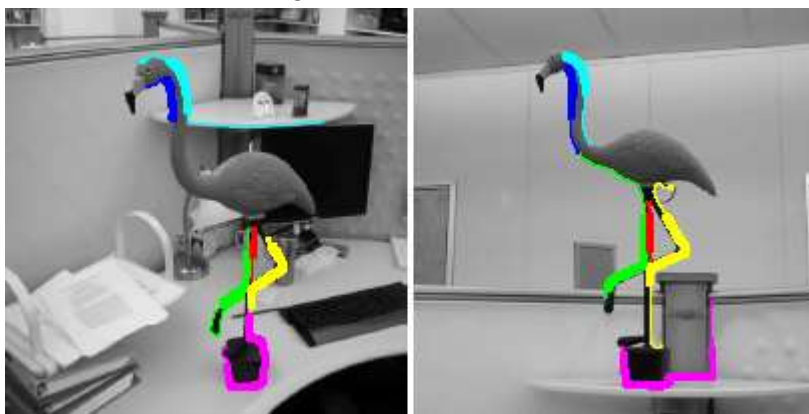


Figura 7. Casamento de arestas baseado em descritores [25].

O casamento de arestas de maneira global através de descritores possibilita o seu uso em imagens não consecutivas (*wide baseline*) e permite que sejam utilizados *keyframes* para a reconstrução de arestas utilizando uma versão modificada do algoritmo SfM. Estes casamentos não exigem uma inicialização como nas técnicas de rastreamento recursivo baseado em arestas, diminuindo a interação com o usuário e automatizando o processo de rastreamento como um todo.

3.2 Baseada em Fluxo Óptico

Métodos baseados em fluxo óptico têm como conceito base a exploração da informação temporal, de forma diferente dos métodos baseados em arestas que exploram as informações espaciais retiradas do modelo e da imagem da câmera. Esta informação temporal é extraída a partir do movimento relativo da projeção do objeto na imagem. Após a inicialização do sistema, que na maioria dos casos é feita manualmente, é capturado o fluxo óptico entre os quadros do vídeo e o método verifica que pontos pertencentes à projeção do modelo ainda se encontram rastreados na imagem.

A mudança de posição destes pontos ao longo do tempo é calculada utilizando algoritmos como o Kanade-Lucas [9]. Com a estimativa de movimento dos pontos é possível também estimar o movimento 3D realizado pela câmera.

Devido à utilização de integração numérica ao longo do tempo, o rastreamento 3D baseado no fluxo óptico apresenta mudanças suaves entre poses consecutivas. Outra vantagem é que a sua carga de processamento é moderada. No entanto, técnicas baseadas em fluxo óptico tendem a acumular erros provenientes das estimativas de poses a cada quadro, resultando em um desvio da pose correta da câmera. Estes métodos também não são robustos a mudanças de iluminação e movimentos rápidos de câmera, gerando erro no rastreamento e forçando uma reinicialização do sistema. A Figura 8 ilustra uma aplicação baseada em fluxo óptico rastreando uma face [5].



Figura 8. Aplicação de MAR baseada em fluxo óptico [5].

3.3 Baseada em Textura

Esta categoria de técnicas leva em conta as informações de textura extraídas da imagem. Estas técnicas podem ser divididas em duas subcategorias: técnicas baseadas em casamento de padrões [6] e em pontos de interesse [7]. As técnicas baseadas em casamento de padrões recuperam a pose da câmera através da descoberta do modelo de distorção que deve ser aplicado à imagem de referência que mapeia o real movimento de um objeto rígido. De acordo com taxonomia de MAR exibida na Figura 2, as técnicas baseadas em textura podem ser classificadas em baseadas em pontos de interesse e casamento de padrões, que serão detalhadas a seguir.

Nas técnicas baseadas em pontos de interesse também são utilizadas características locais (como no fluxo óptico), porém são levadas em consideração informações de textura para auxiliar a busca e o rastreamento.

A abordagem de casamento de padrões é baseada na recuperação de informação global, diferentemente das técnicas baseadas em extração de características. A informação global é dada por um conjunto de descritores ou por trechos (*patches*) da imagem, enquanto as características são pontos genéricos rastreados ao longo dos quadros. Estas informações globais tornam mais fáceis o tratamento de padrões e consequentemente o rastreamento.

Em suma (veja Figura 9), esta técnica estima os parâmetros de uma função de distorção que modifica um padrão para sobrepor em uma imagem, este padrão aparece com um ponto de vista diferente. De acordo com [9], este é o objetivo principal do algoritmo Kanade-Lucas (KL). Em [10], o autor mostra uma abordagem baseada no Jacobiano da função de distorção usada no algoritmo KL para realizar o rastreamento 2D. Contudo, podem acontecer alguns problemas na ocorrência de variação da iluminação e das oclusões parciais. Com relação ao rastreamento 3D, a abordagem com o Jacobiano apresenta algumas dificuldades e não alcança bons resultados.

Em [6], em oposição ao uso do inverso do Jacobiano, a variação das intensidades dos *pixels* foi aproximada utilizando hiperplanos, o que levou a um resultado melhor sem nenhum cálculo adicional. Utilizando a aproximação dos hiperplanos, [11] mostrou como tratar mudanças de iluminação, oclusões parciais e movimentos rápidos de câmera utilizando a normalização das mudanças de iluminação e um modelo estendido de movimento. A Figura 9 mostra uma cena aumentada simples onde uma imagem foi usada como um padrão.

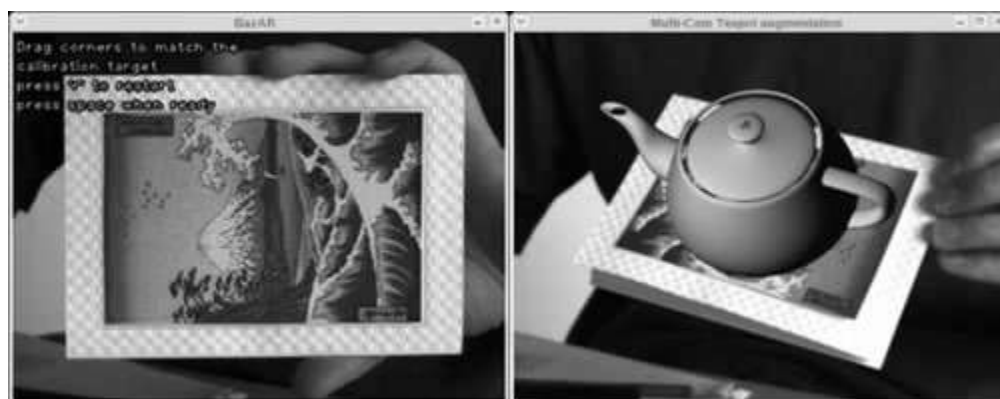


Figura 9. Rastreamento 3D com casamento de padrões. Imagem do padrão no centro do marcador (esquerda) e cena aumentada (direita) [12].

As técnicas baseadas em pontos de interesse utilizam características locais ao invés da busca global que caracterizam as técnicas baseadas em casamento de padrões. Desta forma, esta abordagem demonstra ser computacionalmente mais leve que as outras abordagens, além de lidar melhor com as mudanças de iluminação. Em [13] o autor menciona que na abordagem baseada em pontos de interesse é possível utilizar um *baseline* maior que o fluxo óptico.

De acordo com [7], as características locais podem ser selecionadas manualmente em estágio preliminar da execução da técnica. Por demandar um nível de expertise maior na seleção das características, o ponto forte da técnica descrita em [7] é a substituição da seleção manual por um algoritmo de seleção automática.

É possível identificar somente um subconjunto da imagem através do uso de um operador de interesse. Basicamente os operadores de interesse devem selecionar todos os pontos que possuam um determinado conjunto de características. Em [2] é possível encontrar um levantamento de vários destes operadores.

O rastreamento do objeto de interesse pode ser feito usando KL [9], e uma alternativa ao casamento é a utilização do rastreador Kanade-Lucas-Tomasi (KLT) [14]. Inicialmente é realizada uma fase de extração de características onde as que possuírem os maiores autovalores serão selecionadas. Em uma etapa posterior será realizada a fase de rastreamento baseada no algoritmo KL. Com esta abordagem é mais fácil de atingir uma continuidade no rastreamento, porém o KLT apresenta problemas de desempenho.

Outra técnica de pontos de interesse é baseada em rastreamento de planos [15] ao invés do rastreamento de modelos 3D complexos, veja Figura 10. A ideia principal desta técnica é explorar a homografia formada pelo plano em dois pontos de vista consecutivos. Esta homografia é computada a cada quadro utilizando o algoritmo RANdom SAMple Consensus (RANSAC) [29].

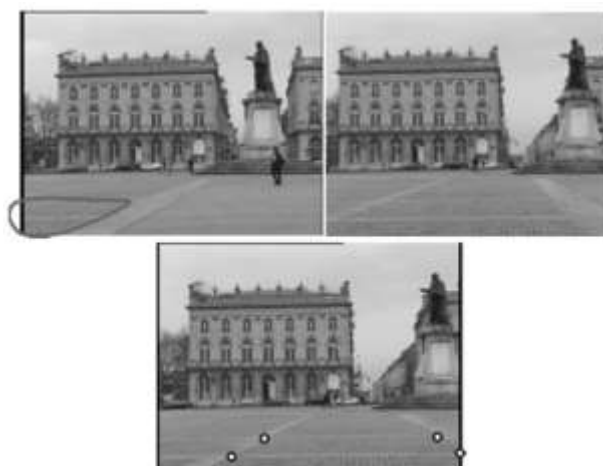


Figura 10. Passos do algoritmo: regiões manualmente selecionadas (superior esquerdo); detecção automática de características (superior direito); rastreamento de plano (inferior) [14].

Como não há um conhecimento prévio da cena, este método acumula erros durante a sua execução ocasionando poses instáveis e falhas de registro. Em [16] o conceito de *keyframes* é introduzido para resolver o problema da falha de registro. *Keyframes* são quadros da cena nos quais os pontos de interesse são previamente calculados. De posse desses quadros, cada frame capturado é comparado com o *keyframe* mais próximo. É possível perceber que esta abordagem não é escalável, uma vez que pode haver grandes diferenças entre os *keyframes* e os quadros capturados ocasionando uma falha de registro.

Em [16] o autor apresenta como utilizar uma abordagem híbrida, abrangendo *keyframes* e informações providas pelos quadros anteriores para reforçar a coerência temporal. Desta forma, o problema de rastreamento de modelos 3D pode ser visto como um problema de minimização conhecido como *Bundle Adjustment*. Em [17] é descrita uma abordagem simples para utilização de informações de arestas para reforçar a robustez do rastreador.

4 Técnicas Híbridas

Uma tendência na implementação de sistemas de MAR é o agrupamento de diferentes técnicas em um único sistema de rastreamento, sendo possível aproveitar o que de melhor cada tipo de técnica oferece. Em [21][22][23] os autores propõem um sistema de rastreamento, onde são integrados algoritmos classicamente utilizados para reconstrução 3D, o SfM [26] e o SLAM [27], em conjunto com técnicas de MAR baseadas em textura ou arestas.

Sistemas de MAR baseados em detecção (*Tracking by Detection* ou TDB) conseguem estimar com precisão a pose de câmera baseando-se na detecção de modelos presentes na cena, para os quais o sistema foi previamente treinado para reconhecer. Podemos destacar dois principais problemas dessa abordagem: as poses de câmera obtidas para cada quadro são independentes, de forma que a imprecisão numérica faz com que a visualização do rastreamento em tempo real fique tremida (*jitter*). No TDB também é necessário que o objeto que foi utilizado no treinamento esteja sempre visível nas imagens capturadas, ocasionando a perda do rastreamento nos quadros onde ele não estiver visível ou parcialmente visível.

Os algoritmos de SfM e SLAM não possuem esses problemas, pois dependem apenas da reconstrução inicial e do rastreamento de características espalhadas por toda a cena. O SfM utiliza as informações de textura da cena capturada, e não apenas de um objeto em particular, logo não existe a restrição de ter-se um determinado objeto visível para que o rastreamento possa ocorrer. Porém, nos algoritmos de SfM a necessidade de uma pose inicial, para que o rastreamento possa ser feito na escala correta é mandatória, e poderão acumular erros com consequente perda de precisão à medida que o rastreamento evolui temporalmente.

Em [22] é proposta uma abordagem que une os algoritmos de SfM com os de detecção baseada em modelo. Essa união pode ser realizada através do algoritmo de minimização de erro, chamado *Bundle Adjustment* (BA) [28]. O algoritmo de BA é utilizado em conjunto com o SfM para realizar a minimização do erro de reprojeção entre os pontos reconstruídos da cena e os pontos rastreados. Os autores propõem que sejam adicionadas algumas restrições no BA considerando as informações de rastreamento obtidas por técnicas de detecção baseadas em modelo. Dessa forma o SfM pode ser utilizado para aplicações de MAR reduzindo o acúmulo de erro, pois as poses obtidas pelo rastreamento baseado em detecção irão auxiliar na minimização do problema.

Em [21] o autor propõe uma extensão ao sistema proposto em [23] para aplicações de MAR em objetos sem textura (*textureless*), através da fusão do SfM com as técnicas de rastreamento baseadas em arestas. As estimativas de câmera são obtidas através do SfM. No entanto, o conhecimento parcial da estrutura da cena (treinamento prévio do algoritmo de rastreamento baseado em arestas, para o objeto a ser rastreado), permite que o acúmulo de erro natural do SfM possa ser minimizado e em contrapartida, consegue-se rastrear um modelo sem textura de forma robusta, pois quando o objeto de interesse não estiver visível ou parcialmente ocluído as estimativas do SfM, que levam em consideração todos os pontos de textura da cena, garantem que o rastreamento continue sendo efetuado.

5 Considerações Finais

Dentro do contexto do projeto, o cenário de utilização possui as seguintes características: tem-se o conhecimento parcial da cena, pois já se tem o modelo 3D da aeronave a ser rastreada; presença de muita reflexão especular o que normalmente dificulta o rastreamento a ser executado; e modelos texturizados para o rastreamento podem não existirem, já que cada empresa aérea utilizará um padrão de pintura diferente para as suas aeronaves.

Portanto, a utilização de uma técnica de rastreamento de modelos baseada em arestas é a escolha mais adequada para este cenário, conforme pode ser visto pela descrição das características deste tipo de técnica na seção 3.1. Porém, como este tipo de técnica não permite um movimento brusco de câmera, nem consegue manter o rastreamento quando o objeto estiver parcialmente ou totalmente ocluído na imagem, a utilização de uma abordagem híbrida com SfM deve apresentar resultados mais precisos, mesmo em um ambiente complexo como o cenário a ser utilizado neste projeto.

Dentre as abordagens híbridas disponíveis, a proposta em [21] apresenta segundo os seus autores um rastreamento robusto num ambiente similar ao cenário a ser utilizado neste projeto. Portanto, esta abordagem será utilizada para a implementação da primeira versão da POC.

6 Referências

- [1] V. Teichrieb, J. P. Lima, E. Apolinário, T. Farias, M. Bueno, J. Kelner, I. Santos. A survey of online monocular markless augmented reality. *International Journal of Modeling and Simulation for the Petroleum Industry*, Recife, v. 1, n. 1, p. 1-7, 2007. ISSN 1982-6028.
- [2] V. Lepetit, P. Fua. Monocular model-based 3d tracking of rigid objects. *Found. Trends. Comput. Graph. Vis.*, Now Publishers Inc., Hanover, MA, USA, v. 1, n. 1, p. 1–89, 2006. ISSN 1572-2740.
- [3] A. I. Comport et al. Real-time markerless tracking for augmented reality: The virtual visual servoing framework. *IEEE Transactions on Visualization and Computer Graphics*, IEEE Computer Society, Los Alamitos, CA, USA, v. 12, n. 4, p. 615–628, 2006. ISSN 1077-2626.
- [4] D. Koller, K. Danilidis, and H.-H. Nagel. 1993. Model-based object tracking in monocular image sequences of road traffic scenes. *Int. J. Comput. Vision* 10, 3, June 1993, 257-281.
- [5] S. Basu, I. Essa, and A. Pentland. 1996. Motion Regularization for Model-Based Head Tracking. In *Proceedings of the International Conference on Pattern Recognition (ICPR '96) Volume III-Volume 7276 - Volume 7276* (ICPR '96), Vol. 7276. IEEE Computer Society, Washington, DC, USA.
- [6] F. Jurie, M. Dhome, A simple and efficient template matching algorithm," *Computer Vision*, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on Computer Vision, vol.2, no., pp.544,549 vol.2, 2001.
- [7] M. Uenohara, T. Kanade. Vision-based object registration for realtime image overlay. *Journal of Cognitive Neuroscience*, v. 3, n. 1, p. 71–86, 1991.
- [8] M. Armstrong, A. Zisserman: Robust object tracking, *Proc. of 2nd ACCV* pp.1-58–62, 1995.
- [9] B. D. Lucas, T. Kanade. An iterative image registration technique with an application to stereo vision (DARPA). In: *Proceedings of the 1981 DARPA Image Understanding Workshop*, 1981. p. 121–130.
- [10] G. D. Hager, P. N. Belhumeur. Efficient region tracking with parametric models of geometry and illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 20, n. 10, p. 1025–1039, 1998.
- [11] F. Dellaert, R. Collins. Fast image-based tracking by selective pixel integration. *Proceedings of the ICCV Workshop on Frame-Rate Vision*. 1999.
- [12] V. Lepetit, P. Laguerre, P. Fua. Randomized trees for real-time keypoint recognition. In: *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2*. Washington, DC, USA: IEEE Computer Society, 2005. p. 775–781. ISBN 0-7695-2372-2.
- [13] S. Ravela et al. Adaptive tracking and model registration across distinct aspects. *Intelligent Robots and Systems 95: Human Robot Interaction and Cooperative Robots*, *Proceedings. 1995 IEEE/RSJ International Conference on*. Vol. 1. IEEE, 1995.

- [14] J. Shi; C. Tomasi. Good features to track. *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*. IEEE, 1994.
- [15] G. Simon, M. Berger. Pose estimation for planar structures. *IEEE Computer Graphics and Applications*, v. 22, n. 6, p. 46–53, 2002.
- [16] L. Vacchetti, V. Lepetit. Stable real-time 3d tracking using online and offline information. *IEEE Trans. Pattern Anal. Mach. Intell.*, IEEE Computer Society, Washington, DC, USA, v. 26, n. 10, p. 1385–1391, 2004. ISSN 0162-8828.
- [17] L. Vacchetti, V. Lepetit, P. Fua. Combining edge and texture information for real-time accurate 3d camera tracking. In: *ISMAR '04: Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'04)*. Washington, DC, USA: IEEE Computer Society, 2004. p. 48–57. ISBN 0-7695-2191-6.
- [18] D. Lowe. Object recognition from local scale-invariant features. *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*. Vol. 2. Ieee, 1999.
- [19] E. Rosten, T. Drummond. Fusing points and lines for high performance tracking. *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*. Vol. 2. IEEE, 2005.
- [20] K. Mikolajczyk, C. Schmid. A performance evaluation of local descriptors, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10, 27, pp 1615--1630, 2005.
- [21] V. Gay-Bellile, S. Bourgeois, M. Tamaazousti, S. N. Collette, S. Knodel. A mobile markerless Augmented Reality system for the automotive field. *ISMAR*, 2012.
- [22] M. Tamaazousti, V. Gay-Bellile, S. N. Collette, S. Bourgeois, M. Dhome. Nonlinear refinement of structure from motion reconstruction by taking advantage of a partial knowledge of the environment. *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011.
- [23] M. Tamaazousti, V. Gay-Bellile, S. N. Collette, S. Bourgeois, M. Dhome. Real-time accurate localization in a partially known environment: Application to augmented reality on 3d objects. In *International workshop on AR/MR registration, tracking and benchmarking*, 2011.
- [24] Y. Tsin, Y. Genc, Ying Zhu; V. Ramesh. Learn to Track Edges. *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on* , vol., no., pp.1,8, 14-21 Oct. 2007.
- [25] J. Meltzer, S. Soatto. Edge descriptors for robust wide-baseline correspondence. *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, vol., no., pp.1,8, 23-28 June 2008.
- [26] D. Nistér. Preemptive RANSAC for live structure and motion estimation. In *Proceedings of the Ninth IEEE International Conference on Computer Vision, ICCV '03*, Washington, DC, USA, 2003. IEEE Computer Society.
- [27] A. J. Davison, I. D. Reid, N. D. Molton, O. Stasse. MonoSLAM: Real-Time single camera SLAM. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(6):1052-1067, June 2007.

- [28] B. Triggs, P. F. McLauchlan, R. I. Hartley, A. W. Fitzgibbon. Bundle adjustment - a modern synthesis. In Proceedings of the International Workshop on Vision Algorithms: Theory and Practice, ICCV '99, pages 298-372, London, UK, 2000. Springer-Verlag.
- [29] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM, 24:381_395, June 1981.