



INTEGRAÇÃO DE SENSORES PARA LOCALIZAÇÃO E MAPEAMENTO SIMULTÂNEOS EM FRUTICULTURA

Patric **Moreto**¹; Luciano Vieira **Koenigkan**²; Thiago Teixeira **Santos**³

Nº 21607

RESUMO – A automação na agricultura consiste na utilização de tecnologias em processos operacionais na produção agrícola, como no monitoramento automatizado de safra, que se faz necessário por possibilitar a maximização do rendimento da colheita enquanto diminui custos. Visando o uso da robótica para tais tarefas, o sistema de localização e mapeamento simultâneos é empregado para auxiliar a navegação segura dos robôs no ambiente agrícola enquanto é realizado um mapeamento tridimensional. Contudo, métodos baseados em sensores visuais tendem a acumular erros ao longo do tempo, comprometendo seu uso. À vista disso, neste estudo preliminar descrevemos como a integração de sensores, através de grafo de fatores, é capaz de otimizar os resultados, obtendo modelos fidedignos de talhões em vinhedos. A partir da biblioteca GTSAM, está sendo desenvolvida uma solução que utiliza dados de câmera, de GPS e de IMU simultaneamente, e também faz uso de otimizadores capazes de diminuir erros acumulados. Espera-se que a integração dos sensores mencionados seja capaz de amenizar o problema relatado, ao ponto do sistema ser considerado livre da influência de tal distúrbio.

Palavras-chaves: SLAM, grafo de fatores, odometria visual, viticultura, integração de sensores.

1 Autor, Bolsista CNPq (PIBIC): Graduação em Engenharia Elétrica, FEEC / Unicamp, Campinas-SP; ptmoreto@gmail.com.

2 Co-orientador: Analista da Embrapa Informática Agropecuária, Campinas-SP; luciano.vieira@embrapa.br.

3 Orientador: Pesquisador da Embrapa Informática Agropecuária, Campinas-SP; thiago.santos@embrapa.br.



ABSTRACT – *Automation in agriculture consists in utilizing technologies in operational processes in agricultural production, such as automated crop monitoring, which is necessary because it makes it possible to maximize crop yields while reducing costs. Seeking the use of robotics for such tasks, the simultaneous localization and mapping system is employed to assist the safety navigation of robots inside an agricultural environment while a three-dimensional mapping is performed. However, methods based on visual sensors tend to accumulate errors over time, compromising its use. Considering this, we describe in this preliminary study how sensor integration, through a factors graph, is able to optimize the results, obtaining reliable models of plots in vineyards. Employing the GTSAM library, a solution is being developed, that uses camera, GPS and IMU data simultaneously, and also makes use of optimizers capable of reducing accumulated errors. It is expected that the integration of the mentioned sensors will be capable of softening the reported problem, to the point that the system is considered free from the influence of such disturbance.*

Keywords: SLAM, factor graph, visual odometry, viticulture, sensor integration.

1. INTRODUÇÃO

A automação na agricultura consiste na utilização de tecnologias em processos operacionais na produção agrícola (EMBRAPA, 2021). Conforme relatado por Dong et al. (2017), o monitoramento de culturas agrícolas é realizado por operadores humanos, um processo caro e laborioso, de modo que o monitoramento autônomo por máquina teria grande impacto no setor. Uma alternativa consiste em mapeamentos tridimensionais da cultura, capazes de prover informações sobre seu desenvolvimento (AGUIAR et al., 2020), viabilizando diversas aplicações possíveis, como a estimação do volume de cachos de uva com alto índice de precisão (SANTOS et al., 2017).

Nesse contexto, a utilização da robótica para desempenhar tais tarefas é um campo em ascensão. Para tanto, busca-se avanços na navegação segura dos robôs em ambientes agrícolas e, em concordância com Aguiar et al. (2020), é possível afirmar que esses avanços estão associados a localizar o agente robótico no espaço. Em razão do mapeamento tridimensional ser capaz de gerar referências posicionais que auxiliam o processo de localização, a união dos conceitos de mapear tridimensionalmente ambientes agrícolas e, ao mesmo tempo, localizar o robô dentro desse ambiente é desejável e será desenvolvida neste trabalho.



Soluções conhecidas demonstram limitações, seja na localização ou no mapeamento. Aguiar et al. (2020) mencionam a escassez de trabalhos que envolvam localização 3D, com a grande maioria optando pela localização 2D. Também é citado que projetos que utilizam odometria visual (VO), ou seja, determinam a posição e orientação do agente através da análise de imagens (SCARAMUZZA; FRAUNDORFER, 2011), não realizam o mapeamento do ambiente, além de não serem escaláveis, dado que existe o acúmulo de erros ao longo do tempo, por conta desse método ser baseado em medições relativas, problema conhecido como *drift*. O acúmulo de erros pode causar grandes distorções no trajeto registrado, impossibilitando a utilização dessa ferramenta em larga escala.

O projeto a ser desenvolvido busca preencher lacunas não comumente exploradas, na qual trataremos do mapeamento tridimensional realizado a partir de ferramentas como a odometria visual em junção a dados de outros sensores, além da realização simultânea da localização tridimensional do robô dentro desse ambiente mapeado. Nosso caso de estudo é a viticultura, um exemplo de cultura frutícola composta de plantas frutíferas agrupadas em corredores. A abordagem escolhida é a utilização de *grafo de fatores* (INDELMAN et al., 2013), modelo matemático que representa o problema de fusão de sensores e estimação de variáveis na forma de um grafo que guia o processo de otimização.

Este trabalho preliminar busca revisar conceitos fundamentais para o SLAM realizado a partir de dados de câmeras. Também será detalhada a sua abordagem a partir dos grafos de fatores e as possibilidades que oferece, permitindo a integração entre sensores de tal forma que suas informações se complementem, inclusive suprindo possíveis indisponibilidades das fontes de dados. Em específico, será visto a integração de câmeras, sistemas de posicionamento global (GPS) e unidades de medição inercial (IMU) dotadas de sensores capazes de prover dados quanto à aceleração linear, velocidade angular e direção obtida a partir de um magnetômetro.

2. MATERIAL E MÉTODOS

2.1. SLAM

A localização e mapeamento simultâneos, ou SLAM, consiste em adquirir um mapa espacial de um ambiente enquanto, simultaneamente, localiza-se um agente móvel nesse modelo projetado (STACHNISS; LEONARD; THRUN, 2016). O mapa é uma interpretação multidimensional concebida pela integração de diversos dados, como posição e velocidade, originários de diferentes sensores que compõe o robô, e que possui o intuito de fornecer informações a respeito do



respectivo ambiente, podendo, então, estabelecer referências para o auxílio do processo de localização (AGUIAR et al., 2020).

É possível definir o problema do SLAM, segundo Stachniss, Leonard e Thrun (2016), como um robô que percorre um ambiente desconhecido, partindo de uma posição inicial e com deslocamento que possui incertezas, dado que os sensores acoplados são ruidosos. Assim, conforme o agente se movimenta, maior é a incerteza sobre sua real posição. Dessa forma, a construção do mapa passa a lidar com incertezas, surgindo a ideia de um mapa estocástico e, portanto, SLAM se define como um problema probabilístico.

Como apresentado por Stachniss, Leonard e Thrun (2016), seja t o tempo e x_t a localização do robô. O trajeto traçado pelo robô é dado por:

$$X_T = \{x_0, x_1, x_2, \dots, x_T\} \quad (1)$$

Sendo T o instante final. Usualmente, a localização inicial x_0 é conhecida e serve como um ponto de referência, sendo as demais posições determinadas probabilisticamente. Considere u_t como a informação de odometria, isto é, a distância percorrida entre duas localizações seguidas, x_{t-1} e x_t . O deslocamento relativo traçado pelo robô é dado por:

$$U_T = \{u_1, u_2, u_3, \dots, u_T\} \quad (2)$$

A ressalva feita por Stachniss, Leonard e Thrun (2016) de que, caso U_T fosse livre de incertezas, ele seria suficiente para a determinação do trajeto traçado dado x_0 , é fundamental para a compreensão da razão de SLAM ser um problema probabilístico.

Sendo m o real mapa do ambiente onde o robô está se movimentando, esse mapa descreve a localização dos pontos de referência existentes. Um ponto de referência é um elemento que possui características observáveis por um ou mais sensores que constituem o robô. Conforme será abordado no decorrer do trabalho, os pontos de referência desempenham papel fundamental na estimação de posição. Assim, o robô realiza medições que relacionam características de m a sua localização x_t , e assumindo, sem perda de generalidade, que é realizada uma medição z_t a cada ponto no tempo, a sequência de medições é dada por:

$$Z_T = \{z_1, z_2, z_3, \dots, z_T\} \quad (3)$$

Finalmente, define-se o problema do SLAM como sendo determinar um modelo do ambiente m e o trajeto X_T a partir da odometria U_T e das medições Z_T . Isso é, estimar *a posteriori*:

$$p(X_T, m | U_T, Z_T) \quad (4)$$



2.2. Visual SLAM

Caso um dos sensores que fornecem dados seja uma câmera, a realização da localização e mapeamento simultâneos utilizando imagens configura-se em um processo conhecido como Visual SLAM. Conforme afirma Stachniss, Leonard e Thrun (2016), sensores visuais provêm informações com grande resolução para a construção de modelos tridimensionais de maior fidelidade.

Uma possível alternativa para a obtenção de informações para o Visual SLAM é através da odometria visual, um processo capaz de determinar a posição e orientação do agente através da análise de imagens (SCARAMUZZA; FRAUNDORFER, 2011). A partir da posição e orientação, é possível realizar uma estimativa do movimento da câmera entre a imagem atual e a anterior de maneira incremental, ou seja, sempre estimando o deslocamento relativo entre apenas duas imagens consecutivas, conforme elas foram capturadas. Segundo Scaramuzza e Fraundorfer (2011), por conta desse método de progressão incremental, os erros introduzidos a cada estimativa de movimento realizada (imagem por imagem) irão se acumular no decorrer do tempo, problema conhecido como *drift*.

Tal como foi abordado na definição do problema do SLAM, para o Visual SLAM também ocorre a detecção da localização de um ponto de referência. Pelo sensor em questão ser uma câmera, o ponto de referência é visual, possuindo características destacáveis por algoritmos de detecção.

A incorporação das informações fornecidas por um sensor visual em um processo de Visual SLAM depende da configuração das câmeras, conforme informa Indelman et al. (2013) e, baseado em seu desenvolvimento proposto, para o caso de um modelo de câmera monocular *pinhole* com uma matriz de calibração K conhecida, é definida uma função não linear π , capaz de realizar a projeção de um ponto tridimensional no mapa para sua observação bidimensional na imagem, dada por:

$$\pi(x, l) = K[R \ t]l \quad (5)$$

Sendo l a localização do ponto de referência tridimensional detectado, R a matriz de rotação e t o vetor de translação obtidos da localização x do robô. Tanto R quanto t representam uma transformação entre a orientação da câmera e as coordenadas globais, de modo que l é transformado para o sistema de coordenadas de câmera antes de ser projetado pela matriz K .

A função π , também conhecida como função de medição por Stachniss, Leonard e Thrun (2016), é uma função de projeção de 3D para 2D, uma projeção esperada pelo modelo. Entretanto,



como se tratam de medições ruidosas, é necessário incluir a diferença entre a projeção observada z e o modelo $\pi(x, l)$, conforme Equação 5, buscando a diminuição do erro de projeção a partir de otimizações. Assim, utilizando a medição z e a projeção $\pi(x, l)$ do ponto l na imagem com câmera posicionada em x , o erro de reprojeção é definido na Equação 6:

$$e_{proj} = z - \pi(x, l) \quad (6)$$

Um sistema de Visual SLAM pode ser dividido em duas partes principais, denominadas *frontend* e *backend*. A primeira é, segundo Dong et al. (2017), responsável por processar imagens à procura de pontos de referência visuais com as quais é possível traçar correspondências entre imagens. O *frontend* do Visual SLAM possui fundamentos compartilhados com um processo de VO, conforme demonstrado por Scaramuzza e Fraundorfer (2011), em que são processadas imagens individualmente, dentro de uma sequência, e, a cada novo processamento, são detectadas e estabelecidas correspondências a partir de características 2D com imagens anteriores. Essas características 2D são ditas correspondentes, pois são reprojeções das mesmas características 3D entre diferentes imagens.

A segunda parte do sistema de Visual SLAM, o *backend*, é a estimação da posição da câmera e dos pontos de referência utilizando as correspondências traçadas anteriormente, da etapa de *frontend*, e informações de outros sensores (DONG et al., 2017), caso não esteja sendo utilizado apenas uma câmera como fonte de dados. Com a utilização do VO, a estimação da posição é realizada ao ocorrer a concatenação de todos os deslocamentos estimados (SCARAMUZZA; FRAUNDORFER, 2011). Durante o *backend*, com base nas otimizações para determinação do modelo do ambiente, que buscam minimizar o erro de projeção, apresentado pela Equação 6, é possível refinar a posição dos pontos de referência visuais e a trajetória do agente.

2.3. Grafo de fatores

Buscando minimizar os efeitos do *drift*, é realizada a junção de dados de diferentes sensores para a criação de modelos fidedignos, procedimento conhecido como integração de sensores. Uma possível abordagem para o problema da fusão de dados ocorre através da utilização do conceito de grafo de fatores, que, segundo Indelman et al. (2013), é um modelo gráfico que consiste em conectar variáveis desconhecidas a fatores definidos por medições realizadas por sensores. Segundo Dellaert (2021), arestas entre fatores e variáveis indicam que o fator em questão depende apenas daquelas variáveis que ele está conectado. Logo, o fator representa uma função sobre a variável que ele está conectado.

Formalmente, um grafo de fatores é definido como um grafo bipartido que possui dois tipos de nós, sendo eles os fatores e as variáveis (DELLAERT, 2021). Arestas podem apenas existir entre os diferentes tipos de nós, ou seja, entre fatores e variáveis (INDELMAN, et al., 2013). Cada fator representa uma função erro que deve ser minimizada, logo, condiz com a ideia apresentada para o erro de projeção no desenvolvimento da Equação 6. Dessa forma, conforme Indelman et al. (2013), define-se uma estrutura padrão para a representação de um fator f_i :

$$f_i(v_i) = d(e_i(v_i, z_t)) \quad (7)$$

Sendo o operador $d()$ uma função custo, e_i uma determinada função erro, v_i uma determinada variável e z_t uma medição realizada em algum tempo t .

Através da Figura 1, é possível observar uma representação gráfica de grafos de fatores. A Figura 1 (a) representa um grafo formado a partir de dois tipos de variáveis, sendo elas a posição do robô x_i e a calibração da IMU c_i , e com fatores de GPS f^{GPS} , de IMU equivalente f^{Equiv} , de calibração f^{bias} e de prior f^{Prior} . A Figura 1 (b) contém todos os elementos já mencionados e é acrescida da variável ponto de referência l_i e do fator de projeção f^{Proj} . O fator de prior é composto de informações disponíveis *a priori*, como uma posição inicial do agente. Demais fatores referidos serão apresentados a seguir.

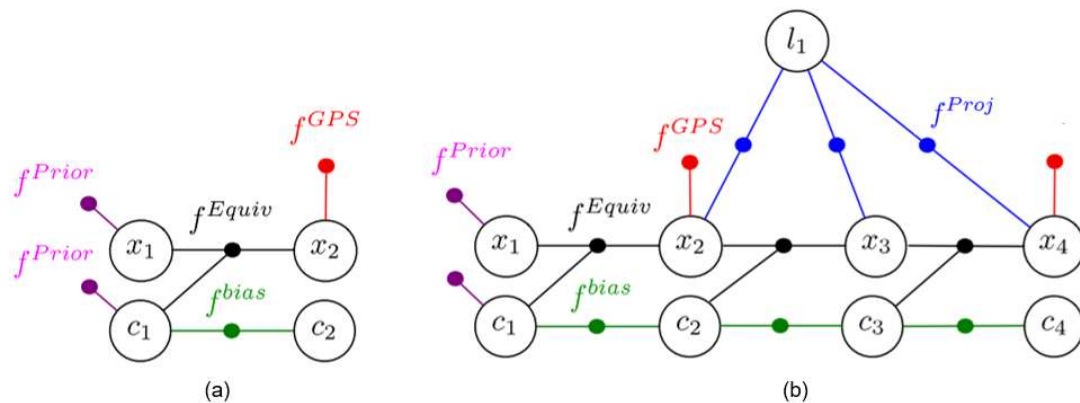


Figura 1. Representações gráficas de grafos de fatores, a partir das variáveis x_i e c_i (a) e acrescida de l_i e f^{Proj} (b), em que $i = 1, 2, 3$ ou 4 .

Fonte: Adaptado de Indelman et al. (2013).

2.3.1. Fator de projeção

Um fator de projeção decorre da incorporação de um sensor visual, como uma câmera. Considerando o objetivo de mapear o ambiente para aplicações como monitoramento de safra, tal qual foi discorrido no tópico de Visual SLAM, sensores visuais são de particular interesse para a



fruticultura, pois são capazes de viabilizar projetos como estimação do volume de cachos de uva com alto índice de precisão (SANTOS et al., 2017).

A formulação desse fator dá continuidade ao desenvolvimento das Equações 5 e 6, utilizando as mesmas ideias e notações já apresentadas. Utilizando como base o erro de projeção, apresentado pela Equação 6, e utilizando a estrutura padrão para a representação de um fator, conforme Equação 7, segundo Indelman et al. (2013), define-se o fator de projeção f^{PROJ} :

$$f^{PROJ}(x, l) = d(e_{proj}) = d(z - \pi(x, l)) \quad (8)$$

2.3.2. Fator de GPS

A utilização de um GPS como um sensor capaz de indicar a localização do agente é possível no ambiente agrícola, por se tratar de um local aberto, mas a incerteza associada deve ser devidamente tratada. Conforme Indelman et al. (2013), uma medição de GPS z_k^{GPS} em determinado tempo t_k é equacionada como:

$$z_k^{GPS} = h^{GPS}(x_k) + n^{GPS} \quad (9)$$

Sendo n^{GPS} o ruído na medição, h^{GPS} uma função de medição do GPS, uma função não linear que realiza uma previsão da medição do sensor, dado a variável x_k , que representa a localização do robô. A função h^{GPS} é a equivalente da função π , da Equação 5, para o caso do GPS. Assim, é possível definir o fator unário (pois depende apenas de x_k) de GPS f^{GPS} :

$$f^{GPS}(x_k) = d(z_k^{GPS} - h^{GPS}(x_k)) \quad (10)$$

2.3.3. Fator de IMU

Uma IMU como possível fonte de dados para sistemas de SLAM é uma ferramenta já bastante difundida. Segundo Indelman et al. (2013), a localização atual do robô, dada pela variável x_{k+1} , pode ser prevista a partir da posição anterior do agente x_k , sendo descrita conceitualmente da seguinte maneira:

$$x_{k+1} = h(x_k, c_k, z_k^{IMU}) \quad (11)$$

Sendo z_k^{IMU} as medições realizadas pela IMU, como velocidade angular e aceleração linear, em determinado tempo t_k entre as posições x_k e x_{k+1} , além de c_k representar os parâmetros de calibração da IMU, utilizados para correção das medições. No caso geral, a propagação de c_k através do tempo t é descrita por algum modelo não linear próprio, sendo representada como:



$$c_{k+1} = g(c_k) \quad (12)$$

A diferença entre a previsão realizada pela Equação 11 e a estimativa real de x_{k+1} é a função erro, dada pelo fator f^{IMU} :

$$f^{IMU}(x_{k+1}, x_k, c_k) = d(x_{k+1} - h(x_k, c_k, z_k^{IMU})) \quad (13)$$

De maneira similar, é possível definir a função erro do parâmetro de calibração da IMU, dada pelo fator f^{BIAS} :

$$f^{BIAS}(c_{k+1}, c_k) = d(c_{k+1} - g(c_k)) \quad (14)$$

A Equação 13 e a Equação 14 são fatores que conectam duas variáveis diferentes entre si a seus respectivos estados futuros, conforme Indelman et al. (2013). Dessa forma, o fator descrito na Equação 13 conecta a variável x_k à x_{k+1} , enquanto a variável c_k é conectada à c_{k+1} através da Equação 14.

Entretanto, na prática, a frequência que a IMU realiza medições geralmente é muito maior do que a frequência que o GPS e a câmera realizam. Desse modo, foi desenvolvida uma ideia de pré-integração dos dados da IMU, que irá fundir diversas medições em um único fator bipartido (DELLAERT, 2021), a ser inserido no grafo de fatores em uma frequência semelhante a outros sensores. Com isso, são integradas consecutivas medições da IMU entre dois tempos t_i e t_j não consecutivos em uma componente $\Delta x_{i \rightarrow j}$, responsável por relacionar as localizações do robô entre os tempos t_i e t_j (INDELMAN et al., 2013). Como a calibração não sofre grandes mudanças no decorrer do tempo, é possível utilizar apenas c_i para reescrever a Equação 13 da seguinte maneira:

$$f^{Equiv}(x_j, x_i, c_i) = d(x_j - h^{Equiv}(x_i, c_i, \Delta x_{i \rightarrow j})) \quad (15)$$

Sendo a função erro agora referenciada como um fator equivalente f^{Equiv} e h^{Equiv} é uma função não linear que realiza uma previsão da posição x_j dado x_i , c_i e $\Delta x_{i \rightarrow j}$.

2.4. GTSAM

Como forma de facilitar a aplicação dos conceitos de Visual SLAM e grafo de fatores, é utilizada a biblioteca de código aberto GTSAM (*Georgia Tech Smoothing and Mapping*), desenvolvida em linguagem C++, com suporte para Python, capaz de implementar fusão de dados para a área da robótica e aplicações como SLAM e VO utilizando grafos de fatores (GTSAM..., 2021). É importante ressaltar que o GTSAM necessita de uma inicialização prévia, realizada em um ambiente externo ao *software*, equivalente à etapa de *frontend* do Visual SLAM. Logo, é possível



afirmar que essa biblioteca realiza a parte de *backend* do sistema: sua função é estimar as variáveis x_i e que I_j que minimizam o somatório das funções expressas nos fatores.

2.5. CONJUNTO DE DADOS

Os dados a serem utilizados nesse trabalho foram coletados na Vinícola Tomasetto. Foram obtidas 1.451 imagens através de uma câmera ZED (Stereolabs, Inc.), na qualidade HD (1280 por 720 pixels). Também foram utilizados dois receptores GPS Reach M+ (Emlid, Inc.) para obtermos 26.493 medições de localização, na qual um atuava como estação base e o segundo era deslocado em conjunto ao agente e demais sensores, recebendo correções da base, configurando em um sistema típico de GPS-RTK. Uma IMU também foi incorporada e foram realizadas 19.372 medições.

3. RESULTADOS E DISCUSSÃO

Através das ferramentas discutidas e do conjunto de dados apresentado, está sendo desenvolvido um projeto de localização e mapeamento simultâneos em viticultura, a partir da integração de dados de câmera, IMU e GPS.

Por meio das informações extraídas de um GPS-RTK, é possível obter o trajeto do agente durante o sensoriamento. A Figura 2 (a) exemplifica parte do percurso realizado, utilizando apenas dados do RTK, e exibe em 2D os valores de latitude e longitude obtidos.

Conforme citado anteriormente, é necessária a realização do *frontend* do Visual SLAM em um ambiente externo ao GTSAM. Dessa forma, utilizando as imagens capturadas como fonte de dados, foi empregado o *software* COLMAP, responsável por localizar pontos de referência visuais para traçar correspondências entre imagens. A partir desse *frontend*, foram extraídas informações para a realização de um processo de odometria visual, em que a Figura 2 (b) apresenta as poses obtidas (exibidas aqui em 2D, para facilitar a visualização). O intervalo do caminho selecionado é o mesmo para ambas, em que a diferença entre figuras evidencia o efeito do *drift* em processos exclusivamente visuais.

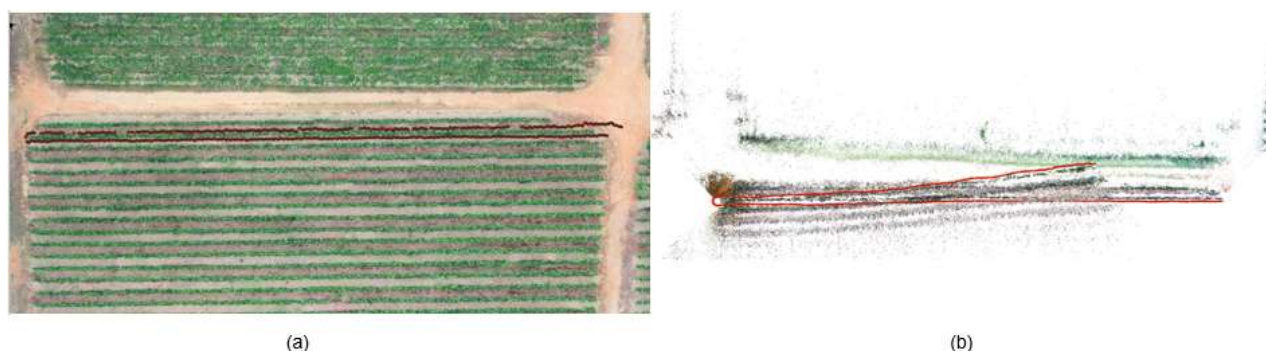


Figura 2. Trajetos destacados em vermelho são obtidos por um agente robótico se movendo em talhões em vinhedo, sendo (a) através do GPS-RTK e (b) através da odometria visual.

Como forma de exemplificar o procedimento de encontrar correspondências entre imagens para realização do VO, duas capturas consecutivas realizadas na viticultura são selecionadas e, após o COLMAP ter executado a rotina de *frontend*, são encontradas correspondências entre as imagens, conforme indica a Figura 3. Essas relações estabelecidas são as informações necessárias para a estimação do deslocamento que ocorreu entre as posições da primeira e segunda imagem.

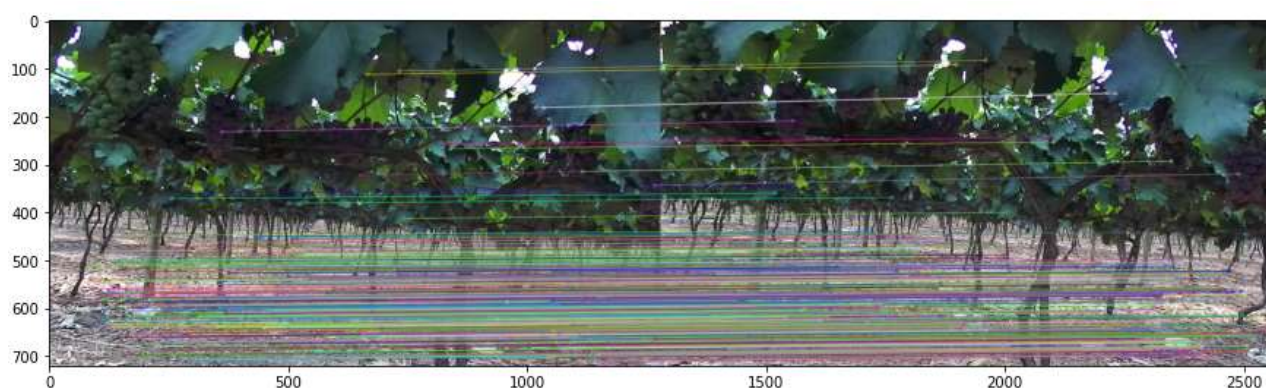


Figura 3. Correspondências visuais encontradas entre as duas imagens a partir de características 2D para a estimação do deslocamento.

Com isso, são utilizados algoritmos de estimação de movimento, como o Algoritmo 3 exposto por Scaramuzza e Fraundorfer (2011) e o algoritmo apresentado por Nistér (2003), para obtenção da odometria visual.

Por se tratar de um trabalho preliminar, são necessários avanços e mais experimentos para a apresentação de resultados concludentes a respeito das técnicas aplicadas. Atualmente, através da biblioteca GTSAM, o projeto busca a implementação de fatores de IMU e fatores equivalentes. Com isso, será possível a criação de grafos que utilizem fatores projetivos, de GPS e de IMU simultaneamente. Também busca-se o uso de otimizadores capazes de diminuir erros acumulados.

4. CONCLUSÃO



Em decorrência da crescente necessidade de monitoramento automatizado de safra, soluções na área da robótica estão sendo desenvolvidas. A utilização do sistema de localização e mapeamento simultâneos se estabelece tanto como uma forma de auxiliar a movimentação de um robô, quanto para providenciar informações sobre a cultura. Dada a dificuldade na aplicação de sensores visuais nesse procedimento, ocasionada pelo *drift*, busca-se a integração de sensores como uma forma de otimizar resultados obtidos a partir de diferentes dados fornecidos. Uma ferramenta que propicia tal conectividade de informações é o grafo de fatores.

Seguem em execução experimentos que utilizam Visual SLAM e grafo de fatores para a união de dados de câmera, IMU e GPS para operação em fruticultura. A partir do *software* COLMAP, foram detectados pontos de referência visuais para definição de correspondências entre imagens, possibilitando a realização de odometria visual através de algoritmos. Utilizando a biblioteca GTSAM, iremos implementar grafos que utilizem fatores projetivos, de GPS e de IMU simultaneamente, permitindo a geração de mapas tridimensionais acurados dos talhões.

5. AGRADECIMENTOS

Patric Moreto agradece ao programa PIBIC e ao CNPq pela bolsa de Iniciação Científica concedida; aos orientadores Thiago e Luciano pelos ensinamentos, incentivos e paciência. Thiago Santos agradece à FAPESP pelo financiamento do projeto *Agricultura ciente de ambiente: raciocínio sobre estrutura tridimensional no campo de cultivo (AACr3)*, processo 17/19282-7, ao qual o presente trabalho se insere.

6. REFERÊNCIAS

AGUIAR, A. S.; SANTOS, F. N. dos; CUNHA, J. B.; SOBREIRA, H.; SOUSA, A. J. Localization and mapping for robots in agriculture and forestry: a survey. **Robotics**, v. 9, n. 4, p. 1-23, 2020. DOI: doi.org/10.3390/robotics9040097.

DELLAERT, F. Factor graphs: exploiting structure in robotics. **Annual Review of Control, Robotics, and Autonomous Systems**, v. 4, p. 141-166, 2021. DOI: 10.1146/annurev-control-061520-010504.

DONG, J. L.; BURNHAM, J. G.; BOOTS, B.; RAINS, G. C.; DELLAERT, F. 4D crop monitoring: spatio-temporal reconstruction for agriculture. In: IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, 2017, Singapore. **Proceedings...** Piscataway: IEEE, 2017. p. 3878-3885. DOI: 10.1109/ICRA.2017.7989447.

EMBRAPA. **Automação e agricultura de precisão**. Disponível em: <<https://www.embrapa.br/tema-mecanizacao-e-agricultura-de-precisao/nota-tecnica>>. Acesso em: 24 jun. 2021.

GTSAM 4.0: factor graphs for sensor fusion in robotics. Disponível em: <<https://gtsam.org/>>. Acesso em: 28 jun. 2021.



15º Congresso Interinstitucional de Iniciação Científica – CIIC 2021
01 a 02 de setembro de 2021
ISBN 978-65-994972-0-9

INDELMAN, V.; WILLIAMS, S.; KAESS, M.; DELLAERT, F. Information fusion in navigation systems via factor graph based incremental smoothing. **Robotics and Autonomous Systems**, v. 61, n. 8, p. 721-738, 2013. DOI: 10.1016/j.robot.2013.05.001.

NISTÉR, D. An efficient solution to the five-point relative pose problem. In: IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2003, Madison. **Proceedings...** Los Alamitos: IEEE, 2003. Não paginado. DOI: 10.1109/CVPR.2003.1211470.

SANTOS, T. T.; BASSOI, L. H.; OLDONI, H.; MARTINS, R. L. Automatic grape bunch detection in vineyards based on affordable 3D phenotyping using a consumer webcam. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 11., 2017, Campinas. **Ciência de dados na era da agricultura digital: anais**. Campinas: Editora da Unicamp: Embrapa Informática Agropecuária, 2017. p. 89-98. SBIAgro 2017. Disponível em: <https://www.alice.cnptia.embrapa.br/alice/bitstream/doc/1083291/1/AutomaticgrapeSBIAgro.pdf>. Acesso em: 28 jun. 2021.

SCARAMUZZA, D.; FRAUNDORFER, F. **Visual odometry**: part I: the first 30 years and fundamentals. **Robotics & Automation Magazine**, v. 18, n. 4, p. 80-92, 2011. DOI: 10.1109/MRA.2011.943233.

STACHNISS, C.; LEONARD, J. J.; THRUN, S. Simultaneous localization and mapping. In: SICILIANO B.; KHATIB O. (Ed.). **Springer Handbook of Robotics**. 2. ed. Berlin: Springer, 2016. p. 1153-1175.